Analysis of Spatial–Temporal Characteristics Based on Mobile Phone Data

Hong-liang Yin and Chang-jiang Zheng

Abstract The most traditional method to collect traffic data is household survey, which is a waste of manpower and material resources. OD matrices estimation using link volumes has been widely studied. But the significant shortcoming is the high cost of detectors. Besides, once installed in the network, the traffic detectors are not easy to be moved. Mobile phones' location data, however, can be acquired over a wider coverage without additional costs. The use of such data provides new spatiotemporal tools for improving urban transportation planning. This paper analyzes the nature and the pre-treatment of data from mobile phone operators in China and highlights the applicability of the data in domain of transportation. It also presents a typology of applications to analyze spatial–temporal characteristics based on mobile phone data.

Keywords Spatial–temporal characteristics · Mobile phone data · Urban transportation planning

Introduction

Recently, mobile communication technology is becoming one of the most widely used technologies, the most convenient and the most influential information and communication technologies (ICTs). The current penetration rates are 128 and 89% in developed and developing countries, respectively [1]. According to the Ministry of industry statistics of China, the number of mobile phone users in China has reached 1.256 billion by the end of May 2014 [2]. Along with the rapid growth of information techniques, the concept of smart city is becoming popular, which attracts the attention of urban planners from construction of cities and economy to

H. Yin
Jiangsu Provincial Department of Communications, Nanjing 210098, Jiangsu, China

C. Zheng (✉)
School of Civil Engineering and Transportation, Hohai University,
Nanjing 210001, Jiangsu, China
e-mail: zhenghhu@sina.com

W. Wang et al. (eds.), *Green Intelligent Transportation Systems*,
Lecture Notes in Electrical Engineering 419, DOI 10.1007/978-981-10-3551-7_79

citizens' activities. Hence, mobile phone users' footprints of their approximate locations, when making a call or sending an SMS, can show the daily routine that would interest urban planners.

Mobile phone data, as a new promising source of location data, consist of individual activities and are acquired almost in real time [3]. The processing, mining, publishing, and applying of such a huge amount of location data can provide the fundamental information of urban traffic demand, and therefore be a crucial for urban transportation planning of China during the rapid urbanization. It would be of great social and economic significance for cities to take advantage of those data and adopt new approach to ease traffic congestion, improve traffic safety and efficiency, reduce air pollution, etc. Data from mobile phone operators are becoming a tool for a smarter city in the future [4]. In recent years, mobile phone data have been used for estimating traffic flow [5, 6], deriving travel modes [7], analyzing interconnections between areas [8], tracking individual travel behavior [9], to name a few.

The main objective of the paper was to analyze how the use of mobile phone data can facilitate urban management and planning. The remainder of this paper is organized as follows: Section "Collection of Mobile Phone Data" is devoted to a description of the nature and the pre-treatment of data from mobile phone operators in China. Section "Analysis of Mobile Phone Data Applying in Urban Transportation Planning" discusses whether and how should such data be applied. Section "Key Technologies of Trip Characteristics Analysis Based on Mobile Phone Data" then gives a broad overview of the applicability of such data. Section "Conclusion" concludes the paper.

Collection of Mobile Phone Data

Interpretation of Mobile Phone Signaling

Original mobile phone data collected by the signaling acquisition system of China mobile consist of information such as service station number and location and switching time, [10]. Data format is shown in Table 1, and the detailed meaning of each field of mobile phone data is shown in Table 2.

Further more, MSID is the code of each mobile phone user given by mobile operators. LAC is usually divided by county or administrative district. CELLID is the code of base station where the mobile station locates. AREA CODE is the code of the region where the mobile station joins the network.

Mobile Phone Data Denoising

Traffic information collection based on mobile phone may be simple and efficient. The impact of communication network and the environment, however, adds much

Table 1 Original mobile phone data format of China mobile

MSID	TIMESTAMP	LAC	CELLID	EVENT ID	CAUSE	FLAG	MSCID	BSCID	CAUSE TYPE
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.30	13208	40482	8	9	1	11573	11483	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.40	13060	24403	8	9	1	11573	11584	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.40	13119	30062	8	9	1	11573	11512	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.40	13061	60331	8	9	1	11573	11563	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.10	13249	40573	8	9	1	11573	11540	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000000.90	33162	10391	8	9	1	11573	11405	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000000.90	33162	10392	8	9	1	11573	11405	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.30	13061	60621	8	9	1	11573	11563	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000000.90	13110	30291	8	9	1	11573	11750	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.60	13119	30582	8	9	1	11573	11512	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.60	33141	31272	8	9	1	11573	11422	4
da4a1baac83a453050449d974ce0e1f50df1add3	20130402000002.40	13119	30473	14	9	0	11573	11512	4
a43ab3c7f4bc16c2de26e924a0d134f1103aef7d	20130402000000.70	13119	30141	5	9	0	11573	11512	4

Table 2 Name and implication of each field of the mobile phone data

Index	Field	No. of characters	Field meaning
1	MSID	32	Only identity code of user
2	TIMESTAMP	14	Signaling time, YYYYMMDDHH24MISS, unit: seconds
3	LAC	5	Location area (LA) code
4	CELLID	5	Zone number
5	EVENTID	3	The type of event, such as startup/shutdown, calling/called, normal location update
6	CAUSE	3	Cause of event
7	FLAG	3	Whether the identification can get IMSI
8	MSCID	8	ID of MSC
9	BSCID	8	ID of BSC
10	AREA CODE	8	Network area code

"noise" to the location data, which would influence the analysis. The noise usually results for the following two reasons [11]:

1. Switching disturbance (such as ping-pong effect), that is to say the frequent switches between two adjacent areas because of the judgment error of user's position switch in base station system. This kind of noise data is ubiquitous.
2. Staying at the same position during a long time. In a GSM network, a periodic location update function was introduced to prevent the loss of contact with a mobile station, which demands a mobile phone to report once every hour its position. Therefore, when the user stays at a fixed place, a large number of redundant data will be produced.

Those noise data are to the disadvantage for determining the track of mobile phone users. It is necessary to process them before analysis. For example, eliminate lines with abnormal value at the field of CELLID and LAC, due to communication failure, and lines with same values of MSID and CELLID, which are redundant data.

After pre-treatment, data need to be sorted by time. Mobile phone data of each user in each day should be extracted in order to calibrate their trajectories, where fields such as MSID, TIMESTAMP, and CELLID must be reserved.

Analysis of Mobile Phone Data Application in Urban Transportation Planning

Analysis of the Application of the Mobile Phone Data

Before applying the data, their application in transportation is analyzed as follows:

1. For traffic applications, the location and time information just fit for the traffic demand analysis. Besides, the characteristics of those data such as high sample rate, continuity, and real time give incomparable superiority to it.
2. The analysis of individual activity is to locate the traffic zone of each users and does not need the accurate location. Hence, the information of location and time satisfies the demand of analysis.
3. Object of data acquisition is mobile terminal, meanwhile most of the travelers in the city are cell phone users. The trip information analyzed from mobile phone data can therefore correctly reflect travel characteristics of habitants.
4. Though errors exist, the huge amount of data will fix the problem according to the theory of Big Data. The growing penetration rate guarantees the quantity of samples.

Research Framework of Urban Traffic Characteristics of Mobile Phone Data

The basic process of travel characteristics analysis based on mobile phone data is shown in Fig. 1. First, collect and extract information desired in the research area from mobile phone data platform of telecom operator. Second, establish the

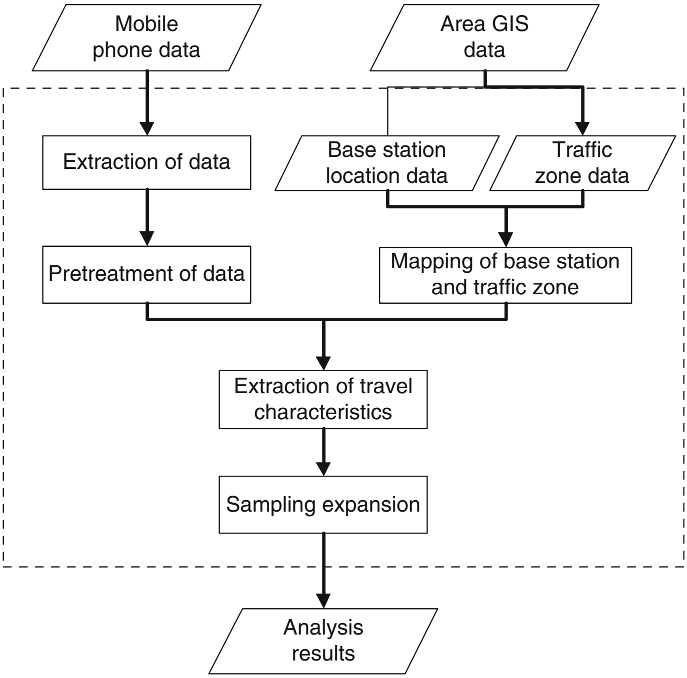

Fig. 1 Framework of travel characteristics analysis based on mobile phone data

mapping of base station and traffic zone data in GIS. Then, analyze travel characteristics with GIS data after pre-treatment of mobile phone data. Finally, get the analysis results by sampling expansion.

Research Area

During the extraction of travel characteristics from original mobile phone data, a mapping of base station zones and traffic zones need to be established. The analysis unit of mobile phone data is based on base station zones, which cover a certain area. The limit of the area, however, can be indefinite because of the instability of communication networks. There is, therefore, a deviation of user's location, which may range from 50 to 300 m in urban area and from 100 to 2000 m in Banlieue.

When the research area is divided into smaller traffic zones, the deviation of location data, especially in city center, will increase because of a greater number of base stations; while a traffic zone consists of base station span of control, this kind of deviation will decrease. In short, it will be more appropriate to apply mobile phone data to a grander range (like several cities) instead of a small area (like a few blocks).

Key Technologies of Trip Characteristics Analysis Based on Mobile Phone Data

Having huge advantages over traditional methods of traffic survey, extracting traffic information from mobile phone data is drawing more and more attention. Since 2005, several empirical studies have utilized data from mobile phone operators to gain insight and knowledge into complex and rapidly changing spatial urban phenomena. Research which utilizes the breadth of data produced by mobile phone operators can be a great opportunity to tackle specific issues. In this section, we will critically analyze several applications and focus on how to use such data.

OD Estimation

A general framework to develop OD matrices using mobile phone data is shown in Fig. 2.

First, generate the mapping relationship between traffic zones and base station zones according to the coverage situation of the telecom networks. Second, record the code of base station zones that a user has entered and left and the timestamp. Third, calculate the time that the user has stayed in every traffic zone. If the time exceeds a certain threshold, the zone can be seen as a stayed zone, while the others

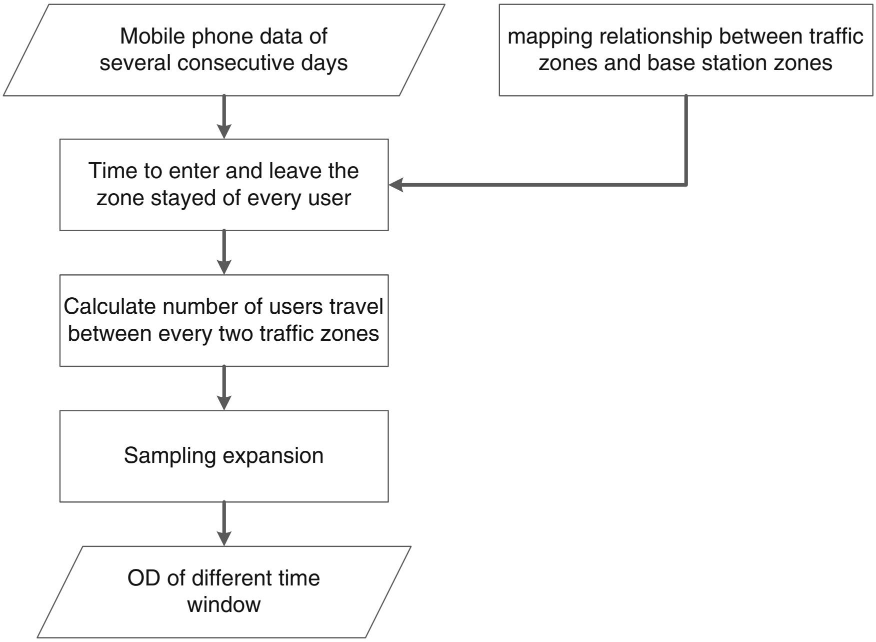

Fig. 2 OD estimation based on mobile phone data

are just passed through. Then, analyze trips occurring within certain time windows to generate transient OD matrices. Finally, considering the penetration rate of mobile phone users, get the OD estimation at different time by expanding samples to whole habitants.

Recognition of the Purpose of Trips

Taking the analysis of place of residence and work as example, the process is shown in Fig. 3.

Step 1: generate the mapping relationship between traffic zones and base station zones according to the coverage situation of the telecom networks.

Step 2: recognize the home zone of users, which should be the place where the user usually stays at night. For example, the zone where a user stays for more than 6 h between 20:00 and 8:00 and for more than 80% of the days studied is the user's home zone.

Step 3: recognize the working zone of users, which should be the place where the user usually stays during daytime. For example, the zone where the

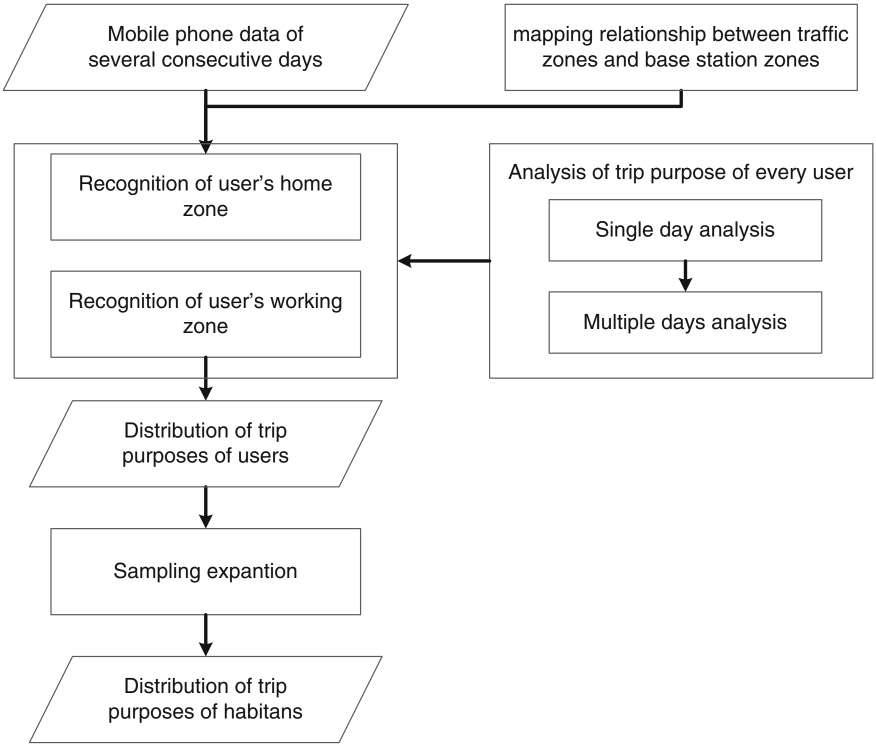

Fig. 3 Analysis of trip purpose based on mobile phone data

user stays for more than 6 h between 8:00 and 20:00 and for more than 80% of the days studied is the user's working zone.

Step 4: count the number of users that live or work in a certain zone to get the distribution of trip purpose of mobile phone users and expand the sample to obtain that of all the habitants.

Links Counts Detection

The process to detect traffic flow on links such as bridges and tunnels is shown in Fig. 4.

First, determine the link to analyze and establish the mapping relationship between telecommunication base stations and the two traffic zones on both sides of the link. Then, record the passing time if the trajectory of a user has passed the link. Finally, get the number of users who passed the link in different time windows.

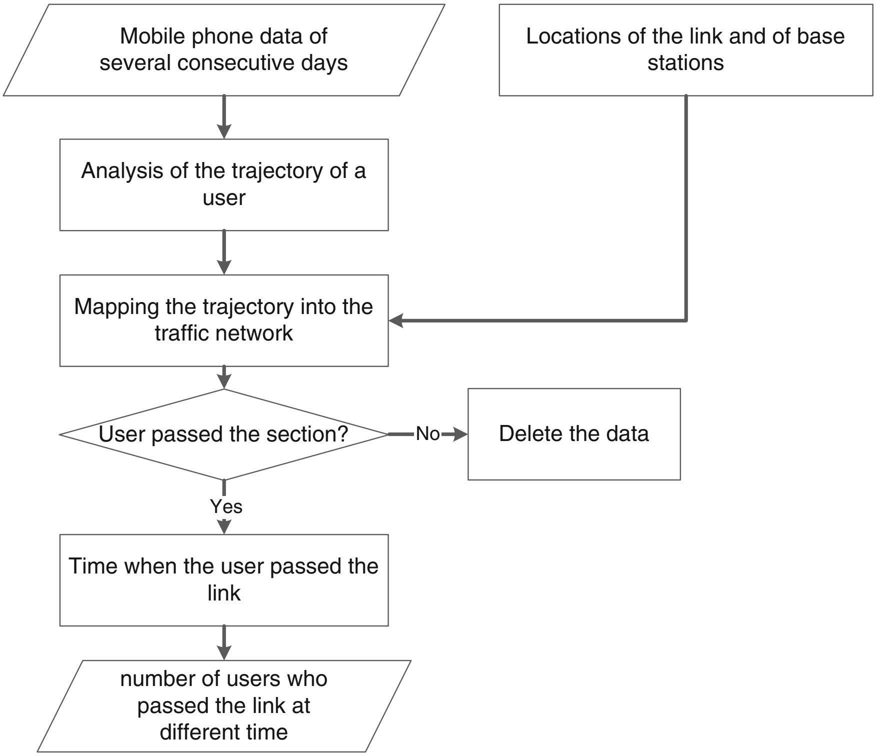

Fig. 4 Link counts detection based on mobile phone data

Conclusion

Better understanding of how, where, and when people move on a daily basis, especially in densely populated areas, could lead to improvement in urban planning, transportation infrastructure design, and assessments of environmental impacts [2, 11]. This paper has demonstrated the wide applicability of mobile phone data and the key technologies of application in achieving smart city objectives. Although such data cannot be approached as a panacea, it is hopeful that, in the near future, they would lead to a better city along with other sources of data.

References

1. Iqbal, M.S., C.F. Choudhury, P. Wang, et al. 2014. Development of origin–destination matrices using mobile phone call data. *Transportation Research Part C Emerging Technologies* 40 (1): 63–74.
2. Ran, B. 2013. Use of cellphone data in travel survey and transportation planning. *Urban Transport of China* 11 (1): 72–81.

3. Steenbruggen, J., E. Tranos, and P. Nijkamp. 2015. Data from mobile phone operators: A tool for smarter cities? *Research Memorandum* 39: 335–346.
4. Demissie, M.G., G.H.D.A. Correia, and C. Bento. 2013. Intelligent road traffic status detection system through cellular networks handover information: An exploratory study. *Transportation Research Part C Emerging Technologies* 32 (4): 76–88.
5. Wang, M., et al. 2013. Estimating dynamic origin-destination data and travel demand using cell phone network data. *International Journal of Intelligent Transportation Systems Research* 11 (2): 76–86.
6. Bohte, W., and K. Maat. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies* 17 (3): 285–297.
7. Trasarti, R., R. Trasarti, A.M. Olteanu-Raimond, M. Nanni, et al. 2014. Discovering urban and country dynamics from mobile phone data with spatial correlation patterns. *Telecommunications Policy* 39: 347–362.
8. Asakura, Y., and H. Eiji. 2004. Tracking survey for individual travel behaviour using mobile communication instruments. *Transportation Research Part C: Emerging Technologies* 12 (3): 273–291.
9. Guan, Z., X. Zhang, B. Hu et al. 2011. Research on decision support system of urban transport planning based on mobile phone data. In *Sixth Annual Conference of China Intelligent Transportation System*.
10. Huang, M., and B. Lu. 2010. Traffic OD data collection technology based on mobile phone location. *Journal of Chongqing Jiaotong University (Natural Science)* 29 (1): 162–166. (in Chinese).
11. Caceres, N., L.M. Romero, and F.G. Benitez. 2013. Inferring origin–destination trip matrices from aggregate volumes on groups of links: A case study using volumes inferred from mobile phone data. *Journal of Advanced Transportation* 47 (7): 650–666.