The Future Role of Social Networks in Urban Planning

Enrique Frías Martínez | March 13th 2015

Citizens are continuously generating information about where they are located and what are they doing through LBSN.

The concept of *smart cities* typically implies the deployment of **intelligent infrastructures used to improve services for citizens**. A number of cities have launched projects that range from optimizing transport to minimizing the impact of urban activity on the environment, being a good example the work done **by Telefonica in** SmartSantander.

Although the deployment of such infrastructures is key, any city already has an implicit sensor infrastructure build using the phones carried by their citizens. In that sense, the increasing capabilities of mobile devices allow individuals to leave behind footprints of their interaction with the urban environment, particularly with the use of **Location-based social networks (LBSN)**.

LBSN have the property of assigning a geolocation to the activity done by an individual, for example, **geolocalized** Twitter **assigns the coordinates where the user was when a comment was made**; or FourSquare gives an indication of where a user is located and what is the activity of that particular place. This geolocation capability is being used by a large and increasing number of applications such as Facebook Places, Urbanspoon, Yelp, Google Places, Loopt, Flickr or Instagram, among others.

As a result, a **city and its citizens are continuously generating information** about where they are located and what are they doing through LBSN. This information provides an unprecedented view of a city, especially regarding how citizens use it and what are their opinions, which has the potential to be used for urban planning applications.

In Telefónica I+D we have been working in **two areas that use the data provided by LBSN** to showcase how they can be **used to improve urban planning**: (1) Inferring User Activity and (2) Understanding the complaints of citizens about their environment.


Using the location, context and information provided by LBSN we can infer user activity in a specific urban area and by extension we can help in the process of urban planning as we can check to which extent an urban environment is being used as it was planned by the authorities. We have **worked in a set of techniques [1][2] that using just the location of geolocalized tweets are able to identify land uses**. Our technique automatically determines land uses in

urban areas by grouping together geographical regions with similar patterns of Twitter activity. Using aggregate activity of tweets, we studied land use in **Manhattan, Madrid and London.** In the first two cases we identified four uses: residential, business, daytime leisure (mainly parks and tourist areas) and nightlife areas. In London, we also established industrial land uses. These results were validated with open data sources. The main advantage of our approach is that it respects individual privacy because no identifier of the actual user is needed nor the content of the twitter is used. As an example, Figure 1 presents the land uses identified in London using geolocalized tweets.
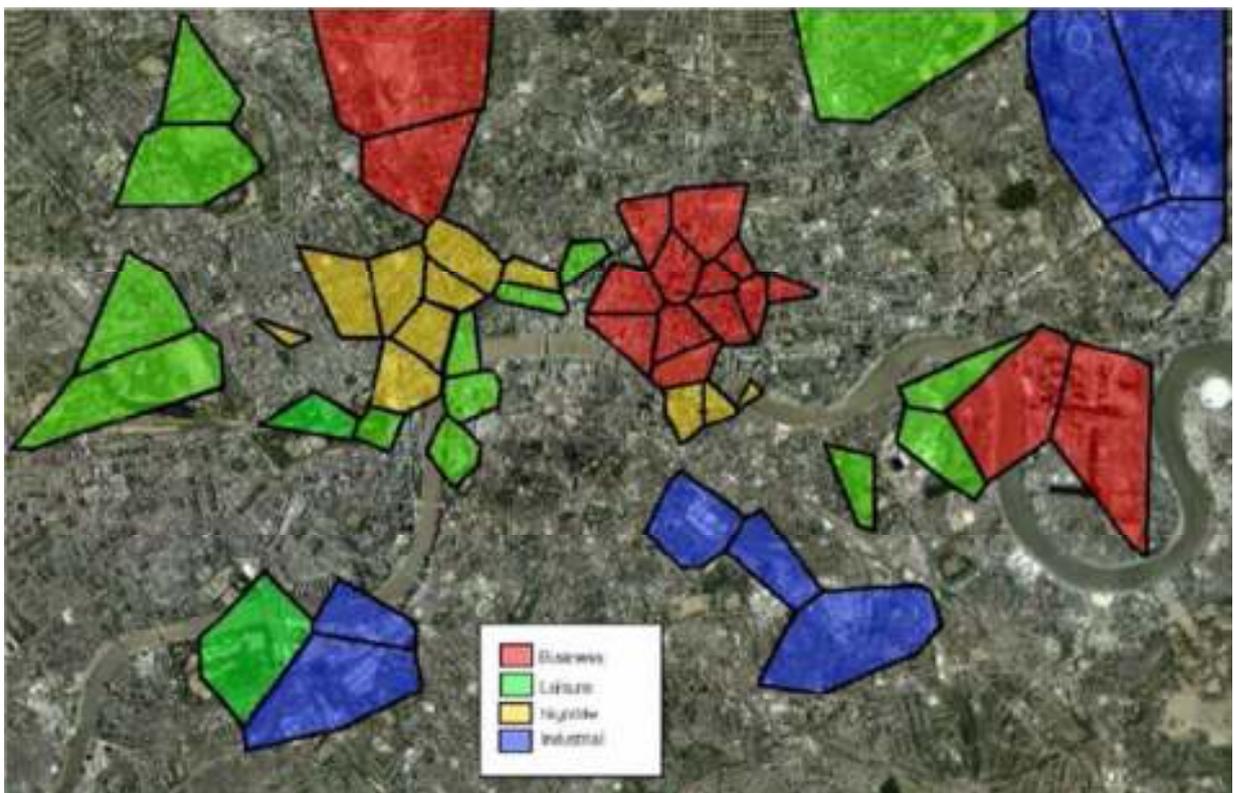
Physical layout of business, nightlife, leisure and industrial clusters in London. Areas not marked with any color indicate residential land use.

Another approach we have taken for the identification of user activities and land uses is using FourSquare [3]. FourSquare already contains information about the activity of the place where a user is located because places are classified according to predefined categories. By considering the distribution of FourSquare categories and their popularity in terms of check-in numbers, we built a set of machine learning features that can be exploited to infer the qualitative characteristics of urban areas. Figure 2 presents a classification of the city of Madrid according to the user activities identified from the FourSquare categories. Each cell at the tessellations is colored with the most popular activity characterized by exploiting the popularity of nearby FourSquare places.
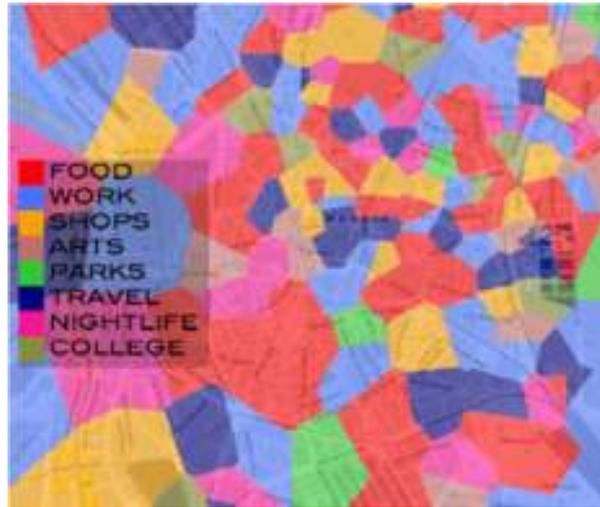
Classification of the city of Madrid according to the user activities identified from the FourSquare categories.

Regarding the identification of complaints in urban environments, **citizens represent the front line when capturing the pulse of the city**. In that sense, they are the first ones to identify problems with their environment. Traditionally city halls have phone-based services where individuals can report these problems. Recently, **Twitter accounts have been created by cities so users can send complaints and contact city hall and its services** (i.e., police, transportation department, metro, cleaning services, etc.) via Twitter.

In [4] we studied to which extent the complaints expressed by users using traditional phone services were similar to the ones they expressed using the specific twitter accounts created by city hall. We focused our study in New York City and used the open data sources available. Our results indicated that phone and specific Twitter complaints are very similar, i.e. the specific city hall twitter account already handles the same volume and type of complaints as the phone service. In any case, **they just provide a partial view of the problems of the city** and focus mainly on transportation and street related complaints.

After that, we evaluated to which extent these formal twitter complaints represent the general view of the city by comparing them with a set of general tweets that we identified as expressing complaints or areas of improvement. In this case, we showed that the use of generic tweets (i.e. tweets not directed at city hall departments) allowed for a global perspective about the areas to improve in the city, like noise complaints and lack of cleanliness.

LBSN are excellent sources of information to capture the pulse of an urban environment and as such can be used for a variety of urban planning applications like land use identification, identification of areas to improve in a city or transport planning. While each one of these data sources are very powerful by themselves the main challenge for the future will be to combine the information provided by different LBSN in a meaningful way.

References

[1] "Spectral Clustering for Sensing Urban Land Use using Twitter Activity", V. Frias-Martinez, E. Frias-Martinez, Engineering Applications of Artificial Intelligence, Vol. 35, October 2014, Pages 237–245, 2014

[2]"Characterizing Urban Landscapes using Geolocated Tweets", V. Frias-Martinez, V. Soto, H. Hohwald, E. Frias-Martinez, 2012 Int. Conference on Social Computing (SocialCom), Amsterdam, The Nederlands, 2012

[3] "Exploiting Foursquare and Cellular Data to Infer User Activity in Urban Environments", A. Noulas, C. Mascolo and E. Frias-Martinez, IEEE Mobile Data Management 2013 (MDM 2013)

[4] "To Call, or To Tweet? Understanding 3-1-1 Citizen Complaint Behaviors", V. Frias-Martinez, A. Sae-Tang, E. Frias-Martinez, Sixth Int. Conf. on Social Computing, SocialCom 2014, Stanford, CA, USA