The real-time Barcelona urban mobility with NoSQL technologies

Marc Pous, Daniel Villatoro, F. Xavier Mercadal and Arol Viñolas

Recently, and with the penetration of the *smart city* paradigm, urban mobility is becoming an active research area. We are specially focused on the real-time urban sensing using citizens as data sources. The explosion of mobile technology allows the transformation of citizens in active and passive sensors increasing the amount of information of our scope of study: the city. The usage of citizens as sensors results in low-cost infrastructureless framework that not only profits from the mobile facet of humans but also from their common sense detecting anomalous situations. These aspects are treated in two different scenarios specific that generates Big Data for the city of Barcelona: one that allows us to obtain information about the state of the public transportation network and the other that detects clusters of activity within the urban environment through social networks.

In this paper we propose two complementary ways to exploit the human sensing capabilities in order to understand their mobility patterns in an urban environment; namely, the information extraction of social networks and the acquisition of more accurate position information, recommendations and network optimizations through the use of gamification techniques. Moreover, we describe our experience and barriers encountered when deploying our solution using NoSQL technologies (based on documents and graphs) with the goal of understanding urban processes such as the identification of transitory areas of activity and mobility patterns in the public transportation network (PTN).

Urban macro vision through social networks

Because of their popularity and geo-positioning capabilities, we opted to use Twitter as the first source of information. This social sensor provides us with information (in the form of geolocalized tweets) to observe and detect alterations of several parameters such as areas of activity, temporal patterns or mobility routes, which define the “Urban Chronotype of the city” [1].

Both processes of tweet-collection and analysis presented certain technical barriers that have been solved with the development of our Urban Sensing Platform. This platform ensures the acquisition in near real-time of all the tweets, without the requirement of any private partnership with Twitter as commonly done by other researchers, allowing deployment in any system. Moreover, it is important to understand the role of people within the city to better understand their behaviour. Our platform has been enriched with a semantic-intelligence module that contextualize the geographical origin of users, obtained from their Twitter profile. Categorizing users as locals or foreigners with respect to the city of study allows for a detailed behavioural

pattern analysis differentiating their urban habits.

The Urban Sensing Platform is technically based on a system that gathers the Twitter real-time streams [https://dev.twitter.com/docs/streaming-apis] and it stores the streaming of tweets into a Hadoop system. The goal of the platform is the persistence but also the extraction of knowledge of these tweets using data mining algorithms with Mahout [http://mahout.apache.org/]. We also tested the query system using Hive [http://hive.apache.org/].

We performed an experiment during the controlled and exceptional situation of the celebration of the Mobile World Congress (MWC) in Barcelona. Our knowledge about this specific event allowed us to compare the average Urban Chronotype of the city (in the normal state without external influences), with that obtained during the event, allowing us to observe different behavioural patterns within the city. The empirical results obtained have shown the behavioural patterns of the city in its normal state and during the event, where substantial differences have been observed in the areas of activity of the city after applying geospatial data mining techniques, determining the success of our urban social sensor.

Urban micro vision using citizens as sensors

However, profiting from the information proactively shared by the citizens in the social media applications only gives us access to a restricted subset of the potential information that represents the Urban Chronotype of the city. Specifically, optimizing the urban public transportation network (PTN) in real-time is a research challenge intensively studied. However, to achieve such optimization it is necessary to have complete knowledge of the state of the network. Unfortunately, the transportation company's existing infrastructure does not provide complete information of the load of the network (in cities such Barcelona it is only controlled the access stations of travellers, but not the exits neither the load of the vehicles or the transportation lines), and the social media applications (as the previously used) do not provide the incentives for users to continuously share their position within the network.

For such reasons, and to obtain more accurate real-time information about the near real-time state of the PTN, we have created an ad-hoc sensor in the form of a app that profits from the sensing capabilities of travellers and, while open, continuously share the citizen geoposition (latitude, longitude and altitude) anonymously to our centralized server. However, we faced two problems: why will users open an app that sacrifices part of their battery life? and how can we obtain enough critical mass of users to make this crowdsourced sensor work? To solve both questions we have created an upper layer using gamification techniques. With a competitive real-time trivia-like game that incentives users to use our app: the game serves as an individual incentive to open the app, and the real-time competition with other users and friends ensures the virality and the long-term usage of the app.

With the gathered real-time data from the commuters of Barcelona we are able to understand the load of the PTN, obtaining also a commuter behavioural profile in the system in various situations (e.g. demonstrations, or unexpected incidents). Our modelization using a dynamic network data structure allows us to apply graph theory algorithms to optimize the behaviour of the system in several aspects and even build a route recommender system.

In this case, our system is a complex real-time architecture with document based databases, in this case MongoDB [http://www.mongodb.org/], to store the user location in real-time and the knowledge inferred from these raw data, stored in another different MongoDB. As a part of the architecture there is also a graph based database, using Neo4j technology [http://neo4j.org/], which models the public transportation network and optimize the shortest path among two stations through the implemented algorithms (e.g. Dijsktra). The system updates the edges properties in real-time through the MongoDB Geospatial Indexing [http://www.mongodb.org/display/DOCS/Geospatial+Indexing/] and the Neo4j API REST in order to offer to the players the most optimized route (depending on the properties chosen by the user, such as, fastest route, less crowded route or less pollutant transportation system) to their destination.

Conclusions

These two specific applications have given us different but complementing visions of the urban flows through the Big Data: a macro vision, captured from the proactively shared social media contents, and a micro vision, with dedicated apps sensing an specific subject such as the PTN.

Nowadays, the concept of smart city appears with a top-down vision where citizens receive the services implemented by the city and the government. As we proposed, the proper usage of the collected information would provide a more complete picture (micro and macro) with near real-time information of the object city allowing for a significant improvement of the urban performance, with no specific investment on hardware infrastructure. This idea is in alignment with the bottom-up paradigm (defended by other researchers such as Jane Jacobs, Usman Haque and Adam Greenfield) where build-up cities need the interaction of the citizens with the widespread use of their mobile devices and geo-positioned applications providing huge amount of information of their environment in real-time.

Acknowledgements

This work has been completed with the support of ACC1Ó, the Catalan Agency to promote applied research and innovation, and by the Spanish Centre for Development of Industrial Technology under the INNPRONTA program, project IPT-20111006, “CIUDAD2020”.

References

[1] Daniel Villatoro, Jetzabel Serna, Víctor Rodríguez and Marc Torrent. “The TweetBeat of the City: Microblogging Used for Discovering Behavioural Patterns during the MWC”, Citisen 2012, ECAI 2012, Montpellier, France, 2012.