ARTICLE IN PRESS

Computers, Environment and Urban Systems xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Human mobility and socioeconomic status: Analysis of Singapore and Boston

Yang Xu[a,b,*], Alexander Belyi[b,c], Iva Bojic[b], Carlo Ratti[d]

[a] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] Senseable City Laboratory, FM IRG, SMART Centre, 1 Create Way, Singapore
[c] Faculty of Applied Mathematics and Computer Science, Belarusian State University, 4 Nezavisimosti Ave, Minsk 220030, Belarus
[d] Senseable City Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

ARTICLE INFO

Keywords:
Mobile phone data
Human mobility
Socioeconomic characteristics
Data fusion

ABSTRACT

Recently, some studies have shown that human movement patterns are strongly associated with regional socioeconomic indicators such as per capita income and poverty rate. These studies, however, are limited in numbers and they have not reached a consensus on what indicators or how effectively they can possibly be used to reflect the socioeconomic characteristics of the underlying populations. In this study, we propose an analytical framework — by coupling large scale mobile phone and urban socioeconomic datasets — to better understand human mobility patterns and their relationships with travelers' socioeconomic status (SES). Six mobility indicators, which include radius of gyration, number of activity locations, activity entropy, travel diversity, k-radius of gyration, and unicity, are derived to quantify important aspects of mobile phone users' mobility characteristics. A data fusion approach is proposed to approximate, at an aggregate level, the SES of mobile phone users. Using Singapore and Boston as case studies, we compare the statistical properties of the six mobility indicators in the two cities and analyze how they vary across socioeconomic classes. The results provide a multifaceted view of the relationships between mobility and SES. Specifically, it is found that phone user groups that are generally richer tend to travel shorter in Singapore but longer in Boston. One of the potential reasons, as suggested by our analysis, is that the rich neighborhoods in the two cities are respectively central and peripheral. For three other mobility indicators that reflect the diversity of individual travel and activity patterns (i.e., number of activity locations, activity entropy, and travel diversity), we find that for both cities, phone users across different socioeconomic classes exhibit very similar characteristics. This indicates that wealth level, at least in Singapore and Boston, is not a factor that restricts how people travel around in the city. In sum, our comparative analysis suggests that the relationship between mobility and SES could vary among cities, and such relationship is influenced by the spatial arrangement of housing, employment opportunities, and human activities.

1. Introduction

The last decade has witnessed an explosive growth of scientific research that characterizes and models how people move around in space and time. The interdisciplinary field — broadly conceived as *human mobility analysis* — has attracted researchers across various backgrounds to tackle questions in epidemiology (Bengtsson, Lu, Thorson, Garfield, & Von Schreeb, 2011), sociology (Lazer et al., 2009) and urban planning (Alexander, Jiang, Murga, & González, 2015), among others. With rapid developments of information and location-aware technologies, researchers nowadays have access to large datasets of different types (e. g., mobile phone records, social media data, public transit records). This allows for acquisition of new knowledge about important aspects of human mobility patterns (De Montjoye, Hidalgo, Verleysen, & Blondel, 2013; Gonzalez, Hidalgo, & Barabasi, 2008; Song, Qu, Blumm, & Barabási, 2010).

Despite the numerous insights uncovered by recent human mobility research, there have been limited studies — especially the ones leveraging new and emerging data sources — that analyze the relationships between movement patterns and socioeconomic characteristics of the travelers. This is partially due to a lack of multimodal data that could reveal both travel behavior and socioeconomic status (SES) of the same population. An improved understanding of the relationship between mobility and SES is very important for many scientific domains and

* Corresponding author.
E-mail address: yang.ls.xu@polyu.edu.hk (Y. Xu).

https://doi.org/10.1016/j.compenvurbsys.2018.04.001
Received 26 August 2017; Received in revised form 7 March 2018; Accepted 4 April 2018
0198-9715/

real-world applications, especially the ones that call for human-centered approaches. For example, knowing how travel patterns vary across social classes could help decision makers to control spread of infectious diseases more effectively by targeting the right population groups (Finger et al., 2016), or improve the performance of transportation systems by providing customized mobility solutions to travelers (Alsnih & Hensher, 2003). It can also shed light on many societal issues such as spatial inequality and social stratification (Echenique & Fryer, 2007; Leo, Fleury, Alvarez-Hamelin, Sarraute, & Karsai, 2016).

Recently, some studies have shown that human mobility patterns are strongly associated with regional socioeconomic indicators such as per capita income and poverty rate (Almaatouq, Prieto-Castrillo, & Pentland, 2016; Frias-Martinez, Soguero-Ruiz, Frias-Martinez, & Josephidou, 2013; Pappalardo, Pedreschi, Smoreda, & Giannotti, 2015). However, these studies are limited in numbers and they have not reached a consensus on what indicators or how effectively they can possibly be used to reflect the socioeconomic characteristics of the underlying populations. Hence, this research proposes an analytical framework — by coupling large scale mobile phone and urban socioeconomic datasets — to better understand human mobility patterns and their relationships with travelers' socioeconomic status. Using Singapore and Boston as case studies, this work aims to answer one important research question: How do people belonging to different social classes move around in a city, and whether they use urban spaces in different ways?

By analyzing large scale mobile phone data in Singapore and Boston, we introduce six indicators — which are (1) radius of gyration, (2) number of activity locations, (3) activity entropy, (4) travel diversity, (5) k-radius of gyration, and (6) unicity — to quantify important aspects of phone users' mobility characteristics. Among these indicators, radius of gyration and the entropy-based measures (e.g., activity entropy and travel diversity) have been widely used in existing studies to quantify two salient dimensions of human mobility patterns (Gonzalez et al., 2008; Pappalardo et al., 2015; Song, Koren, Wang, & Barabási, 2010; Song et al., 2010), namely, the *spatial dispersion* and *predictability* of individual movements. K-radius of gyration and unicity are two measures that were proposed more recently to quantify individual movements among the most frequented locations (Pappalardo, Simini, Rinzivillo, Pedreschi, Giannotti, & Barabási, 2015) and the uniqueness of an individual's activity patterns relative to others (De Montjoye et al., 2013). These six mobility indicators, which have gained considerable attention in human mobility research, can either be derived from raw mobile phone data or meaningful location sequences extracted from mobile phone users' trajectories. They capture a comprehensive picture of phone users' travel behavior, such as the spatial extent of activity space (radius of gyration and k-radius of gyration), the regularity of daily activities (number of activity locations and activity entropy), the diversity of movements among important activity locations (travel diversity), and the re-identifiability of mobility traces (unicity).

By further incorporating several socioeconomic datasets — (1) the sale price of residential properties and household interview travel survey in Singapore, and (2) per capita income estimated at census tract level in Boston — we propose a data fusion approach to approximate, at an aggregate level, the socioeconomic status (SES) of mobile phone users. We then compare the statistical properties of the six mobility indicators in the two cities, and analyze their relationships with the phone users' SES. The comparative analysis reveals the socioeconomic dimensions of human mobility, and suggests whether there exist universal patterns across the cities.

The remainder of this article is organized as follows. Section 2 provides an overview of related work of this research. Section 3 introduces the study areas as well as the mobile phone and socioeconomic datasets. In Section 4, we introduce how the mobility indicators are derived and the data fusion approach for approximating phone users' SES. We then present analysis results in Section 5. Finally, in Section 6, we conclude our findings and discuss future research directions.

2. Literature review

2.1. Dimensions of human mobility

Human mobility analysis is an interdisciplinary field that aims to understand the intrinsic properties of human movements as well as the mechanisms behind the observed patterns. The concept of human mobility is broad in a sense that it encompasses various dimensions of human travel at both individual and group levels. The conceptualization and representation of human mobility also vary depending on the contexts of studies and backgrounds of researchers. One important concept that is widely used in geographical and urban studies is *activity space*. Namely, it denotes the daily environment that an individual is using for his or her activities (Golledge & Stimson, 1997). It is usually conceptualized as the set of locations that a particular person has visited as well as his/her travels among those locations (Schönfelder & Axhausen, 2003). Previous studies have employed various activity space measures, such as standard deviational ellipse (Lefever, 1926; Zehavi, 1981), confidence ellipse and minimum spanning trees (Schönfelder & Axhausen, 2004, 2003), and space-time prisms (Kim & Kwan, 2003; Miller, 2005), to better understand people's travel and daily activity patterns. The *activity space* measures mainly focus on quantifying a person's mobility patterns from three perspectives: (1) the spatial extent of daily activities, (2) one's frequented activity locations (i.e., activity "anchor" points), and (3) movements between those locations (Schönfelder & Axhausen, 2003). They collectively form a geographic representation of individual human mobility, and have been widely used to study household travel behavior (Dijst, 1999; Newsome, Walcott, & Smith, 1998) and individual accessibility to urban facilities (Kwan, Murray, O'Kelly, & Tiefelsdorf, 2003; Sherman, Spencer, Preisser, Gesler, & Arcury, 2005).

Recent advancements in information and location-aware technologies have produced many new datasets (e.g., mobile phone records and social media data) that capture the whereabouts of people in space and time. These new datasets have empowered researchers from a wide range of fields, such as computer science, statistical physics, and transportation engineering, to characterize and model individual mobility for large populations. Using a six-month cellphone trajectories of 100,000 users, Gonzalez et al. found that individual travel distance (i.e., displacement) can be approximated by a truncated power-law and that people tend to return to a few highly frequented locations (Gonzalez et al., 2008). By analyzing a three-month mobile phone trajectories of 50,000 users, Song et al. found that human travel patterns are highly predictable and there is a remarkable lack of variability (in predictability) across the population (Song et al., 2010). In another research (Song et al., 2010), which was also based on mobile phone data, the authors developed a microscopic model (*exploration and preferential return*) that is able to reproduce many intrinsic properties (e.g., jump size, visitation probability) of human travel behavior. By applying eigendecomposition to the *MIT Reality Mining* dataset, researchers were able to reconstruct and predict an individual's travel behavior with a high accuracy based on the principle components of his or her activity diary (Eagle & Pentland, 2009). Several important indicators — such as radius of gyration and entropy-based measures — have been used in these studies to capture the spatial dispersion and regularities of human mobility, respectively (Gonzalez et al., 2008; Song et al., 2010; Song et al., 2010). These studies mark a new wave of scientific efforts to uncover the hidden mechanisms that govern individual movements.

Another strand of research focuses more on analyzing collective human behavior and space-time structures of cities. Topics include, but are not limited to, visual analytics of cellular usage (Calabrese, Colonna, Lovisolo, Parata, & Ratti, 2011; Ratti, Frenchman, Pulselli, & Williams, 2006), community detection in urban population flow (Belyi et al., 2017; Nelson & Rae, 2016), and quantification of urban spatial structures (Louail et al., 2014; Roth, Kang, Batty, & Barthélemy, 2011). Some studies have also taken advantage of big urban datasets to