Transportation
DOI 10.1007/s11116-015-9597-y

Understanding aggregate human mobility patterns using passive mobile phone location data: a home-based approach

Yang Xu[1] · Shih-Lung Shaw[1,2] · Ziliang Zhao[1] · Ling Yin[3] · Zhixiang Fang[2] · Qingquan Li[4]

Abstract Advancements of information, communication and location-aware technologies have made collections of various passively generated datasets possible. These datasets provide new opportunities to understand human mobility patterns at a low cost and large scale. This study presents a home-based approach to understanding human mobility patterns based on a large mobile phone location dataset from Shenzhen, China. First, we estimate each individual's "home" anchor point, and a modified standard distance (S_D') is proposed to measure the spread of each individual's activity space centered at this "home" anchor point. We then derive aggregate mobility patterns at mobile phone tower level to describe the distance distribution of S_D' for people who share the same "home" anchor point. A hierarchical clustering algorithm is performed and the spatial distributions of the

✉ Shih-Lung Shaw
sshaw@utk.edu

Yang Xu
yxu30@utk.edu

Ziliang Zhao
zzhao7@utk.edu

Ling Yin
yinling@siat.ac.cn

Zhixiang Fang
zxfang@whu.edu.cn

Qingquan Li
liqq@szu.edu.cn

1 Department of Geography, University of Tennessee, Knoxville, TN 37996, USA

2 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, People's Republic of China

3 Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, People's Republic of China

4 Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, People's Republic of China

Published online: 26 March 2015

derived clusters are analyzed to highlight areas with similar aggregate human mobility patterns. The results suggest that 43 % of the population sample travelled within a short distance ($S_D' \leq 1$ km) during the 13-day study period while 23.9 % of them were associated with a large activity space ($S_D' \geq 5$ km). The geographical differences of people's mobility patterns in Shenzhen are evident. Areas with a large proportion of people who have a small activity space mainly locate in the northern part of Shenzhen such as Baoan and Longgang districts. In the southern part where the economy is highly developed, the percentage of people with a larger activity space is higher in general. The findings could offer useful implications on policy and decision making. The proposed approach can also be used in other studies involving similar spatiotemporal datasets for travel behavior and policy analysis.

Keywords Passive mobile phone location data · Human mobility · Activity space · Home-based approach

Introduction

Understanding human mobility patterns has been an important topic in transportation research. Traditional studies rely heavily on travel surveys and questionnaires to analyze people's movements and activities (Levinson and Kumar 1995; Dijst 1999; Axhausen et al. 2002; Schönfelder and Axhausen 2003). Such datasets usually consist of detailed information of respondents as well as the trips they made that can facilitate analysis of human travel and activity patterns. However, collection of such datasets can be costly and time-consuming and the sample size is constrained by the human and financial resources available. Recent advancements in location-aware technologies have provided many passively generated datasets for understanding the whereabouts of people in space and time. These datasets refer to a set of data that are passively generated only when particular kinds of human activities occur (e.g., cell phone usage, credit card transaction and check-in of social networks). Mobile phone location data, as one example of such datasets, "enables the study of human mobility at low cost and on an unprecedented scale" (Becker et al. 2013, p. 74). These large-scale anonymized datasets have introduced new opportunities to investigate various aspects of human dynamics from individual activity patterns to collective behavior of the masses (Candia et al. 2008).

In recent years, there have been many studies which used mobile phone location data to understand people's travel and activity patterns as well as the implications to transportation planning, urban design and social dynamics (Ahas et al. 2007; Candia et al. 2008; Gonzalez et al. 2008; Phithakkitnukoon et al. 2010; Song et al. 2010; Cho et al. 2011; Yuan et al. 2012; Becker et al. 2013). Most of them utilized passive mobile phone location data, which is a by-product of business operations of mobile phone companies. Such datasets often contain information about when and where different types of mobile phone activities are conducted (e.g., mobile phone calls, text messages, and location-based services). Unlike traditional travel surveys, these passively generated datasets fall short of collecting individual socioeconomic and demographic information for travel behavior and policy analysis. Moreover, individuals' location records don't explicitly reflect their travel purposes. Hence, novel approaches are needed when using such datasets for a better understanding of human mobility patterns and related transportation problems.

As suggested by Schönfelder and Axhausen (2003), an individual's home location should be treated as important anchor point when analyzing how people move around in their daily lives. Understanding how people move around their home locations could yield novel insights into people's mobility patterns. Moreover, a good understanding of people's use of space around their home locations could potentially benefit urban planning and public transport design around different neighborhoods in a city. Although passive mobile phone location data can be sparse in both space and time, it serves as a valuable data source for uncovering individual activity anchor points such as home locations, and people's use of space around these locations. Hence, this study proposes a home-based approach to studying human mobility patterns using a passive mobile phone location dataset. The main contributions of this research are as follows:

(1) With the use of a 13-day mobile phone location dataset covering more than 1 million people in the city of Shenzhen, China, we estimate each individual's "home" anchor point (which corresponds to a particular mobile phone tower) and use it as the reference point of an individual's mobility pattern. We propose a modified standard distance (S'_D) to quantify the spread of each individual's daily activities centered at the "home" anchor point. Individuals are then grouped based on the locations of their reference points. We then derive aggregate mobility patterns for each mobile phone tower to describe how individuals who share the same "home" anchor point move around in their daily lives.

(2) Based on a multi-level hierarchical agglomerative clustering algorithm, mobile phone towers are grouped into different clusters based on their aggregate mobility patterns. Spatial distributions of the derived clusters highlight distinct human mobility patterns in different areas of the city. We then discuss the socioeconomic and demographic characteristics of the regions covered by different cluster types to gain insights of human mobility patterns in a geographical context.

The remainder of the paper is organized as follows. "Related work" section provides a review of related work of this research. "Study area and mobile phone location dataset" section introduces the mobile phone location dataset and the study area. "Methodology" section describes the methods used in this study to derive reference points of individual activity space and to analyze aggregate human mobility patterns. "Analysis results and discussion" section discusses the analysis results. We summarize our findings and discuss future research directions in "Conclusion" section.

Related work

This section provides a brief review of selected studies related to the understanding of human mobility patterns, urban dynamics, individual activity anchor points using different types of tracking datasets (e.g., travel survey, GPS and mobile phone location data).

Understanding human mobility patterns

Understanding human travel and activity patterns has always been an interest in transportation research. Before mobile technologies pervaded, travel surveys were often used in activity-based studies. Inspired by the time geography framework proposed by Hägerstraand (1970), many studies have used travel surveys to understand how people perform their daily activities that are shaped by various types of constraints. These studies covered many important subjects such as multi-purpose and individual trip chaining behavior (Hanson 1980; Kitamura 1984; Newsome et al. 1998), social roles in travel behavior (Hanson and Hanson 1980; Kwan 1999), impacts of information and communications technologies on human travel (Mokhtarian 1998, 2003; Couclelis 2004), and human activity analysis in the space–time GIS framework (Kwan 2000; Miller 2005; Yu and Shaw 2008; Shaw et al. 2008; Shaw and Yu 2009; Chen et al. 2011; Yin et al. 2011). One important concept related to the understanding of human mobility patterns is individual activity space. Activity space denotes the daily environment that an individual is using for his/her activities (Golledge and Stimson 1997). There are several related concepts such as awareness space (Brown and Moore 1970), action space (Horton and Reynolds 1971), perceptual space (Relph 1976) and mental maps (Lynch 1984). In general, an individual's activity space is usually conceptualized as the locations that have been visited as well as the travel among these locations (Schönfelder and Axhausen 2003). As people's daily activities usually occur at a few locations such as home and workplaces, these locations are often considered as anchor points (Cullen and Godson 1975; Golledge and Stimson 1997) that determine the major characteristics of individual activity space. In order to describe people's daily activity patterns, considerable focus has been placed on measuring the size, geometry and structure of human activity space. For example, Dijst (1999) proposed a model with several discriminant functions to measure the size and shape of individual action space. Schönfelder and Axhausen (2003, 2004) used three measurement approaches (confidence ellipse, kernel density estimates and minimum spanning tree) to describe the structures of people's activity spaces based on travel surveys and GPS tracking datasets. These measures can effectively describe human mobility patterns from different perspectives. However, measures like confidence ellipse often consider the arithmetic mean of individual activity locations as the center of activity space when measuring the spatial dispersion of visited locations. Schönfelder and Axhausen (2003) suggested that individuals' activity anchor points, such as home locations, should be used to substitute the arithmetic mean to "gain a behaviorally more realist measure" (p. 9) on human mobility patterns. This inspires us to study how people move around in their daily lives by considering their home as a key reference point.

The wide adoption of location-aware technologies such as mobile phone positioning, GPS and Wi-Fi have opened up new opportunities to study the whereabouts of people in space and time. Rhee et al. (2011) analyzed individual movements based on GPS trajectories and the study showed that human travel behavior could be described mathematically by levy-walk model. Gonzalez et al. (2008) and Song et al. (2010) conducted their studies based on long-term mobile phone location data and concluded that a high degree of spatiotemporal regularities existed in human movements. Kang et al. (2010) studied individual human mobility patterns among different groups divided by gender and age. Yuan et al. (2012) included other indicators such as mobile phone usage and transportation network densities to study how they impacted human mobility patterns. These studies yield insights into important aspects of human activity patterns. However, few of them have attempted to characterize human mobility patterns by considering important anchor points (e.g., home locations) of people's daily activities. Moreover, little has been discussed in terms of how people's mobility patterns vary geographically, which serves as important information in decision and policy making. With this in mind, our study proposes a home-based approach to studying human mobility patterns within the context of their residential locations. The proposed measure is intended to provide an intuitive way of describing the general characteristics of human activity spaces and shed light on how people move around in a geographical context.

Understanding urban dynamics using GPS and mobile phone location data

As Batty (2009, p. 51) pointed out in his paper, urban dynamics refers to "representations of changes in urban spatial structure through time which embody a myriad of processes at work in cities of different, but often interlocking, time scales ranging from life cycle effects in buildings and populations to movements over space and time as reflected in spatial interactions". It is apparent that human travels and activities play an important role in the manifestation of the dynamics of our cities. The Global Positioning System (GPS) has become widely adopted in studies for understanding various aspects of urban dynamics such as individual commuting patterns (Shen et al. 2013), route choice behavior (Li et al. 2005; Papinski et al. 2009) and spread of disease (Vazquez-Prokopec et al. 2009). Because of its capability to capture human movements with high spatiotemporal accuracy (Richardson et al. 2013), GPS data have been accepted as a valuable source that can help enhance our understanding of human mobility and activity patterns in urban settings (Quiroga and Bullock 1998; Bohte and Maat 2009; Bazzani et al. 2010; Shoval et al. 2011).

As an emerging technique, mobile phone positioning has become a useful way to capture human movements and activities in space and time. Much research has been done to understand various aspects of urban dynamics based on mobile phone location data. In the Real Time Rome project by MIT SENSEable City Lab, large scale anonymous cellular network data was leveraged along with instantaneous positioning of taxis and buses to portray the picture of urban mobility in real time (Calabrese and Ratti 2006; Calabrese et al. 2011). Pulselli et al. (2005, 2006) and Ratti et al. (2006) used mobile phone location data in Milan, Italy to study intensities of human activities and their changes over space and time. Several studies used mobile positioning data to study recurring patterns and daily rhythms of human movements and activities (Reades et al. 2009; Ahas et al. 2010a; Sevtsuk and Ratti 2010). In addition, mobile positioning and GPS data were used for other purposes in urban studies such as tourism analysis (Ahas et al. 2007), traffic state monitoring (Sohn and Hwang 2008), epidemiology (Tatem et al. 2009) and location recommendation (Zheng et al. 2009). In our study, we investigate human mobility patterns using a home-based approach, which can enhance our understanding of people's use of space and shed light on urban dynamics.

Identifying individual activity anchor points

Human beings often spend a significant amount of time at specific places such as home and work locations. These key activity locations serve as important anchor points in people's everyday lives. Individual mobility patterns could be largely explained by the travel activities that occurred around these locations. Since the late 1990s, the GPS has become a popular means of collecting tracking data for studying human travel and activity patterns. Various approaches have been applied to derive trips and important locations from individual GPS trajectories. For example, Wolf et al. (2001, 2004) used GPS to collect travel data and developed several approaches to identifying trip destinations and purposes. Schuessler and Axhausen (2009) developed methods to derive individual trips and activities from GPS data. The analysis results in these studies were compared with travel diaries and demonstrated the feasibility of using GPS for understanding of individual activity patterns. Ashbrook and Starner (2003) and Zhou et al. (2007) developed algorithms to cluster GPS data points into individual activity anchor points. These studies benefited studies of human mobility patterns by uncovering their daily meaningful locations. However, GPS logs pose additional burdens on users and often suffer from problems such as signal loss (urban canyon effect) and short battery life. Mobile phones, on the other hand, become indispensable in people's daily life and various tracking techniques deployed on the devices (GSM and Wi-Fi) make it possible to collect locations of large population in space and time. Nurmi and Koolwaaij (2006) developed four different algorithms (graph clustering, online variant of graph clustering, spectral clustering and duration-based clustering) to derive semantics of individuals' frequently visited locations based on information of GSM cell transitions enriched GPS data. Ahas et al. (2010b) used a passive mobile positioning dataset which covers more than half million anonymous respondents in Estonia to extract individuals' personally meaningful locations, which "offers good potential for the monitoring of the geography and mobility of population" (p. 3). This study takes the advantage of rich information embedded in mobile phone location data to examine how people move around under the context of key activity locations.

Study area and mobile phone location dataset

The study area of this research is the city of Shenzhen, China. Shenzhen is located in southern China (Fig. 1a) and across the border from Hong Kong (Fig. 1b). Shenzhen is China's first special economic zone (SEZ) and covers an area of 1952 km^2. The city has six administrative districts: Baoan, Longgang, Nanshan, Futian, Luohu and Yantian (Fig. 1c). As one of southern China's financial centers, Shenzhen's population has been growing rapidly during recent years. The urbanization process and the rapid economic growth have attracted many immigrant workers who seek job opportunities in Shenzhen. By the end of 2011, the immigrant population accounts for more than 70 % of the total population in the city (see Gazette of the People's Government of Shenzhen Municipality 2011). The unique socioeconomic and demographic status of Shenzhen makes it an interesting area for the study of human mobility patterns.

The mobile phone location data used in this study was collected as call detail records[1] (CDRs) and each location record was generated when a mobile phone user placed or received a phone call/text message. Since such passively generated mobile phone data can be sparse in time, in this study we included 1,219,198 individuals covering a time span of 13 days with a criterion that each individual had at least 5 days with mobile phone location records. For privacy protection, this study did not obtain any personal information and each phone user in the dataset was assigned with an arbitrary user ID. In addition, all mobile phone location data were collected at the mobile phone tower level such that the specific activity locations are not revealed. As shown in Table 1, whenever an individual made or received a phone call or text message, a mobile phone record was generated including the user ID, date, starting time of mobile phone activity, record type (e.g., phone call vs. text message), and the coordinates (latitude/longitude) of the mobile phone tower which handled the mobile transaction. The density of mobile phone towers could vary in different

[1] There have been a number of publications (Phithakkitnukoon et al. 2010; Song et al. 2010; Yuan et al 2012; Yuan and Raubal 2012; Becker et al. 2013) that used mobile phone location datasets for studying human mobility patterns. The mobile phone location dataset used in this study was acquired through research collaboration with Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences and the research was approved through an Institutional Review Board (IRB) process.

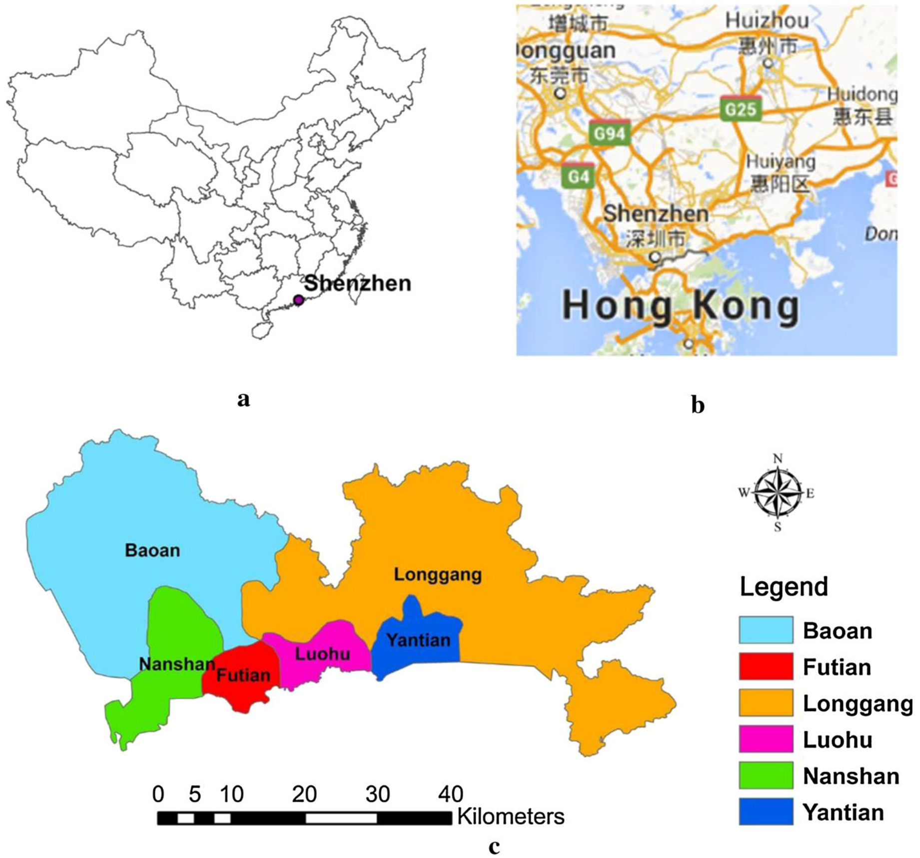

Fig. 1 **a** Shenzhen's location in China; **b** Relative locations of Shenzhen and Hong Kong (from Google Maps); **c** Shenzhen's administrative districts

parts of the study area. In this dataset, the average size of service area covered by a mobile phone tower is 0.67 km^2. Figure 2 shows the geographic distribution of all mobile phone towers included in this study. For each tower, the corresponding Thiessen Polygon is used to denote its service area.

Methodology

In this study, we propose a home-based approach to understanding aggregate human mobility patterns based on the mobile phone location dataset. First, we estimate each individual's "home" anchor point by analyzing his/her mobile phone location records. A modified standard distance measure is proposed to describe the spread of each individual's daily activities centered at the "home" anchor point during the study period. Then, each individual's estimated "home" anchor point is used as a reference point and individuals are grouped together based on the locations of their corresponding reference points. We derive aggregate mobility patterns to describe how individuals who share the same reference point move around during the study period. Finally, a multi-level hierarchical agglomerative

Table 1 Example of an individual user's mobile phone records during the data collection period

User ID	Date	Starting time	Record type	Latitude	Longitude
932********	Day 1	23:02:43	Call	22.*****	113.*****
932********	Day 1	23:11:56	Call	22.*****	113.*****
932********	Day 2	10:58:47	Message	22.*****	113.*****
…	…	…	…	22.*****	113.*****
932********	Day 13	09:22:50	Message	22.*****	113.*****

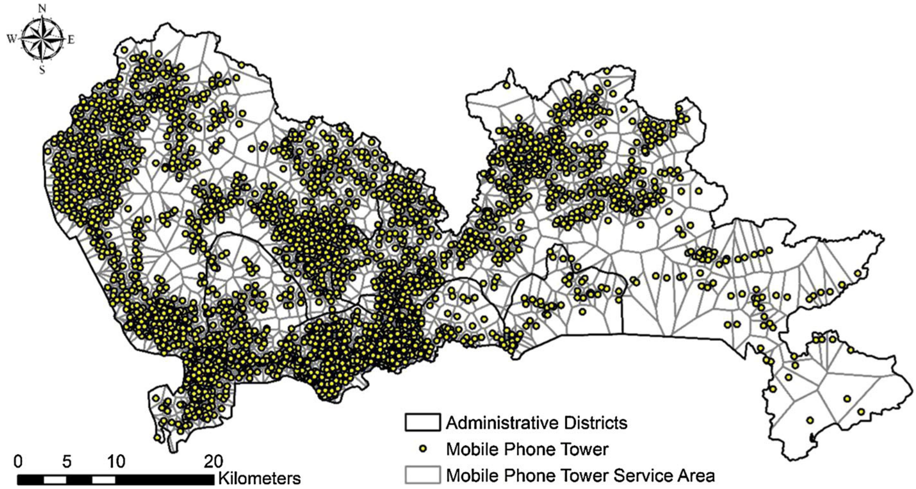

Fig. 2 Spatial distribution of mobile phone towers in Shenzhen (2976 in total). For each tower, the corresponding Thiessen Polygon is used to denote its service area

clustering algorithm is used to group the areas with similar aggregate mobility patterns into clusters. Spatial distributions of the derived clusters are then analyzed to highlight areas with distinctive aggregate mobility patterns.

Estimate individual "home" anchor point

Places such as home and work locations serve as important anchor points of people's daily activities. This information can be acquired from travel surveys and have been used as important reference points when analyzing human activity patterns. As residential locations are not explicitly given in the anonymous mobile phone location dataset, we estimate each individual's "home" anchor point by analyzing his/her mobile phone location records. The purpose is not to pinpoint an individual's residential location, but to provide a way of measuring the spread of each individual's daily activities. Moreover, these reference points can be used as reasonable estimates of individual home locations to further investigate geographic disparities of people's travel and activity patterns.

In this study, we apply an algorithm similar to Ahas et al. (2010b) to estimate individual "home" anchor points. For each individual in the dataset, we first compute the number of days the individual has at least one record at each mobile phone tower. We then choose the top two mobile phone towers with the highest number of days as the candidate towers. The algorithm is based on an assumption that, although many people may not have many mobile phone records at the home location each day, most of them would have some mobile phone uses at the home location on a regular basis. Using the number of days with at least one record rather than the absolute number of phone records during the study period could eliminate popular locations of mobile phone usage such as transportation hubs, restaurants, and shopping malls. Finally, for each of the two candidate towers, we calculate the total number of mobile phone records during the study period and the number of mobile phone records occurred during non-work time (before 6:00 and after 18:00). The tower with higher percentage of records during non-work time is identified as the "home" anchor point.

Calculate modified standard distance

Different measures such as standard distance (Bachi 1962) and radius of gyration (Gonzalez et al. 2008; Song et al. 2010) have been applied in activity-based research to describe the spread of individual activity space. These measures use the arithmetic mean (center of gravity) of each individual's activity locations as reference point. For example, the standard distance of an individual's activity space is calculated as:

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_c)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_c)^2}{n}} \tag{1}$$

where x_i and $y_i (i = 1, 2, \ldots, n)$ denote the coordinates of an individual's sampled activity locations. x_c and y_c stand for the coordinates of the arithmetic mean location.

In this study, we propose a modified standard distance to measure the spread of each individual's activity space centered at his/her estimated "home" anchor point. For a particular individual in the dataset, the modified standard distance is calculated as follows:

$$S'_D = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_h)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_h)^2}{n}} \tag{2}$$

where n denotes the total number of mobile phone records for the individual. x_i and $y_i (i = 1, 2, \ldots, n)$ denote the coordinates of the ith mobile phone record. x_h and y_h are the coordinates of the individual's estimated "home" anchor point. The modified standard distance is used to reflect the spread of an individual's activities centered around the "home" anchor point. Individuals with most of their activities distributed in the vicinity of their homes tend to have smaller S'_D than the people whose activities are more widely spread.

Derive aggregate mobility patterns for mobile phone towers

After extracting each individual's "home" anchor point and modified standard distance, we group individuals based on the mobile phone tower that represents the location of their estimated "home" anchor point. We then analyze the aggregate mobility pattern for each mobile phone tower to distinguish the mobility patterns of people in different geographic locations of the study area.

The aggregate mobility pattern of a mobile phone tower is measured by the modified standard distances (S'_D) of the corresponding individuals. As the modified standard distance may vary significantly from person to person, using measures such as arithmetic mean may not properly reflect the characteristics of people's mobility patterns. We therefore first group the values of S'_D into m classes. Each class $D_i(i = 1, 2, \ldots m)$ corresponds to a particular range of S'_D. The aggregate mobility pattern for a particular mobile phone tower T then can be represented as a vector V, which consists of the percentages of individuals with their S'_D falling in each of the $D_i(i = 1, 2, \ldots m)$ classes:

$$V = (q_1, q_2, \ldots, q_m) \tag{3}$$

where $q_i(i = 1, 2, \ldots m)$ is calculated as:

$$q_i = \frac{N_i}{N} \times 100\ \% \tag{4}$$

where N denotes the total number of individuals with their "home" anchor point at mobile phone tower T. While $N_i(i = 1, 2, \ldots m)$ denotes the number of individuals with their S'_D falling in the class of $D_i(i = 1, 2, \ldots m)$. Note that:

$$\sum_{i=1}^{m} N_i = N \tag{5}$$

$$\sum_{i=1}^{m} q_i = 1 \tag{6}$$

The vector V is useful to distinguish the varying activity patterns of people at different locations in the study area. For example, the mobility pattern of people who live in urban centers may be different from people who live in suburban areas and could show two vectors with different characteristics.

Clustering of mobile phone towers based on aggregate mobility patterns

Clustering multi-dimensional data can be challenging because of computational intensity and outliers. This study uses the agglomerative hierarchical clustering method with average linkage to group mobile phone towers into clusters (Han et al. 2011). Each cluster contains one vector initially. As shown in Eq. (7), the distance I between any two vectors $V(q_1, q_2, \ldots, q_m)$ and $V'(q'_1, q'_2, \ldots, q'_m)$ is calculated based on the Euclidean distance measure. The similarity index between any two clusters is thus calculated as the average value of distance I between vectors from the first cluster and the vectors from the second cluster. At each step, two clusters with the smallest similarity index are merged into a new cluster. The clustering process terminates when it reaches the specified number of clusters.

$$I = \sqrt{\sum_{i=1}^{m} (q_i - q'_i)^2} \tag{7}$$

As the multidimensionality of the vectors is likely to cause the agglomerative hierarchical method to generate several small clusters with unique patterns, these outlier clusters need to be addressed during the clustering process. We adopt a multi-level clustering algorithm similar to Chen et al. (2011) to remove the outlier clusters during each iteration of the agglomerative hierarchical clustering method. As the clustering process progresses, outlier clusters are removed step by step and clusters with distinctive characteristics start to emerge.

Analysis results and discussion

This section first presents the estimation of individual “home” anchor points and their geographic distributions in the study area. Next, the modified standard distance S'_D is calculated for each individual and the distribution of S'_D for 1,219,198 individuals in the dataset is presented to provide an overview of the spread of people’s activity space. We then group individuals based on their “home” anchor points at respective mobile phone towers to find the aggregate mobility pattern of each mobile phone tower using the method discussed in “Derive aggregate mobility patterns for mobile phone towers” section. A hierarchical clustering algorithm is subsequently performed to identify clusters of mobile phone towers that share similar aggregate mobility patterns based on the method described in “Clustering of mobile phone towers based on aggregate mobility patterns” section. Spatial distributions of the derived clusters of mobile phone towers are finally mapped to highlight areas with distinctive aggregate human mobility patterns in Shenzhen.

Estimation of individual “home” anchor point

We estimate each individual’s “home” anchor point and use it as the reference point when measuring the spread of an individual’s activity space. Figure 3 shows the density of estimated “home” anchor points of all 1,219,198 individuals in the dataset. Polygons with darker colors denote the mobile phone tower service areas with higher densities of anchor points. It can be perceived that many polygons with dark colors distribute in the southwest part, which refer to populated areas in Shenzhen. In this study, the estimated “home” anchor points are used to group individuals in space and derive aggregate human mobility patterns discussed below.

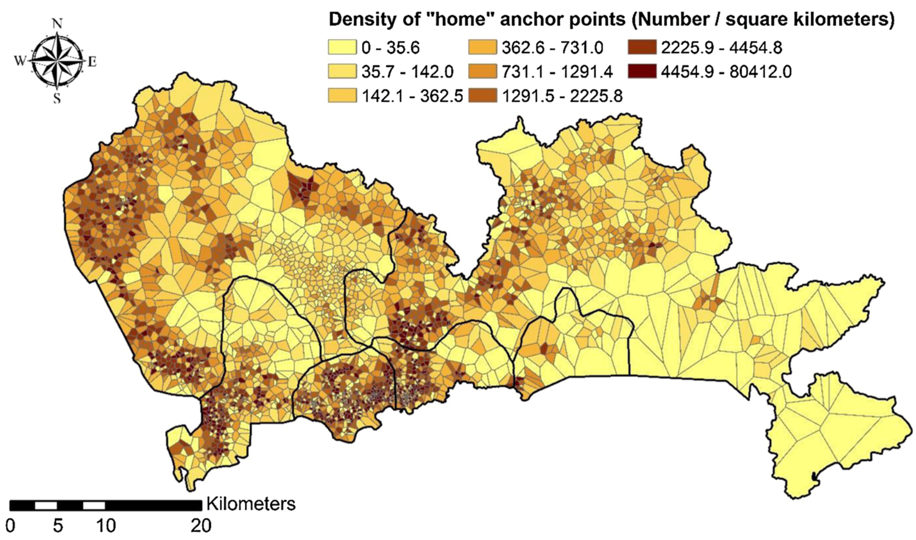

Fig. 3 Density of estimated “home” anchor points by mobile phone tower service area

Distribution of individual modified standard distance

With each individual's "home" anchor point, we compute the modified standard distance S'_D to examine the spread of individual activity space centered at the "home" anchor point. Figure 4 shows the distribution of S'_D for all 1,219,198 individuals in the dataset. The inserted histogram of individual S'_D (*binwidth* = 0.2 km) indicates that a large percentage of the individuals has a low value of S'_D. The histogram has its peak S'_D at 0.4–0.6 km, while the long tail suggests that some people have a large activity space.

The cumulative distribution in Fig. 4 shows that 43 % of the individual S'_D are within 1 km and 58.3 % the individual S'_D are within 2 km. On the other hand, 23.9 % individual S'_D are beyond 5 km. These statistics indicate that most people in Shenzhen have a relatively small activity space around their "home" anchor location during the 13-day study period. This encourages us to further explore the geographic landscape of people's mobility patterns.

Clustering patterns of mobile phone towers

Individuals are grouped based on their estimated "home" anchor points to derive aggregate mobility patterns by respective mobile phone towers. As mentioned in "Methodology" section, we divide S'_D into several classes and apply Eqs. (3) and (4) to measure the aggregate characteristics of human mobility patterns. According to the distribution of S'_D shown in Fig. 4, we divide the values of S'_D into the following 6 classes:

(1) $D_1: 0 \leq S'_D < 1$ km
(2) $D_2: 1\ \text{km} \leq S'_D < 2$ km
(3) $D_3: 2\ \text{km} \leq S'_D < 3$ km
(4) $D_4: 3\ \text{km} \leq S'_D < 4$ km

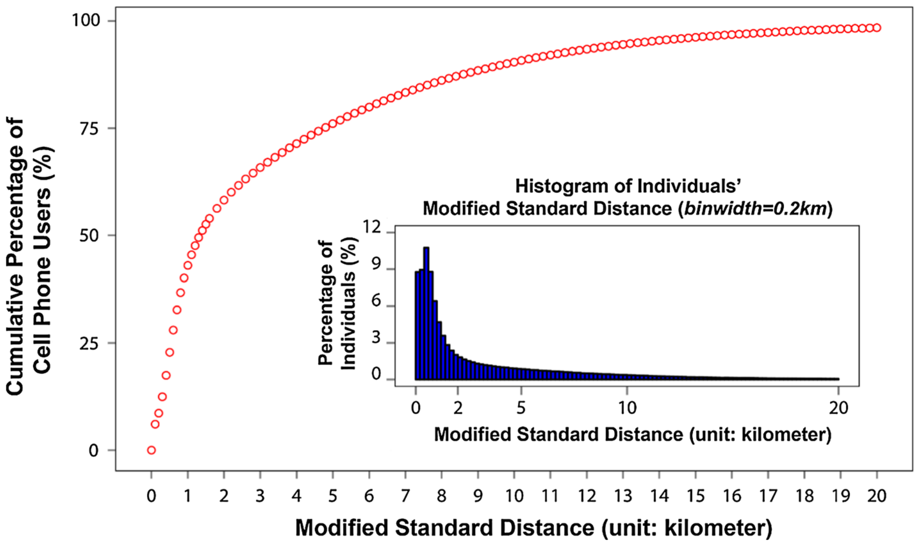

Fig. 4 Distribution of modified standard distance (S'_D) for 1,219,198 individuals in the dataset

(5) D_5: $4\ \text{km} \le S'_D < 5\ \text{km}$

(6) D_6: $S'_D \ge 5\ \text{km}$

By doing so, the aggregate mobility pattern of each mobile phone tower is represented as a vector $V(q_1, q_2, \ldots, q_6)$ recording the percentage of individuals with their S'_D falling in each of the classes $(D_1, D_2, \ldots, D_6)$. Table 2 gives an example of vectors for 5 randomly selected mobile phone towers. For example, the first element (q_1) of mobile phone tower 1 (T_1) denotes that for all individuals with their "home" anchor point at T_1, 27.78 % of them have a S'_D within [0,1 km). The vectors can distinguish different aggregate mobility patterns associated with various mobile phone towers. For example, for T_1, 42.59 % of the individuals have a S'_D equal or larger than 5 km. This number drops to 17.00 % when we look at T_5, where a large proportion of individuals exhibited low value of S'_D.

This study performs a multi-level hierarchical agglomerative clustering algorithm on 2976 mobile phone towers in Shenzhen that groups 2634 mobile phone towers into 9 clusters. The remaining 342 mobile phone towers are not grouped into any cluster either because they are removed as outliers during the clustering process or there is no "home" anchor point detected at those locations. Table 3 shows the clustering result of these mobile phone towers. The average percentages of $V(q_1, q_2, \ldots, q_6)$ for each cluster as well as the average percentages of $V(q_1, q_2, \ldots, q_6)$ for all 2976 mobile phone towers are computed for comparison purpose. The results clearly indicate distinctive characteristics of each cluster type. For example, for mobile phone towers in $C3$, a large percentage of people's S'_D are smaller than 1 km ($q_1 = 60.16\ \%$). Mobile phone towers in $C8$, on the other hand, have most people with S'_D at or larger than 5 km ($q_6 = 92.38\ \%$).

Figure 5 plots the results of these nine clusters in Table 3 to show their distinct patterns. Blue line with square markers in each plot represents the average percentages of $V(q_1, q_2, \ldots, q_6)$ for each cluster and red line with round markers denotes the average percentages of $V(q_1, q_2, \ldots, q_6)$ for all 2976 mobile phone towers. It is evident that, for mobile phone towers in $C1$, the percentage of people with $S'_D \le 1$ km ($q_1 = 27.75\ \%$) is lower than the overall average (43 %) while the percentage of people with $S'_D \ge 5$ km ($q_6 = 33.84\ \%$) is higher than the overall average (23.9 %). $C2$ has a pattern almost identical to the overall average. $C3$ is above the overall average of q_1 and below the overall average of q_6. $C4$ and $C5$ have dual peaks at q_2 and q_6 while having q_1 below the overall average. $C6$, $C7$, $C8$ and $C9$ represent the mobile phone towers with a large proportion of people having $S'_D \ge 5$ km and a low percentage of people having $S'_D < 1$ km. The next section further explores the spatial distributions of these clusters to better understand people's mobility patterns in a geographical context.

Table 2 Vectors representing the aggregate mobility patterns of 5 randomly selected mobile phone towers

Mobile phone tower ID	q_1 (%)	q_2 (%)	q_3 (%)	q_4 (%)	q_5 (%)	q_6 (%)
T_1	27.78	8.56	7.34	6.48	7.22	42.59
T_2	42.46	13.51	8.49	6.58	5.71	23.22
T_3	40.31	12.53	9.90	7.01	4.11	26.11
T_4	38.72	11.08	8.91	6.21	6.74	28.31
T_5	49.58	13.79	8.79	5.82	4.99	17.00
…	…	…	…	…	…	…

Table 3 Clustering results of 2976 mobile phone towers in the dataset

Cluster type	Cluster size	Average percentages of vector for each cluster					
		q_1 (%)	q_2 (%)	q_3 (%)	q_4 (%)	q_5 (%)	q_6 (%)
$C1$	617	27.75	14.4	9.57	7.74	6.71	33.84
$C2$	951	44.3	16.39	7.68	5.45	4.29	21.89
$C3$	496	60.16	14.27	5.51	3.69	2.75	13.61
$C4$	51	28.65	31.66	10.81	4.93	3.41	20.54
$C5$	28	15.75	26.14	8.49	9.32	6.55	33.74
$C6$	81	7.05	6.57	4.15	6.01	6.77	69.45
$C7$	165	15.6	11.86	9.08	7.24	7	49.22
$C8$	227	0.27	0.47	1.06	2.14	3.67	92.38
$C9$	18	7.11	4.56	16.31	3.78	6.93	61.3
Others	342	–	–	–	–	–	–
Overall	2976	43	15.3	7.6	5.6	4.6	23.9

Spatial distributions of clusters with different aggregate human mobility patterns

Spatial distributions of the above 9 clusters are mapped to further examine aggregate human mobility patterns in Shenzhen. To make the maps easier to understand, each of the maps shows the cluster(s) of similar aggregate mobility patterns. Figure 6 presents the spatial distribution of $C1$ with the locations of mobile phone towers represented by their service areas (i.e., Thiessen polygons in Fig. 2). $C1$, which refers to the mobile phone towers with only 27.75 % of people with $S'_D \leq 1$ km and the percentage of people with $S'_D \geq 5$ km is 33.84 %, mainly covers area*s* in the southern part of Shenzhen, with most of the mobile phone towers located in Nanshan, Futian and Luohu districts. Futian and Luohu are two major business districts where many financial and business centers are located. Nanshan, which has the highest per capita GDP among the six administrative districts in recent years, is a district with many universities and high-tech companies. In general, the economy of southern part of Shenzhen is more advanced than that of the northern part. The spatial distribution of $C1$ suggests that regional differences in economic development could be a potential factor affecting people's daily activity patterns in Shenzhen.

Compared with other clusters, $C2$ and $C3$ cover the areas where large proportion of people travelled within short distances during the study period. $C2$ has 44.3 % of people with $S'_D \leq 1$ km and $C3$ has 60.16 % of people with $S'_D \leq 1$ km. The percentage of people with $S'_D \geq 5$ km in $C2$ and $C3$ (21.89 and 13.61 % respectively) is lower than the overall average. Figure 7 shows the spatial distributions of $C2$ and $C3$, which cover mainly the northern part of Shenzhen in Baoan and Longgang districts. Baoan and Longgang are two industrial districts with many immigrant workers. In order to better understand the geographical context of these areas, we also display the major factories in Shenzhen in Fig. 7. We can see that the areas covered by $C2$ and $C3$ are generally co-located with the factories in Baoan and Longgang districts. Many factories in Shenzhen provide workers with dormitories adjacent to their workplace. In addition, many immigrant workers tend to rent apartments near their workplace to save commuting time and cost. We believe the low

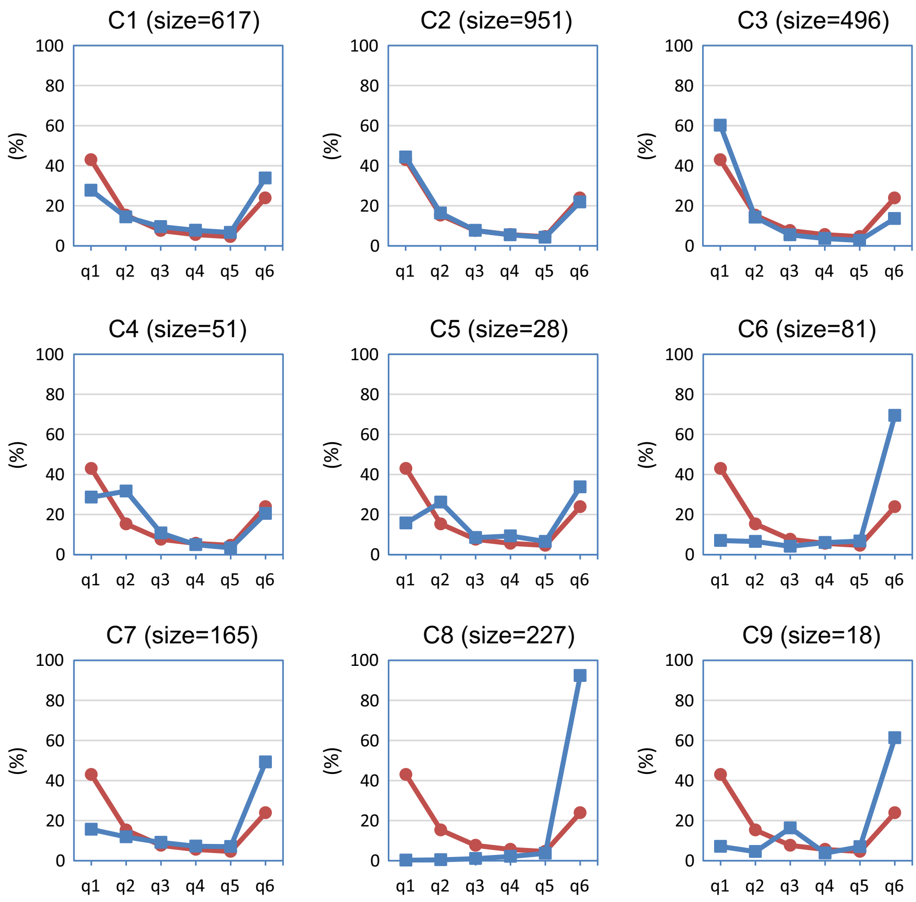

Fig. 5 Clustering patterns of the mobile phone towers. *Blue lines* with *square markers* represent the average percentages of vector for each cluster and *red lines* with *round markers* denote the average percentages of vector of all 2976 mobile phone towers as shown in Table 3. (Color figure online)

level of S'_D in these areas could be partially explained by the daily activity patterns of many immigrant workers in these areas of Shenzhen. There are two interesting findings worth noting here: (1) some areas in $C2$ are located around the boundary between Futian and Luohu districts. It indicates that a large proportion of people in these economically developed areas also have a relatively small activity space; (2) some areas in $C2$ and $C3$ (e.g., southwestern part of Baoan and western part of Longgang) fall along the major subway lines of Shenzhen. Yet people in these areas still experience less mobility.

$C6$, $C7$, $C8$ and $C9$ cover the areas with a large percentage of people with $S'_D \geq 5\,\mathrm{km}$. Figure 8 shows the spatial distributions of these four clusters. $C6$, $C7$ and $C9$ mainly locate in Nanshan, Yantian and some areas in southeast Longgang. The areas covered by $C8$ mainly locate in southeast Baoan. Table 3 shows that the percentage of people with $S'_D \geq 5\,\mathrm{km}$ for $C8$ is approaching 100 %. Displaying these clusters on Google Map shows that the major part of $C8$ is located in a geographically isolated area surrounded by mountains. Also, a large transportation hub (Shenzhen North Station) is located in this area,

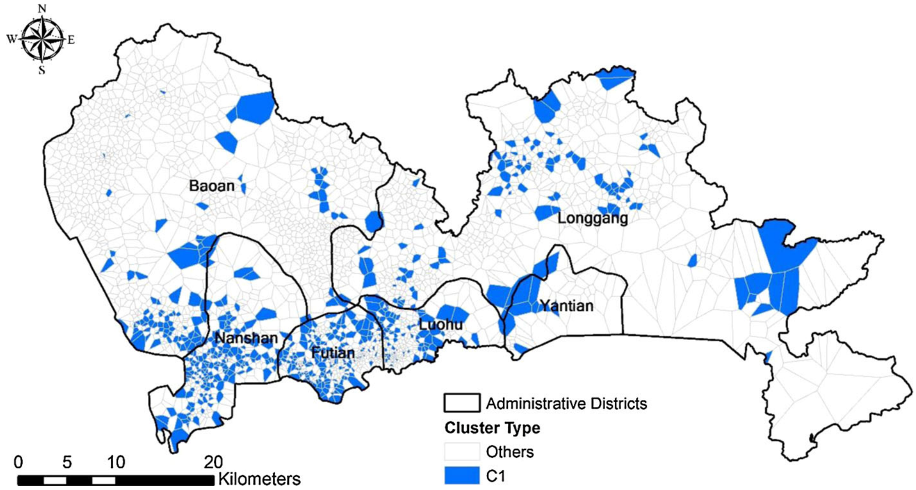

Fig. 6 Spatial distribution of mobile phone tower service areas in $C1$

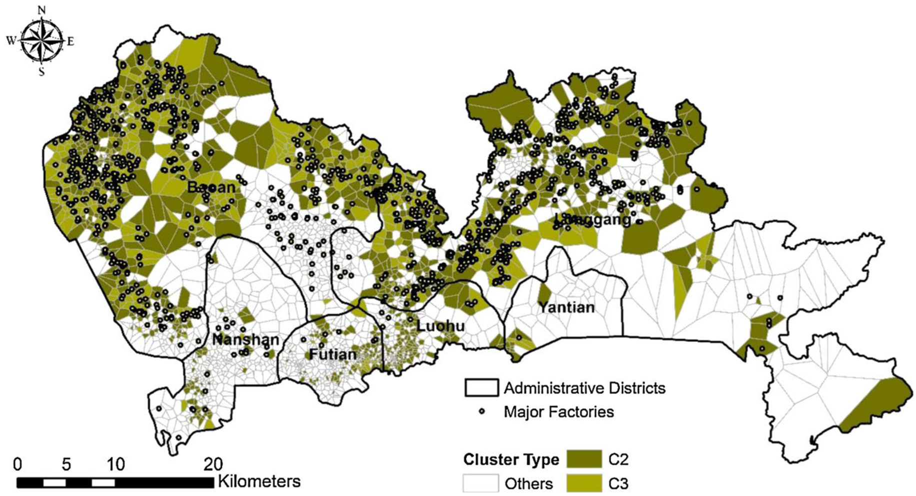

Fig. 7 Spatial distributions of $C2$ and $C3$ and major factories in Shenzhen

which offers transfer services to all kinds of transportation modes such as subway, taxi, bus, among others. Moreover, many inter-city highways such as Longda highway, Meiguan highway and Fulong expressway all go through this area. It should be noted that most areas covered by these four clusters are not highly populated areas in Shenzhen. One possible explanation to the observed spatial pattern is that these areas may have few job opportunities that force people to travel longer distances than people in other parts of Shenzhen. We also investigate the geographic distributions of $C4$, $C5$ as well as the outliers. Most areas covered by $C4$ are located in Baoan and Longgang districts while the areas covered by $C5$ and the outliers scatter across different districts. It appears that these clusters require additional research to identify possible causes to their spatial distribution patterns.

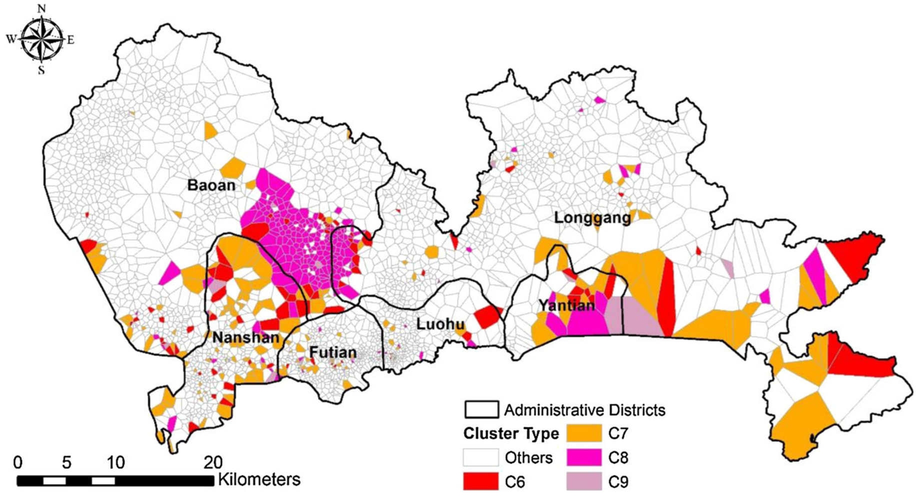

Fig. 8 Spatial distributions of $C6$, $C7$, $C8$ and $C9$

Major characteristics of the clusters derived from this study and their spatial distributions reveal several interesting aspects of human mobility in Shenzhen. For example, people who live around places with industrial parks and factories tend to have smaller activity space than people in other areas. One can also observe the differences of aggregate human mobility patterns between the northern and southern parts, as indicated by Figs. 6 and 7 respectively. The geographic differences are generally in accordance with the socioeconomic divide in Shenzhen. The research findings suggest that aggregate human mobility patterns derived from mobile phone location data could serve as useful indicators for the underlying socio-demographics and land use patterns. Note that several studies have used mobile phone location data for automated land use identification (Soto and Frías-Martínez 2011; Toole et al. 2012; Pei et al. 2014) and the prediction of socioeconomic levels (Soto et al. 2011). The analysis results in our study yield some novel insights into people's mobility patterns in Shenzhen, and serve as an initial effort for us to understand how people use urban space and its relationship with the built environment.

Conclusion

The proliferation of location-aware technologies has created many passively generated datasets that track the whereabouts of people in space and time. Unlike survey or interview datasets, passively generated datasets do not explicitly report every trip and destination of the respondents during a certain time period. The snapshots of people's daily traces included in these emerging datasets encourage novel approaches for studies of human travel and activity patterns. This paper introduces a home-based approach to understanding human mobility patterns based on passive mobile phone location data. The approach considers home location as an important reference point when analyzing people's use of space. The modified standard distance is introduced to describe the spread of an individual's activity space around the home location. Individuals are then grouped in space based on their estimated home locations to derive aggregate mobility patterns at each mobile phone tower. A multi-level hierarchical clustering algorithm is performed to group areas with similar aggregate mobility patterns into clusters. The study provides a unique perspective of examining people's use of urban space with respect to his/her residential location, and complement the behavioral insights (e.g., multi-purpose, trip chaining and route choice behavior) gained from travel surveys and GPS data.

A case study using CDRs of more than 1 million mobile phones over a 13-day period collected in Shenzhen, China is carried out to test the proposed approach and methods. The results of this case study clearly indicate the major characteristics of aggregate mobility patterns at individual mobile phone towers. The methods also identify clusters of mobile phone towers that share similar aggregate mobility patterns. Mapping the service areas of these clusters of mobile phone towers help shed light on the geographic differences of people's use of space among different parts of Shenzhen. These geographic patterns also match well with the different economic and transportation characteristics of the six districts in Shenzhen. Based on the findings, reasonable hypotheses of human travel behavior could be formulated by considering the socioeconomic and demographic characteristics of the built environment. The aggregate human mobility patterns derived at the mobile phone tower level can be further integrated with other datasets with important explanatory variables for travel behavior and policy analysis.

There are several limitations of this study. First, the passive mobile phone location data used in this study can be sparse in time since CDRs are created only when a mobile phone communication occurs. Second, this study removes those individuals who rarely use their mobile phones during the study period since it is not feasible to reliably estimate their home anchor point. The analysis results therefore do not reflect activity patterns of people who rarely use their mobile phones. Nevertheless, the mobile phone location dataset covers many people over a 13-day period in Shenzhen, China. Although some people may not be properly represented in this dataset, the analysis results offer useful insights of understanding aggregate mobility patterns in a geographic context. It is important to point out that different data sources (e.g., travel surveys, GPS, and passively generated mobile phone location data) have their respective strengths and weaknesses for studying human mobility patterns. The mobile phone location data used in this study is a valuable data source which could lead to a good understanding of people's use of space at a very large scale. However, mobile phone location data do not normally provide individual socioeconomic characteristics (e.g., age, gender, income, etc.) Studies based on mobile phone location data therefore can serve as a useful first step for more detailed and targeted follow-up studies. Moreover, while GPS data can capture locations at a finer spatial and temporal granularity but such data are not available for a large population, it would be promising to combine GPS-enabled mobile phones with passive mobile phone location data to further explore human mobility patterns.

In the future, we plan to derive several mobility indices of both spatial and temporal dimensions to describe people's mobility patterns in a more comprehensive way, and further investigate relationships of the derived indices with census (e.g., population density, gender, income, etc.) and transportation data (e.g., road networks and subway lines). This will help us better understand the intrinsic characteristics that drive people's daily travel and activities. In addition, it would be helpful to apply the same approach to similar datasets collected in other cities such that we can empirically test the robustness of the proposed approach and methods. This study demonstrates the potential of using passively generated mobile phone location data to help us gain better understanding of aggregate mobility patterns under the geographic context of a city. We believe this is only one of many possible ways of using passively generated tracking data to study human dynamics in a space–time context.

Acknowledgments This study was jointly supported by National Natural Science Foundation of China (41231171), National Key Technology R&D Program of China (#2012AA12A403-4), Arts and Sciences Excellence Professorship and Alvin and Sally Beaman Professorship at the University of Tennessee, Excellent Talents Funds at Wuhan University, Shenzhen Scientific Research and Development Funding Program (ZDSY20121019111146499, JSGG20121026111056204), and Shenzhen Dedicated Funding of Strategic Emerging Industry Development Program (JCYJ20121019111128765).

References

Ahas, R., Aasa, A., Mark, Ü., Pae, T., Kull, A.: Seasonal tourism spaces in Estonia: case study with mobile positioning data. Tour. Manag. **28**(3), 898–910 (2007)

Ahas, R., Aasa, A., Silm, S., et al.: Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: case study with mobile positioning data. Transp. Res. Part C **18**(1), 45–54 (2010a)

Ahas, R., Silm, S., Järv, O., Saluveer, E., et al.: Using mobile positioning data to model locations meaningful to users of mobile phones. J. Urban Technol. **17**(1), 3–27 (2010b)

Ashbrook, D., Starner, T.: Using GPS to learn significant locations and predict movement across multiple users. Pers. Ubiquit. Comput. **7**(5), 275–286 (2003)

Axhausen, K.W., Zimmermann, A., Schönfelder, S., Rindsfüser, G., Haupt, T.: Observing the rhythms of daily life: a six-week travel diary. Transporation **29**(2), 95–124 (2002)

Bachi, R.: Standard distance measures and related methods for spatial analysis. Pap. Reg. Sci. **10**(1), 83–132 (1962)

Batty, M.: Urban modeling. In: Kitchin, R., Thrift, N. (eds.) International Encyclopedia of Human Geography. Elsevier, Oxford (2009)

Bazzani, A., Giorgini, B., Rambaldi, S., Gallotti, R., Giovannini, L.: Statistical laws in urban mobility from microscopic GPS data in the area of Florence. J. Stat. Mech. **2010**, P05001 (2010)

Becker, R., Cáceres, R., Hanson, K., Isaacman, S., Loh, J.M., Martonosi, M., Rowland, J., Urbanek, S., Varshavsky, A., Volinsky, C.: Human mobility characterization from cellular network data. Commun. ACM **56**, 74–82 (2013)

Bohte, W., Maat, K.: Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: a large-scale application in the Netherlands. Transp. Res. Part C **17**(3), 285–297 (2009)

Brown, L.A., Moore, E.G.: The intra-urban migration process: a perspective. Geogr. Ann. **52**(B), 1–13 (1970)

Calabrese, F., Colonna, M., Lovisolo, P., et al.: Real-time urban monitoring using cell phones: a case study in Rome. Intell. Transp. Syst. IEEE Trans. **12**(1), 141–151 (2011)

Calabrese, F., Ratti, C.: Real time rome. Netw. Commun. Stud. **20**, 247–258 (2006)

Candia, J., González, M.C., Wang, P., Schoenharl, T., Madey, G., Barabasi, A.-L.: Uncovering individual and collective human dynamics from mobile phone records. J. Phys. A **41**, 224015 (2008)

Chen, J., Shaw, S.-L., Yu, H., et al.: Exploratory data analysis of activity diary data: a space–time GIS approach. J. Transp. Geogr. **19**(3), 394–404 (2011)

Cho, E., Myers, S.A., Leskovec, J.: Friendship and mobility: user movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD International Conference on knowledge discovery and data mining, pp. 1082-1090 (2011)

Couclelis, H.: Pizza over the Internet: e-commerce, the fragmentation of activity, and the tyranny of the region. Entrepreneurship Reg. Dev. **16**, 41–54 (2004)

Cullen, I., Godson, V.: Urban networks: the structure of activity patterns. Prog. Plan. **4**, 1–96 (1975)

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

Gazette of the People's Government of Shenzhen Municipality. Issue No. 17, Serial No. 741. http://www.sz.gov.cn/zfgb/2012_1/gb785/201204/t20120423_1844697.htm (2011). Accessed 22 April 2011

Golledge, R.G., Stimson, R.J.: Spatial Behavior. The Guiford Press, New York (1997)

Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.-L.: Understanding individual human mobility patterns. Nature **453**, 779–782 (2008)

Hägerstraand, T.: What about people in regional science? Pap. Reg. Sci. Assoc. **24**(1), 7–21 (1970)

Han, J., Kamber, M., Pei, J.: Data Mining: Concepts and Techniques, 3rd edn. Elsevier/Morgan Kaufmann, Amsterdam/Burlington (2011)
Hanson, S.: The importance of the multi-purpose journey to work in urban travel behavior. Transportation **9**, 229–248 (1980)
Hanson, S., Hanson, P.: Gender and urban activity patterns in Uppsala, Sweden. Geogr. Rev. **70**, 291–299 (1980)
Horton, F., Reynolds, D.R.: Effects of urban spatial structure on individual behaviour. Econ. Geogr. **47**, 36–48 (1971)
Kang, C., Gao, S., Lin, X., et al.: Analyzing and geo-visualizing individual human mobility patterns using mobile call records. In: 18th International Conference on Geoinformatics (pp. 1–7). IEEE (2010)
Kitamura, R.: Incorporating trip chaining into analysis of destination choice. Transp. Res. Part B **18**, 67–81 (1984)
Kwan, M.P.: Gender, the home-work link, and space–time patterns of nonemployment activities. Econ. Geogr. **75**, 370–394 (1999)
Kwan, M.P.: Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. Transp. Res. Part C **8**, 185–203 (2000)
Levinson, D., Kumar, A.: Activity, travel, and the allocation of time. J. Am. Plan. Assoc. **61**(4), 458–470 (1995)
Li, H., Guensler, R., Ogle, J.: Analysis of morning commute route choice patterns using global positioning system-based vehicle activity data. Transp. Res. Rec. **1926**, 162–170 (2005)
Lynch, K.: Good City Form. MIT Press, Cambridge (1984)
Miller, H.J.: A measurement theory for time geography. Geogr. Anal. **37**(1), 17–45 (2005)
Mokhtarian, P.L.: A synthetic approach to estimating the impacts of telecommuting on travel. Urban Stud. **35**(2), 215–241 (1998)
Mokhtarian, P.L.: Telecommunications and travel: the case for complementarity. J. Ind. Ecol. **6**(2), 43–57 (2003)
Newsome, T.H., Walcott, W.A., Smith, P.D.: Urban activity spaces: illustrations and application of a conceptual model for integrating the time and space dimensions. Transportation **25**, 357–377 (1998)
Nurmi, P., Koolwaaij, J.: Identifying meaningful locations. In: 3rd Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services, San Jose, California: IEEE (pp. 1–8) (2006)
Papinski, D., Scott, D.M., Doherty, S.T.: Exploring the route choice decision-making process: a comparison of planned and observed routes obtained using person-based GPS. Transp. Res. Part F **12**, 347–358 (2009)
Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., Zhou, C.: A new insight into land use classification based on aggregated mobile phone data. Int. J. Geogr. Inf. Sci. **28**, 1–20 (2014)
Phithakkitnukoon, S., Horanont, T., Lorenzo, G.D., Shibasaki, R., Ratti, C.: Activity-aware map: identifying human daily activity pattern using mobile phone data. Hum. Behav. Underst. **6219**, 14–25 (2010)
Pulselli, R., Pulselli, F., Ratti, C., Tizzi, E.: Dissipative structures for understanding cities: resource flows and mobility patterns. In: Proceedings of the First International Conference on Built Environment Complexity, (pp. 271–279) (2005)
Pulselli, R., Ratti, C., Tiezzi, E.: City out of chaos: social patterns and organization in urban systems. Int. J. Ecodyn. **1**(2), 125–134 (2006)
Quiroga, C.A., Bullock, D.: Travel time studies with global positioning and geographic information systems: an integrated methodology. Transp. Res. Part C **6**, 101–127 (1998)
Ratti, C., Williams, S., Frenchman, D., et al.: Mobile landscapes: using location data from cell phones for urban analysis. Environ. Plan. **33**(5), 727 (2006)
Reades, J., Calabrese, F., Ratti, C.: Eigenplaces: analysing cities using the space–time structure of the mobile phone network. Environ. Plan. **36**(5), 824–836 (2009)
Relph, E.: Place and Placelessness. Pion Limited, London (1976)
Rhee, I., Shin, M., Hong, S., et al.: On the levy-walk nature of human mobility. IEEE/ACM Trans. Netw. **19**(3), 630–643 (2011)
Richardson, D.B., Volkow, N.D., Kwan, M.-P., Kaplan, R.M., Goodchild, M.F., Croyle, R.T.: Spatial turn in health research. Science **339**, 1390 (2013)
Schönfelder, S., Axhausen, K.W.: Activity spaces: measures of social exclusion? Transp. Policy **10**(4), 273–286 (2003)
Schönfelder, S., Axhausen, K. W.: Structure and innovation of human activity spaces. Arbeitsberichte Verkehrs-und Raumplanung, 258 (2004)
Schuessler, N., Axhausen, K.W.: Processing raw data from global positioning systems without additional information. Transp. Res. Rec. **2105**, 28–36 (2009)

Sevtsuk, A., Ratti, C.: Does urban mobility have a daily routine? Learning from the aggregate data of mobile networks. J. Urban Technol. **17**(1), 41–60 (2010)
Shaw, S.L., Yu, H., Bombom, L.S.: A space–time GIS approach to exploring large individual-based spatiotemporal datasets. Trans. GIS **12**, 425–441 (2008)
Shaw, S.-L., Yu, H.: A GIS-based time-geographic approach of studying individual activities and interactions in a hybrid physical–virtual space. J. Transp. Geogr. **17**(2), 141–149 (2009)
Shen, Y., Kwan, M.-P., Chai, Y.: Investigating commuting flexibility with GPS data and 3D geovisualization: a case study of Beijing, China. J. Transp. Geogr. **32**, 1–11 (2013)
Shoval, N., Wahl, H.-W., Auslander, G., Isaacson, M., Oswald, F., Edry, T., Landau, R., Heinik, J.: Use of the global positioning system to measure the out-of-home mobility of older adults with differing cognitive functioning. Ageing Soc. **31**, 849–869 (2011)
Sohn, K., Hwang, K.: Space-based passing time estimation on a freeway using cell phones as traffic probes. IEEE Trans. Intell. Transp. Syst. **9**(3), 559–568 (2008)
Song, C., Qu, Z., Blumm, N., Barabasi, A.-L.: Limits of predictability in human mobility. Science **327**(5968), 1018–1021 (2010)
Soto, V., Frías-Martínez, E.: Automated land use identification using cell-phone records. In: Proceedings of the 3rd ACM International Workshop on MobiArch. ACM, pp. 17–22 (2011)
Soto, V., Frias-Martinez, V., Virseda, J., Frias-Martinez, E.: Prediction of socioeconomic levels using cell phone records. In: Conejo, J.A.K.R., Oliver, J.L.M.N. (eds.) User Modeling, Adaption and Personalization, pp. 377–388. Springer, Berlin (2011)
Tatem, A.J., Qiu, Y., Smith, D.L., et al.: The use of mobile phone data for the estimation of the travel patterns and imported *Plasmodium falciparum* rates among Zanzibar residents. Malar J **8**, 287 (2009)
Toole, J.L., Ulm, M., González, M.C., Bauer, D. (2012). Inferring land use from mobile phone activity. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing. ACM, pp. 1–8
Vazquez-Prokopec, G.M., Stoddard, S.T., Paz-Soldan, V., Morrison, A.C., Elder, J.P., Kochel, T.J., Scott, T.W., Kitron, U.: Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. Int. J. Health Geogr. **8**, 68 (2009)
Wolf, J., Guensler, R., Bachman, W.: Elimination of the travel diary: experiment to derive trip purpose from global positioning system travel data. Transp. Res. Rec. **1768**, 125–134 (2001)
Wolf, J., SchöUnfelder, S., Samaga, U., Oliveira, M., Axhausen, K.W.: Eighty weeks of global positioning system traces: approaches to enriching trip information. Transp. Res. Rec. **1870**, 46–54 (2004)
Yin, L., Shaw, S.-L., Yu, H.: Potential effects of ICT on face-to-face meeting opportunities: a GIS-based time-geographic approach. J. Transp. Geogr. **19**, 422–433 (2011)
Yu, H., Shaw, S.L.: Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach. Int. J. Geogr. Inf. Sci. **22**, 409–430 (2008)
Yuan, Y., Raubal, M.: Extracting dynamic urban mobility patterns from mobile phone data. In: Xiao, N., Kwan, M.-P., Goodchild, M.F., Shekhar, S. (eds.) Geographic Information Science. Springer, Berlin pp. 354–367 (2012)
Yuan, Y., Raubal, M., Liu, Y.: Correlating mobile phone usage and travel behavior—a case study of Harbin, China. Comput. Environ. Urban Syst. **36**(2), 118–130 (2012)
Zheng, Y., Zhang, L., Xie, X., Ma, W.-Y.: Mining interesting locations and travel sequences from GPS trajectories. In: Proceedings of the 18th International Conference on World Wide Web (pp. 791–800). ACM, Madrid, Spain (2009)
Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L.: Discovering personally meaningful places: an interactive clustering approach. ACM Trans. Inf. Syst. **25**(3), 12 (2007)

Yang Xu is a Ph.D. candidate in the Department of Geography at the University of Tennessee, Knoxville. His research interests include geographic information science (GIScience), space–time GIS, spatiotemporal analysis of human dynamics, and GIS for transportation (GIS-T).

Shih-Lung Shaw is Alvin and Sally Beaman Professor and Arts and Sciences Excellence Professor in the Department of Geography at the University of Tennessee, Knoxville. His research interests include transportation geography, geographic information science (GIScience) and spatio-temporal analysis, especially with topics related to transportation planning and modeling, air transportation, time geography, effects of information and communications technologies (ICT) on human activity and travel patterns, spatiotemporal analysis of human dynamics, space–time GIS, and GIS for transportation (GIS-T).

Ziliang Zhao is a Ph.D. student in the Department of Geography at the University of Tennessee, Knoxville. His research includes space–time GIS data model design, geovisualization, and big data analytics.

Ling Yin is an associate professor at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. Her research interests include spatial–temporal data analysis and GIS for transportation.

Zhixiang Fang is a professor of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. His research focuses on spatiotemporal modeling in transport geography.

Qingquan Li is a Professor with Shenzhen University, Guangdong, China and Wuhan University, Wuhan. His research interests include 3-D and dynamic modeling in GIS, location-based service, surveying engineering, integration of GIS, global positioning system and remote sensing, and intelligent transportation system.