Location Based Information Storage and Dissemination in Vehicular Ad Hoc Networks

Girts Strazdins

Faculty of Computing, University of Latvia, 19 Raina Blvd., Riga, LV 1586, Latvia
gstrazdins@acm.org

Abstract. Vehicular Ad Hoc Networks (VANETs) is an emerging type of information networks in urban areas. A lot of research has been done in the area of increasing the vehicle awareness by disseminating collision and congestion warnings, and parking place availability information. In this paper we propose a novel idea and framework for dissemination of location based information, called digital maps, which are useful not only directly for the drivers and vehicle onboard navigation systems, but also external entities, such as tourists, environmental scientists, emergency services, advertisement companies. Centralized authority defines cooperative knowledge collection tasks and disseminates orders in the network while every vehicle decides which tasks it takes part of, based on hardware equipment, geographical position and individual interests of the driver. The results of preliminary simulation, with vehicles driving in an artificial city, show, that 200 vehicles/km^2 is minimum reasonable density to deploy proposed dissemination system.

Key words: Distributed data bases, vehicular sensor networks, system architecture, framework

1 Introduction

Vehicles are a significant part of everyday lives of urban people. Contemporary vehicles are equipped with onboard computers and wide range of sensors, measuring vehicles position, direction, speed, acceleration as well as various phenomena and events of surrounding environment: temperature, level of light, distance to obstacles and surrounding vehicles.

To improve the safety and efficiency of transportation, Intelligent Transportation System (ITS) paradigm is being intensively developed in recent years. The term ITS describes a range technologies with a common goal - to augment sensing capabilities and therefore enhance the intelligence of vehicles. To share the individual knowledge with other vehicles on the road, wireless vehicle-to-vehicle (V2V) communication support is added. Term *Vehicle-to-Infrastructure* (V2I) communication describes vehicles communicating with roadside units. DSRC and 802.11p (WAVE) are the state of art wireless vehicular communication protocols at the moment.

Using sensing and communication capabilities intelligent vehicular networks are created, on top of which distributed information systems are built. While each vehicle has limited resources and geographical coverage, together the network possesses a great potential for large scale information collection, storage and dissemination. In previous work this potential is mainly used for collision and congestion information exchange.

In this paper we propose a novel application for vehicular networks - collection of digital maps, containing location based information, which includes, but is not limited to: traffic flow speed and congestion; parking place availability; road quality; aggressiveness of drivers (breaking, acceleration, honks); temperature; air and noise pollution; weather indicators: raininess, fogginess, windiness.

Each of the tasks require specific sensing equipment, for example, thermometer, GPS. A vehicle can take a part of all the tasks it has required hardware for.

Our work has just began, the initial results of ongoing study are presented here. More extensive research and evaluation will be done, to find optimal system parameters and bring the proposed solution in deployment.

2 Related work

Most of the previous research has been done in the area of collision detection and congestion information exchange. Authors of [1] propose to detect traffic jam events based on vehicle movement history. System called Nericell [2] explores the capabilities of identifying potholes, bumps and car horns on the city streets using mobile phone with accelerometers and microphone carried in a vehicle.

Physical limitations of wireless media have been studied in both simulations and real world experiments. Simulation results of the paper [3] show that maximum throughput is reached when 150-200 vehicles communicate simultaneously in the same area. A real world experiment from [4] concludes that communication time for vehicles on a highway varies from 15-30s while the average distance, at which the connection is established, is 133m. In [5] Torrent-Moreno et al. conclude that in saturated environments wireless broadcast reception rate can be as low as 20-30%.

Taking the above mentioned physical limitations in account, it is clear that it is not possible to use simple flooding for all the collected data. Various aggregation and diffusion techniques have been proposed in previous work. Authors of [6] propose to divide road in segments relative to each vehicle and aggregate data of each segment. In [7] passive on-demand cluster formation algorithm is advocated for data flooding in ad hoc networks. Authors of Cath-Up aggregation scheme [8] propose to forward messages with a random delay increasing the possibility of multiple reports to meet on their way and become aggregated. TrafficView [9] framework forwards messages in counter-flow direction, the same idea is used in [10].

Local data of multiple vehicles may appear contradictory, merging techniques must be applied to fuse it. Authors of [11] use modified Flajolet-Martin sketch as probabilistic approximation of real data values.

Pan hui et al. [12] propose to classify contacts based on their frequency and duration. The nodes with most frequent and long-lasting contacts form a community, having the greatest potential for data dissemination. Carreras et al. [13] show, that knowing neighbor interests improves the possibility of finding required data in ad hoc networks.

Several prototype systems and frameworks for traffic data exchange have been built, including CarTel [14], TrafficView [9], SOTIS [15] and Mobile Century Project [16]. Authors of SOTIS prove the ability of their system to function even when a tiny part of all vehicles (2%) are equipped with it.

However none of the previous work has been done in the area of distributed digital maps with location based information in VANETs.

3 Problem Description

We examine tasks of collecting spatio-temporal data using vehicles. Let us denote a spatio-temporal map by a tuple $< c, w, h, u_W, u_H, t_F, t_T, t_U, V >$, where:

$c \in N \times N$: top-left corner of the map, described as geographical longitude and latitude, encoded with natural numbers;

$w, h \in N$: width and height of the maps in units u_W and u_H;

$u_W, u_H \in N$: size of one width/height unit of the map, in meters;

$t_F, t_T \in N$: time boundaries of the task (seconds after Jan/01/1970 00:00:00);

$t_U \in N$: time unit, described in seconds;

V: fact value set, map specific.

Example map:

$< (56862487, 24289398), 1000, 800, 10, 10, 1240203600, 1240581600, 60, \{0, 1\} >$

Here top left corner is a point near the city of Riga, the area is 10x8km, the collection task time boundaries are from 20/apr/2009 08:00 until 24/apr/2009 17:00, with a time unit of 60 seconds, and each fact can take value either 0 or 1.

The vehicular network contains a set of maps M. Each map contains a set of facts F. Let us denote fact as a tuple $< m, a, v, t >$, where

$m \in M$: the map;

$a \in N \times N$: coordinates of the fact, offset from top-left corner;

$v \in V$: value of the fact, map specific;

$t \in N$: time, when the fact was registered by its originator.

3.1 VANET specifics

Vehicular ad hoc networks pose a set of specific problems:

1. Rapid changes in topology is the reason, why communication is the critical problem: links are very short and dynamic;
2. Vehicle on-board computers have not so strict resource limitations as conventional sensor network devices or mobile phones. Therefore more computation can be done to minimize and optimize the communication.
3. Vehicle movements follow a common pattern: streets are static. This fact can be used to predict future locations [17] and communication links.

4 Proposed solution

The goal of our application is to merge local and limited information of individual vehicles to build a consolidated map containing the knowledge of the whole vehicular network. Advantages of our approach:

1. The map consists of numerous, exact local measurements with increased accuracy, compared to data collected by several, statically located devices (induction loops on the roads, meteorological stations).
2. The data accuracy increases in more dense regions, having more vehicles.
3. While a central authority is used to define the region and data to collect, and maintain security, it does not break the network scalability: vehicles act as both data collectors and aggregators. It is however advisable to have subset of network nodes connected to the central data base, giving opportunity to build a complete map and provide data for external users.

The potential users of the acquired data: vehicle drivers; tourists, city visitors and local inhabitants; public transport companies; public emergency services; environmental researchers; advertisement companies.

We argue that a completely centralized solution, where each vehicle sends the data and service requests directly to a central data base, is not reasonable because:

- Centralized solution is not scalable [18];
- Direct link with centralized server requires powerful and expensive infrastructure, and more energy is required for long range communication;
- Using cellular networks invokes cost per each transmitted byte.

There are, however, several advantages of central data processing:

- In situations, when city traffic is divided in separate clusters and data is never exchanged between them, having at least one node in each cluster connected to a central database, it would serve as data bridge;
- According to [19] it is unrealistic to assume, that V2V communication will function the same way in real world, as it does in simulations, because of the low V2V-ready vehicle percentage. Therefore even a minimal infrastructure improves the communication probability;
- More resources for data storage and report generation are available to a server compared to individual vehicles;
- To share collected data with third parties, a centralized sercer is necessary;
- To issue, sign and revoke security certificates for encryption and authentication, and define data collection tasks, a central trusted authority is needed.

Based on the above arguments, we propose a hybrid approach, where a central authority (CA) defines data collection tasks and maintains security while the data is disseminated in the network in a decentralized manner, using V2V communication. Subset of network nodes function as bridge-agents between the CA and decentralized network. Requirements, proposed architecture and system components are described in the following sections of this paper.

4.1 Requirements and Assumptions

To participate in the network, each vehicle is required to have:

- Global Positioning System (GPS) and navigation system with street maps;
- Radio communication unit (DSRC);
- Data processing unit;
- Graphical and/or audial user interface to interact with the driver.

The Intermediate Agents (IAs), which are described in the following section, require additionally a connection to the Internet, to communicate with both the vehicles (using DSRC) and the CA. The type of internet connection depends on the type of the IA - it could be a stationary base station on the road side, connected to a wired network, or a mobile taxi cab, which uses temporary connections while passing by WiFi access points. 3G cellular network is also a possible solution for IA connection to the server.

4.2 System Architecture

The distributed data network consists of nodes with the following roles:

- The Central Authority (CA) - a central server, which maintains a data base containing all the gathered data, defines data collection tasks, maintains security procedures and communicates with external data users. Connected to the Internet;
- Vehicles - responsible for data collection and dissemination in the network. Communicate with each other, using short range radio communication;
- Intermediate Agents (IA) - function as a bridge between vehicles and the CA, are connected to both vehicular network and the Internet. Deliver control meta data from the server into the network and transfer data between disconnected clusters of vehicles. A subset of vehicles (taxi cabs, public transport vehicles) or stationary base stations are potential players of the IA role;
- External data users - services and information systems which use the data provided by the CA.

4.3 System Components

System components are shown in Figure 1. They include three Input/Output devices, which communicate with the exterior: Radio, Sensors and User Preferences. *User Preferences* represent graphical/audial user interface providing communication with the driver of the vehicle. All other components are internal.

We do not specify exact algorithm for each component. For example, multiple aggregation techniques, used by Fact validator, are proposed previously, [11], [6] are some of them. Each component is a homogenous part of the system, using predefined I/O interface for communication with other system components. Multiple implementations are allowed for each component, switched at runtime.

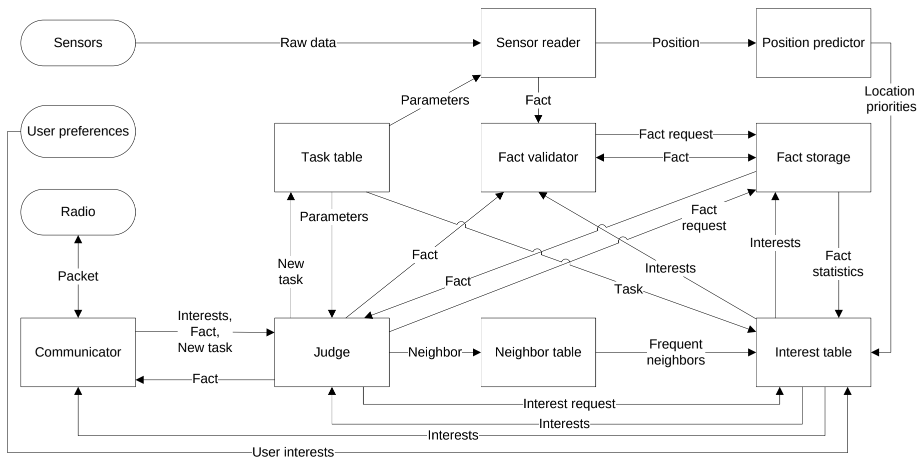

Fig. 1. System components

The remainder of this section describes motivation and function of each component.

Communicator: transforms fact and interest objects into packets and vice versa. Exchanges packets with radio transmitter. Forwards received objects to Judge. Functions as data buffer and bandwidth controller. Performs aggregation techniques, which use data forward delay variation ([8]).

Judge: receives other vehicle interests and facts, as well as CA advertised new tasks from communicator. Answers and ignores requests, based on information received from Task table and Interest table. Also stores security information to check neighbor authenticity. Forwards data to Fact validator and Neighbor table.

Neighbor table: Stores actual and frequently visited neighbor list. Used for data muling: node also stores data, which it is not directly interested in, but which is requested by its neighbors [12], [13]. Modifies local interests by sending statistics to Interest table.

Interest table: Stores local interest information in form of (m, r, p), where $m \in M$ - the map, $r =< c_r, w_r, h_r >, c_r \in N \times N, w_r, h_r \in N$ - region, described by top-left corner coordinates c_r, width w_r and height h_r, $p \in N$ - priority. Makes decisions based on user input, actual and frequent neighbor interests, actual data collection tasks, statistics of actual stored facts (regions with more data become less interesting) and predicted future location.

Task table: stores actual data collection tasks. Receives new tasks, advertised by the CA, from Judge. Provides information to Sensor reader and interest table.

Sensor reader: Reads raw data from sensors and position from GPS, interprets it as facts, based on actual tasks. Forwards position to Position predictor.

Position predictor: Receives position of the vehicle and calculates statistics, which are used to predict future location [17], used to prioritize local interests.

Fact storage: Data base, storing actual knowledge of the vehicle in terms of facts - locally measured and remotely received. Sends facts to Fact validator and Judge upon request. Statistics are reported to Interest table.

Fact validator: validates new facts before storing them. Queries storage and performs aggregation when needed. Follows local interests. Serves as *non-validated data storage* [9]. Data validation takes computational resources and time, therefore this task has a low priority and is performed in background.

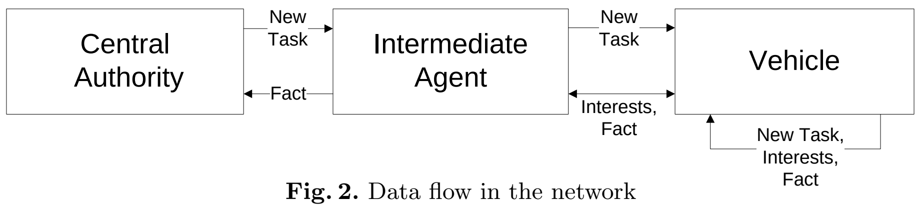

Fig. 2. Data flow in the network

4.4 Communication protocol

We identify the data flow between network nodes, shown in Figure 2, without specific implementation details, which depend on Communicator component used. Whether facts are broadcasted periodically ([8]), in specific directions ([9], [10]) or only sent in response to periodically broadcasted requests containing interests, is up to implementation.

Data flow contains three different objects:

New Task: request to start new data collection task. Created by the CA, sent to all IAs. Disseminated further in the network by each IA and vehicle. Must be signed by CAs private key to assure the authenticity of the tasks origin.

Interests: request sent by IAs and vehicles to let neighbors know facts of which regions should be sent first and in higher resolution. When sending fact reports, vehicles *should* take interests of local neighborhood into account.

Fact: data collected by vehicles. Each vehicle forwards its fact knowledge to neighbor vehicles and IAs. IAs act the same way as vehicles, but additionally they periodically forward received facts to CA.

Interest broadcasts or fact reports serve as beacons for neighbor tracking. Therefore smaller packets containing less interests and/or facts are preferred.

5 Evaluation

We have made a preliminary evaluation of vehicle communication patterns in a simulated city, using C++ simulation application. A city of size $1km \times 500m$ is used, streets forming a grid: horizontally and vertically, every 20×30 meters. Vehicles are placed on streets randomly, start driving in a random direction at a random speed, distributed evenly in interval 5-50 m/s. When crossing an intersecting street, a vehicle decides whether to go straight (8/16 chance), turn left (3/16), turn right (3/16) or turn around (2/16). When turning, vehicle chooses a new speed at which to travel until the next turn.

We run three different simulations with 10, 100 and 1000 vehicles respectively. Simulation time was chosen 10 minutes - enough to catch global trends. The communication time was fixed. Two cars are considered to be in a communication range, if they are on the same street (other area considered to be buildings, blocking signal) and the range between them is less than 200 meters (typical range for DSRC communication). The results are shown in Figure 3 - distribution of communication durations in terms of count and overall percentage.

Several conclusions can be made from the results. Communication possibility increases with vehicle count. However 80% of all sessions are shorter that 2 seconds, even for high vehicle density scenarios. Minimum reasonable vehicle density is at least $\frac{100}{1*0.5} = 200(vehicles/km^2)$.

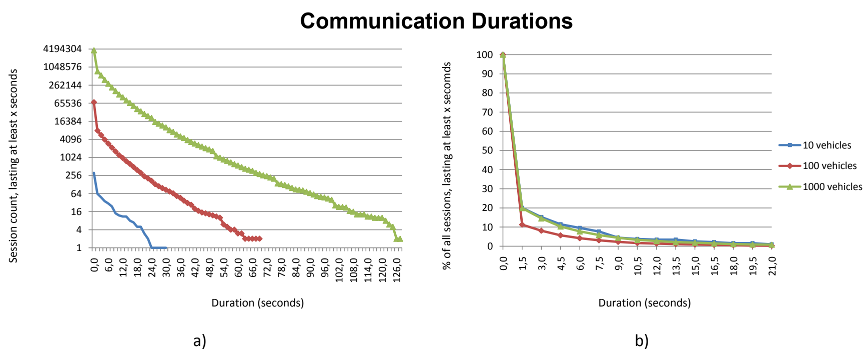

Fig. 3. Communication duration distribution: (a) count; (b) percent. Logarithmic scale

6 Conclusion and Future Work

In this paper we propose a framework for location based information storage and dissemination in VANETs. Hybrid network architecture is used, having a central authority, which serves as centralized data base server, security maintainer and data provider to external users, and decentralized vehicular network for data collection and dissemination. Intermediate agents bridge the central authority and vehicular network together.

We have formalized the problem of data collection, described related work, potential advantages and drawbacks of both completely centralized and completely decentralized architectures. We describe the architecture of our proposed framework, identify data flow, system components and information flows in the network. A preliminary evaluation with simulated vehicular network in a small part of a city has been done. The results show, that even in dense scenarios 80% of communication sessions last only 2 seconds or less, therefore the data exchange must be done in an efficient manner.

In future work more extensive implementation and simulation work must be done. We will focus on simulating data exchange, using real life movement data traces from a set of real vehicles, driving around the city of Riga for multiple hours. A realistic radio signal propagation must also be used. Analysis of communication improvement by using static base stations is also planned.

Acknowledgements

I would like to thank my advisor Leo Selavo for the assistance, comments and suggestions related to this work.

References

1. Yoon, J., Noble, B., Liu, M.: Surface street traffic estimation. In: Proc. MobiSys '07, New York, NY, USA, ACM (2007) 220–232
2. Mohan, P., Padmanabhan, V., Ramjee, R.: Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In: Proc. SenSys'08. (2008) 323–336
3. Eichler, S.: Performance evaluation of the IEEE 802.11 p WAVE communication standard. In: Proc. VTC'07. (2007) 2199–2203
4. Seada, K.: Insights from a freeway car-to-car real-world experiment. In: Proc WINTECH'08. (2008) 49–56
5. Torrent-Moreno, M., Jiang, D., Hartenstein, H.: Broadcast reception rates and effects of priority access in 802.11-based vehicular ad-hoc networks. In: Proc VANET'04. (2004) 10–18
6. Ibrahim, K., Weigle, M.: Accurate data aggregation for VANETs. In: Proc. VANET'07. (2007) 71–72
7. Yi, Y., Gerla, M., Kwon, T., Technologies, T.: Efficient flooding in ad hoc networks using on-demand (passive) cluster formation. contract **14**
8. Yu, B., Gong, J., Xu, C.: Catch-Up: a data aggregation scheme for VANETs. In: Proc. VANET'08. (2008) 49–57
9. Nadeem, T., Dashtinezhad, S., Liao, C., Iftode, L.: Trafficview: Traffic data dissemination using car-to-car communication. ACM SIGMOBILE Mobile Computing and Communications Review **8**(3) (2004) 6–19
10. Mitko Shopov, E.P.: Generation and dissemination of traffic information in vehicular ad-hoc networks. Computer Science (2005)
11. Lochert, C., Scheuermann, B., Mauve, M.: Probabilistic aggregation for data dissemination in VANETs. In: VANET'07. (2007) 1–8
12. Hui, P., Yoneki, E., Chan, S., Crowcroft, J.: Distributed community detection in delay tolerant networks. In: Proc MobiArch'07. (2007)
13. Carreras, I., De Pellegrini, F., Miorandi, D., Tacconi, D., Chlamtac, I.: Why neighbourhood matters: interests-driven opportunistic data diffusion schemes. In: Proc. CHANTS'08. (2008) 81–88
14. Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A., Shih, E., Balakrishnan, H., Madden, S.: Cartel: a distributed mobile sensor computing system. In: Proc. SenSys'04. (2006) 125–138
15. Wischoff, L., Ebner, A., Rohling, H., Lott, M., Halfmann, R.: SOTIS - A Self-Organizing Traffic Information System. In: Proc. VTC'03. Volume 4. (2003)
16. http://traffic.berkeley.edu/
17. Burbey, I., Martin, T.: Predicting future locations using prediction-by-partial-match. In: Proc. MELT'08. (2008) 1–6
18. Wang, W., Lu, W., Mizuta, H., Wang, G., Du, Y., Liu, W., Tang, X.: Probe Car System based Traffic Information Service Experiment. In: Proc. IS Telecommunications. (2006) 1134–1136
19. Misener, J.: To 'V' Or Not To 'V'? Traffic Technology International (April/May 2009) 12

The original publication is available at www.springerlink.com:
http://www.springerlink.com/content/n262002884k4101p/

Bibliography reference:

Strazdins, G: Location Based Information Storage and Dissemination in Vehicular Ad Hoc Networks. In: Advances in Databases and Information Systems: Proceedings of associated Workshops and Doctoral Consortium of the 13th East European Conference (ADBIS 2009), J.Grunspenkis, M.Kirikova, Y.Manopoulos, L.Novickis (Eds.), 2009, Springer, LNCS 5968, pp.169–176.