Vehicular Sensing: Emergence of a Massive Urban Scanner

Michel Ferreira[1], Ricardo Fernandes[1], Hugo Conceição[1], Pedro Gomes[1], Pedro M. d'Orey[1], Luís Moreira-Matias[2], João Gama[5], Fernanda Lima[3], and Luís Damas[4]

[1] Instituto de Telecomunicações, DCC/FC, Universidade do Porto, Porto, Portugal
{michel,rjf,hc,prg,pedro.dorey}@dcc.fc.up.pt

[2] LIAAD-INESC Porto and DEI/FE, Universidade do Porto, Porto, Portugal
luis.matias@fe.up.pt

[3] Instituto Militar de Engenharia, S.E. Engenharia Cartográfica, Praça General Tibúrcio 80, Rio de Janeiro, Brasil
fernandalins@ime.eb.br

[4] Geolink Lda., Avenida de França, 20 Sala 605, Porto, Portugal
luis@geolink.pt

[5] LIAAD - INESC Porto L.A., FEP, University of Porto, Porto, Portugal
jgama@fep.up.pt

Abstract. Vehicular sensing is emerging as a powerful mean to collect information using the variety of sensors that equip modern vehicles. These sensors range from simple speedometers to complex video capturing systems capable of performing image recognition. The advent of connected vehicles makes such information accessible nearly in real-time and creates a sensing network with a massive reach, amplified by the inherent mobility of vehicles. In this paper we discuss several applications that rely on vehicular sensing, using sensors such as the GPS receiver, windshield cameras, or specific sensors in special vehicles, such as a taximeter in taxi cabs. We further discuss connectivity issues related to the mobility and limited wireless range of an infrastructure-less network based only on vehicular nodes.

1 Introduction

Sensor networks build upon the combination of the actions of sensing and networking. Until recently, networking was absent of the vehicular environment and vehicles were not envisioned as nodes of a super-large-scale sensor network. Clearly, the sensing component has always been particularly rich in vehicles, namely in road vehicles such as cars, buses and trucks. Components such as speedometers or more technical mass air flow and oxygen sensors have been part of vehicles for decades. In addition to these myriad of sensors that are factory-installed in production vehicles, the power reserves of these vehicles are also incommensurably larger than that of the typical mote device, allowing to power not only such large set of active sensors, but also the comprehensive number of computers that equip

F. Martins, L. Lopes, and H. Paulino (Eds.): S-Cube 2012, LNICST 102, pp. 1–14, 2012.

modern vehicles. With an estimated number of 800 million vehicles in the world, the advent of vehicular networking can connect the isolated myriad of sensors in each vehicle and create an extremely powerful sensor network.

Wireless networking in vehicular environments has existed based on cellular systems for some decades. Second Generation (2G) systems allowed data communications at the maximum rate of 9.6 kbps. General Packet Radio Service (GPRS)-based systems (or 2.5G) allow for data rates up to 171 kbps and are still heavily used to implemented real-time vehicle tracking based on Global Positioning System (GPS) localization. Third Generation (3G) and Fourth Generation (4G) systems allow for much higher data rates. However, the level of interconnection of such mobile communication devices and the sensors in the vehicles is often very limited or non-existent. In professional fleets, after-market devices are commonly installed in vehicles to allow real-time monitoring of their GPS position, speed or engine mode. Professional telemetry systems designed for trucking companies are even starting to remotely evaluate several parameters that allow quantifying the performance of each driver in terms of eco-driving.

In the mass market of non-professional vehicles, the European Commission (EC) has recently approved a recommendation [6] towards the inclusion of a wireless communication device interconnected to vehicular sensors, such as the deceleration sensors that activate airbags. The proposed emergency Call (eCall) system can automatically dial an emergency telephone number and transmit location data in case of an accident, without any human intervention. This recommendation applies to all new car models launched after September 2011.

Wireless networking for vehicles has also been evolving through a new short to medium range communication technology operating in the 5.9 GHz band, known as Dedicated Short-Range Communication (DSRC) [18]. As these new wireless networking devices are factory-installed in the vehicles, the level of interconnection with the in-vehicle sensors can be very high, namely through the Controller Area Network (CAN) bus. New collaborative Vehicle-to-Vehicle (V2V) warning systems have been proposed, based for instance on the activation of traction sensors over slippery roads, and can create just-in-time and just-in-place virtual road signs displayed on the windshield of nearby vehicles.

In this paper we describe several applications of vehicular sensing, ranging from automatic map construction and updating, to an urban surveillance system based on windshield-installed cameras. Much of the results we present are based on experiments with a large vehicular testbed that has been deployed in the city of Porto, in Portugal. This vehicular testbed comprises 441 taxi cabs from the largest taxi company operating in the city.

The remainder of this paper is organized as follows. In the next section we describe the technological evolution of the variety of sensors that have been installed in vehicles. We then discuss the increased connectivity and sensing range enabled by the mobility of the sensor nodes. Then we describe three concrete applications through vehicular sensing: automatic road map construction and updating; camera-based search of vehicles based on automatic license plate recognition; and the prediction of taxi demand per taxi stand based on machine learning over historical and real-time streamed data of geo-localized taximeter activations. We end with some conclusions.

2 Sensors in Vehicles

Sensors have been installed in vehicles since the early days of the large-scale manufacturing of automobiles. One of the first examples is the speedometer, which measures and displays the instantaneous velocity of the vehicle. In [27] sensors are defined as devices that transform (or transduce) physical quantities such as pressure or acceleration (called measurands) into output signals (usually electrical) that serve as inputs for control systems. This definition has been extended since sensors' outputs are now used as input for a wide range of applications. Modern vehicles are equipped with sensors that allow collecting information on the vehicle state, its surrounding environment and, more recently, on the driver condition.

In the following, a short survey of the main in-vehicle sensing devices relevant for networked vehicular sensing is given. These devices are categorized according to the measuring principle (e.g. radar)or the primary usage application (e.g. automotive). The most common devices are *Automotive* sensors, which measure motion and pressure characteristics. For instance, the Anti-lock Braking System (ABS) is able to detect rapid decelerations, traction control allows to infer varying road surface conditions and belt pressure sensors can be used to determine vehicle occupancy. *Environmental* sensing devices allow to monitor the external surroundings of the vehicle. Light and rain sensors detectors are good examples of this class. In [16] and in the URBISNET [1] project the authors propose the deployment of sensors (measuring the concentrations of gases) in the public transportation system for monitoring urban pollution levels. In [28] Rodrigues et al. present a system that monitors the heart wave and data from other in-vehicle sources to better understand driver behavior. *Radar* use radio waves to detect a variety of objects (e.g. pedestrians or other vehicles). Many modern cars are equipped with short range radars (e.g. for parking assistance) and with long range radars (e.g. for anti-collision). *Cameras* are used as sensors for many applications. Image processing and pattern recognition techniques allow the development of a variety of applications, such as, lane departure warning, traffic sign/light recognition and object recognition. Lastly, *Communication* devices can be used to add geographical information and communicate measurement of other sensors but also as sensors for vehicular sensing. For example, GPS devices associated with communication capabilities can be used to perform real time monitoring of traffic conditions (e.g. [22]) or of a fleet of vehicles. Bluetooth devices available at most cars (e.g. hands free) can be used to capture traffic flow and to predict travel times and origin/destination matrix estimation [2].

With the evolution of the automotive industry, the number of in-vehicle sensors is increasing rapidly. The interconnection of these sensors is a complex task due to the distributed wired infrastructure and stringent latency times. The issue of intra-vehicle communication has been considered in [25,12]. The autonomous vehicle may be the next big thing in the automotive industry.

Several companies like Volkswagen, Mercedes, General Motors and Google are working on their own version of an autonomous vehicle [1]. These new vehicles incorporate new sensing devices, such as, laser-beam range finders and position estimators.

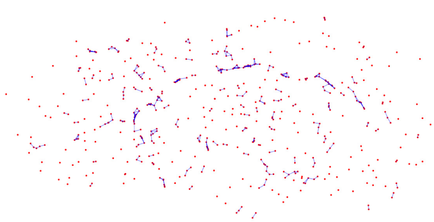
3 The Increased Range of Mobile Sensors

An interesting characteristic of vehicular sensing is that the sensors are highly mobile. While this mobility can create some difficulties regarding the wireless networking aspect of inter-vehicle communication, it can greatly amplify the sensing range of each sensor. In [8] and [9] we have studied the delay-tolerant connectivity of a vehicular sensor network in an urban scenario. We have defined this delay-tolerant connectivity as a function of the time interval considered. Instead of considering just the instantaneous connectivity of the vehicular sensors based on short-ranged inter-vehicle communication, we considered the achieved connectivity after allowing the vehicles to move for a period of time. Computing the transitive connectivity over such an interval of time has to be done considering only inter-node paths that maintain a certain time order between the travelled wireless links. Following this, the transitive closure $\mathcal{L}'*$ of the resultant network is as follows:

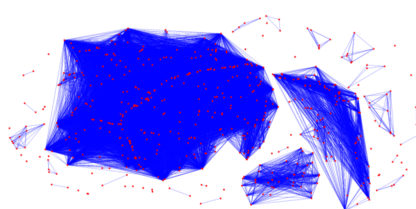
$$\forall u_i, u_j \in \mathcal{U}, t' \in [0...t],$$
$$\{u_i, u_j, t'\} \in \mathcal{L}'^* \iff \begin{cases} u_j \in \mathcal{N}(u_i)_{t'} \\ \exists u_k \in \mathcal{U}, \{u_i, u_k, t'\} \in \mathcal{L}' \wedge \{u_k, u_j, t''\} \in \mathcal{L}'^* \wedge t' \geq t'' \end{cases} \quad (1)$$

where $\mathcal{U}$ is the set of nodes in the network, $\mathcal{N}(u_i)_{t'}$ is the set of one-hop neighbors of node u_i at time t', and $\mathcal{L}'$ is the set of links.

Using the traces generated by Development of Intervehicular Reliable Telematics (DIVERT) traffic simulator [7], we were able to understand the evolution of the connectivity in time in a large scale scenario, which was, in this case, the city of Porto, Portugal. This evolution is evident in Fig. 3. For more detailed results, we refer the reader to [8] and [9].



(a) Instant connectivity at time t.



(b) Transitive connectivity $\mathcal{L}'^*$ at $t' + 30s$.

Fig. 1. Evolution of the connectivity in time

[1] Recently, Google presented their driverless Toyota Prius to the media [15].

Fig. 2. Comparison of the instantaneous vs. mobile (1 hour interval) sensing range on a vehicular sensor network in the city of Porto, Portugal

In addition to the increased (delay-tolerant) connectivity of a vehicular network over a period of time, the sensing range of each sensor is also clearly amplified. In Fig. 2 we show a snapshot of the sensed area of a set of vehicular sensors over the city of Porto, in Portugal. In frame A, we have considered a sensing range of 100m per node (depicted by a red circle at each node). In frame B, we allowed nodes to move for one hour and we depict the sensed area covered by the sensors. It translated into and increase of 0.7% of the area of the city, in frame A, to 92.6% in frame B.

4 Sensing the Road Network

Since vehicles move on roads, a straightforward application of vehicular sensing is exactly related to the sensing of the road network. Based on the data collected by a large network of vehicles, it is possible to build detailed maps of highly dynamic geographic information related to this road network, such as the presence of potholes on the road [13]. Much less dynamic, but also interesting, is the information that allows mapping and updating the actual road network based on a vehicular infrastructure of remote sensing. The goal is not only the acquisition of road axle geometries, but also their characterization in terms of topological connectivity, traffic rules and speed patterns, in an accurate and permanently up-to-date manner. Several projects have been developed with the goal of making a better use of the data collected through GPS receivers [4,29]. One of the most important of these projects is the OpenStreetMap that hosts a collaborative network of GPS traces for the assisted construction of road maps. Despite the increasing research around this area, very few references relax the need of a base map in a non-assistive approach [4,11]. Most of the work presented in the literature has focused more on refinement issues and updating of existing cartography [29]. In [23] we have used a set of vehicular probes to collect data oriented to a completely automated generation of vectorial roads maps. Such automatic construction of road maps from GPS traces requires the availability of a large data set collected over the area of interest. In our implementation, we used more than 30 millions of GPS points, collected in real-time by a vehicle tracking company, using a temporal detail of one point per second. Such level of detail is particularly important for the representation of the road network, since it allows keeping the geometry continuity of the vehicular trajectories. The position of each point is defined by its geographic coordinates: longitude, latitude and altitude. To meet our purpose of accurate road map construction, we extended the protocol to also include information about the number of satellites and the Horizontal Dilution of Precision (HDOP). We also stored additional relevant information, such as speed of the vehicle, its azimuth and the time of the position reading. We collected a total of 371,600 km of vehicular traces, spatially distributed in our zone of relevance, constituted by a small city, Arganil, in the middle of a rural area in Portugal. In Fig. 3, frame A, we depict the spatial distribution of these GPS traces.

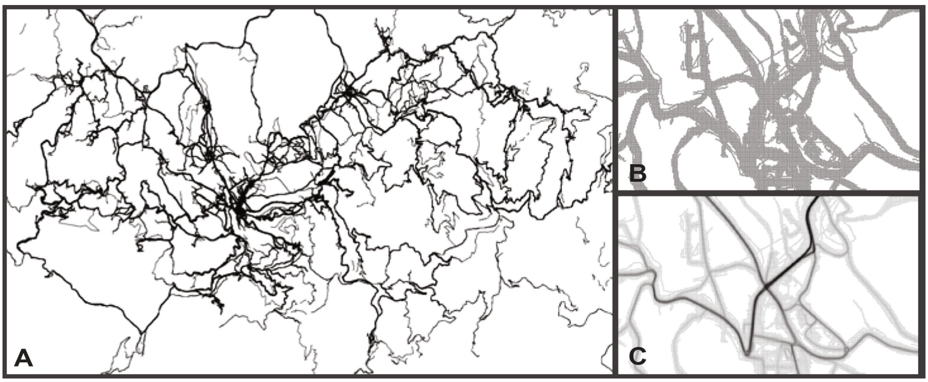


Fig. 3. Frame A depicts the set of GPS traces collected in the area of Arganil. Using such a small scale, we can easily perceive that it depicts a road network. On frame B, we highlight a small part of these road network at a much larger scale. Clearly, the noisy data reported by the GPS sensor results in a blurred description of the road network, particularly in areas where this road network is more complex. On frame C, we depict the same area as in frame B, but we use the counting of the number of GPS traces that intersects each raster cell to define the grey color intensity of the cell. Clearly, a much more defined road network starts to emerge.

Because of GPS errors, it is necessary to rely on several processes that allow the elimination of inconsistent data, aiming at obtaining higher quality input data to our algorithm. We have thus established three filters, where the first one is based on speed information. Points collected at speeds lower than 6 km/h were not considered to be sufficiently accurate for the automatic construction of road maps. As a result of this filter, we have eliminated 15,31% of the collected points. Our second filter is based on the HDOP value, which is a measure quantifying the degradation level of the horizontal positioning accuracy of the GPS (2D-based positioning). This value is mainly determined by the relative geometry of the visible satellites when the positioning reading was taken. A low HDOP value means a more accurate horizontal positioning. The filter has been configured to eliminated points with an HDOP value higher than 2.

Our third filter in the pre-processing phase of our data is based on the number of satellite used in the positioning. When more than 4 satellites are used for the positioning, the redundant satellites can serve for the detection of erroneous readings, thus increasing the accuracy of the positioning. The filter is set to eliminate points collected using a number of satellites lower than 5.

In addition to the filters mentioned above, the pre-processing phase of the data collected from the vehicles has an extra module that is responsible for the identification of large intervals of time between consecutive points of the same vehicular trace, which are either caused by obstacles to the reception of the signal broadcasted by the satellites, or by the elimination of points from previous filters. Such large intervals can erroneously affect the geometry of the road network and we thus divide vehicular traces where two consecutive points are separated by more than 7 seconds into two distinct traces. The last step in our pre-processing phase consists in the simplification of the GPS traces in order to cope with performance issues of our algorithm and minimize the amount of memory required to store the vehicular traces. Our traces have been simplified using the Douglas-Peucker algorithm [10], resulting in the elimination of 67% of the collected points. The maximum threshold distance allowed for the elimination of points through this simplification algorithm was of 1 meter.

The main goal of our algorithm is the construction of a graph representing the road network, where the roads are represented by edges and the intersections are represented by nodes (known as a network model of the road network). The algorithm is divided in five steps: rasterization, centroid generation, geometric connectivity of the centroids, topologic connectivity (node-edge topology) and turn-table construction.

4.1 Dealing with Noise from the Positioning Sensor

Until May, 2000, the real-time positioning of a point, through a navigation GPS receiver, provided a planimetric accuracy better than 100 meters. Since then, with the ending of the Selective Availability (technique used to degrade the accuracy of the positioning), such value became, on average, better than 15 meters. Even with this significant improvement, the attained accuracy is not considered to be sufficient for a valid geometric representation of the road network. As can be seen in Fig. 3, frame B, the unprocessed image that results from plotting each of the GPS points received from the vehicles, even after the filtering phase described above, creates a fuzzy image where the perception of roads is not easy. Aimed at an accurate and automatic construction of road maps, we proposed the spatial aggregation of a large set of GPS traces through a rasterization process.

The term rasterization is used in the context of the transformation of a vectorial representation into a matrix-based representation. In the work presented in [23], the rasterization process enables the transformation of the vectorial layer of GPS traces into a raster layer of 5-meter-resolution cells. For each cell, we assign a value that translates the number of GPS traces that intersect it. Using this value to vary the color intensity of each cell (depicted using different level of gray in Fig. 3, frame C), it becomes possible to easily identify the road axles. The probability of existing a road in a given cell is proportional to the value of the attribute of the cell. Similarly, cells holding a low counting value of intersecting traces represent disperse vehicular trajectories or low-travelled roads. The rasterization step thus performs an highly refined filtering of our data set, using a process based on spatial aggregation together with sampling correction. This approach becomes particularly important to the representation of small roundabouts, nearby roads and other complex parts of the road network. Furthermore, the rasterization process allows a better representation of wide roads (e.g. roads with two ways separated by a central structure), becoming possible the identification of a road axle in each of the directions.

After generating an accurate geometric map of the road network, we use again the vehicular traces to infer the connectivity layer of the identified road segments (segments between intersections. We are also able to infer the the turn-table that reflects the allowed maneuvers.

4.2 Evaluation

The evaluation phase of the results obtained was performed by comparing the geometric and topological layers of the extracted road network with those from vectorial maps provided by the map-making company InfoPortugal, S.A., which are constructed by manually processing ortho-rectified aerial images, with a resolution of $25cm^2$ per pixel. Coverage evaluation is done through three main metrics: total number of kilometers of roads generated by the algorithm in the zone of relevance; total number of kilometers of roads to which a match is found in InfoPortugal's map; total number of kilometers of roads that are not present on InfoPortugal's map (cartographic updating). Using vehicular sensing, we were able to build a road map comprising 421.82 km of roads. InfoPortugal's map has a total of 522.778 km of roads. Note that some of the roads of InfoPortugal's map, which are built based on aerial photography, correspond in fact to pedestrian-only zones and could never be detected based on vehicular sensing.

One of the most important results from the process of automatic road network extraction based on vehicular sensing consists in the identification of roads that are still non-existent in current maps. This aspect shows the ability of vehicular sensing to provide an inexpensive and highly accurate way of constantly performing cartographic updating. Our results present a correspondence of 82.94% between the kilometers of roads that were extracted and those from the base map. Hence, the remaining 17.06% represent the percentage of cartographic updating, i.e., the extraction of non-existing roads in the base map. Regarding the correspondence of our map with InfoPortugal's map, our results present a correspondence of 82.94% between the kilometers of roads that were extracted and those from the base map. Hence, the remaining 17.06% represent the percentage of cartographic updating, i.e., the extraction of non-existing roads in the base map.

The evaluation of the accuracy of the geometry of the extracted road network is done in a continuous manner: the average distance between the extracted roads and the corresponding roads in the base map is obtained through the computation of the area between the two, divided be the average length of the two geometric representations. This method presents a more accurate evaluation as compared to discrete based measurements. The resulting average distance was of 1,43 meters between the two representations of the same roads. The evaluation of the topology of the extracted road network is done through two main metrics: the number of nodes that have been extracted that match nodes in the base map; and the traffic direction of each extracted edge, compared to the direction of the associated edge in the base map. We were able to extract 596 nodes (road segment intersections) that matched with 458 nodes of the base map. Regarding the traffic direction on the matched road segments, vehicular sensing correctly identified 1158 of the 1206 road segments. The nodes which have no matching node on the base map do not necessarily represent false intersections. Such nodes can result from junctions of wide roads into more narrow segments, where the algorithm of automatic extraction usually created a non-existent node in the base map. Such nodes can also result from cartographic updating and the identification of intersections with non-existent roads in the base map. The last evaluation concerns the traffic rules between the generated roads. The extracted enforcements between connectivity of specific edges, stored in the direction table, were checked against the Do Not Enter signs of the base map, and reported an accuracy of 98%.

5 Vehicular Sensing Based on Windshield Cameras

With the introduction of cameras in production cars, and the wide dissemination of wireless inter-vehicle communication devices in all new vehicles in the foreseeing future, a potential application of vehicular sensing is the localization process of a particular vehicle, as the result of a broadcasted search warrant from some particular authority.

While these cameras have been designed to collect information to assist the driver and improve driving safety, with the support of a V2V communication infrastructure, all the in-car cameras disseminated in a city can provide a visual scope of tremendous proportions. A license-plate-based localization of vehicles, is an example of a delay-tolerant application that profits highly from the increased

Fig. 4. License plate recognition in a traffic jam situation

connectivity (radio and visual) of a set of mobile vision and communication enabled sensors. The automatic license-plate recognition is currently deployed in a number of fixed points, e.g. parking lots, gas stations and highway toll collectors, which can be globally connected to a police database of search warrants and instantaneously report the match. However, the visual scope or visual connectivity of such fixed sensors is clearly not comparable to the visual perception supported by in-car cameras distributed massively over the road network, which allow the localization of vehicles in large cities in a matter of seconds. Moreover, the vehicular sensor network empowers not only a distributed collection of data from the vehicles, but a more powerful distributed computation, that resorts to the processing power available in cars that are able to identify traffic signs or even pedestrians on the road. If care is taken to guarantee the secure and authorized usage of such an infrastructure, then vision-enabled vehicular sensor networks will possibily become critical monitoring tools.

In [14] we have described a car search protocol implemented on top of a vision-enabled vehicular newtork, where police cars issue search warrants based on license plates and the nodes of the network use the camera sensor to capture images, process them to recognize license plates, match the license plates against the search warrant and report the position through the Vehicular Ad Hoc Networks (VANET) until it reaches a police car 4. The objective was to understand how fast a query to this vision-enabled network, would return. This has many applications such as locating stolen vehicles, or lost children. In this work we focused on the first example.

We designed an experiment in which after the search warrant is issued, 4 minutes are allowed to find the stolen vehicle. Otherwise, it is considered that it was not found. In case the stolen vehicle leaves the map, or parks in a garage, it is considered that the vehicle escaped. Through Development of Intervehicular Reliable Telematics (DIVERT) traffic simulator [7], we have thoroughly evaluated this application using a realistic urban scenario (highly detailed map of the city of Porto) with thousands of vehicles, divided in several categories (police cars, vision-enabled cars, communicating cars, normal cars).We considered a very-sparse network of only 5000 vehicles, fixing the number of "police cars" at 50 (a realistic value in the city of Porto), and varied the percentage of "communicating cars" (10% and 20%), and of those, the percentage of "vision-enabled cars" (50% to 100%). We also considered diferent ranges of vision, that is the distance at which it is possible to "read" the license-plate (5 and 15 meters). The results are summarized in Table 1. For further details on the experiment, and results for different scenarios, please refer to [14].

The obtained results clearly are affected by the sparse network, which however may be accurate in representing the night period in a city like Porto, when more vehicles are stolen. Still, in a scenario where only one fifth of the vehicles are able to communicate, the results were very good and in at least 50% of the tests, the stolen vehicle was found within the allowed 4 minutes. These results can be extrapolated to other application such as the ones already mentioned, or others such as locating free parking slots in a city [5].

Table 1. Results of the locating stolen vehicles experiment

Rear and front plates results								
range of view	5m				15m			
# communicating sensors	500		1000		500		1000	
% vision enabled sensors	50%	100%	50%	100%	50%	100%	50%	100%
delay (seg)	143.53	143.28	141.16	136.41	144.58	125.59	119.80	111.15
not founded	57.5%	44.0%	40.1%	28.0%	46.0%	19.0%	18.0%	7.00%
escaped	41.0%	39.5%	40.0%	35.5%	38.5%	30.5%	29.0%	27.0%

6 Sensing Origin/Destination in a Taxi Network

Taxis are an important mean of transportation which offers a comfortable and direct transportation service. In the last decade, the real time GPS location systems became a key player on every taxi networking all over the world. All the vehicles are equipped with sensors continuously transmitting its GPS coordinates and instant speed, among others.

The streaming data provided for such network can highlight useful insights about the mobility patterns of an entire urban area over time and space like the following:

1. A cubic matrix origin/destination/time describing the mobility demand;
2. The seasonality of mobility problems such as non-forced traffic jams;
3. The characterization of the traffic flow patterns;

We focused into explore the insights described in the bullet 1 to improve the taxi driver mobility intelligence (i.e. to pick up more passengers and therefore, increase their profit). There exist just a few works on this specific topic [21,19,20,24]. Mainly, they suggest offline strategies (i.e. intelligent routing) to improve the passenger finding depending on time/space in scenarios where the demand is greater than the offer. However, there are cities with inverse scenarios (i.e. multiple companies with fleets larger than the actual demand competing between it selves) where just a fast and online strategy can actually speed the gains up.

In [26] we present a ubiquitous model to predict the number of services on a taxi network over space (taxi stand) for a short-time horizon of P-minutes. Based on historical GPS location and service data (passenger drop-off and pick-up), time series histograms are built for each stand containing the number of services with an aggregation of P-minutes. Our goal is to predict at the instant t how many services will be demanded during the period [t, t+P] at each existent taxi stand, reusing the real service count on [t, t+P] extracted from the data to do the same for the instant t+P and so on (i.e. the framework run continuously in a stream).

This model stands on three distinct pillars: 1) periodicity; 2) frequency and 3) a sliding time window. The first two correspond to two distinct forecast methods cleverly aggregated by the third one. The demand on taxi services exhibit, like other transportation means, a 1) periodicity (see Fig. 5) in time on a daily basis that reflects the patterns of the underlying human activity: so we used both seasonal and non-seasonal time varying Poisson models [17] to predict the service demand. When the frequency is distinct from the expected one (i.e. weekly), we have to handle it. The Autoregressive Integrated Moving Average (ARIMA) [3] models were used to such task by its well-known regressive properties. Finally, the models are ensemble using a 3) sliding time window: both models are used to build a weighted average where the weights are the models accuracy in the last X time periods (i.e. where X is a predefined time window that will slide forward constantly).

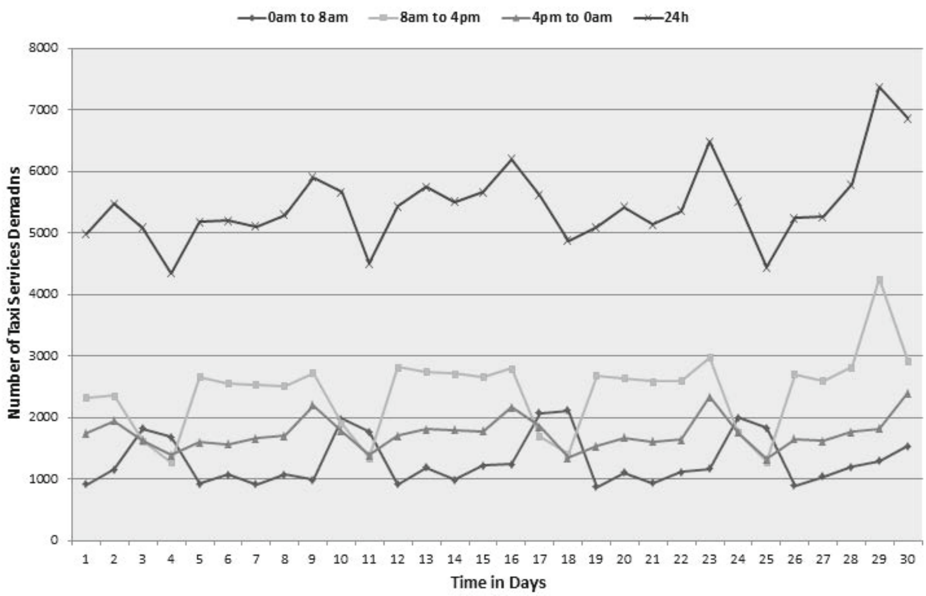


Fig. 5. One month taxi service analysis (total and per driver shift). There is a clear week periodicity in the data.

We applied this model to data from a large-sized sensor network of 441 taxi vehicles running on the city of Porto, Portugal. In this city, the taxi drivers must pick a route to one of the existing stands after a passenger drop-off. Our test-bed was a computational stream simulation running offline using 4 months data. The results obtained were promising: our model accurately predicted more than 76% of the services that actually emerged. This model will be used as a feature of a recommendation system (to be done) which will produce smart live recommendations to the taxi driver about which taxi stand he should head to after a drop-off.

7 Conclusions

The emergence of wireless networks in vehicles combined with the computation capabilities, mobility and energy-autonomy of modern in-vehicle sensor networks has the potential to create an ubiquitous platform for a wide range of sensing applications. Vehicular sensing allows to collect massive and varied information through sensors available at vehicles and to disseminate it using wireless communications, either using V2V/V2I or cellular networks. In this paper we have presented and demonstrated the feasibility/performance of three applications of vehicular sensing, namely automatic map construction and updating based on GPS traces, license-plate monitoring using windshield-installed cameras and taxi demand prediction per taxi stand over time, which allows a better understanding of urban dynamics.

References

1. Urbisnet: Urban pollution monitoring using a public transportation infrastructure for networked sensing (2012), http://users.isr.ist.utl.pt/~jpg/proj/urbisnet/urbisnet_main.html
2. Barcelo, J., Montero, L., Marques, L., Carmona, C.: Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring. Transportation Research Record: Journal of the Transportation Research Board 2175, 19–27 (2010)
3. Box, G., Pierce, D.: Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association, 1509–1526 (1970)
4. Bruntrup, R., Edelkamp, S., Jabbar, S., Scholz, B.: Incremental map generation with GPS traces. In: Proceedings of IEEE Intelligent Transportation Systems, pp. 574–579. IEEE (2005)
5. Caliskan, M., Graupner, D., Mauve, M.: Decentralized discovery of free parking places. In: Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks, pp. 30–39. ACM (2006)
6. Commission, E.: Commission Recomendation of 8 September 2011 on support for an EU-wide eCall service in electronic communication networks for the transmission of in-vehicle emergency calls based on 112 (eCalls). Official Journal of the European Union 54(L303), 46–47 (2011)
7. Conceição, H., Damas, L., Ferreira, M., Barros, J.: Large-scale simulation of v2v environments. In: Proceedings of the 2008 ACM Symposium on Applied Computing, pp. 28–33. ACM (2008)
8. Conceição, H., Ferreira, M., Barros, J.: On the Urban Connectivity of Vehicular Sensor Networks. In: Nikoletseas, S.E., Chlebus, B.S., Johnson, D.B., Krishnamachari, B. (eds.) DCOSS 2008. LNCS, vol. 5067, pp. 112–125. Springer, Heidelberg (2008)
9. Conceiçao, H., Ferreira, M., Barros, J.: A cautionary view of mobility and connectivity modeling in vehicular ad-hoc networks. In: IEEE 69th Vehicular Technology Conference, VTC Spring 2009, pp. 1–5. IEEE (2009)
10. Douglas, D., Peucker, T.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization 10(2), 112–122 (1973)
11. Edelkamp, S., Pereira, F., Sulewski, D., Costa, H.: Collaborative map generationsurvey and architecture proposal. In: Urbanism on Track Application of Tracking Technologies in Urbanism. Research in Urbanism Series, vol. 1. IOS Press, Amsterdam (2008)

12. ElBatt, T., Saraydar, C., Ames, M., Talty, T.: Potential for intra-vehicle wireless automotive sensor networks. In: 2006 IEEE Sarnoff Symposium, pp. 1–4. IEEE (2006)
13. Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., Balakrishnan, H.: The pothole patrol: using a mobile sensor network for road surface monitoring. In: ACM MobiSys (2008)
14. Ferreira, M., Conceição, H., Fernandes, R., Reis, R.: Locating cars through a vision enabled vanet. In: 2009 IEEE Intelligent Vehicles Symposium, pp. 99–104. IEEE (2009)
15. Google: Self-driving car test: Steve mahan (2012), `http://goo.gl/k5K9Q`
16. Hu, S.C., Wang, Y.C., Huang, C.Y., Tseng, Y.C.: A vehicular wireless sensor network for CO_2 monitoring. In: IEEE Sensors, pp. 1498 –1501 (October 2009)
17. Ihler, A., Hutchins, J., Smyth, P.: Adaptive event detection with time-varying poisson processes. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 207–216. ACM, New York (2006)
18. Jiang, D., Delgrossi, L.: IEEE 802.11 p: Towards an international standard for wireless access in vehicular environments. In: IEEE Vehicular Technology Conference, VTC Spring 2008, pp. 2036–2040. IEEE (2008)
19. Lee, J., Park, G.-L., Kim, H., Yang, Y.-K., Kim, P., Kim, S.-W.: A Telematics Service System Based on the Linux Cluster. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4490, pp. 660–667. Springer, Heidelberg (2007)
20. Lee, J., Shin, I., Park, G.L.: Analysis of the passenger pick-up pattern for taxi location recommendation. In: International Conference on Networked Computing and Advanced Information Management, vol. 1, pp. 199–204 (September 2008)
21. Li, B., Zhang, D., Sun, L., Chen, C., Li, S., Qi, G., Yang, Q.: Hunting or waiting? discovering passenger-finding strategies from a large-scale real-world taxi dataset. In: 2011 IEEE International Conference on Pervasive Computing and Communications Workshops, pp. 63–68 (March 2011)
22. Li, X., Shu, W., Li, M., Huang, H.Y., Luo, P.E., Wu, M.Y.: Performance Evaluation of Vehicle-Based Mobile Sensor Networks for Traffic Monitoring. IEEE Transactions on Vehicular Technology 58(4), 1647–1653 (2009)
23. Lima, F., Ferreira, M.: Mining spatial data from GPS traces for automatic road network extraction (2009)
24. Liu, L., Andris, C., Biderman, A., Ratti, C.: Uncovering taxi drivers mobility intelligence through his trace. IEEE Pervasive Computing 160 (2009)
25. Mahmud, S., Alles, S.: In-vehicle network architecture for the next-generation vehicles (2005)
26. Moreira-Matias, L., Gama, J., Ferreira, M., Damas, L.: A predictive model for the passenger demand on a taxi network. In: Intelligent Transportation Systems Conference (to appear, September 2012)
27. Norton, H.N.: Handbook of Transducers. Prentice Hall (1989)
28. Rodrigues, J., Vieira, F., Vinhoza, T., Barros, J., Cunha, J.: A non-intrusive multisensor system for characterizing driver behavior. In: International IEEE Conference on Intelligent Transportation Systems, pp. 1620 –1624 (September 2010)
29. Schroedl, S., Wagstaff, K., Rogers, S., Langley, P., Wilson, C.: Mining GPS traces for map refinement. Data Mining and Knowledge Discovery 9(1), 59–87 (2004)