Research on Collective Human Mobility in Shanghai Based on Cell Phone Data

Xiyuan Ren, Tongji Universiity, Shanghai, China

De Wang, Tongji Universiity, Shanghai, China

ABSTRACT

The high-frequency mobility of a massive population has caused an enormous influence on the urban internal structure, which is unable to be described by traditional data sources. While recent advances in location-based technologies provides new opportunities for researchers to understand daily human movements and the structure as a whole. The article aims to explore human spatial movements and their aggregate distribution in Shanghai using large-scale cell phone data. The trajectory of each individual is extracted from cell phone data after data cleansing. Then, an indicator system which includes mobility intensity, mobility stability, influential range, and temporal variation is developed to describe collective human mobility features in census tracts scale. Finally, spatial elements are extracted using the indicator system and the structure of human mobility in Shanghai is discussed.

KEYWORDS

Cell Phone Data, Human Mobility, Indicator System, Shanghai, Urban Structure

INTRODUCTION

With the development of traffic information technology and acceleration of life and work rhythm, the high-frequency mobility of massive population has caused enormous influence on the urban internal structure, which has brought a new challenge to urban spatial policy. How to measure the intensity of human mobility and explain its dynamic structure is of great significance to understand the law of intra-urban dynamic structure and support real-time population management. Traditional population surveys such as national censuses are conducted every ten years, which reflect only long-term changes. While human activities often result in large-scale population movements within a short period. Such rapid and irregular changes in urban space exceed the scope of traditional census data, and thus challenge population management in recent years. How to describe, measure and evaluate collective human mobility in urban space? Does the distribution of human behavior fit into the urban spatial form? These two problems are difficult to be solved with the traditional urban structure model and static spatial analysis method.

The developing succession of the urban spatial structure is influenced not only by the external influences of the region's natural and humanistic environment but also by factors such as internal division of functions and land use. Relative researches have reached a considerable depth in the evolution and morphology of urban spatial structure with a series of empirical research by like

Zhang (2001), Wang and Sun (2011), and Lee (2007). However, limited by the research data and methods, there is always a certain distance between macroscopic scale structure and the real human behavior. Whether our urban structure conforms to the real law of urban life, there isn't a quantitative explanation that can be generally accepted.

In the era of big data, an increasingly broad array of user-generated data like cell phone data derived from location-based services and the Global Positioning System (GPS) provides new possibilities to analyze human behavior and their spatial distribution. Mobile positioning big data have opened up the interaction between human behavior and urban spatial structure, which not only enables quantitative analysis of urban structure under greater spatial and temporal granularity, but also breaks the long-term research barrier of "interpreting space with space", revealing the spatial law behind complex form of the city from the perspective of collective human activity. Relative research has proved that the cell phone data can be used to identify these daily activities, such as commuting, traveling, and the whole space-time path (Ahas et al., 2015; Widhalm et al., 2015; Phithakkitnukoon et al., 2010; Song et al., 2010; Doyle et al., 2014). However, cell phone data still have great potential that needs to be exploited to advance human behavior studies further.

Therefore, the paper aims to explore urban spatial structure from the perspective of collective human mobility. Two-week cell phone data of 2G users in Shanghai are used. By revealing the intensity, stability, spatial influence and temporal variation of daily human mobility, we extracted spatial elements and concluded them into a spatial structure. The remainder of this paper is organized as follows. "Literature Review" gives a panorama of relevant research on the urban structure and human behavior. In the part of "Methodology", a detailed introduction to how cell phone data can be applied to urban spatial structure research is described. Section 4 illustrates the results of the case study performed, which shows collective human behavior in Shanghai. In Section 5, the paper discusses about the results and possibilities for application. The final part presents a conclusion of this article and provides insights for further research. The results can provide a reference for the formulation of urban space policy and population management strategy.

LITERATURE REVIEW

Definition of Human Mobility

In China, human mobility tends to be confused with population mobility since the Chinese household registration system makes the movements of people more complicated than other countries. Definitions in relevant researches can be divided into three types. First, "Population Migration", which is defined as movement that crosses the boundaries of certain jurisdictions, usually brings changes in household registration (Zhao, 2005). Then, "Population Movement", which is referred to the change of residence or working place within a certain period, includes short-term, long-term and permanent relocation (Zhao et al, 2013). Thirdly, "Human Flow/Mobility", which is defined as residents' locational changes in a short period, usually occurs inside the city and continuously changes over time (Wen et al., 2000; Kwan et al., 2008). The last concept is more closed to "human mobility" in this research.

Human Mobility Researches Based on Traditional Data

Spatial features of human mobility are always an essential topic in the field of population geography and urban planning. Gaining spatial-temporal information of residents' travel behavior is the basic step of these researches. Trip survey data and activity-daily-record data are the two main data sources. For example, Pan et al. (2009) and Zhang et al. (2008) used trip survey data to explore residents' travel patterns in Shanghai and Beijing. Kwan (1998) and Shaw et al. (2000) used activity-daily-record data to show specific individual travel information within the city like commuting and shopping.

Based on these two types of data, researchers focus on the description of individual travel patterns, the explanations of decision-making mechanisms, and the build of a comprehensive activity

space system. For example, Chen et al. (2011) applied 3D visualization technology based on GIS technology and geographic calculation method to mined spatial-temporal behavior patterns of residents in Beijing. He then divided the spatial patterns of residents' daily activities into six categories; Golledge (2003) analyzed the spatial structure of big cities in the United States from three aspects: declarative, procedural and configuration information, and found that spatial information affect residents' perception of urban space; Chai et al. (2015), analyzed the spatial-temporal behavior of residents in Beijing, and drew four kinds of ideal urban life circle including basic life circle, commuting life circle, expansion life circle and collaborative life circle. He suggested that these circles show a comprehensive human mobility structure.

To sum up, human mobility researches based on traditional data are mainly based on behavioral theory, time geography and activity analysis method. Relative results and discussions provide abundant theorical and practical foundations, which paves way for further researches. However, research gap is also obvious due to the spatial and temporal limitation of traditional data. First, most studies focus on residents' specific activity space (such as commuting, shopping, etc.) while few of them can conclude a comprehensive human activity structure except for some primary research on "life circle". Second, limited by the sample coverage of survey data, most researches are conducted under resident-scale. When it comes to human mobility under urban scale, few researches can cover.

Human Mobility Based on Cell Phone Data

With the increasing availability of big data, cell phone data is widely used in the human behavior and urban spatial studies. The reliability of cell phone data on human mobility research has also been verified by many scholars from different aspects.

In terms of identifying behavioral activities, Ahas et al. (2010) used one-year cell phone data of Estonia to construct a stop point recognition model so that to identify individuals' stop points (including residence, workplace, other places, etc.), after which the population registration data were utilized for comparison and verification. Based on communication data from 100,000 people in Portugal, Csáji et al. (2013) simplified data through clustering and principal component analysis. He identified their residences and workplaces and compared them with census data. Widhalm et al. (2015) explored cell phone data and called detail (CDR) data in Vienna, Austria, and Boston. Considering low accuracy of cell phone data, he also constructed Markov network model and proposed a method for identifying the behavior patterns of residents. Wang et al. (2018) provided a review of existing travel behavior studies that have applied mobile phone data, presented the progress that has been achieved to date, and summarized basic methods in stays and trips identification.

When it comes to the description of human behavior, Phithakkitnukoon et al. (2010) visualized and analyzed nearly one million cell phone data of users in Boston central region. He found that people from the same workplace have strong correlations in their daily activities. Ahas et al. (2015) defined four indicators: "midnight", "morning start-time", "noon" and "day length" based on the call details (CDR) data of three cities -- Harbin, Paris, and Tallinn. By comparing those indicators within and among cities, he revealed the differences in the spatial and temporal behaviors of residents in different cities and the central and suburban areas within cities. Yuan et al. (2012) used cell phone data of Harbin to study the correlations among mobile phone usage and radius, eccentricity and entropy. The results proved that characteristics such as age, gender, social time and built-up environment had impacts on the usage of mobile phone and residents' activities. Ran (2013) identified user's travel trajectory through call detail (CDR) data and analyzed the distribution of resident population and employment population, commuting pattern, OD matrix, the commuting pattern in specific regional and commuting characteristics of floating population. Zhong et al. (2017) used cell phone data to build a dynamic "population-time-behavior" analysis framework and then explored the dynamic spatial structure of population and activities in Shanghai.

At the same time, some research topics have gradually shifted from the identification and description of behavior to the simulation of human mobility. Song et al. (2010) explored the possibility

and limitations of cell phone data for predicting human mobility. It is considered that irregular human mobility (except for commuting and leisure travel) is inherently unpredictable. In the same year, Song et al. (2010) used cell phone data to construct a random walk (CTRW) model so that to stimulate human mobility quantitatively. Doyle et al. (2014) visualized population movements across Ireland with call details (CDR) data, by using Markov chain model to rank population hotspots, he found that the results are strongly correlated with census data.

Existing researches have proved the feasibility that cell phone data can be used to identify daily activities, such as commuting/leisure travel, temporal and spatial patterns, etc. Some studies have further explored the possibility of simulating human mobility with cell phone data. Main indicators in these kinds of researches contains activity intensity, activity range, activity stability and the change according to time period. Based on existing explorations, a comprehensive indicator system can be concluded which shows collective human mobility features from mobility intensity, mobility stability, influential range, and temporal variation.

METHODOLOGY

Data Introduction

In the paper we use cell phone data generated by 2G subscribers of Shanghai Mobile from March 15 to March 28, 2014. The data record 16-18 million different mobile phone users (about 70% of the 24.15 million population in Shanghai), and the total number of daily records is about 600-800 million. The data is recorded by the lac cell towers, which mainly contain five fields. They are user MSID (unique identification code for each mobile user), TStamp (time stamp), LAC (general location area number) and CellID (detail location area number), and EventID (type of the record). Table 1 shows an example of the original cell phone data.

The total number of lac cell towers in Shanghai is about 37,000, with 17,000 in the main city, accounting for 47.5% of the total. The distribution of these towers is shown in Figure 1, which shows obvious regularity in space. The average distance between two lac cell towers is about 100-300m in the main city, while in outskirt areas the value increases to 1000-3000m. In Shanghai, more than 70% of the lac cell towers are within 500m from their nearest neighbors, so the spatial accuracy of these lac cell towers can support the study of collective human mobility well.

Data Process: Identification of Individual Travel Chain

The whole processes are showed in Figure 2, where three steps are necessary to convert raw data into individual travel chain data in order to support indicator calculation.

First, raw cell phone data have intrinsic problems such as noise and positioning uncertainty, and thus need to be cleaned (Calabrese et al., 2015). Signal drifts mainly cause noise in raw cell phone data, which refers to a movement from one place to another that is almost impossible to reach in that time interval for mobile phone users. Accordingly, the research eliminates the piece of data if it shows the speed of a user is faster than 100m/s. Positioning uncertainty usually occurs when the position of a user is recorded by several lac cell towers at the same time. So, if several lac cell towers closed to each other record a same user in a short time, the research merge these points and use position with the latest timestamp.

Table 1. An example of cell phone data

MSID	Timestamp	LAC	CellID	EventID	Flag
6F7ABDD2ECF3C941FAC1EFC884692112	20140303000001	6216	58547	061	0

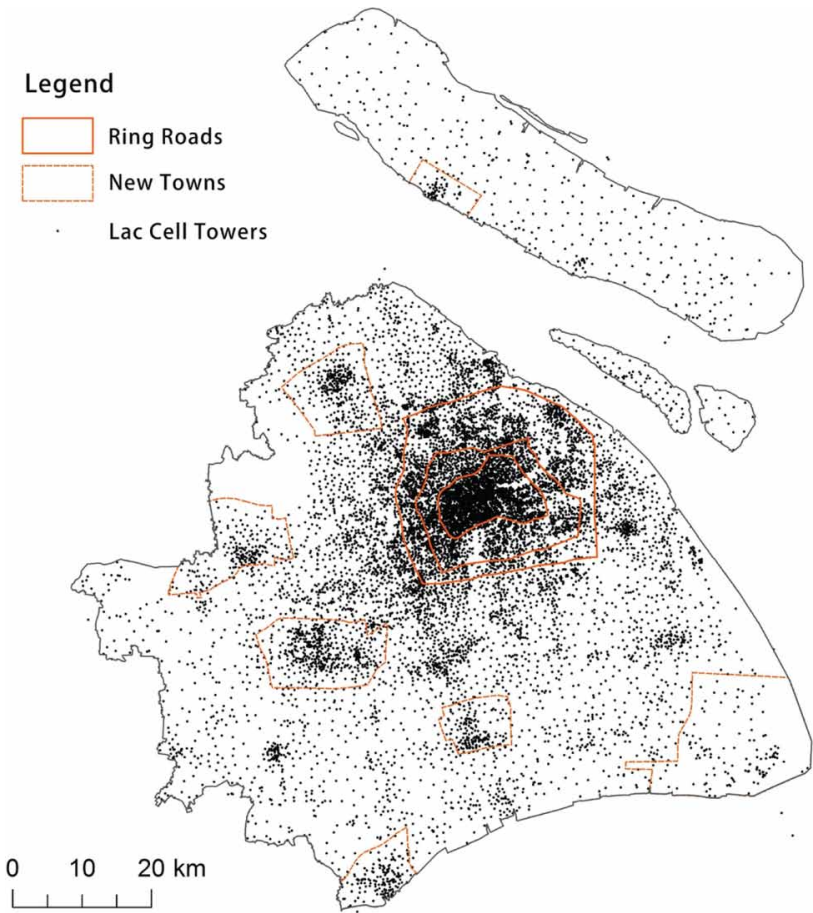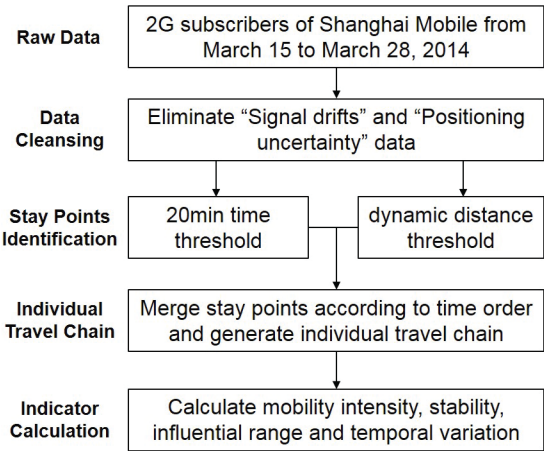Figure 1. Distribution of lac cell towers in Shanghai

Figure 2. A flowchart of data process

Secondly, the distance and time thresholds are set to identify the stay points of mobile phone users (Chen et al., 2014; Alexander et al., 2015). If users spend a period longer than time threshold in a range whose radius is shorter than distance threshold, they will be identified stays in the basic lac cell towers. Time threshold varies from 600s to 2700s according to the definition of "stay" (Shen et al., 2014; Feng et al., 2014). Since 1200s is a time threshold commonly used, the paper set time threshold to 20 minutes (duration of two nearest timestamps should be longer than 20min). When it comes to distance threshold, considering the average distance between lac cell towers increases form the main city to suburban areas in Shanghai, the paper uses a dynamic distance threshold which changes according to the density of lac cell towers in certain areas. Figure 3 shows the calculation of distance threshold: 1) divide six sectors with the basic lac cell tower as the center, every sector is 60 degrees; 2) calculation linear distance from the nearest lac cell tower in each sector, and get six distances; 3) choose the maximum value in these six distances as distance threshold. The advantage of dynamic distance threshold is that it can weaken the error brought by disequilibrium of lac cell tower density between city center and outer suburbs, and thus improve the accuracy of stay point identification.

Finally, stop points of the same users are merged according to the time order, and the interval between two stop points are identified as one trip. Individual travel chain data include six fields: msid (user's identification number), olaccell (original location), dlaccell (destination location), otime (start timestamp), dtime (end timestamp) and type (0 represents a stay, 1 represents a trip). Table 2 shows

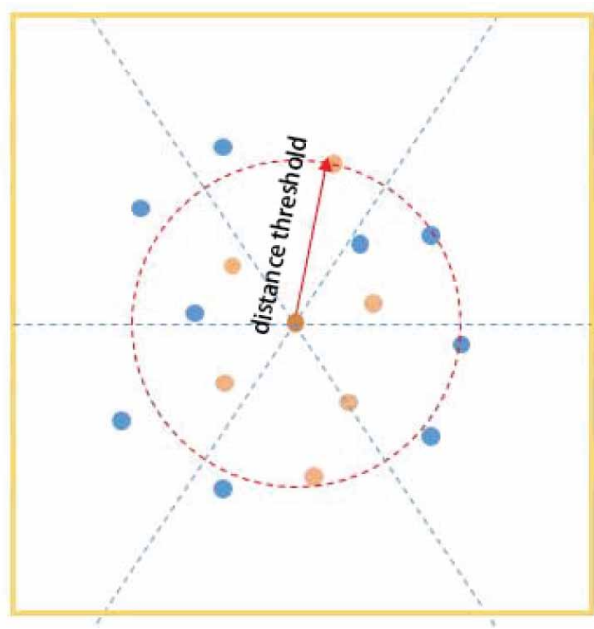**Figure 3. Calculation of dynamic distance threshold**

Table 2. An example of individual travel chain data

msid	olaccell	dlaccell	otime	dtime	Type
4B6…8C54	614613090	615825235	20140318064516	20140318072616	1
4B6…8C54	615825235	615825235	20140318072616	20140318091907	0
4B6…8C54	615825235	615824947	20140318091907	20140318092401	1

an example of individual travel chain data. If we accumulate individual trips into census tracts scale (Figure 4), we can find the most active period are during 8:00-10:00 in the morning and 18:00-20:00 in the evening, which means commuting compose most of human mobility. During other periods, human mobility in the main city and the centers of several new towns are more vigorous.

Indicator System of Collective Human Mobility

In Shanghai, which place has the highest human activity vigor? Are people come these places randomly, regularly or even daily? How large is the influential range of human activity? What's the difference of activity vigor between day and night, workdays and weekends? These questions show the basic concern on collective human mobility. Thus, this research describes collective human mobility from four dimensions: mobility intensity, mobility stability, influential range, and temporal variation. Figure 5 shows the indicator system of collective human mobility.

Mobility intensity shows the vigor of human activities in a certain area, including crowd flux volume and crowd flux rate. We aggregate users' destination points into 5432 census units and calculate their density, so crowd flux volume reflects how many people per square kilometer choose a certain census unit as a destination every day. Since crowd flux volume is strongly influenced by population density, we divide it with the basic population of each census unit (stay more than 3 hours here from 0:00 to 6:00) and get crowd flux rate which reflects how many times is every-day visiting population larger than basic population here.

Mobility stability shows the stability of human mobility by classifying residents visited here. We divide visiting population of each census unit into three parts according to their visiting frequency. If a user visits a census unit only once in two weeks, he/she will be identified as the random population since the frequency he/she come here is relatively low. If a user chooses a census unit as his/her destination between twice and five times in two weeks, he/she will be identified as the stable population since he/she come here regularly. If a user chooses a census unit as his/her destination more than six times in two weeks, he/she will be identified as the daily population since he/she almost come here every day and has a strong relationship with this place.

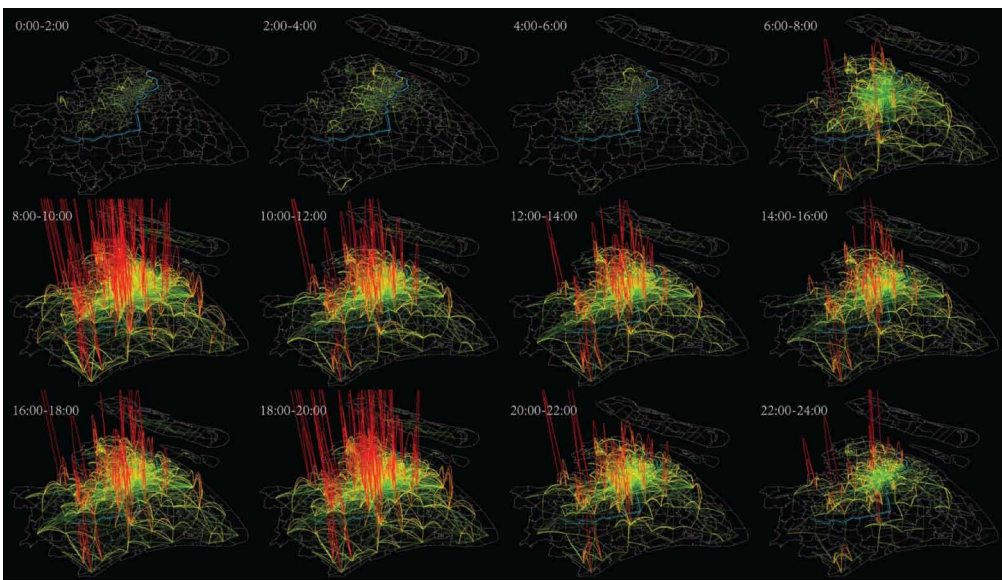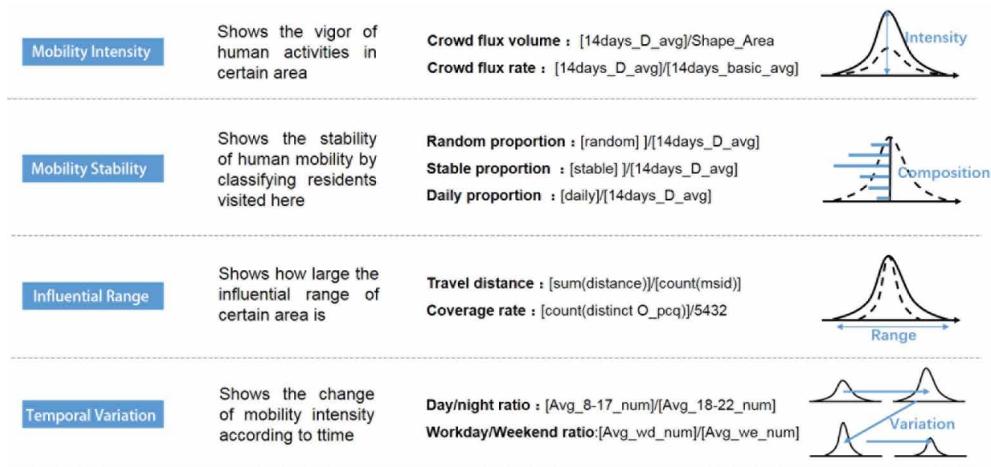Figure 4. Spatial visualization of individual trips

Figure 5. Indicator system of collective human mobility

Influential range shows how large the influential range of a certain area is. We calculate the average travel distance of people visiting each census unit. The longer travel distance means the census unit can attract visitors from a larger range of area. We divide the number of census units that have strong relationship with a certain census (more than 50 OD links every day) with the total number of census units in Shanghai to calculate coverage rate. The indicator shows the importance of a certain census in Shanghai from human mobility view. The higher coverage rate means the census unit is connected with more other census units and thus has a larger influential area.

Temporal variation shows the change in human mobility vigor according to the time. We calculate average crowd flux volume per hour from 8:00 to 17:00 as day volume and from 18:00 to 22:00 as night volume, thus day/night ratio equals day volume divided by night volume. A day/night ratio larger than 1.0 means human mobility during the day is more vigorous than at night. Also, we calculate the average crowd flux volume in workdays and weekends, thus workday/weekend ratio equals workdays volume divided by weekends volume. A workday/weekend ratio larger than 1.0 means human mobility in workdays is more vigorous than in weekends.

RESULTS

Mobility Intensity and Mobility Stability

Figure 6 shows the distribution of crowd flux value in Shanghai. The main city along with several new towns like Jia Ding, Qing Pu, Song Jiang, etc., have an obviously higher crowd flux value than other areas, which reflects that central concentration of human mobility is the basic feature in Shanghai, especially to the city centers and some important nodes. Census units can be divided into five grades according to crowd flux value using Natural Discontinuity Point method (Jenks). The first grade enjoys a visiting population larger than 80 thousand people per square kilometer such as the People's Square, Jing An Temple. The second grade has a visiting population between 45 to 80 thousand per square kilometer such as sub centers like Da Ning, Wu Jiao Chang. The third grade has a visiting population of 25 to 45 thousand per square kilometer such as some community centers and several new town centers like Song Jiang, Jia Ding, etc. The fourth and fifth grade have a relatively lower human mobility vigor, which is below 25 thousand per square kilometer.

Figure 7 shows the distribution of crowd flux rate in Shanghai, which shows a different feature compared with crowd flux volume. Census units in the main city are no longer high value in Shanghai,
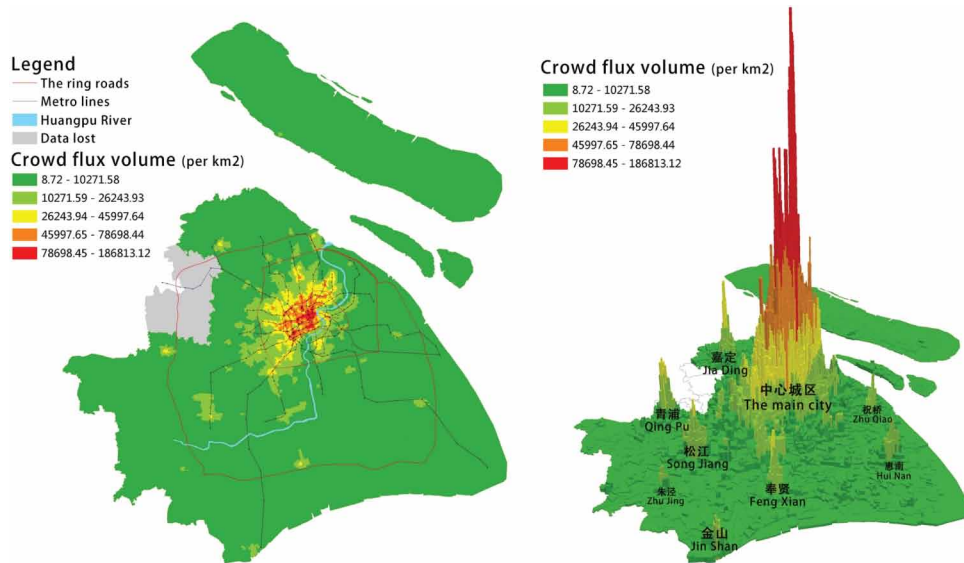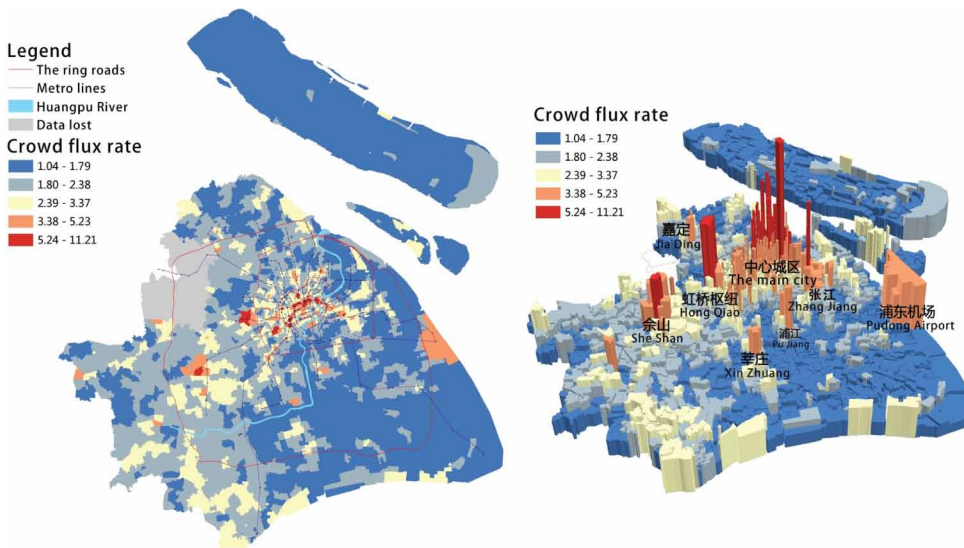
Figure 6. Crowd flux volume in Shanghai

Figure 7. Crowd flux rate in Shanghai

some areas have low crowd flux volume while high crowd flux rate, which means human mobility also influence these areas a lot. Areas along metro line 2 (including Hong Qiao Transportation Hub, People's Square, Zhang Jiang, Pu Dong Airport etc.) enjoys higher crowd flux rate than others. We also divide census units into five grades according to crowd flux rate using Natural Discontinuity Point method (Jenks). Employment centers like Lu Jia Zui have the largest crowd flux ratio of nearly 10.0, indicating that the number of visiting people in these places are 10 times that of the basic population. The second grade are Pu Dong Airport, Hong Qiao Transportation Hun and areas along line 2 inside the inner ring road, with a value higher than 3.4. Chong Ming and the southeast part of

Shanghai have a relatively lower value, which means that crowd flux here are mainly composed by local people, few people visit here from other areas.

Combine crowd flux volume and ratio we can get comprehensive mobility strength of these census units in Shanghai (Figure 8). We can conclude that the areas with the highest mobility intensity in Shanghai are in turn: city center > subcenter (high volume) and transportation hub (high ratio) > regional center > new town. Song Jiang has higher mobility than other new towns, followed by Jia Ding, Bao Shan, and Jin Shan, while Nan Hui new town has low mobility intensity.

In terms of mobility stability, Figure 9 shows the distribution of three kinds of visiting population. We can find that most of the random population locate in the city centers, the stable population distributed around the main city, and the distribution of daily population is more homogenous.

Furthermore, we compare Random, stable, and daily population proportion of different census units in Shanghai (Figure 10). The result indicates that west-east axis connected by metro line 2 and south-north axis connected by metro line 1 have a extremely high random population. Among these areas, People's Square and Hong Qiao Transportation Hub have the highest random proportion, more than
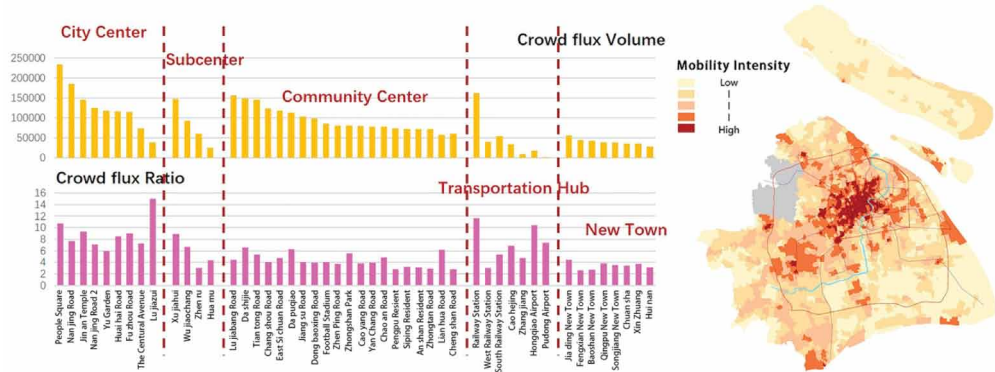
Figure 8 Comprehensive mobility intensity

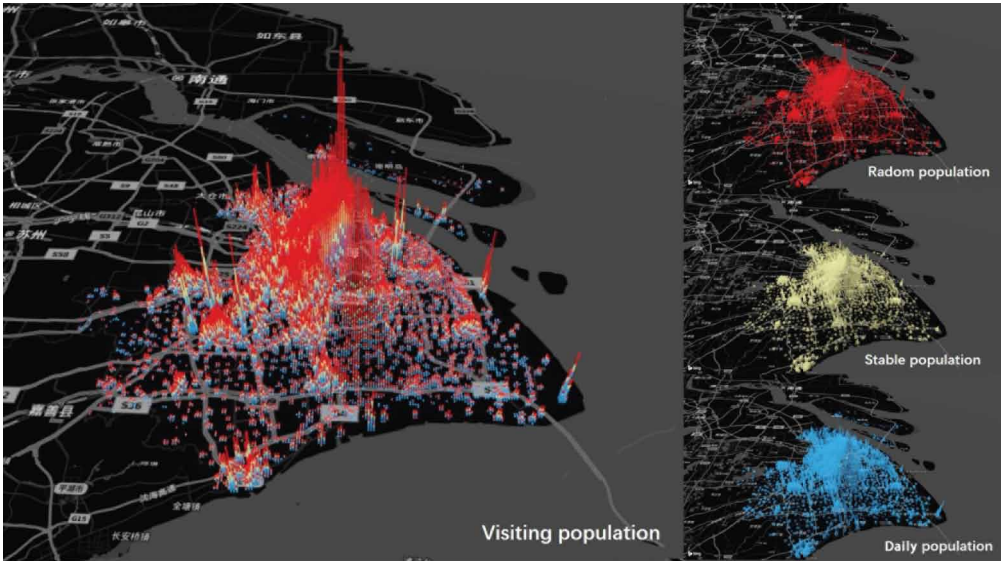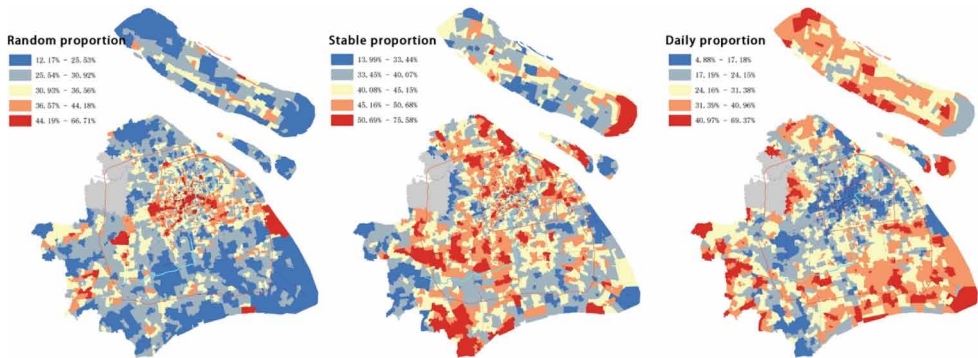**Figure 9. Distribution of different population**

**Figure 10. Random, stable, and daily proportion in Shanghai**

60% of the visiting population here are random population. The northern part of Shanghai (Yang Pu and Bao Shan district), and the southern part of Shanghai (Jin Shan, Ping Qu, Song Jiang, new towns) have the highest stable population. Residential areas and outer suburbs with small visiting population have relatively high daily proportion. For example, Nan Hui new town has an extremely high daily proportion of 60%. According to the proportion of the three groups of people, 5432 census units can be divided into random oriented type, stable oriented and daily oriented three types, the classification will be used in the abstraction of collective human mobility structure.

Influential Range and Temporal Variation

Form the distribution of average travel distance we can found that the value increases gradually from the northern part of the main city to outskirts (Figure 11). The lowest average distance is about 3 kilometers while the highest is about 26 kilometers. Several new towns like Jia Ding, Ping Pu etc. have lower average distance compared with census units around, which shows that these new towns are self-sufficient to certain extent. Divide census units into five grades, we found that the average travel distance of the transportation hub is longer, with Pudong airport reaching 18.8km. Areas inside

Figure 11. Average travel distance in Shanghai
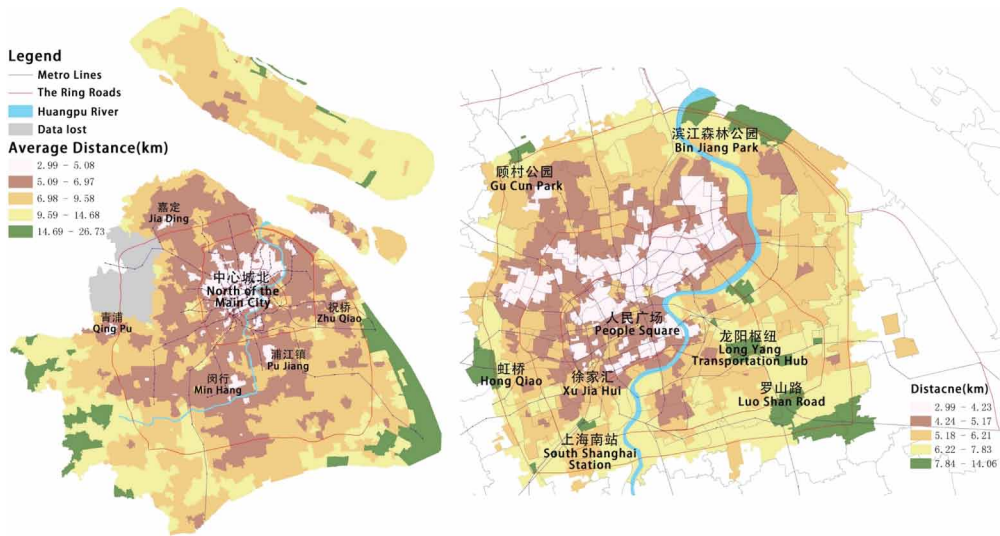
the inner ring road have relatively short travel distance, while there are some exceptions such as the People's square, Tian Tong Road, Lu Jia Zui, and Xu Jia Hui. These areas can attract visitors from an average distance of 5.5-6.5 kilometers. In addition, parks around the outer ring road like Gu Cun Park and Bin Jiang Park have a long average distance for about 8 kilometers.

When it comes to the coverage rate (Figure 12), west-east "horizontal axis" connected by metro line 2 and the south-north "vertical axis" connected by metro line 1 form a "cross high-value area", whose coverage rate are more than 8%, reflecting their central position in Shanghai. Among them, Lu Jia Bang Road has the highest coverage rate reaching 25.7%, which means it connects a quarter of the census units in Shanghai. Several important nodes like People's Square, Xu Jia Hui, Shanghai Railway Station etc. also have high values, excessing 15%. Chong Ming Island and most of the outskirts in Shanghai have a low coverage rate for lower than 3%, indicating that the crowd flux volume here is low or has a clear direction.

Combine average travel distance and coverage rate we can define whether the influential range of a certain census unit is large or small. We can see from Figure 13 that influential range of suburban
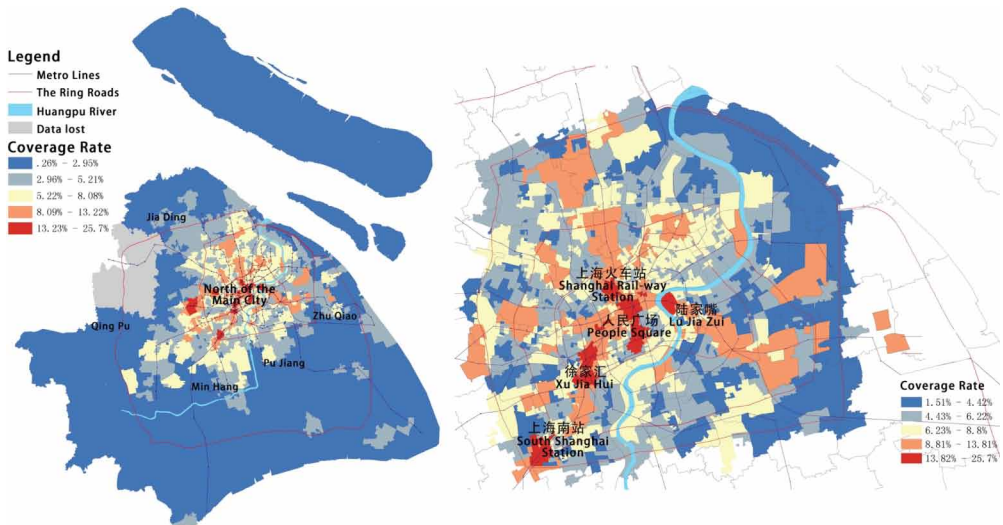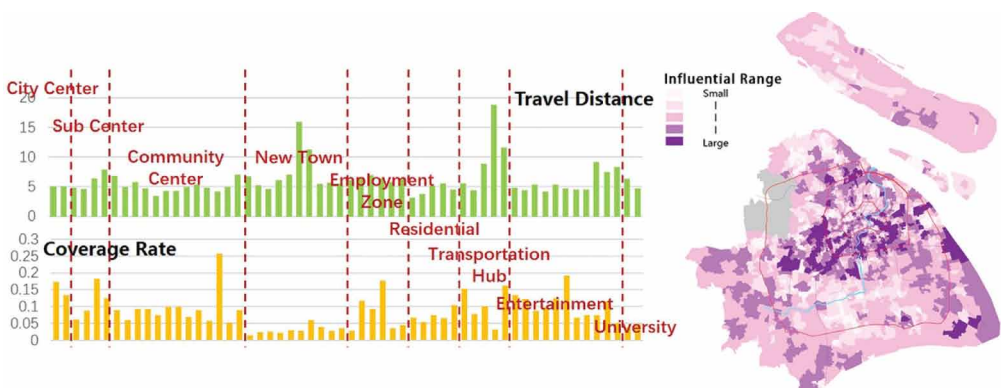
Figure 12. Coverage rate in Shanghai

Figure 13. Comprehensive influential range

areas is the most extensive. People's Square, Lu Jia Zui, Xu Jia Hui, Shanghai Railway station inside the inner ring road have large influential ranges. The high-value zone is banded and has a strong coupling with metro lines, indicating that the metro station can indeed expand the influential range of areas nearby. Moreover, transportation hub and employment zones have relatively larger ranges, while the residential areas and community centers have smaller ones.

Day/night ratio shows the comparison of crowd flux volume during the day and at night. A value higher than 1.0 means the vigor of human mobility during the day is higher than at night, vice versa. Figure 14 is the distribution of day/night ratio in Shanghai, which shows a "high-low-high" pattern from the central city to outer suburbs. Divide census units into five grades according to their day/night ratio, we find that employment centers like Lu Jia Zui, Jin Qiao, Wai Gao Qiao, Cao He Jing, Hong Qiao, and Century Avenue have the highest day/night ratio, with an hourly crowd flux of more than 1.5 times that at night. Some residential areas in the northern part of the central city and in Pudong are low, with an hourly crowd flux of 0.6 times that at night. In addition, the day/night ratio of Gu Cun Park is relatively high, which is related to the opening period of tourist spots.

Workday/weekend ratio shows the comparison of crowd flux volume in workdays and in weekends. A value higher than 1.0 means the vigor of human mobility in workdays is higher than in weekends, vice versa. From the distribution of workday/weekend ratio in Shanghai we can find that employment areas and city centers enjoy a relatively high value, indicating that human flows mainly happen in workdays (Figure 15). Also, we divide census units into five grades. It can be seen that the high value areas are almost all employment centers, including Lu Jia Zui, Zhang Jiang, Cao He Jing, Wai Gao Qiao, etc. On weekdays, the number of visiting people is 1.6-2.0 times that of the weekend. Most areas in the main city have values more than 1.0, but shopping centers such as East Nanjing Road and Yu Garden are significantly lower on weekdays. Residential areas like Xin Jiang Wan, Cheng Shan Road, Zhong Tan Road and Chang Zhong Road have a workday/weekend ratio generally between 0.7 and 1.0.

Combine day/night ratio and workday/weekend ratio we can define whether the crowd flux volume of a certain census varies obviously according to different periods. The overall temporal variation of the central city and outer suburbs is relatively high, that is, which means the mobility intensity of these regions varies greatly with time (Figure 16). As to different functional areas, human

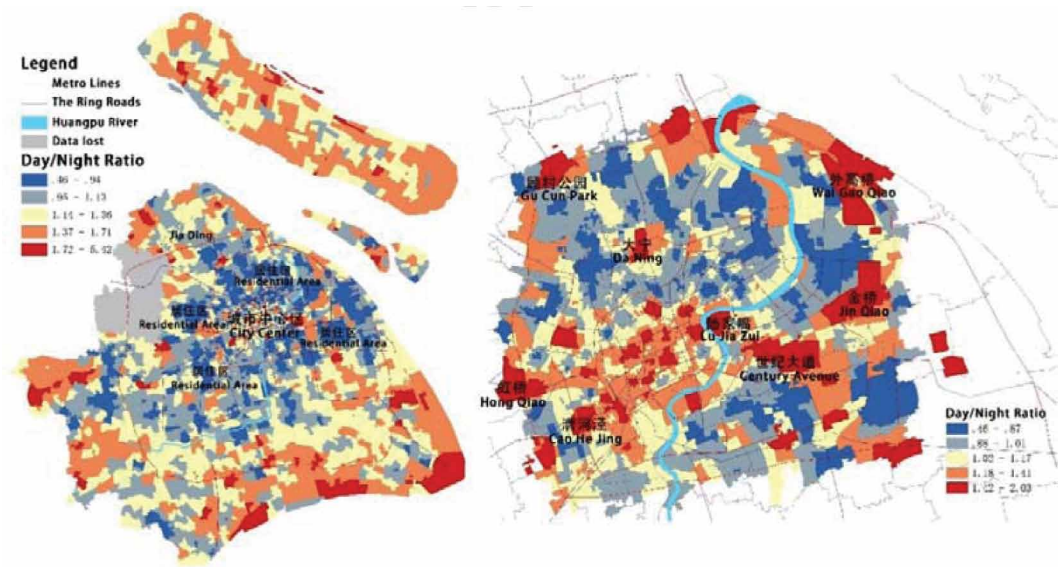**Figure 14. Day/night ratio in Shanghai**

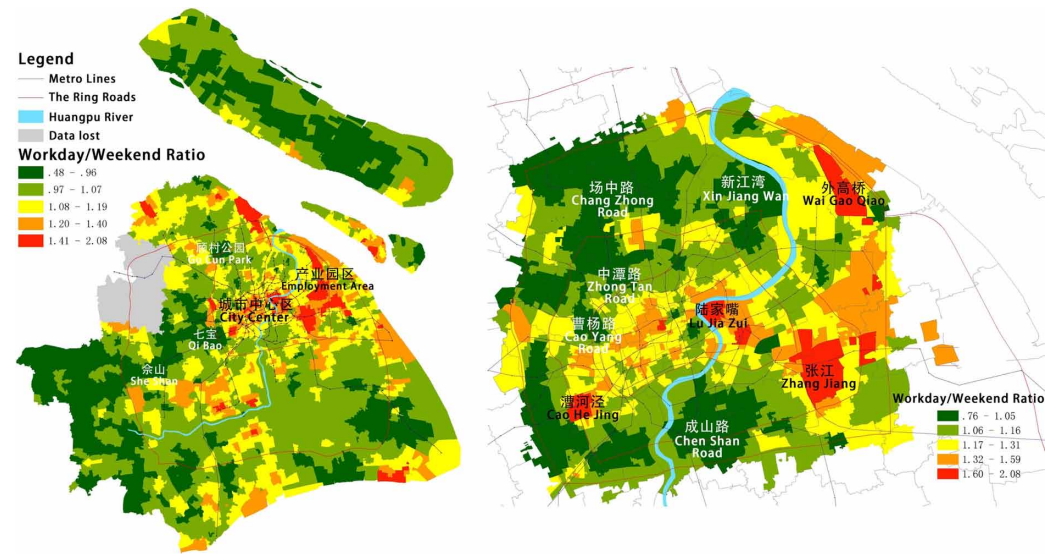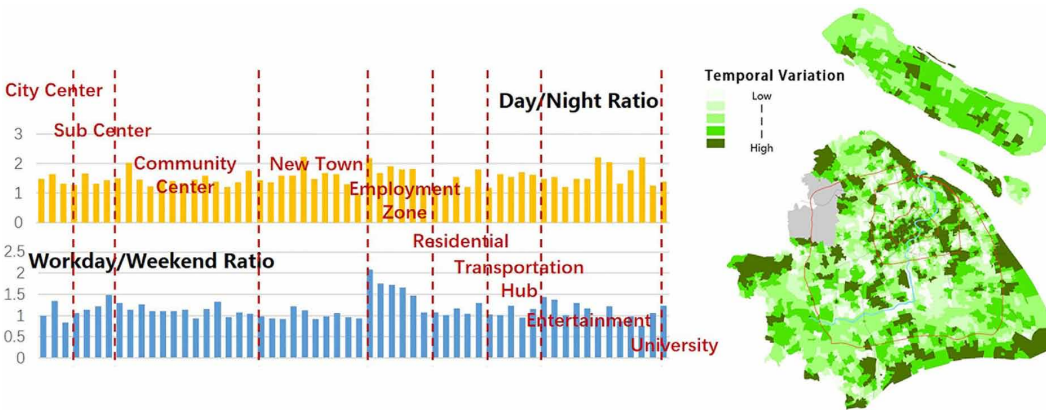**Figure 15. Workday/weekend ratio in Shanghai**


Figure 16. Comprehensive temporal variation

mobility vigor in employment centers and local centers varies obviously according to time, while in residential areas and new towns, human mobility vigor changes slightly between different periods.

Spatial Structure Under View of Human Mobility

Figure 17 shows how we extract spatial features from our four-dimensional indicator system. Four kinds of spatial elements can be abstracted from collective human mobility features. Dynamic surface, from the city center to subarea, shows the intensity level of daily human mobility, including core area, dense area and shadow area. Dynamic center, including five centers and eleven subcenter, becomes the most influential areas in the flowing network besides the core area. Dynamic cluster can be divided into six groups according to mobility stability, influential range, and temporal variation, including distributing zone (high random proportion), daily one (high stable and daily proportion), wide-spreading zone (large influential range), isolated zone (small influential range), fluctuating

Figure 17. Spatial elements identification
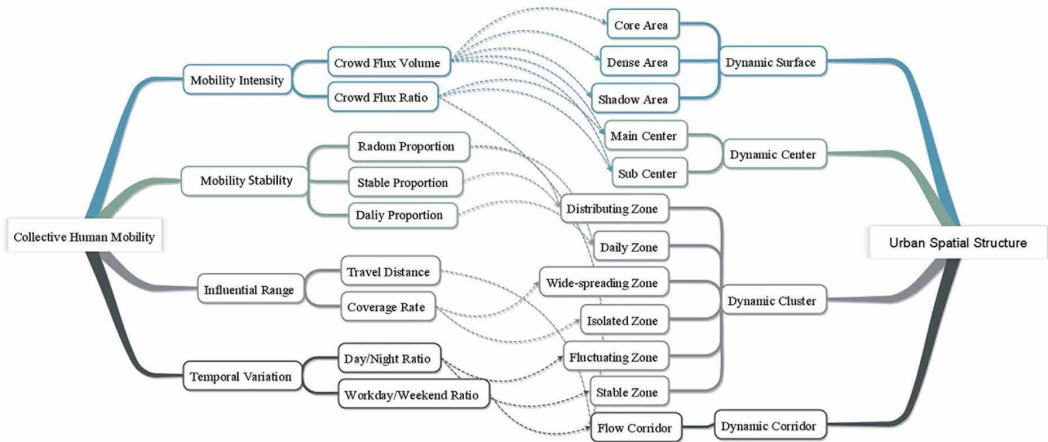
zone (obvious temporal variation), and stable zone(slight temporal variation). Dynamic corridor, banded areas of high or low values presented by different indicators, strongly related to metro lines.

Figure 18 shows the final structure of collective human mobility in Shanghai, which is different from the layer-form in the traditional urban spatial form. It can be seen that the core area mainly locates inside the inner ring road, the centers of new towns like Jia Ding, Song Jiang, and Qing Pu are main mobility centers, and other local community centers in the suburbs compose sub mobility centers. Six kinds of dynamic clusters scattered in suburbs, which contains most of big residential areas and industrial park in Shanghai, indicating that although the functions of these areas are simply residence or working, their collective human mobility features varies largely from one to another. The spatial distribution of dynamic corridor has a strong relationship with metro lines, which is shown by thin black lines. Especially, metro line2 and line 2 along with line 8,9 have great influence on the collective human mobility form.

Figure 18. Structure of collective human mobility in Shanghai
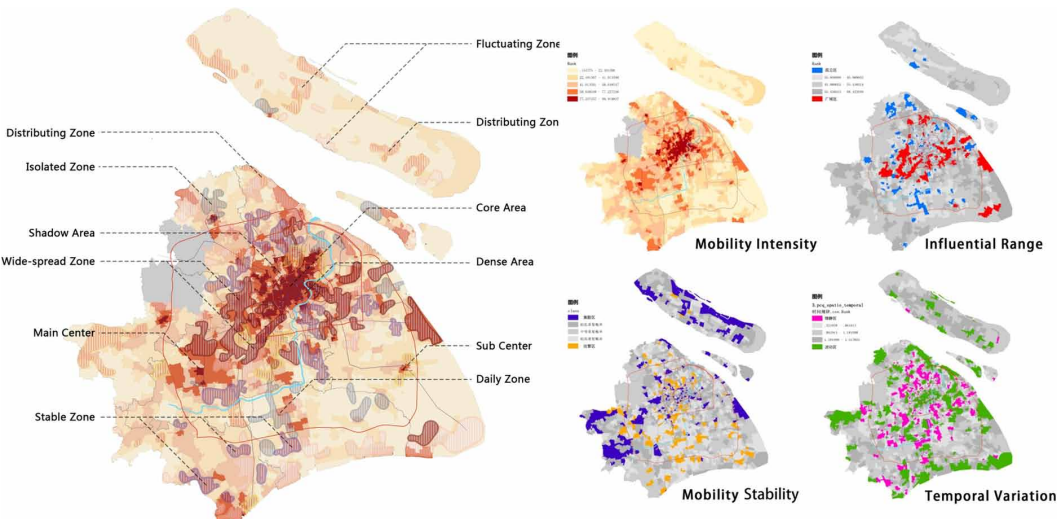
DISCUSSION

The urban spatial structure has been widely discussed from the view of land use function, industry, and built environments. While there is an increasing importance for scholars to explore how the urban structure look like from human mobility view. Although the activities of people in the city can't fully represent the characteristics of urban structure, the functional travel of large numbers of people can be seen as the most intuitive representation of urban structure (Figure 18).

The most innovative point of this study is to propose a technical method to observe the spatial aggregation characteristics of human activities through mobile phone data. Referring to the corresponding indicators in the indicator system (Figure 5), we can quickly grasp the characteristics of collective human mobility caused by frequent human activities within the city.

Mobility intensity feature is the most important among the four dimensions. It explains under the view of human mobility; how many people are daily associated with a certain space unit. These people can be called "visiting population" or "actual population", their volume in high mobility areas maybe several times of resident population counted by the population census (Figure 7). The distribution of crowd flux volume (Figure 6) shows human activities have a strong preference to the main city, which means numbers of resident's lives outside the main city tends to get here for job and entertainment at daytime. So, the population scale is underestimated in the main city according to tradition population concept.

Meanwhile, mobility stability shows how strong the relationships are between urban space and their visitors. Areas with high mobility stability means visiting population everyday are the same, largely related to local residents. So relative planning strategies should mainly focus their daily life. On the other hand, if a certain area has a high random proportion (Figure 10), its daily visiting population are mainly tourists or business population. Correspondently, the planning should take their demand into consideration.

In addition, influential range and temporal variation shows more detailed information that how far away are visiting population come from and when they come. This information is helpful to transportation planning. For example, some areas have a relatively long average travel distance (Figure 11), which means a high volume of traffic demands. Some areas have obviously higher traffic volume in certain periods (Figure 14, Figure 15), planning should get prepared for their "rush hours".

CONCLUSION

The study summarizes relative research on human mobility using cell phone data and constructs a four-dimensional indicator system to describe the collective human mobility features in Shanghai. By abstracting the spatial structure from our indicator system, an exploration is conducted to make up present research gap and evaluate Shanghai's urban space from a different view. Findings of this study suggest that:

1. The visiting population associated with a census unit in city center, transportation hub and other places with high human activity vigor can be many times that of the local resident population. The characteristic of collective human mobility can be revealed from mobility intensity, mobility stability, influential range, and temporal variation;
2. The spatial distribution of mobility intensity in Shanghai looks like "central hardcore + suburb mobility centers". Radom and stable population larger than the daily population in high mobility areas, so for these high mobility regions visiting population should be taken into consideration as well as resident population;
3. The influential range of suburb areas is the most extensive, and relationship with metro lines is strong. The distribution of day/night ratio in Shanghai presents "high-low-high" tendency from the central city to outer suburbs. Human mobility vigor in employment centers and local centers

varies obviously according to time, while in residential areas and new towns, human mobility vigor changes slightly between different periods;

4. Spatial structure can be extracted from our indicator system, which is consist of four kinds of spatial elements: dynamic surface, dynamic center, dynamic cluster, and dynamic corridor. The results can be further drawn into a spatial structure under mobility view showing a different form compared with the traditional urban spatial structure.

However, there are still some limitations exist and could be discussed in future work. First, the predefined rule used to identify individual travel chain need further proved. Although the study has tried to ensure these thresholds commonly used and reasonable, test of different time and distance threshold are necessary to see whether different threshold influence the results a lot. Secondly, the market share of Shanghai Mobile Company is not 100 percent (actually is about 70 percent), that means the volume calculated in the research is a "relative volume" instead of "accurate volume". Since existing "sample enlargement" methods are still an exploration, such question is unable to be solved in a short time. Finally, the study uses cell phone data collected five years ago (2014), the latest changes in Shanghai can't be captured. Actually, further studies are in progress, in which data from different years in the same city and data from different cities in the same year are compared to get more findings.

REFERENCES

Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2010). Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: Case study with mobile positioning data. *Transportation Research Part C, Emerging Technologies*, *18*(1), 45–54. doi:10.1016/j.trc.2009.04.011

Ahas, R., Aasa, A., Yuan, Y., Raubal, M., Smoreda, Z., Liu, Y., & Zook, M. et al. (2015). Everyday space–time geographies: Using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn. *International Journal of Geographical Information Science*, *29*(11), 2017–2039. doi:10.1080/13658816.2015 .1063151

Alexander, L., Jiang, S., Murga, M., & González, M. C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C, Emerging Technologies*, *58*, 240–250. doi:10.1016/j.trc.2015.02.018

Bin, R. (2013). Use of Cellphone Data in Travel Survey and Transportation Planning. *Urban Transport of China*, *11*(1), 72–81.

Calabrese, F., Ferrari, L., & Blondel, V. D. (2015). Urban sensing using mobile phone network data: A survey of research. *ACM Computing Surveys*, *47*(2), 25.

Chai, Y.-W., Xue, B., & Sun, D.-S. (2015). A Study on Life Circle Planning Based on Space Time Behavioral Analysis: A Case Study of Beijing. *Urban Planning Forum*, *3*, 61-69.

Chen, C., Bian, L., & Ma, J. (2014). From traces to trajectories: How well can we guess activity locations from mobile phone traces? *Transportation Research Part C, Emerging Technologies*, *46*, 326–337. doi:10.1016/j. trc.2014.07.001

Chen, J., Shaw, S. L., Yu, H., Lu, F., Chai, Y., & Jia, Q. (2011). Exploratory data analysis of activity diary data: A space–time GIS approach. *Journal of Transport Geography*, *19*(3), 394–404. doi:10.1016/j.jtrangeo.2010.11.002

Csáji, B. C., Browet, A., Traag, V. A., Delvenne, J. C., Huens, E., Van Dooren, P., & Blondel, V. D. et al. (2013). Exploring the mobility of mobile phone users. *Physica A*, *392*(6), 1459–1473. doi:10.1016/j.physa.2012.11.040

Doyle, J., Hung, P., Farrell, R., & McLoone, S. (2014). Population mobility dynamics estimated from mobile telephony data. *Journal of Urban Technology*, *21*(2), 109–132. doi:10.1080/10630732.2014.888904

Golledge, R. G. (2003). Reflections on recent cognitive behavioural research with an emphasis on research in the United States of America. *Australian Geographical Studies*, *41*(2), 117–130. doi:10.1111/1467-8470.00199

Kwan, M. P., & Ding, G. (2008). Geo-narrative: Extending geographic information systems for narrative analysis in qualitative and mixed-method research. *The Professional Geographer*, *60*(4), 443–465. doi:10.1080/00330120802211752

Lee, B. (2007). "Edge" or "edgeless" cities? Urban spatial structure in US metropolitan areas, 1980 to 2000. *Journal of Regional Science*, *47*(3), 479–515. doi:10.1111/j.1467-9787.2007.00517.x

Miller, H. J. (2006). Social exclusion in space and time. In *AXHAUSEN KW Moving through nets: The Physical and Social Dimensions of Travel* (pp. 353-380). Elsevier Science Publishing Company.

Min, Z., Chen, C., & Yu, H.-W. (2013). On Urbanization Driven by Rural-Urban Migration and Its Policy Implications. *Urban Planning Forum*, *2*, 1-9

Pan, H.-X., Qing, S., & Ming, Z. (2009). Impacts of Urban Forms on Travel Behavior: Case Studies in Shanghai. *Urban Transport of China*, *7*(6), 28–32.

Phithakkitnukoon, S., Horanont, T., Di Lorenzo, G., Shibasaki, R., & Ratti, C. (2010, August). Activity-aware map: Identifying human daily activity pattern using mobile phone data. *Proceedings of the International Workshop on Human Behavior Understanding* (pp. 14-25). Springer. doi:10.1007/978-3-642-14715-9_3

Shaw, S. L., & Wang, D. (2000). Handling disaggregate spatiotemporal travel data in GIS. *GeoInformatica*, *4*(2), 161–178. doi:10.1023/A:1009824122914

Song, C., Koren, T., Wang, P., & Barabási, A. L. (2010). Modelling the scaling properties of human mobility. *Nature Physics*, *6*(10), 818–823. doi:10.1038/nphys1760

Song, C., Qu, Z., Blumm, N., & Barabási, A. L. (2010). Limits of predictability in human mobility. *Science*, *327*(5968), 1018–1021. doi:10.1126/science.1177170

Wang, X.-H., & Sun, B.-D. (2011). The Economic Performance of the Polycentric Spatial Structure of Mega-cities: Based on the Models of Urban Economics. *Urban Planning Forum*, *6*, 20-27.

Wang, Z., He, S. Y., & Leung, Y. (2018). Applying mobile phone data to travel behaviour research: A literature review. *Travel Behaviour and Society*, *11*, 141–155. doi:10.1016/j.tbs.2017.02.005

Wen, C. H., & Koppelman, F. S. (2000). A conceptual and methodological framework for the generation of activity-travel patterns. *Transportation*, *27*(1), 5–23. doi:10.1023/A:1005234603206

Widhalm, P., Yang, Y., Ulm, M., Athavale, S., & González, M. C. (2015). Discovering urban activity patterns in cell phone data. *Transportation*, *42*(4), 597–623. doi:10.1007/s11116-015-9598-x

Yuan, Y., Raubal, M., & Liu, Y. (2012). Correlating mobile phone usage and travel behavior–A case study of Harbin, China. *Computers, Environment and Urban Systems*, *36*(2), 118–130. doi:10.1016/j.compenvurbsys.2011.07.003

Zhang, T.-W. (2001). The Urban Restructuring of Chinese Cities in 1990s and its Dynamic Mechanism. *City Planning Review*, *07*, 7–14.

Zhang, W.-J., & Chai, Y.-W. (2008). Theories and Confirmed Model of Urban Resident's Travel Demand: Considering Intra-household Interaction. *Acta Geographica Sinica*, *63*(12), 1246–1256.

Zhao, L.-D. (2005). A Statistical Study on the Population Migration and Transient Population in Recent Time in China. *Economic Survey*, *6*, 80–83.

Zhong, W.-J., De, W., Xie, D.-C., & Yan, L.-X. (2017). Dynamic characteristics of Shanghai's population distribution using cell phone signaling data. *Geographical Research*, *36*(5), 972–984.

Ren Xiyuan received his master's degree in urban planning from Tongji University in 2019. He writes and presents widely on issues of human spatial behavior, urban population, and transportation using location-based service (LBS) data.