Design of a Real-time Crowdsourced Mobility Sensor for Public Transportation Networks

Marc Pous, Daniel Villatoro, F. Xavier Mercadal and Arol Viñolas
Barcelona Digital Technology Center, Barcelona, Spain

Keywords: Smartcities, Public Transportation Network, Crowdsensing.

Abstract: Recently, and with the ever-increasing interest in *Smart-city* paradigm, urban mobility optimization is becoming an active research area. We are specially focused on the real-time optimization of the Public Transportation Networks, and the passenger flows. In order to achieve such task, it becomes essential to know the state of the network and its load of passengers in real-time. To the best of our knowledge, no existing infrastructure is able to keep track of the amount of public transportation users (and their position) in real time. In this paper we propose a framework architecture that transforms the passengers in active sensors and profit from the information they provide to estimate the actual usage of the network. In order to create an engaging experience for passengers (and incentivize the activation of the mobility sensor) we have implemented a higher level layer which offers users with a trivia-like game, where users compete against other users in real-time, while providing their position on the network.

1 INTRODUCTION

Catalyzed by the Industrial Revolution, cities have become the acting scenario for the economic tradings and exchanges that have derived into the modern economic system. Due to several factors (such as a reduction of commuting times and increase the interaction capabilities), the population within the cities suffered an outstanding growth and becoming such organizational structure the predominant and preferred for human interactions.

With the specific objective of improving urban mobility, it is necessary to have a complete understanding of the mobility patterns of the citizens. However, and up to the best of our knowledge, there is no complete dataset that allows us to obtain such patterns. With the penetration of mobile technologies, citizens have acquired the capability to continuously share information anytime anywhere. This continuous sensing capability allows us to combine our sensorized object (the citizens) and the sensor its tracks its behavior (its mobile device)

Our research focuses on the exploitation of the human sensing capabilities to understand their mobility patterns in an urban environment, and for the sake of specification we target commuters as our specific case of study. Specially, we present the design of a real-time crowdsourced mobility sensor in public transportation networks whose usage is incentivized through gamification. A trivia-like game serves as an incentive for public-transportation users to activate the designed sensor, and provide our platform with real-time position information of users, obtaining a complete picture of the state of the network with up-to-date information. The specific case of study that we focus on is the city of Barcelona. Nowadays the Barcelona transportation agency, among others, does not have methods neither infrastructure to know the exact number of people travelling in real time in its transportation network (composed by bus, metro, tram and train). With the raise of the computing capabilities in mobile devices and the appearance of the Big Data paradigm, it seems appropriate to profit from the multidisciplinar and interrelated sources of real-time data and process them in an unified intelligent core, maximizing the information that can be obtained from that data. One specific and interesting advance proposed is the usage of Case-Base Reasoning (CBR) to infere the correct location of users from the GPS position provided by them when travelling in the Metro. This is a challenge as the GPS positioning return poor positions when the application is used underground, and therefore, the CBR will correct that position to what it should really correspond after an adequate training.

The paper is organized as follows: in section 2 we present a brief summary of the problem and a discussion related on the improvement of public trans-

Pous M., Villatoro D., Mercadal F. and Viñolas A..
Design of a Real-time Crowdsourced Mobility Sensor for Public Transportation Networks.
DOI: 10.5220/0004224402310235
In *Proceedings of the 2nd International Conference on Sensor Networks* (SENSORNETS-2013), pages 231-235
ISBN: 978-989-8565-45-7

portation and available public metrics. In section 3 we describe the modular solution proposed to detect commuters' real-time location. Section 4 presents the actual intelligent analysis tasks that are performed to overcome some of the problems we have faced. In section 5 we present some related works and their contributions to the field. In section 6 we summarize our on-going work for analysing the commuters locations and state of the Barcelona's public transportation network in real-time. Finally, in section 7 we sketch the next steps that will be taken to improve the performance of the platform

2 THE PROBLEM: THE LACK OF REAL-TIME PUBLIC DATA

The optimization of urban processes such as the public transportation system has been subject of interest of researchers for many years/for decades (Lynch, 1960; Mandl, 1980; Daganzo, 1997). Some public transportation networks such as "Transport for London" or Seoul metro, have information about the users entrances and exits controlled through RFID cards or similar systems, used for optimizing the payment costs using a *"pay-as-you-go"* approach. However, and to the best of our knowledge, there is no existing infrastructure that provides the load of the network in real-time. In this paper we sketch the structure of the architecture that would allow public transportation users to become active sensors and provide this data while they travel.

This information is of clear importance to any public transportation system, as it has been proven by the numerous costly and time-consuming interviews performed to passengers asking about their origins, destinations, preferred routes and other parameters. Moreover, providing real time information about the state of the network would increase the awareness of the urban flows within cities through public transportation systems, and pop up the weakness of the networks and the behaviours of the commuters in front of the incidences or big events, amongst many others.

In this paper, we propose a platform of an integrated solution that allows the real-time tracking of participants within the transportation network. To achieve this task, we rely on the sensing capabilities of passengers (acquired through their mobile devices such as smartphones) and the continuous provision of information. However, we face two problems with this approach: (1) this sharing implies a cost for the individual user in terms of battery life and bandwidth that users are not always willing to sacrifice, and (2) we need to critical mass of users to obtain significant sample of information. The individual voluntary provision of information (for the intelligent analysis and process in our platform) would represent a long-term benefit for the users, despite the associated analyzed costs. However, battery life is a valued resource that users are always reluctant to share for communal goods, therefore, we need to provide short term positive incentives to motivate humans to provide with this data. Moreover, we need to ensure that a critical mass of users will participate providing information. To ensure both problems we opt for the usage of gamification techniques, translating these incentives into our platform in terms of a competitive game, that provides users with distraction during their trips, compensating for the resources consumed, and it is attractive enough to engage a sufficient amount of users.

As an specific case of study we want to study the commuting trends of Barcelona, being the commuters a representative portion of the users of the public transportation networks. Once we obtain real-time sample data of mobility, we will be able to understand the commuting trends and urban flows within the city, and consequently we would be able to analyze the weak points that can be reached by us to improve the system. However, this intelligent processing imply also the creation and usage of a common language that refers to each of the units of information treated in this process.

3 FRAMEWORK ARCHITECTURE

The solution proposed for the cooperative GPS tracking in the Barcelona's public transportation network is based on the gamification concept to incentivize users. We defined and developed a game users play while an underlying application gathers the geopositon (latitude, longitude and altitude) of the gamers with the aim of being aware of the Barcelona's public transportation usage in a city in real-time. The goal of the application is the cooperative anonymous real-time monitorization of the commuters, in the public transport network. Commuters seems to be ideal candidates for this research as they have internalized the commuting route that they transit everyday, and their cognitive capabilities can be re-focused on playing our game. While commuters play on their mobile device, they will be competing in real-time with other players, while sending information to our mobility server. This solution would allow us to track the whole Barcelona's public transportation system in real-time creating an incentive for the commuters to

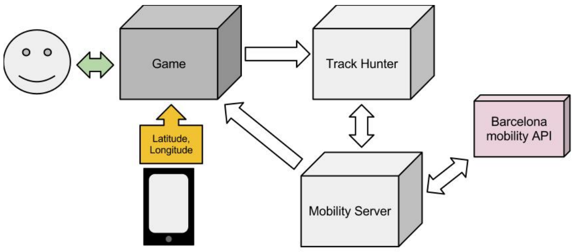

Figure 1: Crowdsourced Public Transportation Sensing Framework Architecture.

use this game.

The complete architecture (shown in Figure 1) is formed by three essential interacting components, which are described below.

- **Track-Hunter.** The mobile application, that runs on the users device, gathers the GPS position and sends it to our server. This component creates a message that contains the specific GPS position (obtained with the native GPS libraries of the mobile device), the timestamp for which that position was obtained, and the user introduced nickname, that will allow the continuous track of an specific user without compromising his privacy.
- **Game Layer.** Designed as a incentive mechanism, the Track-Hunter application is complemented with a trivia-like game, where players compete in real-time with other players. The game itself, combined with the real-time competition factor creates an engaging and satisfying interface for which users obtain a trade-off for the sacrificed battery-life consumed by the Track-Hunter, as well as the attraction to open the application every time they commute.
- **Mobility Server.** The client application sends (using a REST API[1]) the gathered information to a server that recollects all the provided information in real-time. This server has been designed to handle concurrent requests on real-time, implemented with *Node.js*[2].

These explained components are essential for the data acquisition process. However, once the information is acquired we need to process it in order to obtain knowledge about it. This knowledge extraction capabilities are integrated in what we have defined as the *Intelligent Layer*. This intelligent layer will be able to provide specific information and detailed analysis of the commuting trends, specifically obtaining information such as the most used routes, transportation modes, or peak hours.

[1]Due the simplicity and easiness of the REST protocol over SOAP and WSDL, most of the new web services are being build using REST.

[2]http://nodejs.org/

4 INTELLIGENT LAYER

In this section we will analyze the different problems we have faced after the design of the general architecture. These problems deal with the treatment of a huge amount of raw data obtained in real-time from the sensors and its transformation into manageable information in order to extract knowledge.

4.1 Data Model and Ontology: A Network Approach

Commuters provide us with a continuous feed of locations and timestamps, which is almost straightforward that are associated to movements. However, these movements are constrained by the actual transportation infrastructure. This restriction is understood as a positive point in terms of information storage optimization. Moreover, the public transportation network can be easily transported to the mathematical paradigm of networks and profit from all the theory and operators developed: stations (where passengers get in and out) are the nodes of the network, and the connections amongst stations are the edges. As users are mobile, each node or edge might have a different weight. This weight would be changing dynamically, depending on different interests, such as, distance, time spent for travelling from one node to another, or the number of people travelling, amongst other measures.

These first decisions are part of an specific problem that we have faced in this challenge, which is the development of a common specific vocabulary that fixes the different types of units and measures that the system will handle. The objective of the public transportation mobility ontology is twofold: (1) characterize the units necessary for our designed optimization processes, and (2) separate the information acquisition process from the actual calculations, allowing for more transparent and differentiated processes that do not interfere.

Moreover, the proposed ontology is a passenger-centric ontology as opposed to the majority of the analyzed ontologies in the literature which are mainly transport centric models (Houda et al., 2010). The ontology designed for our platform focuses on the real-time location of the commuters, their journeys and the different paths (changes in transportation modes or transportation lines), within the public transportation network. From those journeys it is possible to extract the stop points and edges, and on the other side, it is possible to infer the real-time position of the sample of vehicles with the commuters in them but not taking into account the schedule or fares.

4.2 Dealing with Poor GPS Signals

Given the precision of the GPS signals obtained from the users devices, it seems difficult that a mobile device will provide us with the accurate exact position that is saved in the reference repository, becoming therefore an obstacle the association of certain GPS location to an specific mobility resource. The most straightforward solution is to construct a simple Euclidean Distance Filter that given a certain GPS position and a tolerance level (which determines the radius of acceptability of distance to the mobility resource), it will return the specific station (or edge) associated. This initial solution seems to work fine for the overground transports, but obtaining poor performance in underground transportation, given the lack of precision of the GPS signal.

Our initial test have proven that despite obtaining a significant reduced number of GPS positions when using underground transportation, some positions are obtained. Moreover, we do not only have to deal with the reduced amount of positions, but also with problems in the precision of the data obtained (e.g. when being in a certain subway station, the GPS system will return a GPS position that is 50m. away from that position). In order to solve this problem we propose a semi-supervised learning approach, using a Case-base Reasoning (CBR) technique.

Initially, a knowledge base needs to be gathered, by transiting the whole network, attaching tags with the real semantic location (e.g. Metro Station: *Diagonal*, Bus Station: *574*, or Edge: *Diagonal - Verdaguer*) to every GPS position sent. This knowledge base will allow the training of the intelligent system, that will later associate any received GPS position to the most similar case obtained from the knowledge base.

4.3 Detecting the Transportation Mode

Strictly related with the previous problem, we will need to identify the type of transportation that the user is using, to infer its real location in the network. Given the different rates of information arrival from the client application depending on the transport used (low rates in subway and higher rates in overground), we propose a simple rule-based classifier, that will intelligently categorize the transportation mode of a certain user from a subset of its provided positions. One rule example generated by this classifier would be:

```
IF(Samples with GPS Position > 70 %)
THEN Bus
```

5 CONCLUSIONS AND FUTURE WORK

Our contribution proposes a gamificated cooperative methodology for tracking urban mobility footprints, contrasting with traditional survey methods, to improve the understanding of commuters behaviour and public transportation mobility dynamics.

In the designed system, commuters run an application with a competitive game while they are anonymously sending their real-time location, allowing us to know the state of the network and the weak points. On the other side the application would report real-time incidences to the users and would recommend alternative routes to arrive to their usual destination (in case the system have historical data of user's destination).

At the moment we only focus on capturing the information from the commuters and we are aware that we will be gathering only a subset of potential commuters to play our game. Given the openness of the system, and the nature of any crowdsourced application, it is hard at the moment for us to evaluate which is the validity of our sample. This issue remains as an open problem to be solved in the future.

As one of the main tasks for the future work, the server will be given with an intelligent multi-modal transportation recommender engine which will calculate personalized routes for users in order to optimize his commuting journey depending on the real-time status of the public transportation network. As an ambitious scientific challenge, it will be interesting to develop a behavioural model of commuters that could be integrated with the discoveries provided by other researches about the mobility pattern knowledge of humans, resulting in a predictive system. This information would be highly valuable to understand the overall behaviour of such complex system, specially when facing special events that might produce an overload of the network.

ACKNOWLEDGEMENTS

This work has been completed with de support of ACC1Ó, the Catalan Agency to promote applied research and innovation.

REFERENCES

Daganzo, C. F. (1997). *Fundamentals of Transportation and Traffic Operations*. Elsevier Science.

Houda, M., Khemaja, M., Oliveira, K., and Abed, M. (2010). A public transportation ontology to support user travel planning. In Loucopoulos, P. and Cavarero, J.-L., editors, *Proceedings of the Fourth IEEE International Conference on Research Challenges in Information Science, RCIS 2010, Nice, France, May 19-21, 2010*, pages 127–136. IEEE.

Lynch, K. (1960). *The Image of the City*. The MIT Press.

Mandl, C. E. (1980). Evaluation and optimization of urban public transportation networks. *European Journal of Operational Research*, 5(6):396–404.