The TweetBeat of the City: Microblogging Used for Discovering Behavioural Patterns during the MWC2012

Daniel Villatoro, Jetzabel Serna, Víctor Rodríguez, and Marc Torrent-Moreno

Barcelona Digital Technology Centre, Spain
{dvillatoro,jserna,vrodriguez,mtorrent}@bdigital.org

Abstract. Twitter messages can be located in a city and take the pulse of the citizens' activity. The temporal and spatial location of spots of high activity, the mobility patterns and the existence of unforeseen bursts constitute a certain Urban Chronotype, which is altered when a city-wide event happens, such as a world-class Congress. This paper proposes a Social Sensing Platform to track the Urban Chronotype, able to collect the Tweets, categorize their provenance and extract knowledge about them. The clustering algorithm DBScan is proposed to detect the hot spots, and a day to day analysis reveals the movement patterns. Having analyzed the Tweetbeat of Barcelona during the 2012 Mobile World Congress, results show that a easy-to-deploy social sensor based on Twitter is capable of representing the presence and interests of the attendees in the city and enables future practical applications. Initial empirical results haven shown a significant alteration in the behavioural patterns of users and clusters of activity within the city.

1 Introduction

In the last decades, the number of inhabitants in urban spaces has enormously increased. In addition to being a place where people dwell, cities have become the center of human activities, the place where people move, work, play, learn, buy, sell and experience emotions. From a high-level perspective, a certain *beat* of the city can be perceived, and its realization as measurable figures and precise facts has always been of the highest interest for sociologists, entrepreneurs, urban planners, policy makers, or just for mere observers.

To achieve the understanding of citizens' behaviour in the cities, information has been traditionally gathered via random polling and indirect measures, when not directly from the mere intuition. These approaches however, imply the main disadvantage of obtaining biased information (i.e. polled participants not providing real information, having partial sampling), and often require expensive procedures that often require long execution time.

Fortunately, in the last few years, mobile technologies have gained massive spread among the citizens, and these can now be used as proactive sensors. Device capabilities improve constantly, prices drop and wireless connectivity infrastructures are becoming each time more universal. Indeed, device-holders have now the facility to continuously share all kind of information anytime and anywhere, and much of this information is public. Moreover, the proliferation of social networking has generated an incomparable

J. Nin and D. Villatoro (Eds.): CitiSens 2012, LNAI 7685, pp. 43–56, 2013.

and incentivizing framework for users to share any type of information about friendships (e.g. Facebook), work-colleagues (e.g. Linkedin), pictures (e.g. Flickr) or even the favourites dishes in the restaurants around (e.g. Foodspotting).

In this paper, we propose to capture the beat of the city and its alterations (provoked by the visitors activity) in the face of a public event by exploiting the public data offered by Twitter (i.e. a microblogging platform with a relative abundance of messages, geolocation capabilities and good temporal stamp for each point). Moreover, the portmanteau *Tweetbeat* conveys the idea that the beat of the city is reflected in the beat of Twitter.

Therefore, the scope of this research is twofold: (1) to develop a hardware-infrastructure-less social sensor whose observation target is the city (fed with the microblogging information individually provided by their users), and (2) to evaluate the viability of the newly developed sensor to build urban-behavioural models (e.g. HotSpots identification, mobility patterns and unforeseen events detection, etc), which from now on will be referred as *Urban Chronotypes (UC)*. To achieve this, Twitter seemed to be the ideal candidate because Twitter-users can attach the GPS position of the device from where they *tweet*, allows information public access, and this access is obtained in "near" real-time. Thus, we highlight the role of Twitter as a Social Sensor for Urban Chronotypes identification. Gathered information will allow us to identify the average UC, and therefore, be able to detect potential disturbances within a city.

Our case of study focuses on the city of Barcelona, host of the international event Mobile World Congress (MWC). To cover the necessary phases in the UC creating process, we implemented the Social Sensing Platform (SSP), which gathered all tweets in Barcelona during 3 weeks: one week before to the MWC, the MWC week and one week after the event. Once data was captured, the framework was capable of performing several types of knowledge extraction through statistical analysis, and clustering techniques, using the DBScan algorithm. The SSP allowed us to observe the average Chronotype of the city of Barcelona, and how the UC was affected by a major event such as the MWC.

The rest of the paper is organized as follows: Section 2 reviews the state of the art and similar contributions; in Sec. 3 we state the problem, and the case of study is determined in Sec. 4. In Sec. 5 we describe the modular architecture of our social sensor, and some initial statistical results are presented in Sec. 6. Later in Sec. 7 we present the results obtained after applying clustering techniques, and finally we draw some conclusions and sketch the future work in Sec. 8.

2 Related Work

The idea of examining the mobility patterns in a city during a certain event by observing microblog posts has not been directly considered, but a number of related experiments have been described.

The first piece of information about an event is its mere existence. Microblogging has actually been used as a sensor to detect both natural phenomena (like earthquakes [1], levels of pollen in the air [2] or even weather events [3]) and events of human nature (crime and disaster events [4], those that gather crowds [5] or general events in [6]).

Beyond the mere detection, these events have been also further characterised in order to extract useful information. Among this information, the spatial and temporal coordinates of the users have been a key aspect. The temporal description of microblogging posts or social media in general, together with a sentiment analysis can be used to anticipate events of any sort, like the commercial success of a movie [7] or even the stock market [8]. The spatial description of social media can be used to detect points of interest, like geolocated *flickr* photos for the tourist case in [9], which actually reaveal trajectories when it combines this information with the time stamps. The joint analysis of temoporal and spatial description of Twitter messages produces indeed richer results and it has been used to anticipate music popularity [10], political alignment [11] or general Twitter themes [12].

The rich information provided by telecommunication networks used to characterize the city dynamics in front of a certain event, as in [13], can be thus replaced by *public* information and far more sparse; as an example the regular beat of New York was well characterized from Tweets collected by [14].

From another point of view, [15] characterized the nature of different regions regarding mobility by proposing a first step of identification of the relevant areas (applying a k-means clustering algorithm to Tweets), and then a second step tracking the movement pattern with new users coming into a cluster or existing users leaving it. These movements were shown to be predictable with a semi-Markov model by [16], although in this case data was acquired from Mobile Social Networks where the posting frequency was much higher than that of microblogging.

3 Problem Statement

The Smartcity paradigm has recently received an ever-increasing level of attention from the scientific community. Optimizing urban processes is a research goal that gathers different scientific areas such as Policy-making, Computer Science, Urbanism and Sociology. Opposed to the classical passive continuous sensing approach (where system owners had to study special parameters such as where to locate the sensor or the sensing-frequency, and then detect anomalies in the normal behavior of the observed object/phenomenon), in this work we plan to profit from the human pro-active sensing capabilities (enabled by their mobile devices such as smartphones).

As we have seen in Sec. 2 other researchers have profitted from Social Media platforms to obtain information related to the urban behavior of users. Although, until now, experiments done with Twitter as a source of information have been performed by advantaging of the full Twitter support (i.e. not considering only publicly available information via the standard API, but a full opened connection which is not commonly granted to the general public). Thus, in this work we emphasize in, how this Social Sensor can provide us information to observe and detect alterations in the Urban Chronotype of the city, and, the viability of achieving it with a given dataset that is based on information publicly available without the requirement of any private partnership or special Social Domain participation. Moreover, we demonstrate that, the developed platform can be considered as a low-cost social sensor. Finally, by applying intelligent analyses (such as geospatial clustering), our main goal is to support decision making processes and identify important changes in the city in near real-time.

To achieve this goal, we profit from our knowledge and expertise in Smartcities and apply it to the city of Barcelona, together with the celebration of a worldwide event such as the MWC. This specific case of study is interesting, as it provides us of a controlled event (such as the MWC) where we have already a background knowledge inferred from previous editions (e.g. the amount of participants or location of the venue), and makes it an ideal candidate for the SSP initial experimentation. The controlled situation that the MWC offers, serves as a control test, allowing us to evaluate the performance of the platform even when facing unexpected results.

4 Case of Study

The city of Barcelona has been the host for the Mobile World Congress for six editions celebrated yearly (from 2006 to 2012) at *Fira de Barcelona - Plaza España Pavilion*. From the previous editions, we know that this event brings to the city a numerous amount of people (an average of 50.000 people in the last 6 editions, and 65.000 in the 2012 edition) from all over the world, with a common interest: mobile devices. Because of this common technological interest, we hypothesized that the infiltration level of Social Media applications within this community should be high. Therefore, the MWC attendees (as any other visitor) have an effect on the average Urban Chronotype of the city, with the slight difference that this effect might be reflected on the activity recorded by Twitter with a higher impact.

In order to observe this variation on the Barcelona Urban Chronotype, first we need to obtain the average Urban Chronotype that would work as a control case for us. The city of Barcelona was the host for the Mobile World Congress from February 27th to March 1st 2012, and as control cases we have decided to use the week before and after the event.

5 Social Sensing Platform Architecture

The proposed social sensing architecture is composed of 3 independent but interacting elements, as shown in Figure 1, each of them described below.

This platform receives as an input the City to be "sensed". Other parameters are optional, such as the hashtags of the potential event to track.

5.1 Tweet Hunter

This module is in charge of the information acquisition of the City, which in this specific case is gathered from the microblogging site Twitter (further versions of this platform will include other social media sources). Twitter allows us to query via a single streaming connection per user and IP by using the public Streaming API, obtaining near real-time information. Queries to the Twitter Streaming API can be of several types, but we focused on the geospatial type of queries, where, given a bounding box (delimited by the south-west and north-east coordinates) Twitter will return 'all' the Tweets generated within that area. Moreover, according to the Twitter streaming API Documentation[1], it

[1] How are rate limits determined on the Streaming API? on `https://dev.twitter.com/docs/faq`. Accessed on May 31st 2012

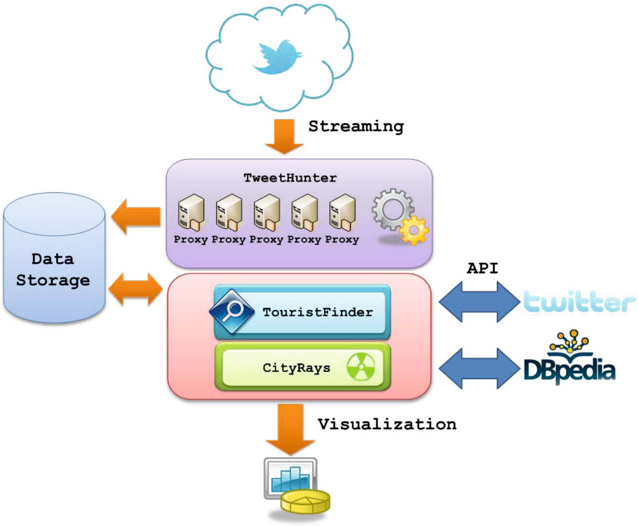


Fig. 1. Social Sensing Platform Architecture

provides a maximum of 1% of the global Twitter streaming (returning a message if this limits is exceeded), meaning that we could risk information loss, if the query response overpasses this limit. Thus, to reduce the probability of information loss considering the specific case of Barcelona, we have setup five proxies, each of them with an opened stream connection, targeting particular area of the city and configured with the targeted area related parameters. We have strategically selected four different and highly important spots in Barcelona (i.e. airport, main train station, tourist center and Fira Barcelona - epicenter of the MWC2012 event) and setup four of the collectors with each of the spot's bounding box. Additionally, a fifth collector with a bounding box that covered the whole city of Barcelona. By taking advantage of the bounding box configuration, we assumed that, *1)* the targeted spots (bounding boxes) were small enough and therefore not able to produce more than 1% of the global twitter streaming, and, *2)* information not captured by the fifth streaming - targeting the whole city, could be completed with the other four streaming connections (targeting the most crowded places), and ultimately the data loss could be neglected.

5.2 Tourist Finder

Some information about the users is handled by the Tourist Finder, and its main task is to determine the origin of the captured Tweet's users and place them in any of the following categories: *Local*, *Tourist* or *Unknown*. The content of the Tweet, which in some occasions has been used to locate the message ([17]) has been neglected in our experiment.

The Tourist Finder performs two important interactions:

1. With the Twitter REST API: to query about user's location. As each gathered tweet is accompanied with the user id that have originated that tweet, we can query

Twitter (via REST API) about the user's specified location in his own Twitter profile. In that way, our platform obtains a list of users and their locations.
2. With the DBpedia REST API: to query the set of keywords that identifies a city or region of the world. With the origin's dataset, this module builds up an ontology based on *1)* automatic queries to identify a set of keywords related to a particular city/place of the world, this is done thanks to the DBpedia API (which for this specific case returned us a complete list of all the *Populated Places*, such as cities, towns or villages, in the Catalonia Region), and *2)* a semi-automatic classifier that automatically extracts a set of unrecognized keywords (not identified by the previous process), and that, needs manual interaction to be able to identify when a location input is clearly undefined (e.g. "somewhere in the world", "in this planet", "Gaga's Heart", etc). Note that, since the location parameter in Twitter is an input to be entered by the user, an inherent truthfulness error probability will always exist.

After these two steps the Tourist Finder will be able to classify all the users in one of the following categories:

1. Locals: all users with a location parameter within Catalonia region.
2. Tourists: all users with a location parameter outside Catalonia.
3. Unknown: all users with no location or location undefined (location not recognized).

5.3 City Rays

This module is in charge of actually analyzing the data obtained from the gathered Tweets and extract knowledge out of them. It is also in charge of analyzing temporal and geospatial analysis. It is capable of extracting average behaviours in different temporal ranges, combining the information obtained from the *Tourist Finder* module, and it also performs spatial clustering techniques. This module however can be easily extended with new functionalities (e.g. spatio-temporal analysis).

6 Geostatistical Analysis

As previously specified, the experiment lasted for three uninterrupted weeks from Feb 20th (00:00:00) 2012 to March 11th (23:59:59) 2012. Table 2 describes the date ranges covered by each of the experiment weeks and its corresponding identifier. Week 1 and Week 3 are the control cases, although we understand that both weeks might be slightly affected by the event.

During these three weeks we gathered around 250,000 tweets from the Twitter Streaming API (generated by 15,911 different users), where the 43.10% of them contained the GPS coordinates associated to the user's geolocation when posting those tweets.

Figure 2 shows the total number of different Twitter Users (that tweeted at least once) in Barcelona with respect to those whose Tweet's contained specific geopositions. It is easy to observe the behavioural variation during Week 2. We hypothesize that the peak reached the first day of Week 2 is affected by the studied event (the MWC). Moreover,

Table 1. Experiment Weeks Identifiers and Coverage

Week Id.	Initial Date	End Date	Exp. Category
Week 1	02/20/2012	02/26/2012	Control Case
Week 2	02/27/2012	03/04/2012	Subject of Study
Week 3	03/05/2012	03/11/2012	Control Case

we can observe that the ratio of Geolocated tweets remains constant at around 40% during the three weeks[2] as it can be seen in Fig.2. This is an interesting result that lead us to think that the individual behavioural trends with respect to geo-positioning are not affected by the external event.

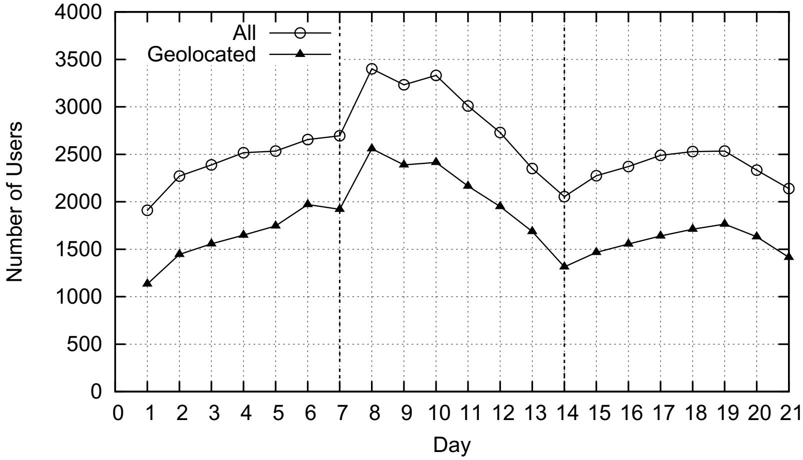


Fig. 2. Twitter Users and Geopositioned Users

However, we were interested in observing the specific trends of the event attendees. To do so, we extract the tweets that contained any information related to the event that for the sake of readability will referred as #MWC (although it refers to a set of hashtags and keywords related to the event such as, #MWC2012, #MWC12, MWC, "Mobile World Congress", etc).

Figure 3 compares the amount of different users per day that used at least once one of the #MWC related hashtags in a geolocalized tweet with those that did not geolocalize the tweet. We can observe that the event possesses a strong geolocalized facet wrt the average city trends shown in Fig. 2, since users actively provided their geoposition when they tweeted about this event. Moreover, we can observe that the curve reaches its peak during the event (as predicted by [18]), showing an initial activation effect two days previous to the event.

[2] We would like to remark that this ratio improves the the state of the art analysis performed in the literature, where the most successful identified case worked with a dataset that presented at geopositioning ratio of the 0.41%.

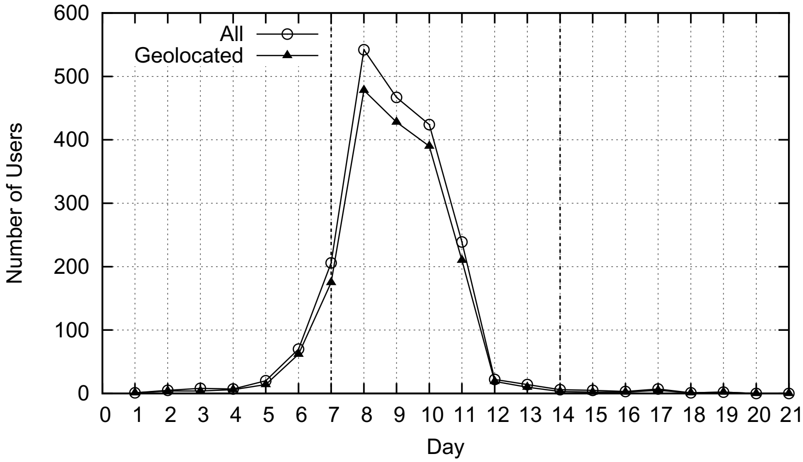


Fig. 3. #MWC User distribution

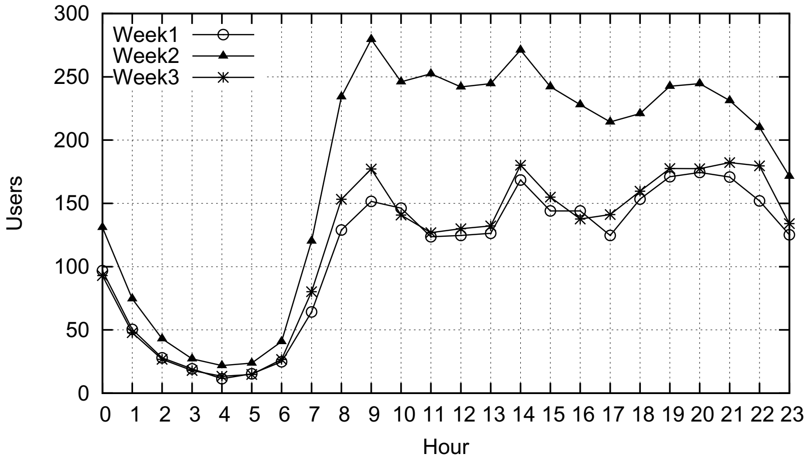


Fig. 4. Week comparison of Geopositioned Users

In order to assure the geoposition character of the event, we extract the average daily pattern of the amount of the different geopositioned twitters (users that tweeted with the specific geolocation at least once in a period of time) and compare the results of the three different weeks. In Figure 4 we can observe the substantial difference observed during Week 2. This result combined with the previous result lead us to hypothesize that users interested in the MWC event were also active geopositioning users, which can therefore provide us with relative information about their behavior within the host city.

In order to know more about the type of users that are influencing the city during the event, the *Tourist Finder* module help us to determine the origin of users. We focus only on the geopositioned users. Figure 5 shows how during the event, the number of *Tourists* increases up to the point of exceeding the number of *Locals* during the opening day.

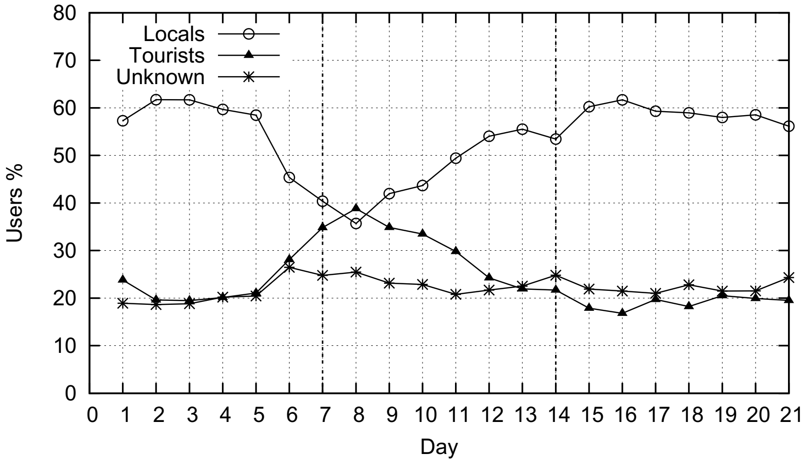


Fig. 5. Geopositioned Users Origin's Distribution

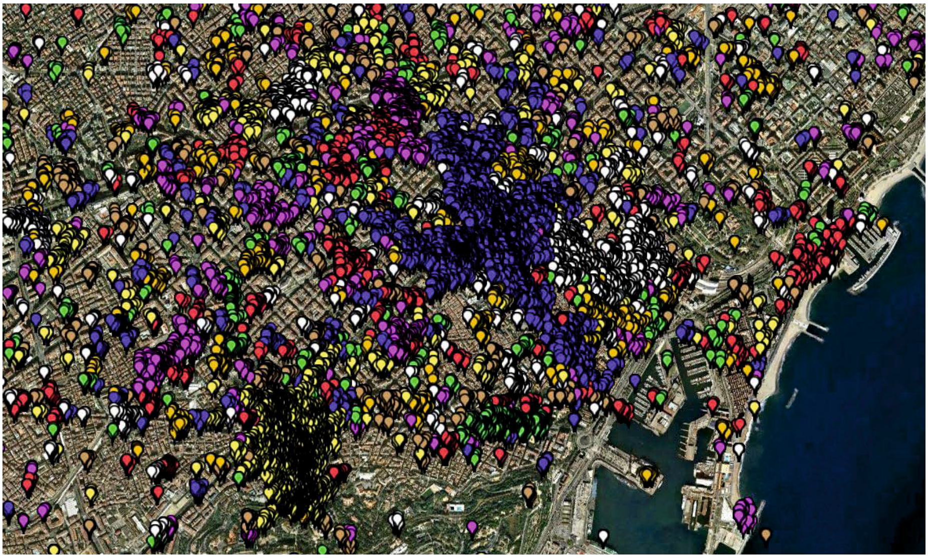
7 Geospatial Clustering Analysis

After a detailed comparative analysis of clustering algorithms with spatial data [19], we opt for DBScan, rather than other well-known clustering algorithms such as CLARANS, EM or *k*-means. The DBScan Algorithm [20] is a clustering algorithm that posses a number of characteristics that differentiates it from the other standard algorithms in the literature and makes it the ideal candidate for our scientific scope:

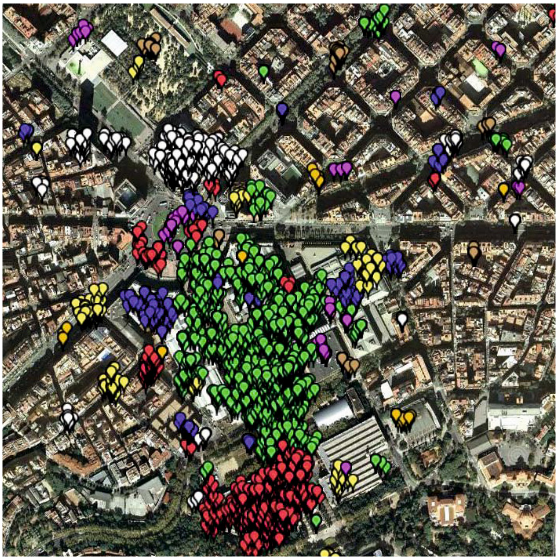
1. It is based in the concept of *density reachability*, producing satisfying results identifying arbitrarily shaped clusters.
2. The number of clusters is not given a priori.
3. The algorithm tolerates noise, allowing for some data points not to be assigned to any cluster.

Table 2. DBScan ε-sensitivity Results

Epsilon	m.	# Clusters	Noise
0.025	2000 m.	7	63
0.0125	1000 m.	46	229
0.00625	500 m.	273	1248
0.0025	200 m.	1385	6899
0.00125	100 m.	2773	17223
0.000625	50 m.	3452	31298



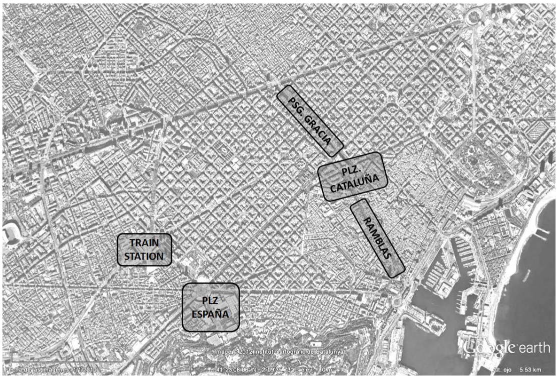
(a) $\epsilon = 0.00125$ **Downtown Barcelona Close-up**



(b) $\epsilon = 0.000625$ **Plaza España Close-Up**

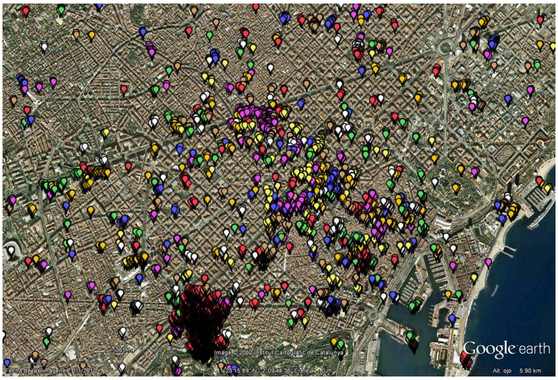
Fig. 6. ϵ-Sensitivity

The DBScan is sensible to two input parameters: the minimum amount of points (intuitively fixed to 5 for the results presented), and the ϵ value. The ϵ value determines the minimum distance amongst point to be part of a cluster, and therefore, also determines the granularity of the cluster (higher values of ϵ results in coarser clusters).



(a) **Downtown Barcelona Areas of Interest**

(b) **Clusters obtained in Week 1 and 3**



(c) **Clusters obtained in Week 2**

Fig. 7. Geospatial Clustering Results

Determining *a priori* the correct value of ϵ is problem dependent and almost unfeasible for the type of scenario we face. Therefore we perform a search space of this parameter. For the complete dataset gathered from the 3 weeks, we can observe in Table 2 that with lower values of ϵ the number of clusters increases, and therefore the number of noise generated. As the bounding box that we defined included some surroundings towns of Barcelona, we can see how certain values of ϵ ($\epsilon = 0.0125$ or $\epsilon = 0.00625$) create clusters that differentiate cities, although the whole city of Barcelona remain as one unique cluster, being hard to determine the areas of interest. When we apply lower values of ϵ ($\epsilon = 0.00125$), we start detecting realistic clusters to be considered in a city such as Barcelona (e.g. in Figure 6(a) we can see how the Ramblas are perfectly clustered, as well as Plaza España). However, with smaller levels of ϵ ($\epsilon = 0.000625$) the clusters become very granular (e.g. in Figure 6(b), focused in Plaza España, we can see different clusters for the different pavilions and surrounding areas). Determining the correct value of ϵ becomes an interesting problem to be solved, although for our initial experimental set we use empirically obtained values.

Having empirically tested the effects of different values of ϵ in our dataset, we decided to use $\epsilon = 0.00125$. This value represents a distance of 100m, which in the city of Barcelona has special sense, as it is the regular measure of one block in Eixample neighborhood (dominating a substantial area of the city-center).

The clusters generated in each of the three weeks are substantially different in the city after executing the DBScan algorithm (with $\epsilon = 0.00125$). Figure 7 shows a snapshot of these results, accompanied with a reference guide of the city of Barcelona (in Fig. 7(a)). We can easily observe that in the control-case weeks the clusters generated in the city (partially shown in Fig.7(b)) would be part of the Urban Chronotype of the city, showing the clusters of activity in the city in its normal state. However, we can spot how a substantial cluster appears in the Plaza España area (host of the MWC) in Fig. 7(c).

8 Conclusions and Future Work

In this work we have presented a modular social sensing architecture for urban environments. This architecture is fed with information obtained from the Streaming Twitter API, and has resulted satisfactory since we have obtained a 40% of the generated tweets with specific coordinates. This architecture can be seen as a low-cost sensor of the city, and allow us to construct the urban chronotype. This urban chronotype serves to compare the current behavior of the city and try to detect anomalous behaviour in the city in near real-time. However, before taking a real-time approach, we have used a controlled scenario that have an impact in the city, such as the Mobile World Congress, which in the 2012 edition had around 65.000 participants.

Along this paper we have shown the behavioural patterns of the city in its normal state and during the event, where we can see trends in the amount of users, geospatial information generated or distribution of the population depending their origin. The differences between the control case weeks and the event week are substantial enough to determine the success of our urban social sensor.

Moreover, we have used clustering algorithms to extract the areas of high-activity in the city. Amongst the existing clustering algorithms we opted using the DBScan algorithm, even though, it is extremely sensible to the ϵ value. Unfortunately, there are no existing techniques to approximate the value of ϵ for specific problems.

As future work, we will extend the DBScan algorithm, by profiting from the nature and context of the information that we are handling and improve the efficiency of it.

Specifically, there is some urban information that is publicly available in the form of open data (such as the average cost of housing square meter or the population density per neighborhood). We plan to dynamically adapt the ϵ value with respect to the the average population density of the dataset; in that way, when one instance is selected to be evaluated, the adapted algorithm will obtain the neighborhood or city to which that instance pertains and then obtain the population density to dynamically adapt the value of ϵ: coarser when evaluating points out of the observed city and, fine grained when evaluating points within the city.

Moreover, we plan to evaluate the efficiency of our platform when facing real-time detection of anomalous behaviour within the city, which will imply the adaptation of the algorithms to perform in real-time.

Finally, and as part of our long-term research, we will evaluate the mobility patterns of those active users, considering each of their geolocalized tweets as digital footprints that can be evaluated as part of a track. Combined with the classification of the user's origin we will be able to extract mobility patterns within the city depending on the users origin.

Acknowledgements. This research is partially supported by the Spanish Centre for the Development of Industrial Technology under the INNPRONTA program, project IPT-20111006, "CIUDAD2020" (`www.innprontaciudad2020.es`).

References

1. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes twitter users: real-time event detection by social sensors. In: Proceedings of the 19th International Conference on World Wide Web, WWW 2010, pp. 851–860. ACM, New York (2010)
2. Takahashi, T., Abe, S., Igata, N.: Can Twitter Be an Alternative of Real-World Sensors? In: Jacko, J.A. (ed.) HCI International 2011, Part III. LNCS, vol. 6763, pp. 240–249. Springer, Heidelberg (2011)
3. Jeff Cox, B.P.: Improving automatic weather observations with the public twitter stream. Technical report, Indiana University Computer Science Program (February 2011)
4. Li, R., Lei, K.H., Khadiwala, R., Chang, K.C.-C.: Tedas: a twitter based event detection and analysis system. In: Proceedings of the IEEE International Conference on Data Engineering, ICDE (April 2012)
5. Fujisaka, T., Lee, R., Sumiya, K.: Detection of unusually crowded places through microblogging sites. In: 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 467–472 (April 2010)
6. Weng, J., Yao, Y., Leonardi, E., Lee, F.: Event Detection in Twitter. Technical report, HP Labs (2011)
7. Asur, S., Huberman, B.A.: Predicting the future with social media. In: Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2010, vol. 01, pp. 492–499. IEEE Computer Society, Washington, DC (2010)
8. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. Journal of Computational Science 2(1), 1–8 (2011)
9. Girardin, F., Fiore, F.D., Ratti, C., Blat, J.: Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. J. Locat. Based Serv. 2(1), 41–56 (2008)

10. Schedl, M.: Analyzing the potential of microblogs for spatio-temporal popularity estimation of music artists. In: Proceedings of the IJCAI 2011: International Workshop on Social Web Mining (2011)
11. Conover, M., Gonçalves, B., Ratkiewicz, J., Flammini, A., Menczer, F.: Predicting the political alignment of twitter users. In: Proceedings of 3rd IEEE Conference on Social Computing, SocialCom (2011)
12. Nagarajan, M., Gomadam, K., Sheth, A.P., Ranabahu, A., Mutharaju, R., Jadhav, A.: Spatio-Temporal-Thematic Analysis of Citizen Sensor Data: Challenges and Experiences. In: Vossen, G., Long, D.D.E., Yu, J.X. (eds.) WISE 2009. LNCS, vol. 5802, pp. 539–553. Springer, Heidelberg (2009)
13. Martino, M., Vaccari, A., Ratti, C.: Pulse of the city: Visualizing urban dynamics of special events. In: GraphiCon - Internation Conference on Computer Graphics and Vision (2010)
14. Ferrari, L., Rosi, A., Mamei, M., Zambonelli, F.: Extracting urban patterns from location-based social networks. In: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks, LBSN 2011, pp. 9–16. ACM, New York (2011)
15. Fujisaka, T., Lee, R., Sumiya, K.: Discovery of user behavior patterns from geo-tagged micro-blogs. In: Proceedings of the 4th International Conference on Uniquitous Information Management and Communication, ICUIMC 2010, pp. 36:1–36:10. ACM, New York (2010)
16. Du, Y., Fan, J., Chen, J.: Experimental analysis of user mobility pattern in mobile social networks. In: 2011 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1086–1090 (March 2011)
17. Cheng, Z., Caverlee, J., Lee, K.: You are where you tweet: a content-based approach to geo-locating twitter users. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM 2010, pp. 759–768. ACM, New York (2010)
18. Lehmann, J., Gonçalves, B., Ramasco, J.J., Cattuto, C.: Dynamical classes of collective attention in twitter. CoRR abs/1111.1896 (2011)
19. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proc. of 2nd International Conference on Knowledge Discovery and Data Mining, pp. 226–231 (1996)
20. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Simoudis, E., Han, J., Fayyad, U.M. (eds.) Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231. AAAI Press (1996)