Spatio-social Metrics for Information Diffusion in Online Social Networks

School of Computer Science

Antonio Lima and Mirco Musolesi
{axl162,musolesm}@cs.bham.ac.uk

UNIVERSITY OF BIRMINGHAM

1. Introduction

In online social networks, ranking users by importance can be useful for understanding and driving information diffusion processes [1]. Until now, this has mainly been done using various definitions of *centrality* [2,3].

Social networks websites now allow users to attach spatial data to shared content. The concept of importance with respect to geographic area in social networks is still largely unexplored. We define metrics of *spatio-social centrality* that quantify the importance of an individual or a group of individuals with respect to a certain geographic area.

This study can be helpful for a variety of practical applications such as location-targeted advertisement and local emergency alert dissemination.

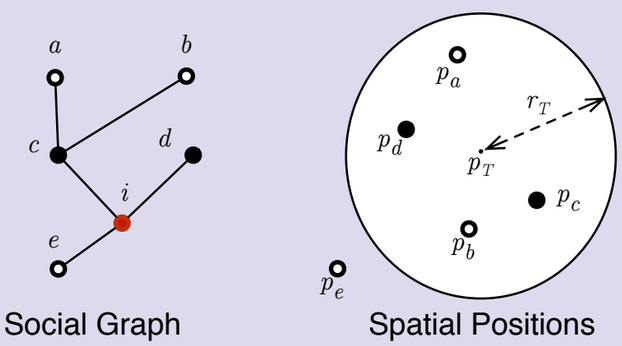

Figure 1. A small spatial network is represented by a social graph (at the left) and the geographic positions of nodes (at the right). Full circles represent the friends of user i (in red) that have their points inside the area of center p_T and radius r_T (in white).

2. Metrics

We represent a *spatial social network* as a social graph, representing friendships between users. Each node also has at least a significant location (their home, their workplace, their favourite cafe, etc.), as in Fig. 1. Given this model, we define two metrics:

- *Spatial degree centrality* measures how many friends of the user are within the considered geographic area, and is defined as follows:

$$C_i^D(r_T, p_T) = |\{j \in N_i : d_G(p_j, p_T) < r_T\}|$$

where i is the user we are considering, N_i is the set of his/her friends, p_T and r_T are the center and the radius of the considered area, d_G is the geodesic distance.

- *Spatial closeness centrality* computes for the user the average of geographic distances between every friend's significant location and the center of the area taken into consideration:

$$C_i^C(p_T) = \frac{1}{k_i} \sum_{j \in N_i} d_G(p_j, p_T)$$

where k_i is the number of friends of i.

3. Results

We studied the metrics in a real-world datasets obtained from Twitter, a popular social network website. Our dataset contains 657,777 users, obtained through a snowball sampling seeded with 1375 seed users. Results are shown in Fig. 2.

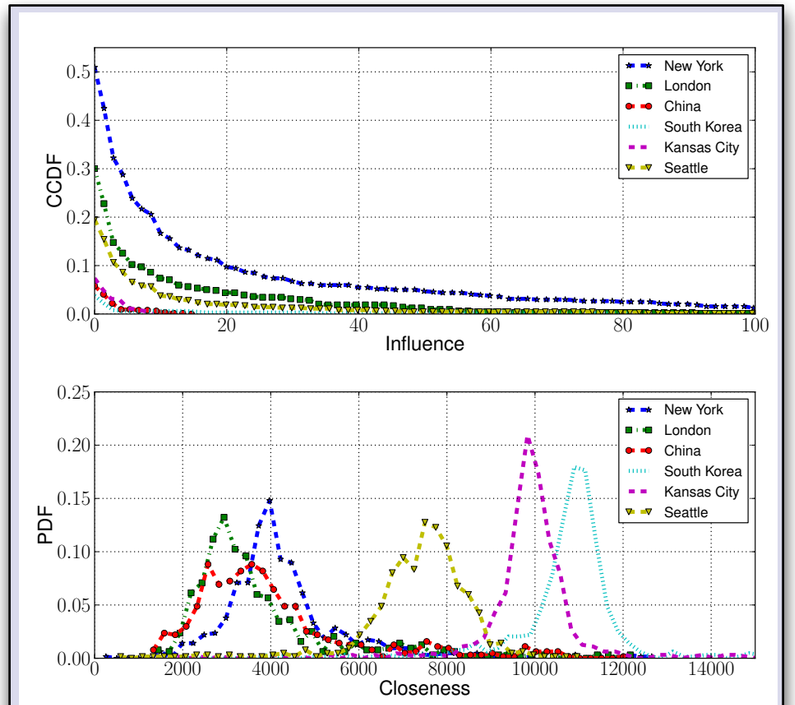

Figure 2. Spatial degree centrality (top) and spatial closeness centrality (bottom) for Twitter users in London with respect to the geographic areas of New York, London, China, South Korea, Kansas City, Seattle.

[1] D. Kempe, J. Kleinberg, É. Tardos. Maximizing the Spread of Influence through a Social Network. In *Proceedings of KDD '03*. Washington, D.C., USA. 2003. ACM, pp. 137-146.

[2] M. Newman. *Networks: An Introduction.* Oxford University Press, 2010.

[3] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, 1994.

http://goo.gl/tDsRn

Download me!