International Journal of Innovative Research in Information Security (IJIRIS) ISSN: 2349-7017(O)
Volume 2 Issue 5 (May 2015) ISSN: 2349-7009 (P)
www.ijiris.com

STUDY OF VESPA (VEHICULAR EVENT SHARING WITH A MOBILE P2P ARCHITECTURE) ON DATA EXCHANGE

Rajashekara H.G
Research Scholar/CS
PRIST University,Thanjavur

Dr.A.Arul Lawrence S.K
Professor & Head /CSE
Bangalore -32

ABSTRACT- ***Vehicular ad-hoc Networks is to share the information between the vehicles through Vehicular Event Sharing with a mobile P2P Architecture (VESPA). The VESPA is used to distribute any type of event like parking spaces, accidents, emergency braking, etc and evaluating their relevance once received in order to determine, for instance, whether the driver should be warned or not. This paper deals with the knowledge extraction in VESPA that it focuses on how to exploit data exchanged among vehicles. Existing systems only use exchanged data to produce warnings for drivers when needed. Then, data is considered outmoded and deleted. 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.***

Keywords: data aggregation, data storage, VANETs, VESPA and

1. INTRODUCTION

Today, the car is the most familiar mode of transportation vehicle. 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 injuring on the road, this number remains quite high, mainly due to human factors (low response time). To reduce the number of accidents, the "Intelligent Transportation Systems" have been initiated in Japan, Europe and the United States. These wireless networks rely on the use of short-range networks (about 100 mtrs), like IEEE 802.11 or Ultra Wide Band (UWB) standards for vehicles to communicate and provide bandwidth in the range of Mbps. Using such communication networks, a car driver can receive information from their neighbors. Many pieces of information may be exchanged in the context of inter-vehicle communications such as accident, emergency braking, obstacle on the road, etc.

This paper presents our approach that considers both fresh data for warning drivers but that also maintains and aggregates data histories for distribute 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 parking with available places. Such knowledge can be determined using previously received warnings even if no new event has arrived during a period of time. There are two approaches used to aggregate events. These are 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 decentralized approach, vehicles construct their own aggregates. This paper presents a fully decentralized approach improved by aggregate exchanges between vehicles.

2. RELATED WORK

2.1. Inter-vehicle Communications

CarTalk worked to exploit inter-vehicle communications to make driving safer. Both FleetNet and CarTalk used multi-hop communication techniques, but while Fleet- Net was supported by a partial fixed infrastructure, CarTalk used no existing infrastructure. The TrafficView system proposes dissemination protocols for broadcasting data periodically. To determine the dissemination area, which is the area where the data should be broadcast. It 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, 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.

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 [4]. Mobi-Dik provides a solution for the problem of information-sharing inside a restricted spatio-temporal area. VESPA aims at proposing a single data sharing approach valid for all types of events, in order to deploy a generic system in cars. So, VESPA uses an encounter probability to determine the relevance of an event.

2.2. Data Aggregation in Vehicular Networks

Data aggregation proposed as a way of optimizing bandwidth or storage usage. Many data aggregation strategies developed in the context of sensor networks, 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.

Some works related to data aggregation in the specific context of vehicular networks:

1. Proposes the Region - based Location Service Management Protocol (RLSMP)
2. Focusses on security aspects – Here, two types of aggregation are proposed. In syntactic aggregation used to reduce the overhead of message headers. In semantic aggregation is used to save bandwidth at the cost of loss of information. Similarly, re-aggregation is not considered & the omission of records by malicious cars cannot be detected.
3. Distinguishes between data compression and data aggregation - Only data aggregation considers the semantics of the data. This method utilized in the ratio-based aggregation algorithm and the cost-based aggregation algorithm.
4. Considers vehicles which aggregate warning data - It 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.
5. Hierarchical data aggregation - The aggregation hierarchy is pre-defined in the data map, grouping areas according to their natural relation.
6. Proposes a Location Based Aggregation (LBAG) protocol - In this protocol, data aggregation relies on a hierarchy of static locations. A geocast protocol is used to deliver a message to a target area.

3. EVENTS AGGREGATION

When the data is received by a vehicle, its relevance is evaluated, generally using both spatial and temporal criteria. Once used, the data produced to describe an event is considered outdated and deleted. The goal is to minimize the amount of data to be exchanged between vehicles. In this paper, 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 definitely possible to use the data collected to produce, at the vehicle level. 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. In another context, it is possible to use the different messages received about accidents, emergency braking, 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 [5]. Thus, a vehicle may produce and receive messages describing road events.

We consider an event described as a tuple <type of event, timestamp, location>. When disseminated in the network, the events generally have a more complex representation. The aggregation process verify the properties are summarize the events according to the fundamental dimensions that are location and time, incremental and the volume of data stored should be limited, each driver should be able to choose the types of events driver is interested in, as well as the spatial and temporal scales to be used to aggregate the data, to exchange the aggregated data between vehicles in order to improve their respective knowledge and the aggregation process should not be costly, neither in terms of computational cost nor of memory usage.

3.1. Aggregation structure used for the events

In addition to the “active” events received, each vehicle stores a summary of the previously received events. In this summary, each type of event is associated with a 2D-matrix V. Each spatio-temporal cell of matrix contains a no. of events and a confidence value ranging between 0 & 1. 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 and the size of the rectangles is not necessarily homogeneous. So, each matrix used to aggregate the event is associated to a table S representing the set of defined rectangles.

International Journal of Innovative Research in Information Security (IJIRIS) ISSN: 2349-7017(O)
Volume 2 Issue 5 (May 2015) ISSN: 2349-7009 (P)
www.ijiris.com

Each cell of the table contains the left low point (llp) and the right high point (rhp) of the rectangle. In temporal dimension, there are two items specified. The first one is the observation window determining the events to aggregate [7]. The time representation may also be non-homogeneous and it 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).

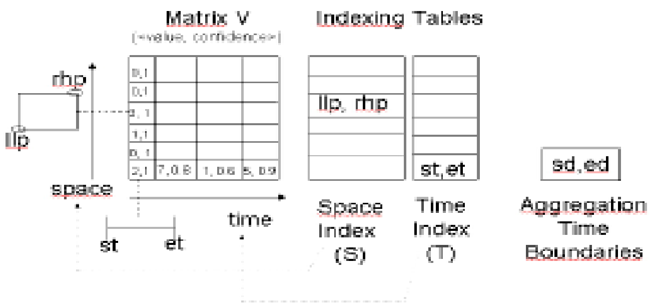

Figure 1: Storage structure of the events

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. 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. 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. 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 used for both the temporal and spatial dimensions may be determined either statically.

3.2 Mapping cells

The aggregation matrices are used to store and manipulate summaries of known events. The operation on these matrices is the reduction of a cell according to another one. The reduction operation 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 spatiotemporal frames of reference. The main method is **function reduction (c_1, c_2) returns sub-cell** where c_1 & c_2 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. Clearly, the following cases have to be considered:
• if c_2 is included in c_1 then c_2 is returned;
• if c_2 & 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 and up to 3 for the temporal one. 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 $subc_2$, is given by **value ($subc_2$) = value (c_2) / nb, conf ($subc_2$) =conf (c_2) / nb**. Here, nb corresponds to the number of cells created.

The sub-cell $subc_2$ is returned.
• if c_1 & 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 is estimated using both the number of observed events and the confidence value. Accurately, we consider that the filter may be defined by a simple query retrieving the number of events observed in a given spatio-temporal area. Therefore, the functions are 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 union cells function is used to compute the union of those sub-cells.

These function is 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 of the confidence factors [9].

3.4. Exchanging summaries among vehicles

These exchanges may be performed either by using relays or directly. Each vehicle may decide to publish the summaries it manages to make them available to other vehicles. Mutually, a vehicle may be interested in subscribing to some of the summaries published by the others. To simplify, the vehicles publish or subscribe towards all the other vehicles. The publication implies defining which summaries are published. 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 [10]. 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.

In this purpose, we use the methods of procedure fusionmatrix (m_1, m_2)**.** Here, m_1 is the target matrix in the fusion process. All the cells of matrix m1 are 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 m2 and the sub-cell returned by reduction. The method of function cellfusion (c1, c2) returns c3, c1 & c2 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. We approximate the aggregated value with a reduction by using the max of both values3 are value(c3) = max(value(c1), value(c2)) and conf(c3) = value(c1) * conf(c1) + value(c2) * conf(c2) / value(c1) + value(c2).

Table - 1. Cellfusion Max. value

Max (c1)	**Max (c2)**	**(c3)**
100	125	125
125	190	190
190	250	250

Table - 2. Confidence factors

conf (c1)	**conf (c2)**	**conf (c3)**
100	125	124.61
125	190	189
190	250	249.56

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, the timestamp of the last update LT(v, e), together with its type (local or by exchange). Furthermore, it maintains the timestamp ET(v, vi, e) of the last exchange for all known vehicles vi. Thus, a vehicle V1 can decide whether it is interesting or not to consider the summary of a vehicle V2 for an event e, comparing LT(V 2, e) with ET(V1, V2, e). The amount of data exchanged for the summaries is significant 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

The main principle of our approach based on data aggregation in VANETs to generate additional knowledge for drivers. This approach allows to exchange summaries among vehicles in addition to information about traditional events. In this paper that it consists in defining scenarios in order to evaluate our solution and refine the choice of scales used for the generation of summaries.

REFERENCES

[1] Bruno Defude, Thierry Delot, Sergio Ilarri, José-luis Zechinelli, and Nicolas Cenerario. International Conference on Mobile and Ubiquitous Systems: Networking and Services - MOBIQUITOUS , 2008 DOI: 10.1145/1594978.1594995
[2] N. Cenerario, T. Delot, and S. Ilarri. Dissemination of information in inter-vehicle ad hoc networks. In IEEE Intelligent Vehicles Symposium, 2008.

[3] 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.
[4] S. Eichler, C.Merkle, & M. Strassberger. Data aggregation system for distributing inter-vehicle warning message. In 31st Conf. on Local Computer Networks, pages 543–544, 2006.
[5] 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.
[6] T. Nadeem, P. Schankar, and L. Iftode. A comparative study of data dissemination models for VANETs. In 3rd Annual Int. Conf. on Mobile & Ubiquitous Systems, pages 1–10, 2006.
[7] S. Orlando, R. Orsini, A. Raffaet`a, 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.
[8] 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.
[9] 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.
[10] J. Zhao and G. Cao. VADD: Vehicle-assisted data delivery in vehicular ad hoc networks. IEEE Transactions on Vehicular Networks, To appear, 2008.

