Chapter 36
Review on Application of Call Details Records (CDRs) Data to Understand Urban Mobility Scenarios for Future Smart Cities

Namrata Ghosh, Udit Sarkar, and Prakash Nagesh

Abstract Rapid urbanisation, pollution and inadequate public transit have made mobility more complex, plaguing people worldwide. Urban and Transport planners face a considerable challenge with the vision of future smart cities as it focuses on creating a more cohesive transportation ecosystem for congested cities. The new advancements in mobility with digital innovations and updated real-time data sources supported by data and models will help design an efficient transport system through a thorough understanding of human mobility. However, conducting conventional travel surveys is expensive, with limited sample sizes. Detailed information on travel patterns and the actual demand for travel is hard to get today. Cellular network data collected using the existing infrastructure of mobile operators is a promising new data source and an optimal source to analyse the individual's mobility pattern. Researchers have utilised passively collected data, such as vehicle global positioning system (GPS), mobile network data including call details record (CDR) and Google location history, to define individual travel behaviour patterns. The chapter produces condensed reviews of previous case examples that have adopted similar analytic approaches that involve mobile data aggregation to glean travel information. This study may help researchers and transport authorities understand the potential of mobile phone data as an alternate and more frequently updated data source for future smart cities with several key inferences and the challenges associated with the data. This chapter recommends the framework for data processing and their associated algorithms to understand the mobility pattern using mobile phone data.

Keywords Mobile phone data · GPS data · Urban mobility · Future smart cities

N. Ghosh (✉)
Public System Group, Indian Institute of Management (IIM), Ahmedabad, Gujarat 380015, India
e-mail: namrataghosh033@gmail.com

U. Sarkar
Data and Management Unit, Ministry of Housing and Urban Affairs (MoHUA), Government of India, New Delhi, Delhi 110011, India
e-mail: usarkar@niua.org

P. Nagesh
Centre for Transportation and Logistics, Indian Institute of Management (IIM), Ahmedabad, Gujarat 380015, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

U. Chatterjee et al. (eds.), *Urban Commons, Future Smart Cities and Sustainability*, Springer Geography, https://doi.org/10.1007/978-3-031-24767-5_36

Introduction

With the advent of 'big data era', many huge avenues for understanding human mobility using significant digital footprints, such as geo-tagged social media information, have emerged (Batty 2010) and have been influenced by urban morphology. Urbanism is a way of thinking about how people communicate with the built environment's technologies. The transportation modes (walking, bus, train, flight, etc.), different population groups (age, racial, occupation, etc.), different geographical environments (road network, region, geology, etc.) and different spatial scales are generally considered to be the causes of disparities in the observed patterns of human mobility (intra-urban, inter-urban, etc.). It is well known that mobility resources play several roles in transportation networks and people's travel patterns in urban areas. Data associated with various modes of transportation typically represents human movements for various purposes and at different spatial scales. People from various socio-economic backgrounds use public transit and urban spaces from distinctive perspectives. Geographical factors, population group, workforce like land use distribution and accessibility to resources significantly affect human mobility patterns. For human mobility modelling, it is crucial to understand the constraints of the identified components on observed human mobility patterns. The use of mobile phone data has many advantages in urban planning, such as the activities from people's movements and the intensity of communication at different times of the day.

Understanding human mobility patterns has received a response due to the fast development of location acquisition technologies, complex network sciences and human dynamics (Kang et al. 2013). The hidden patterns of human activity in space and time are important for traffic forecasting (Jiang et al. 2009), disease spread (Wesolowski et al. 2012), tourism management (Ahas et al. 2007), smart city framework (Ratti et al. 2006), urban computing (Zheng et al. 2011) and prediction of future human mobility based on social network structure (Cho et al. 2011). Some studies also track individual spatio-temporal activities for transportation intelligence (Gao et al. 2012; Li et al. 2011; Liao et al. 2010). We revisited research by Hägerstrand (1970) on time geography, which is well known in human mobility studies where researchers use travel diaries to analyse human mobility patterns in space and time due to a lack of updated data collection ability. Detailed information on travel patterns and the actual demand for travel is hard to get today, but due to the introduction of the Internet of Things (IoT) and Information and Communication Technology (ICT) in the cities, real-time feed on individual spatio-temporal movements can be easily extracted. In 2008, González et al. (2008) used mobile phone data to analyse mobility patterns and found that human mobility patterns emerged from the complexity of population heterogeneity and individual Levy Techniques. Due to their inconsistent spatio-temporal data qualities, various techniques often create uncertainty in the concept of human movements.

Why should the urban planning sector be concerned about mobile phone data? Compared to travel survey data, CDR data provides researchers with new opportunities to explore individual mobility from an alternate view, with low cost, a larger sample, greater update frequency and broader spatio-temporal coverage in the cities. Service providers frequently collect mobile phone locations for network management-specific purposes, so databases are technically free of cost to researchers. The databases enable the research of the individual mobility of millions of people across the major city over a prolonged period of time compared to a few thousand household movements within 1–2 days, usually collected through travel surveys. They are continuously updated in real time, which could lead to more efficient and trackable urban performance indicators and promote more prompt policy responses to growing urban problems.

The rapid mobile communication infrastructure installation, accompanied by personal smartphones, greatly influenced city life and technological advancements made people change their social and working habits. As a result, urban dynamics have become much more complex, and data based on the location of smartphones could potentially become one of the most promising new sources for urban analysis. Scientific literature largely lacks areas such as mapping mobile phone operations in cities or simulating urban dynamics based on smartphone movement. Research efforts in the field are limited as obtaining raw data has hindered academic study.

This study takes the first step to understanding the use of CDRs in human mobility characterisation in the Indian context, using a mobile phone location dataset like CDRs footprints. We hypothesise the CDR records to evaluate the effectiveness of the use CDR dataset in human mobility analysis from an individual perspective. For our study, we develop a methodology to establish OD matrices using mobile phone trajectory of individual's continuous and longitudinal movements, though deviating from real human trajectory is captured using call record data (CDR). This chapter facilitates a better understanding of the use of CDR datasets in India and prompts the transport researchers and authorities to utilise real-time data for analysing human mobility patterns as reported in the literature and case examples that have adopted similar analytic approaches that involve mobile data aggregation to glean travel information and explore human movements.

The remainder of this chapter is organised as follows: The next section discusses the existing research related to this study. Section "Case Studies" presents case studies of countries that have used CDR data in their transport studies. Sections "How Location Information of CDR Data Can be Obtained" and "Privacy Issues of Telephonic Regulatory Authority of India (TRAI)" show how the location information is obtained from the CDR data and the privacy concerns in accessing CDR data in India. We then propose CDR data processing framework and interpretation of the CDR data in sections "Data Processing Framework" and "Data Interpretation" based on individual user CDR data to examine the effectiveness of CDR data using origin–destination (OD) matrix analysis to understand the travel pattern. We conclude and discuss this research in section "Conclusion".

Related Work on Previous Applications of Mobile Phone Data to Mobility Studies

This section discusses the relevant research on the applications of CDR datasets in transport studies. As technological advances have evolved, emerging data sources from passive data collection techniques have shown promise in assisting transportation practitioners in better understanding people's movements through space and time. Limited response rates, a big burden on survey participants and high implementation costs plague traditional travel survey results (Wolf et al. 2003). Over the last two decades, mobile phone penetration rates have amplified in both developed and developing countries, and we have seen a surge in the significant research effort. They aim to implement pervasive technologies such as mobile network data, GPS devices and WiFi hotspots in collecting large amounts of real-time data about individuals and cities to investigate various aspects of mobility. Traditional household travel surveys may be supplemented or enhanced with passively collected data, such as GPS, mobile network and cell phone GPS data, to solve relevant problems. Mobile network data, primarily call detail records, were incorporated in other efforts to improve travel surveys.

Due to the unprecedented scale of digital footprints it offers, mobile phone location data obtained by mobile network operators (MNOs) has also emerged as an appealing data source. Call detail records (CDRs) generated by mobile communication activities (i.e. make/receive a phone call, send/receive a text message) are the type of mobile phone location data used in most existing studies. CDRs keep records of the relevant details (e.g. caller/callee, time, duration) of each occurrence and a unique identifier of the cell tower that manages the communication for billing purposes. Related to GPS, a cell phone's location is obtained by measuring its distance from nearby towers. The distance between the towers, the number of them and the signal strength all have an impact on how accurate the data is. Simply stated, data is only captured when the phone is in use, such as when making or receiving a call or sending a message. This technique can be used to find a phone within 50–300 m. When a cell phone moves, the signal from the nearest and strongest tower is used.

On the other hand, a phone does not need to move to shift towers. A phone can 'oscillate' or switch between towers due to network policies on performance optimisation or vicinity to a competing mobile network with similar strength. Oscillation can cause records in travel studies to demonstrate false movements; real mobility can also be wrongly interpreted as an oscillation due to the repetitive nature of the motion. The caller ID, the timestamp, the length of the call or other activity, the longitude and the latitude are all common components of a CDR dataset. There may also be access to additional information, like the call recipient's ID. These IDs are always made anonymous for security reasons, and the formatting differs across carriers.

Consequently, mobile phone data has become a very promising data source for transportation researchers and has strengthened our knowledge of human mobility in recent years. Since individuals are entities in an urban setting, the spatio-temporal aspects of an urban system can be perceived as a generalisation of individual behaviour; thus, CDR data can be used to analyse aggregated mobility patterns of mobile users in cities. Mobile phone location positioning can be done in several ways, for instance, the combination of GPS, signal strength (SS), finger point, ray trace, time difference of arrival (TDOA) and angle of arrival (AOA). Using these data sources would enable researchers to understand better the laws governing human movements and increase the effectiveness and reliability of urban policies.

The future smart cities can be considered complicated mechanisms composed of various processes and components. The rapid growth of IoTs has resulted in inevitable changes in the spatio-temporal aspects of urban mobility, in addition to providing a rich data source for modelling urban systems. Researchers focused on two elements when examining the development of urban and regional planning based on CDR data.

- The spatial structure where the emphasis is given to how mobility patterns in cities are affected by the compactness and size of the cities (Kang et al. 2012).
- Spatial clustering in which hotspot clustering and activity distribution pattern has been addressed (MIT Senseable Lab 2006).

Since CDRs became popular in the research community in recent years, many valuable results about human behaviour and interactions with the urban environment have been observed (González et al. 2008; Song et al. 2010a, b). However, most past studies did not address the representativeness of CDR data in the population-wide context. A large body of literature shows the potential of acquiring geographical location from cell phone data based on mapping the phone usage for urban analysis (Ratti et al. 2006). Calabrese et al. (2011) study describes a new real-time evaluation of the urban dynamics system in Rome based on data from the location of buses, taxis, traffic conditions and pedestrians to understand urban mobility by anonymous monitoring of cellular network data. Bolla and Davoli (2000) suggested a framework for predicting traffic using an algorithm that determines traffic parameters based on location data from cellular phones. Various studies were carried out by Lovell (2001) and Wideberg et al. (2006) on cellular data of mobile devices in terms of transport applications such as traffic data and journey times or speed. To determine the origin–destination flows, different cell phone signalling datasets like billing data consisting of cell phone tower information have been considered and modelled to evaluate the validity of estimating trips (White and Wells 2002). For example, Sohn and Kim (2008) used cell phone tower handover data obtained every time a phone switches a tower it is connected to during a call. While these findings show great potential for using cellular probe trajectory information to understand mobility patterns, all methods have a number of disadvantages that must be resolved before they can be adopted.

There are some other researches on the usage of mobile phone data/CDR data for human travel pattern visualisation (Asakura and Hato 2004; Reades et al. 2009; Phithakkitnukoon et al. 2010; Phithakkitnukoon and Ratti 2011), mobility pattern extraction (Candia et al. 2008; González et al. 2008; Sevtsuk and Ratti 2010; Song et al. 2010a, b; Wang et al. 2012), route choice modelling (Schlaich et al. 2010) and traffic model calibration (Gundlegård and Karlsson 2011; Demissie et al.

2013). The potential for assessing OD matrices using CDR data which are stored by service providers for billing purposes and thus more easily accessible has also been explored by various researchers like (Mellegård et al. 2011), who developed an algorithm to extract CDR data to traffic nodes and Calabrese et al. (2011) focus on the proposing methodology to reduce noise in CDR data. These studies do not discuss the relationship between mobile phones and traffic OD in detail.

Road traffic congestion continues to manifest and propagate in cities around the world. The research by Nair et al. (2019) looks into drivers' route choice behaviour for route selection, with the help of Google Maps API, by gathering traffic speed data from 29 cities worldwide over 40 days. The study compared traffic conditions across global cities on a common datum using crowdsourced data, which has now become readily available for research purposes. Some studies are based on Global Positioning System (GPS) data primarily from navigation devices (Bierlaire and Frejinger 2008; Broach et al. 2012; Hess et al. 2015). In certain areas of the world, navigation system devices are still not generally used.

Certainly, mobile phone datasets can capture partial trajectories of frequent phone users; the sample is huge, thus opening the way to a new paradigm in urban planning. According to researchers, urban structure significantly affects urban-scale mobility patterns, showing that different parts within a city are correlated with different individuals' movement patterns. As a result, the previous study has concentrated on extracting aggregate trends from cell phone data from different urban areas, such as hubs, nodes and places of interest (Phithakkitnukoon et al. 2010). Thus, we discuss some of the case studies where CDR data has been used to understand mobility patterns in different cities.

Case Studies

Dhaka Joint Trip Generation Model Using CDR

The CDR data has been utilised to identify individual mobility pattern using multiple techniques (Bwambale et al. 2020). His chapter explains CDR data's use to predict the trip generation model. The study uses CDR data to verify the reported individual trips in household survey data in Dhaka city. A total of 16,750 household data and 600 million CDR records have been used in the study. The CDR data is analysed to identify users' home locations by identifying a maximum frequency cell tower for each unique ID. Then, home-based trip is extracted for all the users by arranging each user record in order of its timestamp. Using QGIS software, the total trip is calculated for each tower zone.

The proposed framework has significantly tested accurately for trip generation modelling, and it has the potential to model other transport choices, i.e. trip distribution, route choice, etc. This study is a full-proof concept for using mobile phone data fused with the traditional data source to improve the model's temporal and spatial transferability.

Boston Individual Mobility Analysis

The study by Calabrese et al. (2013) compares mobile phone data with odometer reading data to characterise individual mobility. The study helps identify the intra-urban variation of mobility and the non-vehicular component of overall mobility. An algorithm is developed to calculate the users' total trips and the home location at a 500 m-by-500 m grid level. The trip's length is calculated by measuring the distance between the user's consecutive tower location, and the user's tower location estimates the home location at night. The tower location's repetitiveness helps evaluate the home location's accuracy. While comparing the data with census, data showed that 40% of the users had estimated home location with an accuracy greater than 0.5. The result shows that total trip length has been impacted significantly by the spatial distribution of activities like job accessibility. It has been noted that the increasing intersection density negatively impacts total trip length.

When these figures are coupled with vehicle safety inspection results, the study shows that mobile phone traces are a fair proxy for individual mobility and can provide valuable insights into intra-urban mobility trends and the non-vehicular aspect of overall mobility.

Singapore Taxicab and Mobile Phone Data Analysis

In an insightful study by Kang et al. (2013), taxicab trips and mobile phone data are used to understand individual mobility patterns. The datasets consist of 15,000+ GPS logs of taxicabs and 2,000,000+ telecom users' CDR data. The trip length of taxicabs is calculated by measuring the distance between the passenger on board (POB) and the taxicab payment status. To calculate the trip using CDR data, change in two consecutive mobile activity locations is measured. The weekly spatial distribution of taxicab trips and mobile phone movement shows that both trips are highly homogeneous on Sundays. The results also show that it is possible to identify the land use of a particular location by analysing the temporal pattern of incoming and outgoing taxicab trips. The study clarifies the variations in observed human mobility patterns based on taxicab and cell phone data. It also suggests a combined approach to taxicab and cell phone use obtain more in-depth insights into population dynamics, transportation and urban configuration.

How Location Information of CDR Data Can Be Obtained

An urban area is divided into hexagonal cells, or ‘cells’, with transceiver antennas or ‘cell towers’ at their centres. When a user travels from cell to cell, the network transfers the call to the next cell tower. Cellular technology enables radio channels to be replicated in non-adjacent cells, allowing a device to accommodate a much larger number of users. Each cell within a cellular network is geographically defined by the range of RF signals perpetuated to multi-dimensional space. When a mobile phone user moves and enters a service cell, the network base stations are designed to recognise that the user is located in the vicinity of the station’s area.

In contrast to the latter, utilising data from cell phone networks has many benefits (Townsend 2002). First, travel activity can be explicitly mapped to where it occurs, while on the Internet, activity is linked to the nominal, often fictitious location where a domain is defined. Second, data is not geographically static and, therefore, can account for individual movements as well as the level of communication activity at different times of the day. Third, mobile phones have a very high penetration in most developed countries, making them a perfect technology for acquiring large amounts of statistically relevant data more than the Internet or WiFi; CDR data may include details like the number of calls a person has made to which number and its time and duration of calls to which tower the individual received network while making the particular call. Then, the tower is identified by its unique cell ID number, but it is not always precise as the location is under a sphere of 500 m. CDRs are different from phone tapping in that they contain information, for example, time, date and duration, instead of recording what the person talked about on a phone call. Call data is stored by telecom companies for a period of 6 months, in accordance with government guidelines.

Privacy Issues of Telephonic Regulatory Authority of India (TRAI)

What about privacy? This is the first concern to anyone using location-based devices or apps that may be individual travel movements can be tracked, thus breaching the individual’s privacy. Therefore, it is crucial to emphasise that the CDR data mentioned in the chapter does not violate cell phone users’ privacy. Sample personal data was received from a single GitHub user (Agarwal 2016) and was treated and analysed in an aggregated and anonymous form, in accordance with the consultation chapter on privacy, security and ownership of the data published by the Telephonic Regulatory Authority of India (TRAI) on July 2018. Thus, it is not possible in any way to link the location data with real individuals or have any access to any personal information like age, gender and phone number.

The chapter aims to create a code of conduct that addresses any privacy concerns that may arise during the review, which is far from encouraging individuals to be tracked and to see the potential of CDR data that could provide researchers with useful travel information to enhance the quality of life in urban communities. Several research findings have shown that sharing such data carries benefits and risks.

The mapping of cell phone data can reveal mobility patterns and interactions in urban centres that are likely to have profound value to urban planning and design. Till date, however, no use of CDR data has been used in India. This may be due to the difficulties in establishing a partnership between academics and network operators and mainly due to government regulations. However, we can see that the mapping of CDR data has been a fertile field of research in many cities across the globe.

Data Processing Framework

The sample of the CDR dataset of a single user is extracted from GitHub. These CDR data consist of anonymous SMS, call log details and other telecommunication transactions of telecom users. The data processing framework for our study is divided into two parts, first, to identify the Base Transceiver Station (BTS) location and, second, to generate the OD matrix from the CDR data, which has been described in the following section. Figure 36.1 represents the summary of the data processing framework. Each attribute in CDR consists of a unique ID (which in this case is the IME number), timestamp, latitude and longitude of the (BTS), which has been used by the telecom user (see Table 36.1).

Tower data conversion

The overall idea is to map the geocoordinates of the BTS and subsequently identify the user's travel pattern. The BTS address is completely unstructured and expressed in a string format. To visualise BTS's location, we first converted them into latitude and longitude values using Google Geocoding API. The BTS address in CDR data enables the capture of the transient origin and destination matrix (Wang et al. 2011), which uses mobile phone data to efficiently and economically capture the travel pattern.

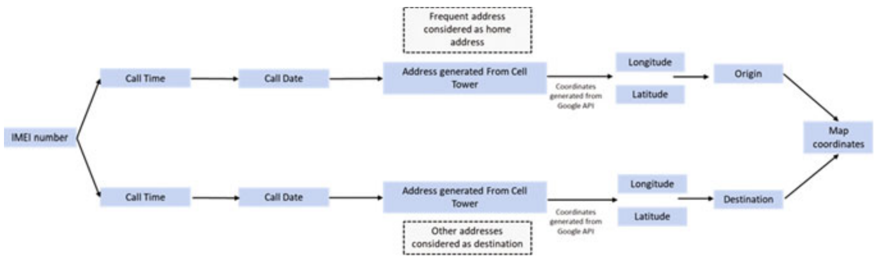

Fig. 36.1 Data processing framework. *Source* Generated by author

Table 36.1 Excerpt of the CDR data

CALL_DIREC	ESN_or_IME	MIN_or_IMS	START_DATE	CALL_TIME	BTS_ADDRESS
IN_CALL	3.54E+14	4.05E+14	01-Jul-12	00:25:04	Old No.36 New
IN_SMS	3.54E+15	4.05E+15	02-Jul-12	01:0S:20	No.28,EVR Rd,
IN_SMS	3.54E+16	4.05E+16	03-Jul-12	05:11:03	Old No.36 New
OUT_SMS	3.54E+17	4.05E+17	04-Jul-12	10:18S7	RAJE5HBATI
OUT_CALL	3.54E+18	4.05E+18	05-Jul-12	10:58:51	No.3/1,Basin
OUT CAIL	3.54E+19	4.05E+19	06-Jul-12	11:28:51	RAJE5HBATI

Source Compiled by author

Tower data to OD matrix

The plotted BTS address has been considered as the user's active location. It is assumed that the serviceable area of a BTS is a traffic analysis zone. To map the OD matrix, initial BTS is considered the origin, and subsequent BTS is considered the destination. For the next set of OD matrix, Destination BTS is changed to Origin BTS, the subsequent BTS is considered the destination, and the process continues. OD cost matrix tool in GIS has been used to calculate the travel pattern.

One limitation of the CDR data is that the geo-tracing of the information depends solely on the unique ID's BTS location. Data can only be traced if the phone has received a call, SMS, or Internet on a certain BTS location. We have used individual CDR data for the study to analyse the travel pattern as it is challenging to extract public CDR in India. The data is extracted from an online source, i.e. GitHub, uploaded by Agarwal (2016). The data comprises 775 call/SMS records, dated between 1 June 2012 and 19 June 2012, of a single user. The CDR data is generated in the following three scenarios,

1. when a user receives or places a call;
2. when a user receives or sends a SMS;
3. when a user connects to the Internet.

Figure 36.2 illustrates the temporal variation recorded over one week. In general, the user generates a few activities on Friday and Saturday. As a result, the number of active pinging from the tower these days shows a decline. The peak is on Tuesday and Wednesday, corresponding to the significant trips generated by the user, respectively.

Data Interpretation

The CDR footprint generated by the single user data is used to identify the user's home location, which is the most frequently used BTS throughout the observation period. From CDR data Fig. 36.3, BTS address, Choolai High Road, Chennai, is the user's home location, and Murugappa Street, Chennai, was the user's regular occupational destination. Other frequent destination travelled by the user includes

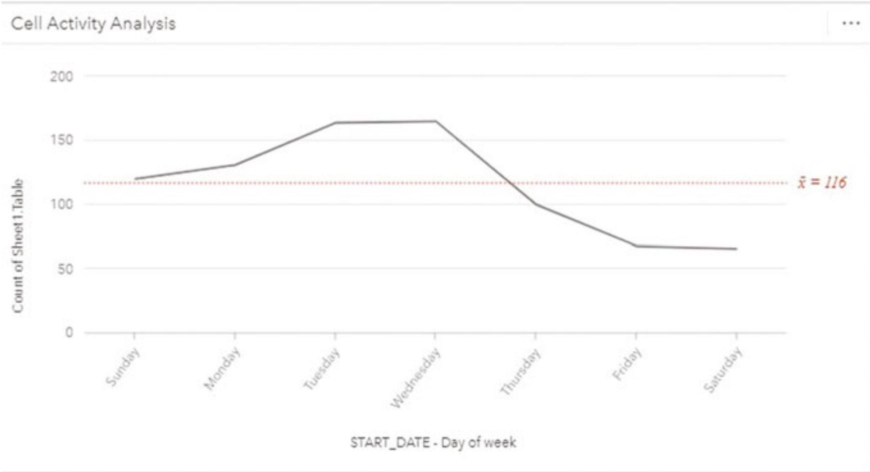

Fig. 36.2 Temporal variation of CDR data over one week. *Source* Generated by author

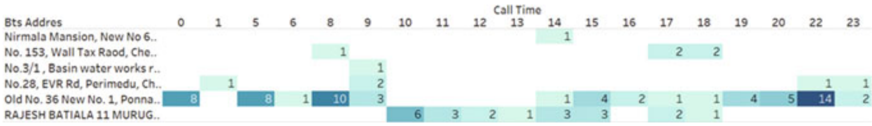

Fig. 36.3 Distribution of user's BTS location. *Source* Generated by author

No. 28 EVR Rd and No. 153 Wall Tax Rd (see Fig. 36.3). The assumptions were based on the number of evidence deduced from the user's call timings. Here, the user's morning and evening calls are deduced as home location, and subsequently, post 10 a.m. till 5 p.m., calls are assumed to be a call from the user's occupation location.

The map shown in Fig. 36.4 attempts to analyse by assessing the mobile behaviour of a single user travelling between different states of India. The OD matrix calculated to analyse the user's travel pattern indicates that the user originated his journey on 1st June 2012 from Chennai and ended in Agra on 19th June 2012. Out of 775 total trips, the user has travelled to 58 unique destinations, including the origin location. In this study, the mobility pattern is analysed based solely on the movement of the person carrying a phone and making calls over 15 days. This mental map determines how an individual moves around cities in India and the travel choices made in the process. Likewise, it builds on the prevalence of cell phones to capture extensive urban dynamics and how it reflects urban interaction patterns.

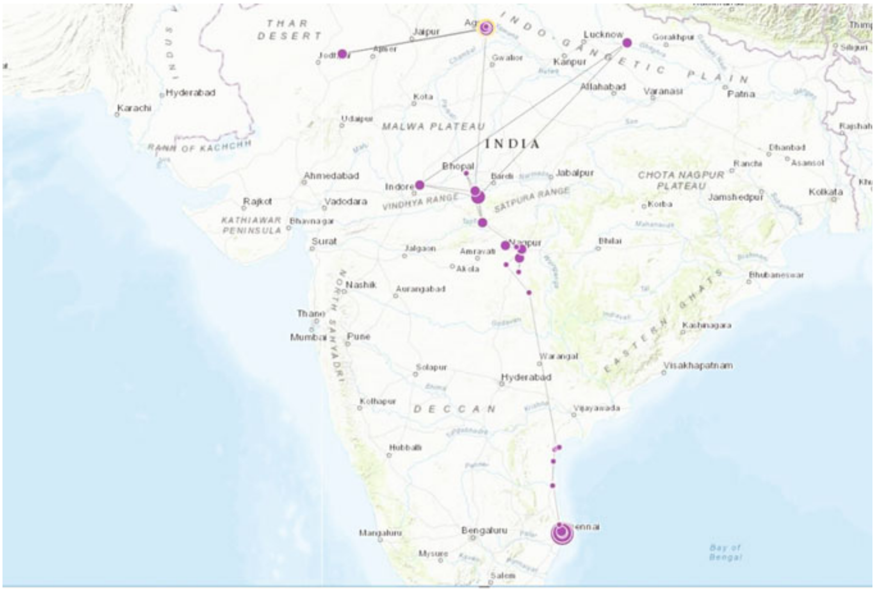

Fig. 36.4 Travel pattern of the user. *Source* GitHub, Sharad Agarwal

Conclusion

This chapter is based on the notion that, despite the boom in the mobile communications market, cellular data is not used in urban analytics in the Indian context. CDRs have been recognised as valuable raw data for human mobility studies. Thus, the aim of the chapter is twofold. This study takes the first step to establish some concepts, demonstrates applications of CDR data in mobility research, summarises state of the art in mapping CDR data and describes some implications for privacy concerns in India. Second, after identifying a lack of research in the application of CDR data in urban studies in India, we present the first findings of an individual user mobility pattern across different cities in India. The important outcome of the research is the data processing framework proposed to understand the travel pattern using CDR data which is interpreted by the author. The methodology is generated based on the individual user mobility analysis. Our objective of this study is to assess the potential of CDR data in mobility research as it has a wide coverage of the population across geographies and can be extracted in real time with relatively low cost. Future research could develop methods to use the CDR data with traffic count data to validate the data, which could be the further scope of the study. Since mobile phone companies already record CDR for billing purposes, the approach is more economical than traditional ones, which rely on expensive household surveys and/or extensive traffic counts. It is also convenient for periodic updates of the OD matrix and extendable for dynamic OD estimation. This method is particularly effective for generating a complex OD matrix where the land use pattern is heterogeneous and asymmetry in the travelling pattern prevails throughout the day. However, there is still a limitation of traditional data sources.

In general, the movements of mobile phones serve as a proxy for all types of human mobility. A brief discussion of potential works based on mobile phone usage is included in this chapter. In recent years, the extensive range of handheld devices and the cell phone infrastructure can provide unlimited real-time information on any geographic area at a very low cost. The lack of understanding of the potential of cell phone data to study urban dynamics is one of the main issues faced by urban and transport researchers. The data processing framework and the interpretation of the CDR dataset analysed in this chapter deals only with an individual user's mobile phone movement. Therefore, access of the CDRs database of all subscribers in any selected study area would have given a deeper understanding of the travel behaviour and the route choice of individuals. Such validation attempts still need to be made in India.

In addition, research has so far been limited to available datasets collected by service providers and the privacy issues in India. The most successful next step will come from ad hoc experiments developed in partnership with India's cell phone operators and government agencies. Such findings seem to open up a new and promising line of urban and transport researchers in India and help calibrate and validate land use and transportation models. In the urban setting, making sense of the infinite flow of data from mobile phone networks is still unknown territory. By analysing data from base stations, urban planners can obtain the ability to monitor quickly evolving urban dynamics that are hard to get with traditional surveys. Furthermore, using data from mobile phones can also deepen our understanding of integrating transportation and land use inference. One promising approach is to use datasets on human movement patterns linked to different modes of transportation to reveal the hierarchical community structure of the city at different scales. In this big data era, understanding the efficacy of CDRs in addressing various research questions can be important to many applicable areas ranging from urban design to transportation planning to air pollution and smart cities.

Acknowledgements The research in this chapter used mobile phone data made available by a GitHub user, Sharad Agarwal.

References

Agarwal S (2016) Comprehensive CDR analyzer—CCAT : call data record analyser for forensic investigation. Available at: https://github.com/sharad1126/ccat (Accessed: 1 Mar 2021)

Ahas R et al (2007) Seasonal tourism spaces in Estonia: case study with mobile positioning data. Tour Manage 28(3):898–910. https://doi.org/10.1016/j.tourman.2006.05.010

Asakura Y, Hato E (2004) Tracking survey for individual travel behaviour using mobile communication instruments. Transp Res Part C: Emerg Technol 12:273–291. https://doi.org/10.1016/j.trc.2004.07.010

Batty M (2010) The pulse of the city. Environ Plann B Plann Des 37:575–577. https://doi.org/10.1068/b3704ed.
Bierlaire M, Frejinger E (2008) Route choice modeling with network-free data. Transp Res Part C: Emerg Technol 16(2):187–198. https://doi.org/10.1016/j.trc.2007.07.007
Bolla R, Davoli F (2000) Road traffic estimation from location tracking data in the mobile cellular network. In: 2000 IEEE wireless communications and networking conference, pp 1107–1112. https://doi.org/10.1109/wcnc.2000.904783
Broach J, Dill J, Gliebe J (2012) Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transp Res Part a: Pol Pract 46(10):1730–1740. https://doi.org/10.1016/j.tra.2012.07.005
Bwambale A et al (2020) Getting the best of both worlds: a framework for combining disaggregate travel survey data and aggregate mobile phone data for trip generation modelling. Transportation 1–28. https://doi.org/10.1007/s11116-020-10129-5
Calabrese F et al (2011) Estimating origin-destination flows using mobile phone location data. IEEE Pervasive Comput 10(4):36–44. https://doi.org/10.1109/MPRV.2011.41
Calabrese F et al (2013) Understanding individual mobility patterns from urban sensing data: a mobile phone trace example. Transp Res Part C: Emerg Technol 26:301–313. https://doi.org/10.1016/j.trc.2012.09.009
Candia J et al (2008) Uncovering individual and collective human dynamics from mobile phone records. J Phys a: Math Theor 41:1–11. https://doi.org/10.1088/1751-8113/41/22/224015
Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: user movement in location-based social networks. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining, pp 1082–1090. https://doi.org/10.1145/2020408.2020579
Demissie MG, de Almeida Correia GH, Bento C (2013) Intelligent road traffic status detection system through cellular networks handover information: an exploratory study. Transp Res Part C: Emerg Technol 32:76–88. https://doi.org/10.1016/j.trc.2013.03.010
Gao Y et al (2012) 'Visualization of Taxi Drivers' Income and Mobility Intelligence'. In: Advances in visual computing, Springer, pp 275–284. https://doi.org/10.1007/978-3-642-33191-6_7
González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. Nature 453(7196):779–782. https://doi.org/10.1038/nature06958
Gundlegård D, Karlsson JM (2011) Road traffic estimation using cellular network signaling in intelligent transportation systems. Wireless Technol Intell Transp Syst
Hägerstrand T (1970) 'Reflections on "what about people in regional science?"'. Papers of the Regional Science Association, pp 7–24. https://doi.org/10.1007/BF01954291
Hess S et al (2015) Developing advanced route choice models for heavy goods vehicles using GPS data. Transp Res Part E: Logistics Transp Rev 77:29–44. https://doi.org/10.1016/j.tre.2015.01.010
Jiang B, Yin J, Zhao S (2009) Characterizing the human mobility pattern in a large street network. Phys Rev E 80(2):1–11. https://doi.org/10.1103/PhysRevE.80.021136
Kang C et al (2012) Intra-urban human mobility patterns: an urban morphology perspective. Phys A 391(4):1702–1717. https://doi.org/10.1016/j.physa.2011.11.005
Kang C et al (2013) Exploring human movements in Singapore: a comparative analysis based on mobile phone and taxicab usages. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining. https://doi.org/10.1145/2505821.2505826
Li B et al (2011) 'Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset. In: IEEE International conference on pervasive computing and communications workshops, PERCOM Workshops 2011, pp 63–68. https://doi.org/10.1109/PERCOMW.2011.5766967
Liao Z, Yu Y, Chen B (2010) 'Anomaly detection in GPS data based on visual analytics'. In: VAST 10—IEEE conference on visual analytics science and technology 2010, Proceedings, pp 51–58. https://doi.org/10.1109/VAST.2010.5652467
Lovell DJ (2001) Accuracy of speed measurements from cellular phone vehicle location systems. ITS J 6(4):303–325. https://doi.org/10.1080/10248070108903698

Mellegård E, Moritz S, Zahoor M (2011) Origin/destination-estimation using cellular network data. Proceedings—IEEE international conference on data mining, ICDM, pp 891–896. https://doi.org/10.1109/ICDMW.2011.132

MIT Senseable Lab (2006) Mobile landscape graz in real time. Available at: http://senseable.mit.edu/graz/ (Accessed: 28 Feb 2021)

Nair DJ et al (2019) Characterizing multicity urban traffic conditions using crowdsourced data. PLoS ONE 14(3):1–16. https://doi.org/10.1371/journal.pone.0212845

Phithakkitnukoon S, Ratti C (2011) Inferring asymmetry of inhabitant flow using call detail records. J Adv in Inf Tech 2(4):239–249. https://doi.org/10.4304/jait.2.4.239-249

Phithakkitnukoon S et al (2010) Activity-aware map: identifying human daily activity pattern using mobile phone data. Hum Behav Underst 6219(3):14–25. https://doi.org/10.1007/978-3-642-14715-9_3

Ratti C et al (2006) Mobile landscapes: using location data from cell phones for urban analysis. Environ Plann B Plann Des 33(5):727–748. https://doi.org/10.1068/b32047

Reades J, Calabrese F, Ratti C (2009) Eigenplaces: analysing cities using the space—time structure of the mobile phone network. Environ Plann B Plann Des 36(5):824–836. https://doi.org/10.1068/b34133t

Schlaich J, Otterstatter T, Friedrich M (2010) Generating trajectories from mobile phone data. TRB 89th annual meeting compendium of papers, vol 339, pp 1–18. Available at: http://www.isv.uni-stuttgart.de/vuv/publication/downloads/201001_JS_TO_Fr_TRB-Generating_Trajectories_from_Mobile_Phone_Data.pdf

Sevtsuk A, Ratti C (2010) Does Urban mobility have a daily routine? Learning from the aggregate data of mobile networks. J Urban Technol 17(1):41–60. https://doi.org/10.1080/10630731003597322

Sohn K, Kim D (2008) Dynamic origin-destination flow estimation using cellular communication system. IEEE Trans Veh Technol 57(5):2703–2713. https://doi.org/10.1109/TVT.2007.912336

Song C, Koren T et al (2010a) Modelling the scaling properties of human mobility. Nat Phys 6:818–823. https://doi.org/10.1038/nphys1760

Song C, Qu Z et al (2010b) Limits of predictability in human mobility

Townsend AM (2002) Mobile communications in the 21st century city. In: Diaper D, Sanger C (eds) The wireless world: social and interactional aspects of the mobile age. Springer, pp 62–77

Wang P et al (2012) Understanding road usage patterns in urban areas. Sci Rep 2:1–6. https://doi.org/10.1038/srep01001

Wang J et al (2011) Dynamic OD expansion method based on mobile phone location. In: Proceedings—4th international conference on intelligent computation technology and automation, ICICTA 2011, pp 788–791. https://doi.org/10.1109/ICICTA.2011.204

Wesolowski A et al (2012) Quantifying the impact of human mobility on malaria. Science 338(6104):267–270. https://doi.org/10.1126/science.1223467

White J, Wells I (2002) Extracting origin destination information from mobile phone data. Road Tran Sporl Inf Cont pp 30–34. https://doi.org/10.1049/cp:20020200

Wideberg JP, Caceres N, Benitez FG (2006) Deriving traffic data from a cellular network. In: 13th world congress on intelligent transport systems and services, pp 1–8

Wolf J, Oliveira M, Thompson M (2003) Impact of Underreporting on mileage and travel time estimates: results from global positioning system-enhanced household travel survey. Transp Res Rec 1854:189–198. https://doi.org/10.3141/1854-21

Zheng Y et al (2011) Urban computing with taxicabs. In: UbiComp'11—Proceedings of the 2011 ACM conference on ubiquitous computing. New York, NY, pp 89–98. https://doi.org/10.1145/2030112.2030126