Understanding Individual Daily Activity Space Based on Large Scale Mobile Phone Location Data

Yang Xu[1], Shih-Lung Shaw[1 2 *], Ling Yin[3], Ziliang Zhao[1]

[1] Department of Geography, University of Tennessee, Knoxville, TN, USA 37996
Email: {yxu30; sshaw; zzhao7}@utk.edu

[2] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, 430079, PR China

[3] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, PR China
Email: yinling@siat.ac.cn

* Corresponding author: sshaw@utk.edu

1. Introduction

Activity space has been an active research topic in Geographic Information Science during recent years. It aims at understanding people's use of space in their everyday lives. Many studies have used travel/activity surveys to collect people's activity data in support of studies of human mobility, spatial behaviour, and related socio-economic implications (Dijst 1999; Kwan 1999; Schönfelder and Axhausen 2003). Due to advancements of information, communications and location-aware technologies, mobile phone location datasets now offer opportunities to study people's use of space at relatively high spatial and temporal granularity levels.

Measuring the size, geometry, and structure of human activity space can be important to travel behaviour study and policy analysis. Various measures such as standard deviational ellipse (Lefever 1926), standard distance (Bachi 1962), and daily potential path area (Kwan 1999) have been used to measure individual activity space for different research purposes. The standard deviational ellipse (SDE) method could describe the average location, dispersion and orientation of spatial point patterns (Yuill 1971). Thus it has been used extensively in activity-based research (Zahavi 1979; Sherman et al. 2005; Buliung and Kanaroglou 2006). However, the practice of using an ellipse to represent an individual's activity space has not been empirically tested. This study uses SDE to empirically evaluate individual daily activity spaces with a mobile phone location dataset that covers 5.8 million people in Shenzhen, China. The big mobile phone dataset enables us to empirically assess to what extent that SDE can properly reflect activity space of individuals in an urban area.

2. Study Area and Mobile Phone Location Dataset

The mobile phone location dataset used in this study was collected on a workday in the city of Shenzhen, China. The dataset actively tracked the mobile phone locations of 5,812,305 individuals with a time span of 24 hours. Each individual's location was recorded at the mobile phone tower level approximately once every 60 minutes. Table 1 shows an individual's mobile phone location records during the study period. Each record includes user ID, date, time, and coordinates (latitude/longitude) of the mobile phone tower to which the mobile phone was assigned. Figure 1 shows the geographic distribution of all mobile phone towers included in this study. Thiessen polygons are created to denote the service area of each mobile phone tower. The density of mobile phone towers varies across the city. The average service area of all mobile phone towers is 0.34 km^2.

User ID	Date	Time	Longitude	Latitude
583********	2012-03-22	23:02:32	113.*****	22.*****
583********	2012-03-23	00:02:34	113.*****	22.*****
583********	2012-03-23	01:02:36	113.*****	22.*****
...	**…**	**...**	113.*****	22.*****
583********	2012-03-23	23:06:09	113.*****	22.*****

Table 1. Example of an individual's mobile phone location records (we used asterisks to protect personal IDs and locations).

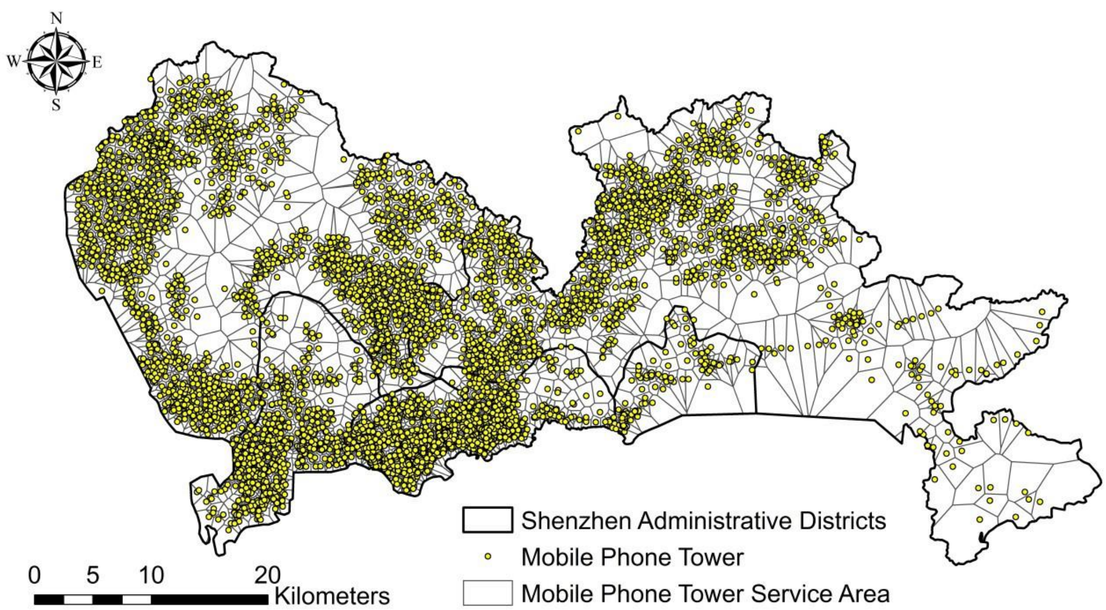

Figure 1. Spatial distribution of mobile phone towers in this study. Thiessen polygons are created to denote the service area of each mobile phone tower.

3. Individual Activity Space Represented by A Standard Deviational Ellipse

A standard deviational ellipse (SDE) can summarize a spatial point pattern by deriving the average location, spatial dispersion and orientation of the point distribution. It can be used to investigate spatial point patterns at both individual (e.g., individual activity space) and aggregate level (e.g., group activity space). In this study, we apply the measure of SDE to understand important aspects of individual daily activity space in Shenzhen, China by using a one-day mobile phone location dataset. We then examine the difference between the derived SDE and the original mobile phone locations for each individual to assess to what extent that SDE can properly reflect an individual's activity space.

We first derive a standard deviation ellipse (SDE) for each individual based on the equation introduced by Yuill (1971). As each individual's location was recorded approximately once every 60 minutes, all locations recorded for an individual are assigned the same weight to compute the SDE. Individuals with only 1 or 2 unique locations are treated as special cases since their locations represent a single point or a line. This dataset has 9.87% of the individuals with one unique location and 14.89% of the individuals with 2 unique locations.

We then construct SDEs for the remaining 4,373,039 individuals who have 3 or more unique locations in the dataset. Figure 2 shows the cumulative distribution of areal size of these SDEs. It indicates that 62.59% of the individuals have an SDE with $area \leq 1km^2$ and 80.82% of the individuals have an SDE with $area \leq 5km^2$. The results indicate that a large proportion of individuals in Shenzhen travelled within short distances during this workday.

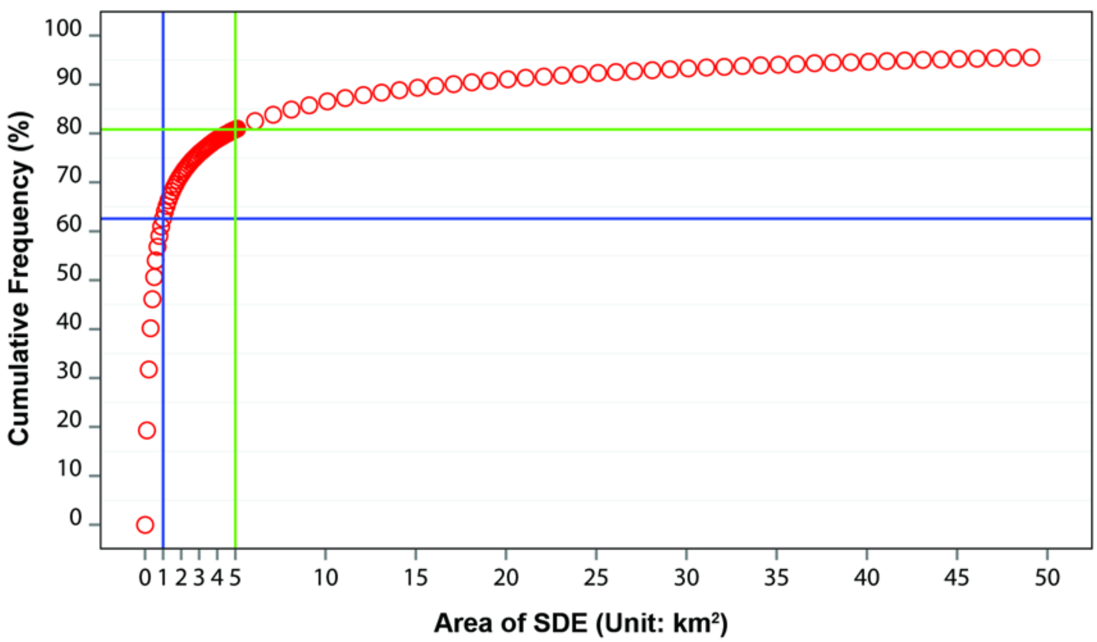

Figure 2. Cumulative distribution of the area of SDEs for 4,373,039 individuals.

Next, we examine the shape of each individual's SDE with its eccentricity. The eccentricity ranges from 0 (ellipse = circle) to 1 (ellipse = line) depending on the shape of the ellipse. Figure 3 shows the cumulative distribution (with inserted histogram) of the eccentricity of SDE for the same subset. The inserted histogram shows that a large percentage of individuals have an SDE with its eccentricity between 0.9 and 1.0. According to the cumulative distribution, only 23.86% of the individuals have an SDE with its eccentricity smaller than 0.9. When we compute the SDEs, the observed mobile phone locations for each individual are assigned the same weight. Thus the eccentricity of SDE is mainly determined by each individual's activity anchor points, which refer to the locations where he/she stayed for a significant amount of time. Applying other weighting schemes might produce different distribution patterns. For example, assigning equal weight to the unique locations of each individual might result in an overall decrease of the eccentricity of SDEs.

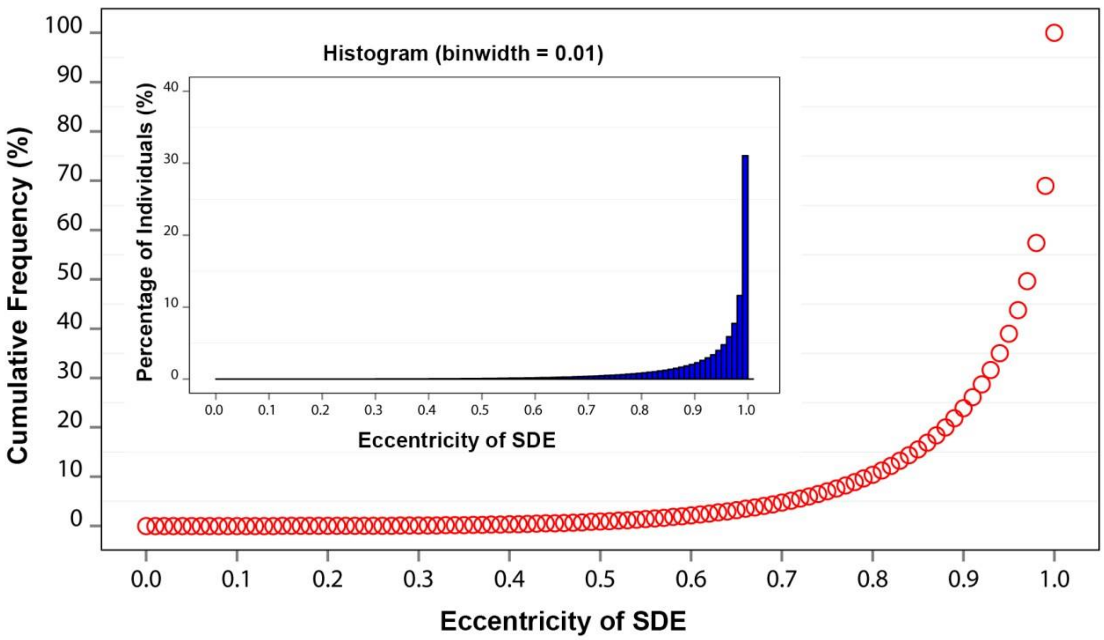

Figure 3. Distribution of eccentricity of SDEs for 4,373,039 individuals.

4. Difference between SDE and Spatial Distribution of Observed Mobile Phone Locations of Each Individual

The standard deviational ellipse is a measure of geographic concentration of point distribution patterns. Hence, the derived SDE for an individual may only cover part of his/her observed mobile phone locations. Figure 4 shows the histogram for the same subset of 4,373,039 individuals (with 3 or more unique mobile phone locations in their activity spaces) that summarizes the percentage of each individual's observed mobile phone locations falling outside of his/her SDE. The results indicate that 27.3% of the individuals have more than 30% of observed mobile phone locations falling outside of their SDE, and 11.2% of them have more than 40% of observed mobile phone locations falling outside of their SDE. Such mobile phone locations include individual activity anchor points (e.g., home and workplace) and other mobile phone locations traversed by the individuals during the study day. The mean value of the histogram (green line) is 22.18%.

As suggested by Golledge and Stimson (1997) and Schönfelder and Axhausen (2003), an individual's activity space is mainly determined by frequent activity locations as well as travel between and around these pegs. Based on our analysis results, an individual's daily activity anchor points are not usually covered by his/her SDE. Figure 5 illustrates an example of our exploratory analysis. The concept of space-time path proposed by Hägerstrand (1970) is used here to represent the spatial movements of an individual over time. The dots denote the observed mobile phone locations for the individual at particular time points of the day, and the extruded ellipse denotes the corresponding SDE. The green and red dots denote the observed mobile phone locations which fall outside of and inside of the SDE, respectively. We can observe that segments with consecutive green dots stand for two activity anchor points for the individual on this workday. Although the SDE is able to describe the spatial dispersion and orientation of this individual's daily activity space, it fails to include these two important activity anchor points.

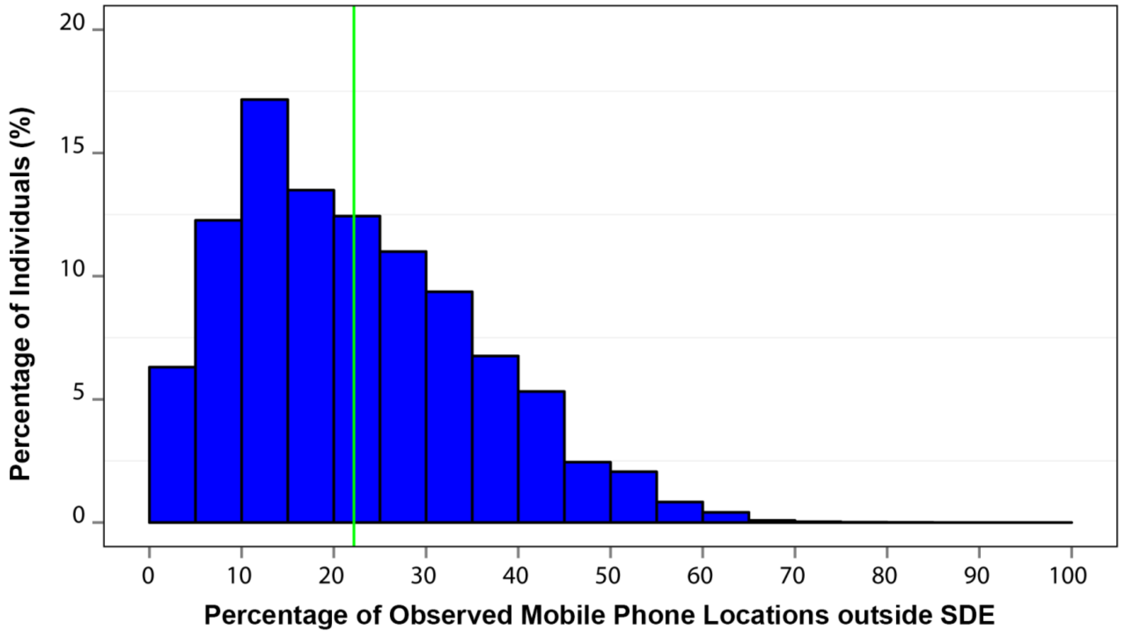

Figure 4. Histogram showing the percentage of each individual's observed mobile phone locations outside the SDE (4,373,039 individuals).

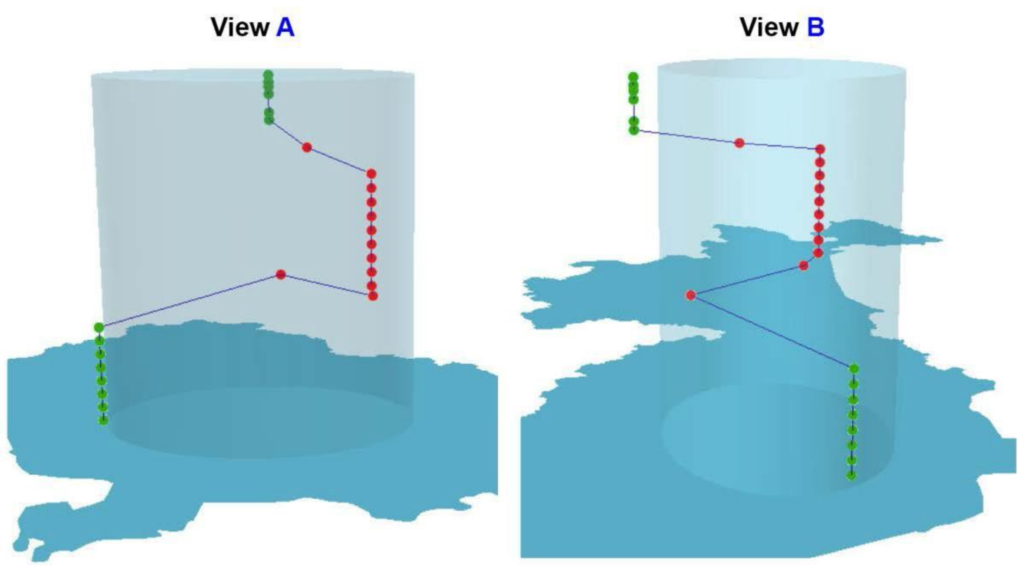

Figure 5. Difference between an individual's observed mobile phone locations and his/her SDE (Note: The vertical dimension represents time.)

5. Summary

This study evaluates how well a standard deviational ellipse can properly reflect an individual's daily activity space based on a large mobile phone location dataset collected in Shenzhen, China. According to our analysis results, the standard deviational ellipse is useful to reveal some characteristics of an individual's activity space such as its spatial dispersion (ellipse size) and shape (eccentricity). However, as individuals' daily activities tend to concentrate in a few anchor points such as home and work places, the standard deviational ellipse often does not cover these key activity locations. This raises concerns when an individual activity space is used for particular research purposes. For example, it will introduce bias in health-related studies if standard deviational ellipse is used to examine individual access to health care services or to estimate individual-level pollution exposure. The analysis results suggest that we need to pay special attention to individuals' activity anchor points when standard deviational ellipse is used to represent individual activity space. In the future, we will identify activity anchor points for the individuals in this dataset and further investigate the spatial relationship between each individual's standard deviational ellipse and the activity anchor points.

6. Acknowledgements

This study was jointly supported by National Natural Science Foundation of China (41231171), National Key Technology R&D Program of China (#2012AA12A403-4), and Alvin and Sally Beaman Professorship and Arts and Sciences Excellence Professorship at the University of Tennessee.

7. References

Bachi R, 1962, Standard distance measures and related methods for spatial analysis. *Papers in Regional Science*, 10(1), 83-132.

Buliung R. N. and Kanaroglou P. S, 2006, A GIS toolkit for exploring geographies of household activity/travel behaviour. *Journal of Transport Geography*, 14(1), 35-51.

Dijst M, 1999, Two-earner families and their action spaces: A case study of two dutch communities. *Geojournal*, 48(3), 195-206.

Golledge R.G. and Stimson R.J, 1997, *Spatial Behavior*, The Guilford Press, New York and London.

Hägerstrand T, 1970, What about people in regional science? *Papers of the Regional Science Association*, 24(1), 7-24.

Kwan M. P, 1999, Gender, the home-work link, and space-time patterns of nonemployment activities. *Economic Geography*, 75(4), 370-394.

Lefever D, 1926, Measuring geographic concentration by means of the standard deviational ellipse. *American Journal of Sociology*, 32, 88-94.

Schönfelder S. and Axhausen K. W, 2003, Activity spaces: measures of social exclusion? *Transport Policy*, 10(4), 273-286.

Sherman J. E, Spencer J, Preisser J. S, Gesler W. M. and Arcury T. A, 2005, A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics*, 4(24), 1-21.

Yuill R. S, 1971, The standard deviational ellipse: an updated tool for spatial description. *Geografiska Annaler, Series B. Human Geography*, 53(1), 28-39.

Zahavi Y, 1979, *The 'UMOT' Project Report*. US Department of Transportation DOT-RSPA-DPB-20-79-3.US Department of Transportation, Washington.