Detecting daily commuting distance from GPS trajectory

Wei Jiang[1], Yang Yue[2]

[1]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,

Wuhan University

Email: jiangweiyjs@126.com

[2] Shenzhen Key Laboratory of Spatial-temporal Smart Sensing and Services,

Shenzhen University

Email: yueyang@whu.edu.cn

1. Background and objective

Most people have certain acceptance of daily travel distance, which determines where they choose to live, to work, and to go for leisure. Such information is an important input for facility location-allocation, urban planning and transport management. Therefore, daily commuting distance has been adopted as an indicator to measure the distance acceptance, the rationality of land use structure, and livability of a city. At the same time, daily commuting distance also can reveal many personal characteristics. However, such information traditionally is obtained by questionnaire which is both labor-intensive and time-consuming. And most importantly, the sampling could not be very comprehensive.

With the popularization of mobile phones, it has become possible to obtain individual movement trajectories to understand personal travel behavior. Related studies have shown that, human activities, ranging from communication to entertainment and work patterns, follow non-Poisson statistics (Barabási, 2005). This study uses GPS trajectories generated by smart phone to detect daily travel pattern, especially the daily commuting distance.

2. Methodology

The data we used in this study is a set of testing assisted GPS (A-GPS) data recorded by a telecommunications service provider, which were generated in every 2 min, and the time span is one month. Each positioning data also associated with a time stamp. Our previous study (Yue et al., 2012) using another set of cell-tower based mobile phone positioning data, has shown that, even the very sparse positioning data can reveal many personal characteristics, such as locations of homes and work places, their daily routine, incoming level, and even consuming preferences, by taking Point of Interest (POI) and other web-based public data into consideration. Other related studies also demonstrated the potentials of human trajectory data (Licoppe et al., 2008; Huang et al., 2010; Sevtsuk et al., 2010).

Although the spatial distribution of these data is very wide, there are still many areas with dense trajectory points. In order to find out the daily commuting distance, we first identified meaningful places for each individual using DBSCAN clustering algorithm on their daily traces. We determined the clustering parameters, *Minpt* and

Eps, by controlling the number of meaningful places. Since each personal activity is driven by some certain factors (Candia et al., 2008), most people have limited number of meaningful places on a daily basis, for example, 2-3 places. This is also constrained by both space and time dimensions. We used a visualization-based interactive approach, and determined the upper threshold of the daily meaningful places as 7. Then, the minimum number of points in a cluster, *Minpt*, is 60 and the distance between the points, *Eps*, is 30 meters, respectively. The Fig. 1 shows the percentage of the number of meaningful places based on the experiment data.

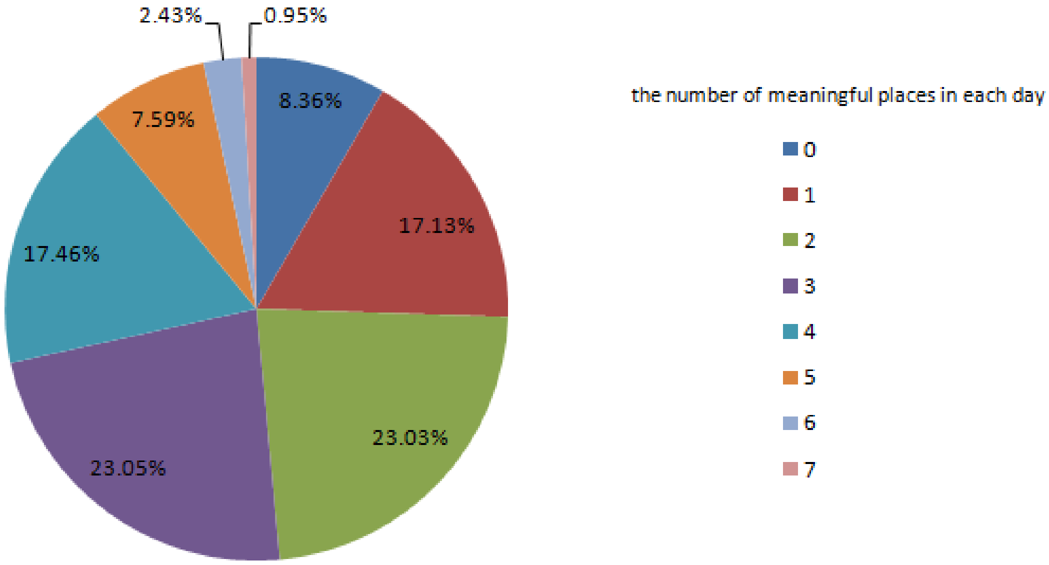

Figure 1: Percentage of the number of meaningful places (daily)

We found that most of the people travelled following some fixed routines which have a pronounced periodic feature. We assumed that the cluster with the maximum number of points collected during 0:00-6:00 is the place where a person lives, i.e., HOME. We further calculated the distances of the other identified clusters with the HOME. Fig. 2 shows the probability distribution of the distances. Most meaning places are 2-4km to the HOME for this person. The result is the same with the finding of Jahanbakhsh (2012).

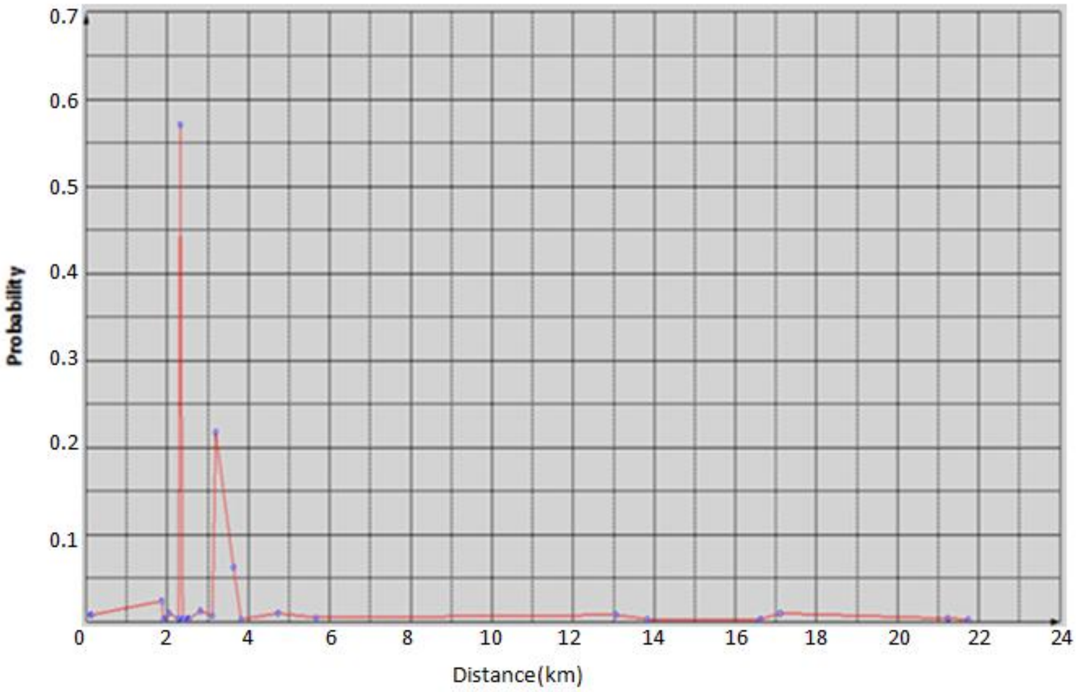

Figure 2: A probability distribution of a person's activity range

Because there are also weekend travels, or holiday travels recorded in the trajectories, in order to obtain the typical daily commuting distance, we adopted the entropy of the meaningful places to reflect the strength of the regularity（Cho et al., 2011):

$$\mathrm{Entropy} = -\sum \mathrm{P(x)} \log_2[\mathrm{P(x)}] \quad (1)$$

Where, P(x) is the probability of cluster and the x is the distance between the cluster and the HOME. The correlation between the entropy of the position and the strength of the regularity is negative, so the day in which the entropy is the lowest is regarded as the day of strongest regularity. We then used the trajectory data recorded in that day to calculate the daily commuting distance.

3. Result and implication

Result shows that in the experiment dataset, most people commuted within 4 km, and almost half of the persons' daily commuting distances are within 2 km, as shown in Fig. 3.

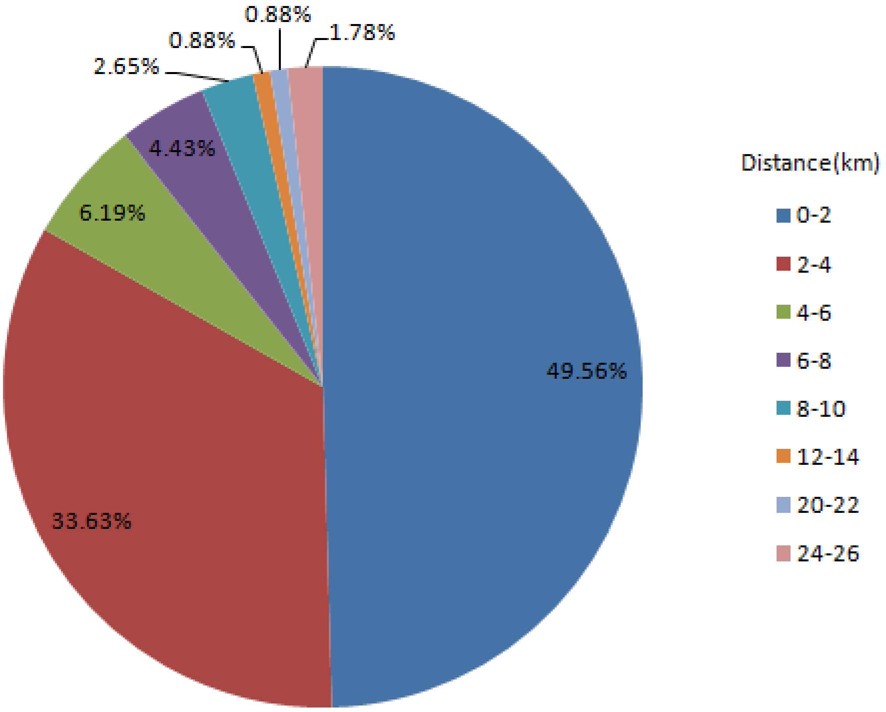

Figure 3: Distribution of daily commuting distances

In our future study, we will further analyze the corresponding commuting time, and the proportions of different travel modes. We are also interested in how the land use structure and urban form influence the personal commuting distance. The results may vary in different cities, but the proposed methodology and findings have importance implications for understanding human travel behavior, transport and urban planning.

4. Acknowledge

This work is supported by China NSFC 41171348 and 41061043.

Reference

Barabási A L, 2005, The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211.

Candia J, González M C and Wang P, 2008, Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical*, 41(22):1-11.

Cho E, Myers S and Leskovec J, 2011, Friendship and Mobility: User Movement in Location-Based Social Networks. *Proceedings of the 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining(KDD'11)*, California, USA, 1082-1090.

Huang L, Li Q Q and Yue Y, 2010, Activity Identification from GPS Trajectories Using Spatial Temporal POIs' Attractiveness. *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks 2010*, San Jose, USA, 27-30.

Jahanbakhsh K and King V, 2012, They Know Where You Live! *preprint arXiv:1202.3504*.

Licoppe C, Diminescu D, Smoreda Z and Ziemlicki C, 2008, Using mobile phone geolocalisation for 'socio-geographical' analysis of co-ordination, urban mobilities, and social integration. *Tijdschrif voor economische en sociale geografie*, 99 (5):584-600.

Sevtsuk A and Ratti C, 2010, Does Urban Mobility Have a Daily Routine? *Learning from the Aggregate Data of Mobile Networks, Journal of Urban Technology*, 17(1): 41-60.

Yue Y, Chen J, Hu B, Xie R, Zuo X Q and Xie X, 2012, Labeling Personal Characteristics from Mobile Phone Traces. *in the 2nd International workshop on mobile sensing, IPSN'12 and CPSWeek*, Beijing, China.