New City Landscapes

Tracking Location Based Social Networks Using Twitter Data

F.NEUHAUS
90 Tottenham Court Road, W1T 4TJ
, United Kingdom
0044 0 3108 3877, Fabian.neuhaus@ucl.ac.uk

Keywords: Social network, Location based information, Tracking, Geographical visualisation, Urban studies

Introduction

Where is the city active and does it physically change over time? Urban areas are no static artefacts as they are preferably described in texts and theories. Urban areas are buzzing hot spots of human activity that, to some extend, manifest themselves as built structure, but are largely temporal and ephemeral. Meaning, that no constant being of this 'artefact' is present, but merely a past aggregate is telling tales of memories and rumours.
In an attempt to listen to these stories and narrative as they unfold through the streets, alleyways, in courts, buses, on roof terraces or in swimming pools the social networking platform, twitter was employed to reconstruct the cities activity hotspot as a chunck of time-frozen 'New City Landscape' drawing out the ever changing locations of people's presence and the power of spatial creation through narratives and activity.
This paper discusses the emerging potential of social media data used for urban area research and city planning. Working with crowd-sourced data in a web 2.0 manner as described for example by Hudson-Smith et al. (2009). Specifically we look at the connections between the emerging social network, as for example described by Boccaletti et al. (2006), and the local physical surrounding and conditions. Also aspects of visualisation as well as privacy and ethical implications are discussed.
From the collected data a new landscape based on density is generated. The features of this landscape of digital activity correspond directly with the physical location of their origin, but at the same time represent with hills the peaks of locations from where the activity tales are submitted. The flanks and valleys stand for areas with lesser activity and vast plains and deserts of no twitter tales stretch across the townscapes that lay dormant. These New City Landscape (See Figure 1) maps don't represent any physical features, but the interaction with physical features on a temporal basis. The digital realm has become as much part of the urban environment as the physical and with these tweetography maps they are made visible for the first time. The maps allow us to make a direct comparison between real word activity, physical location and digital message. In a globalised world this local reference develops an increased importance as a sense of place, a source of identity and memory. The digital social media data allows us to investigate into this realm of

peer groups' social location interaction, combining the global scale with its local source.
The information from the social media network twitter, is gathered directly of the platform via the provided API. This methods provides us with termporal location data as individual tracks similar to for example employed by Eagle et al. (2009) in their study of social networks using mobile phones. The twitter data can be associated with a physical location if the GPS of the smart phone has writen it into the tweet's meta data. Research investigating such location-based technology together with a temporal structure has been carried out for example by Reads et al. (2009).
These local activity are analysed and visualised based on networks of interaction. Who knows whom and get in touch with whom? However the social networks in the sense of specific interest are these datasets in relation to place and how this location based network enable the individual to shape a distinct sense of place.

Methods and technology

The technology to collection twitter data is based on the technology developed for the Tweet-O-Meter (tom). This service was developed at CASA by Steven Gray. Similar to the tom service the data is collected using the twitter API.

Process

The API offers a built in spatial search function according to which the querry to Twitter is designed. In additona we can filter the incoming messages as for specific geo location details. For the NCL maps we have defined the urban areas consistently as an area with a 30km radius around an urban centre.

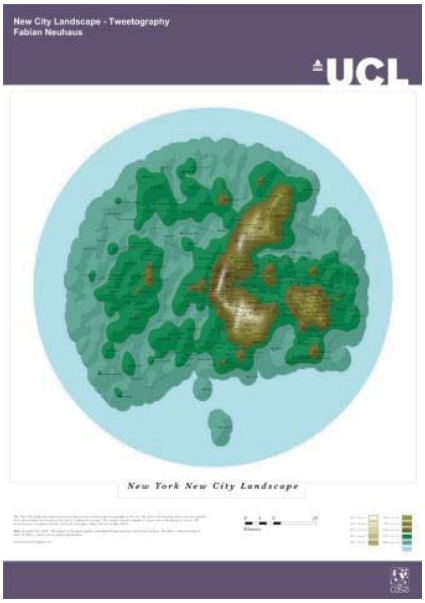

Figure 10. - New York New City Landscape Map showing the twitter activities over the period of one week as a density surface within a radius of 30 km of the NY urban area.

What we are using for the processing of the visualisations are all messages containing real GPS information as Latitude and Longitude coordinates. With this information a more accurate mapping is possible. The accuracy of this information lies within the normal range of GPS accuracy between some 5 to 15 metre.
In a second step a social network is computed, based on the interactions of users in the dataset. To do this especially re-tweets (RT), twitter messages that have been resent by other users and at-tweets (@), messages specifically addressed at selected twitter users are employed to establish links between individual tweeters as well as a direction of interaction.
Together with the emerging social network and the location as well as the temporal information contained in the data a location based temporal social network is visualised.

Data

The amount of data collected varies dramatically between the different locations. There are the very actively tweeting cities such as New York and London with more than 800'000 location based messages sent over the course of one week. On the other hand there are a lot of places especially non-English speaking countries with far less activity, down to a few hundred. Furthermore, there is not a simple, more messages result in more GPS tagged messages, equation that applies. It can well be that a very actively tweeting place turns out very few Latitude/Longitude tagged tweets. As it appeared for example in the case of Sydney, Jakarta or Sao Paulo, where the percentage of geotweets is below 1 % of all location based messages.

Results

For the mapping the individual points are being aggregated as a density surface.
Throughout the emerging landscape features have been renamed to reflect these conditions. The new names are constructed using the real world names of the location, in combination with a landscape description of the virtual surface overlaid. This is 'Mountain' or 'Peak' for high points, 'Slope' or 'Valley' for descending features or 'Desert' and 'Meadow' for average and consistent areas. Inactive areas are termed for example 'Quarry' or 'Ditch'. Together with the familiar real world element the locations become tangible and memorable points of orientation and maybe identification.
The defining landscape features in the virtual NCL map are the hot spots of twitter activity, the peaks. Here the morphology varies between the urban Areas dramatically. How the twitter traffic structures the NCL is unique for each city. There are however some characteristics that can be pointed out. The different groups could be described as central, where one main location towers over the whole urban region, the multi, where different hotspots appear as peaks across the landscape and the featuring, where one or more features draw out as shapes, groups of peaks or ridges.
Linking this to the social network (see Figure 2), the activity pattern in the temporal sense, gain of importance. The variety of different pattern displayed by different groups is very distinct from activity pattern we normally see in everyday activities.

There is more scope for the individual to jump in and out of activities, but connections on hold and reactivate others than what we know from real world interaction.

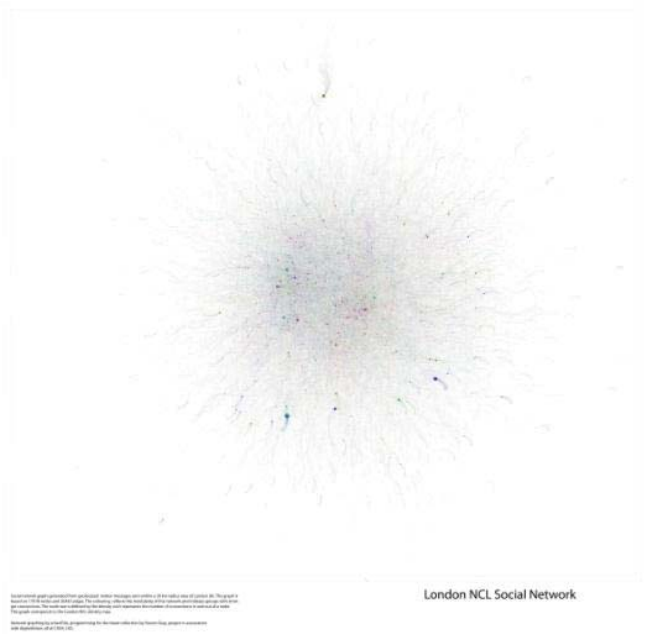

Figure 11 - Twitter activity based social network using the London NCL map data collected in a radius of 30 km within the London urban area.

On the city side the transformation of network activity hubs through out different time periods are striking and offer a new perspective on urban area usage as well as sense of place.
Application for this can be found ranging from urban planning to transport management and modelling to health and safety as in the spatial spreading of information or infections, but aso for the plnning of urban struture and development of density and morphology models this is interesting information.

Acknowledgement

A collaboration between Dr Andrew Hudson-Smith, Steven Gray and Fabian Neuhaus initiated this project. The code to query, collect, store and output the twitter data is developed by Steven Gray.

Bibliography

[1] **Boccaletti, S. et al**., 2006. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5), pp.175-308.

[2] **Eagle, N., Pentland, A. & Lazer, D.**, Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*. Available at: http://www.pnas.org/content/early/2009/08/14/0900282106.abstract [Accessed January 28, 2011].

[3] **Hudson-Smith, A. et al.**, 2009. Mapping for the Masses. *Social Science Computer Review*, 27, pp.524–538.

[4] **Reades, J., Calabrese, F. & Ratti, C.,** 2009. Eigenplaces: analysing cities using the space – time structure of the mobile phone network. *Environment and Planning B: Planning and Design*, 36(5), pp.824 – 836.