Data Management in Vehicular Networks

Thierry DELOT, LAMIH – University of Valenciennes
Thierry.Delot@univ-valenciennes.fr
Valenciennes:  
50,000 habs

Valenciennes and its suburbs: 350,000 habs

University:  
10,000 students

60 km from Lille

200 km from Paris

100 Km from Brussels

230 km from London
LAMIH – University of Valenciennes
DIM team: Decision, Interaction & Mobility

3 main research groups

- RAIHM
  - Artificial Intelligence, Multi-agent systems and HMI
- ROAD
  - Operational research, decision aid systems and logistics
- SyME
  - Mobile and Embedded Systems
General context

- Recent development of:
  - Mobile devices
    - Sensors, smartphones, navigation devices, etc.
  - Wireless networks
    - Wi-Fi, 3G, etc.
  - Global Navigation Satellite Systems (GNSS)
    - GPS system

These mobile devices produce and/or store data!
They may also be used to access data!
Impact of such mobile devices

- **New types of queries**
  - Location-based Queries, Continuous queries, Nearest neighbor queries, Range queries, Spatio-temporal queries, etc.

- **New processing techniques**
  - Traditional techniques are no more suited

- **Optimisation objectives**
  - Non classical optimisation
    - Energy, financial cost, etc.
  - Local vs. Global optimisation

- Even the notion of query result is different!
Example with a centralized architecture
Issues with centralized solutions

Scalability but also response time!
Accessing data in vehicular ad hoc networks
PhD thesis N. Cenerario, def. 2010

- No more fixed infrastructure (data servers)
  - Direct interactions between mobile nodes
  - Use of short range communication networks (e.g., IEEE 802.11, UWB, etc.)
  - Use of positioning systems (e.g., GPS)

- Driver assistance systems designed on top of inter-vehicle networks

Be careful accident!

Accident!!

Accident - 100m

Accident - 200m

Accident - 300m

I don’t care, I’m driving in the opposite direction!
Why such assistance systems?

- The car is indisputably the most heavily used mode of transportation

- Nevertheless, numerous problems remain:
  - Number of accidents, environment, importance of the human factor (e.g., accident-prone behavior or low response time).
Inter-vehicle networks

- **Objective:**
  - Intelligent Transportation Systems
  - Assist drivers by providing them information

- **Exchange information between vehicles**
  - Ephemeral relevance of exchanged data
  - Different from data (e.g., POI) managed today by navigation systems

- **Numerous interesting projects:**
  - FleetNet, CarTalk, TrafficView, MobiDik, PreDrive C2X, CICAS, VESPA…
MobiDik [Wolfson04]
Sharing information about available parking spaces

- Diffusion and Replication of the information between close vehicles
Example... [Wolfson04]
Sharing information about available parking spaces
Example... [Wolfson04]
Sharing information about available parking spaces
Example… [Wolfson04]
Sharing information about available parking spaces
Main characteristics of the solution

- **Objective:**
  - Maintain an information in a determined spatio-temporal area

- **Replication Techniques**
  - Use of a relevance function:
    \[ F = - \alpha t - \beta d \] (\( \alpha, \beta \geq 0 \))
Mobile query processing

Routing results towards the moving object in a highly dynamic network is a (very) difficult task!
Strategies based on the use of a fixed infrastructure

\[ Q = \text{query} \]
\[ A = \text{answer} \]

Send answer directly to the client if it is possible.
Strategy for rare resources

I have the answer

Q = query
A = answer

Static peer
Client

Trajectory
Data dissemination among vehicles

- **Objective:**
  - Push data towards (interested) mobile nodes

- **Challenges:**
  - Inform a potentially interested driver/vehicles
  - Avoid network flooding
Many existing protocols for VANETs

I had an accident

Accident in 100m

Accident in 200m

???

Very low node density

Ok i will carry and forward the information?

Carry and forward protocols
How to adapt data dissemination to the type of event considered?

**Direction dependent events**

Accident!!

Be careful accident!

Accident - 100m

Accident - 200m

Accident - 300m

I don't care, I'm driving in the opposite direction!
How to adapt data dissemination to the type of event considered?

Non direction dependent events
How to adapt data dissemination to the type of event considered?

Non direction dependent events
Encounter Probability

Objective:
  ▶ Determine whether a vehicle will encounter an event or not
  ▶ Computed on spatio-temporal criteria:
    ▶ Using digital maps:
      ▶ Shortest path algorithm
    ▶ Using geographic vectors

\[
EP = \begin{cases}
1 & \text{if } TTR < TTL \\
0 & \text{otherwise}
\end{cases}
\]
Spatio-temporal criteria

- 3 new relevance criteria
  - Colinearity coefficient (c)

Accident!!
Spatio-temporal criteria

- 3 new relevance criteria
  - Colinearity coefficient ($c$)
  - Minimal distance over time ($\Delta d$)
Spatio-temporal criteria

- 3 new relevance criteria
  - Colinearity coefficient (c)
  - Minimal distance over time ($\Delta d$)
  - Temporal distance ($\Delta t$)

**Accident!!**

The difference between the current time and the time when the vehicle will be closest to the event
Encounter Probability

- How to evaluate whether a vehicle will encounter an event or not?
  - Rely on:
    - The 3 values defined previously
    - The age of the event ($\Delta g$)

$$EP = \frac{1}{\alpha \times \Delta d + \beta \times \Delta t + \gamma \times \Delta g + \zeta \times c + 1}$$

**Penalty Coefficients**
Data Dissemination

Introduction of a latency in the dissemination process

\[ t_i = D \times (1 - \frac{d_i}{r}) \]
Data dissemination

Diffusion canceled!

Relays the message

Diffusion canceled!
Data dissemination
The VESPA system

- (Generic) driver assistance systems relying on the use of vehicular networks
  - Exchanges of information between vehicles
    - Accidents
    - Emergency brakings
    - Available parking spaces
    - Traffic information
    - etc.

Transportation Research Part-C (Emerging Technologies) 2010, IF: 1.7
Simulator
Simulator
Simulator
Open issues…
Information is now everywhere!

- May not only be provided by cars
  - Sensors, remote web services or databases
- Considering multi-level access to data
  - Different types of networks
    - Short range but also telephony networks
  - Push and pull techniques to access data
- Examples
  - What are my nearest hotels for the next hour?
  - Among close petrol stations which one is the cheapest?
Multi-scale mobile query processing

Where should I refuel?

Which are the rest areas that will be close to me in next hour proposing a gas station, lodging facilities for two persons and a restaurant?

Available parking space, Traffic congestions, emergency brakings, ...

IEEE 802.11

Int. Conf. on Ambient Systems, Networks and Technologies (ANT), 2010
Example

- Retrieve the list of petrol stations located in a radius of 10 Km around me where fuel prices are less than 1$

- Let us consider several « data sources »
  - `<dataSource id="DS1" type="local cache">`<description>Gas station addresses and prices broadcasted by neighboring vehicles</description>`
  - `<dataSource id="DS2" type="web service">`<description>returns information about gas stations (prices and addresses) in a radius of n kilometers around a point represented by a symbolic address</description>`
  - `<dataSource id="DS3" type="positioning service">`<description>returns mobile user position (GPS coordinates)</description>`
  - `<dataSource id="DS4" type="positioning service">`<description>returns mobile user position (symbolic representation)</description>`
  - `<dataSource id="DS5" type="web service">`<description>allows to translate gps coordinates to symbolic address or to translate symbolic address to GPS coordinates</description>`
Candidate queries

- **Candidate Query 1**
  ```sql
  SELECT GS.Name
  FROM DS2.GasStation(DS5.LatLngToAddress(DS3.myloc), 10) GS
  WHERE GS.Price < 1 ;
  ```

- **Candidate Query 2**
  ```sql
  SELECT GS.Name
  FROM DS2.GasStation(DS4.myloc, 10) GS
  WHERE GS.Price < 1 ;
  ```

- **Candidate Query 3**
  ```sql
  SELECT StS.Nom
  FROM DS1.StationService StS
  WHERE StS.Prix < 1
  AND inside(10, DS3.myloc, StS.Adresse);
  ```

- **Candidate Query 4**
  ```sql
  SELECT StS.Nom
  FROM DS1.StationService StS
  WHERE StS.Prix < 1
  AND inside(10, DS5.AddressToLatLng(DS4.myloc), StS.Adresse);
  ```

- …
Query optimization

- **Objective:**
  - Determine the best execution plan

- **Costs can be computed over several QoS dimensions:**

\[
C_{Time}(Q) = C_{Time_{QueryDelivery}}(Q) + C_{Time_{Processing}}(Q) + C_{Time_{ResultDelivery}}(Q)
\]

\[
C_{Money}(Q) = C_{Money_{QueryDelivery}}(Q) + C_{Money_{Processing}}(Q) + C_{Money_{ResultDelivery}}(Q)
\]

\[
C_{Energy}(Q) = K \times n
\]

- **Global cost (to minimize):**

\[
C(Q) = \sum_{i=1}^{3} \omega_i \times C_i
\]
GeoVanet: another approach to process queries in VANETs

Objective:
- Provide an access to « rare » information in VANETs

Example
- A tourist is arriving in a city and s/he is searching for information about the interesting places to visit

Principle
- Queries are spread in the network
- Results are routed to a (stationary) mailbox in a bounded time

Collaboration with POPS team / INRIA LNE
Algorithm

- QUERY = [request, exp-date, key], in which:
  - ‘request’ is the core of the request i.e., ‘what are the interesting sites to visit in Paris?’.
  - ‘exp-date’ is the date by which the answer is expected.
  - ‘key’ is used to determine the location where the answer should be sent and retrieved (i.e., the rendezvous point).

- Exploits both the mobility of moving objects (carriers) and hops in the network to reach the target

- 2 versions
  - Using digital maps or geographic vectors
Other protocols may be useful!

- How to deal with competition?
- Some elements exchanged are rather resources than events
  - e.g., an available parking space
- Competition among the drivers
  - First arrived, only served…
Reservation Protocol

- **Objective:**
  - "allocate" a resource to only one driver

- **Principle:**
  - Rely on a coordinator
    - In charge of the allocation of the resource
    - e.g., the vehicle leaving the parking space
Main steps…

- A vehicle leaves a parking space and becomes coordinator for that resource
Main steps...
Main steps...

Interested!
Main steps...
Main steps…
Main steps…

Acknowledgment
Extending the notification range...

- What if no interested vehicle is in the communication range of the coordinator?
  - Need to communicate the information farther from the resource
  - Change of coordinator
    - Difference with the usual dissemination in VESPA
Allocation issues...

- How to choose:
  - The winning vehicle to which the resource is allocated?
    - Distance to the resource, Time elapsed since the vehicle started searching, Highest Encounter Probability, etc.
  - A new coordinator

- No way to prevent drivers from taking an available parking space they see, even if it was allocated to another one...
Problem:
- The lifetime of an event is low
  - e.g., an available parking space

What to do if no event relevant in the vicinity is produced?

Our approach
- Compact and store events previously observed
- Use the aggregated data to extract information
Our Approach

- Store and aggregate (summarize) data
  - Do not destroy them once used to warn the driver
- Use the summaries generated to extract additional knowledge usable by drivers
  - Estimate the probability that an event occurs in a spatio-temporal area
  - Detect areas where the probability of finding a parking spot is the most important (depending on the day and time).
  - Detecting dangerous areas by correlation of messages received accident and emergency brake
- Exchange aggregated data between vehicles to generate a more complete information
Properties of the aggregation process

- Promote fundamental dimensions that are the location and time;
- Be incrementally constructible and inexpensive in computing time and storage space;
- Allow each driver to choose which types of events is interested in and the spatial and temporal dimensions s/he wants;
- Allow an exchange of summaries between vehicles to enrich their knowledge.
Model space with two levels

Spatial Model
Spatio-temporal Model

- **Spatial Model:**
  - Physical level:
    - cutting the space into squares of fixed size which form a complete partition (exchange level)
  - Logical level:
    - Rectangle corresponding to a set of physical squares (user level)

- **Temporal Model:**
  - Divide time into daily and hourly segments
  - Goal: obtain a fine grain representation to extract useful information
Example

Interest areas

<table>
<thead>
<tr>
<th>1</th>
<th>80</th>
<th>70</th>
<th>110</th>
<th>90</th>
</tr>
</thead>
</table>

Parking spaces

Event types

- Parking spaces
- Event types

Temporal granularities

- Interest areas
- Event types
Principles of inter-vehicle exchange

- Each car/driver decides what to exchange and his/her preferences
  - Preferences with priorities
- Duplicate detection is important (I might observe the same events as my neighbor!) → sketch
- Need to know the vehicles with which exchanges have taken place recently (list of identifiers)
Data management in mobile environments is a very interesting and challenging topic.

- Very recent research field.
- At the border between several communities:
  - Network protocols, Data management, Distributed systems, etc.
- A lot of contributions are still necessary if you are interested!
Multimedia data in vehicular networks

- Poor event representation
  - Only a type is used to describe the event
  - Additional information may help the driver

- Benefit of multimedia contents:
  - Improve the description of the event
    - Avoid disturbing the driver
  - Easy to generate
    - Using camera
    - Using microphone

- Difficulty to disseminate multimedia content!
  - More than one packet are needed for one multimedia content
    - Time to disseminate
    - Packet loss with wireless communication
    - High mobility of vehicles
  - But new communication standards are coming! (IEEE 802.11p)
Example

On event reception

Available parking space

Enlarge picture
Query processing in vehicular networks

- How to route a (set of) result(s) towards a mobile target in highly mobile networks?
- New routing protocols are needed
- How to adapt to such a changing context?
Selected publications

- T. Delot, N. Mitton, S. Ilarri, T. Hien, Pull based information gathering in vehicular networks using GeoVanet, Int. conf. on Mobile Data Management (MDM), 2011