Methods of Social Sensing for Urban Studies

CHAPTER 4

Yu Liu[1], Song Gao[2], Yihong Yuan[3], Fan Zhang[1], Chaogui Kang[4], Yuhao Kang[2], and Keli Wang[1]

[1]Institute of Remote Sensing and Geographical Information Systems, School of Earth and Space Sciences, Peking University, Beijing, China

[2]Department of Geography, University of Wisconsin-Madison, Madison, WI, USA

[3]Department of Geography, Texas State University, San Marcos, TX, USA

[4]Department of Geographic Information Engineering, Wuhan University, Wuhan, China

Abstract

Analyzing large volumes of big geo-data through social sensing provides new research opportunities in urban studies. Such big geo-data include mobile phone records, social media posts, vehicle trajectories, and street view images. They can be used to extract human behavior patterns and infer the geographical characteristics of cities. This chapter discusses a number of analytical methods for big geo-data in social sensing studies, such as temporal signature analysis, text analysis, and image analysis. These methods can be used for various applications such as estimating urban vibrancy, formalizing place semantics, and modeling intraurban human mobility patterns. We structure the chapter sections from a perspective of first- and second-order properties in spatial statistics.

Urban Remote Sensing: Monitoring, Synthesis, and Modeling in the Urban Environment,
Second Edition. Edited by Xiaojun Yang.

4.1 INTRODUCTION

The emergence of various types of big geo-data provides new opportunities to understand individual mobility, urban dynamics, and the complex socioeconomic environment. Social sensing refers to using user-generated spatiotemporal data and corresponding analytical methods to better understand human dynamics and human–environment interactions (Liu et al., 2015b). While remote sensing has traditionally been used to extract physical characteristics of the Earth's surface, social sensing complements remote sensing by revealing human dynamics and the underlying socioeconomic characteristics. Data from social sensing can be used to extract both individual-level characteristics and collective-level features of urban systems. As shown in Figure 4.1, on the one hand, social sensing can uncover individual-level human behavior patterns, such as particular activities and movements, emotions and perceptions, and social ties. On the other hand, it can also reveal geographic environmental patterns, such as land uses, spatial interactions, and place semantics. More discussions on these perspectives are given in this chapter.

Due to the high penetration rate of mobile phones, actively and passively collected mobile phone data are a major data source for modeling human activity behaviors in social sensing studies (Gonzalez et al., 2008; Song et al., 2010; Pei et al., 2014; Gao, 2015; Xu et al., 2015; Jiang et al., 2017; Peng et al., 2019). An example of passively collected mobile phone data is call detailed records (CDRs). CDRs are only generated when a user initiates or receives a phone call or a text message, so these datasets often have a low temporal resolution (Kang et al., 2012a; Yuan et al., 2012a; Gao et al., 2013). Unlike CDRs, actively collected mobile phone data, such as mobile signaling data, are continually produced by mobile devices through

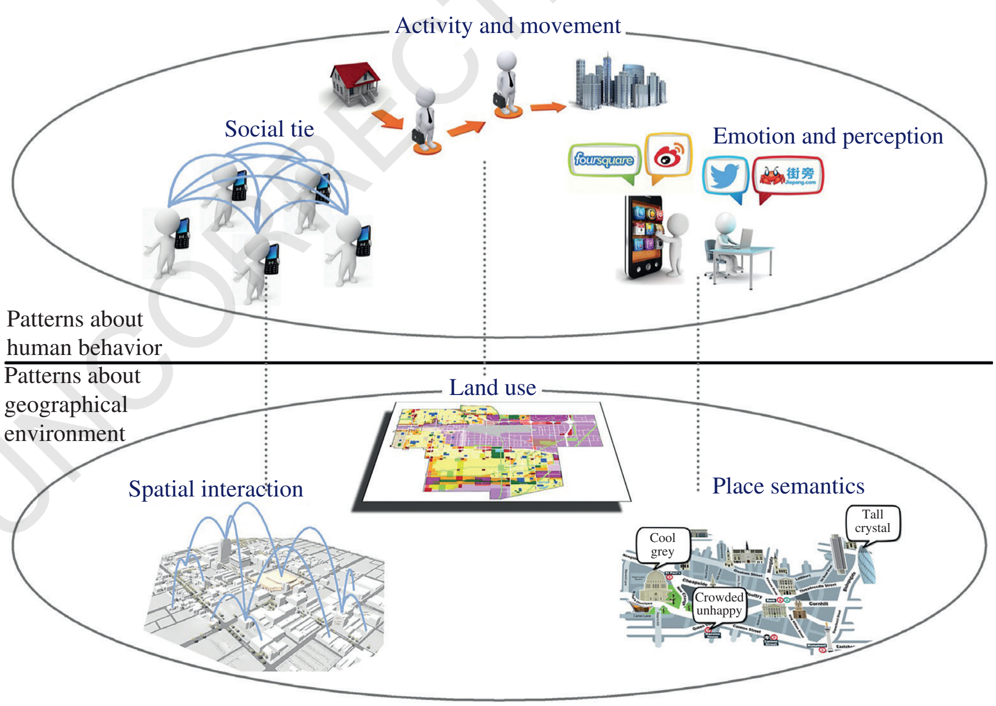

FIGURE 4.1 Social sensing framework at the individual and geographical aggregation levels.

the automatic generation of telematic logs. The telematic logs include data on communications, regular ping updates, cellular handovers, power-ons, and power-offs (Zhao et al., 2016; Xu et al., 2016; Li et al., 2019a). Actively collected mobile phone data are updated more regularly. Both active and passive mobile phone data include locations of the connected cell tower, so they can be used to geo-locate mobile phone users. Based on mobile phone data, three categories of research have drawn much attention in the field: (i) identifying large volumes of individual trajectories and investigating the underlying human mobility patterns; (ii) connecting human activity patterns with urban land uses and urban structures; and (iii) constructing social networks from phone records and investigating how geographic factors affect the formation of such social networks.

Geotagged social media data are another prominent data source for social sensing. Many social media platforms (e.g. Facebook, Twitter, Weibo, Foursquare, Flickr, Yelp, and Instagram) provide a geotagging function that allows users to attach GPS locations or place names to their posts or enable locations for searches (Sui and Goodchild, 2011; Li et al., 2013; Liu et al., 2014a; Ye et al., 2016; Yuan et al., 2020). Therefore, it is feasible to extract the locations of social media users and analyze their movement patterns. Social media data often have data sparsity and representativeness issues, which may affect the generalizability of the results. Additionally, we can also extract semantics and sentiments of places based on social media data using natural language processing (NLP) and content-based image analysis techniques.

In addition, transportation smart card transactions (Gong et al., 2012; Yue et al., 2014; Long and Thill, 2015; Gao et al., 2019a) and vehicles equipped with GPS and cameras also provide rich trajectory- and vision-based datasets to study urban environments (Liu et al., 2012a,b, 2015a; Gebru et al., 2017; Zhu et al., 2017; Zhang et al., 2019a,b).

Given that social sensing provides an alternative and complementary approach to modeling geographical environments in addition to remote sensing methods, researchers have proposed various methods to analyze big geo-data in social sensing studies. A related concept is "semantic signatures" proposed by McKenzie et al. (2015b) and Janowicz et al. (2019). It represents high-dimensional features extracted from places and helps to understand how humans interact with places. Semantic signatures include three types of features: spatial, temporal, and thematic. This chapter discusses in detail the analytical methods and applications of big geo-data in social sensing studies.

The following sections are organized as follows. Section 4.2 focuses on the first-order effect (spatial heterogeneity) of human activities and its applications in understanding place characteristics. In Section 4.3, we discuss the second-order effect (spatial dependency and spatial interactions) of human activities in social sensing studies. Section 4.4 addresses the relation between place characteristics and spatial interactions in urban studies. Section 4.5 concludes this chapter and discusses future research directions.

4.2 SENSING FIRST-ORDER PLACE CHARACTERISTICS

The geographic distribution of human behaviors can be modeled as a mapping between attributes and locations: $y = f(x)$, where x denotes a location, and y represents the value of a particular attribute. In spatial statistics, first-effect properties refer to the attributes directly associated with a location. Social sensing data make it possible to quantify first-order attributes of places from a human perspective. Recently, many analytical methods are available to derive urban land-use types and vibrancy, understand urban environments and human perceptions, and sense human sentiments and emotions toward places.

4.2.1 SENSING URBAN LAND USES AND VIBRANCY

As mentioned in Section 4.1, the availability of big geo-data allows us to investigate various first-order characteristics of places, such as population distributions, land-use types, and urban vibrancy. This subsection introduces two main categories of methods for extracting first-order place characteristics, temporal signature analysis and topic modeling techniques.

4.2.1.1 Temporal Signature Analysis

As shown in Figure 4.2, different land uses (e.g. residential, commercial, and working places) are associated with varying temporal rhythms of human activities under different timescales (e.g. hourly, daily, and weekly). Researchers define such temporal patterns as temporal signatures that are indicative to certain place types (Ye et al., 2011; McKenzie et al., 2015a,b; Janowicz et al., 2019). Therefore, temporal signature analysis is crucial for understanding human dynamics and how different place types or land-use types affect the spatiotemporal distributions of human activities.

Temporal signatures of human activities can reflect population distribution, diurnal movement patterns, and urban structures. Kang et al. (2012a) analyzed a CDR dataset that consists of nearly two million mobile subscribers and found a linear relationship between the number of calls and the number of active mobile subscribers. They further estimated urban population distributions based on mobile phone user activities. One important component of estimating population distribution is to identify users' home and work locations through visit frequency or a density-based spatial clustering technique in different time windows (Xu et al., 2015). We can also infer other types of activity anchor points (e.g. shopping, recreation, and entertainment) using points of interest (POI) data and social media data (Huang and Wong, 2016; Tu et al., 2017). Liu et al. (2012b) investigated the temporal variations of taxi pickups and drop-offs in Shanghai and found an association between taxi pickups/drop-offs and different land-use types (e.g. commercial, industrial, residential, institutional, and recreational) using k-means clustering. Pei et al. (2014) applied a semisupervised fuzzy c-means clustering approach to infer the land-use types in Singapore using mobile phone data. They also found that the accuracy rate of land-use detection decreases as the spatial heterogeneity of land uses increases; however, the accuracy rate increases with the rise of the spatial density of cell phone towers. Jia et al. (2019) used individual trajectories extracted from mobile phones to measure the vibrancy of urban neighborhoods based on the movement patterns of residents. In addition, temporal signatures may also help improve the quality of location-based services, such as place recommendations. For example, we can improve the accuracy of reverse geocoding by incorporating the temporal signatures of places in different cities (McKenzie and Janowicz, 2015).

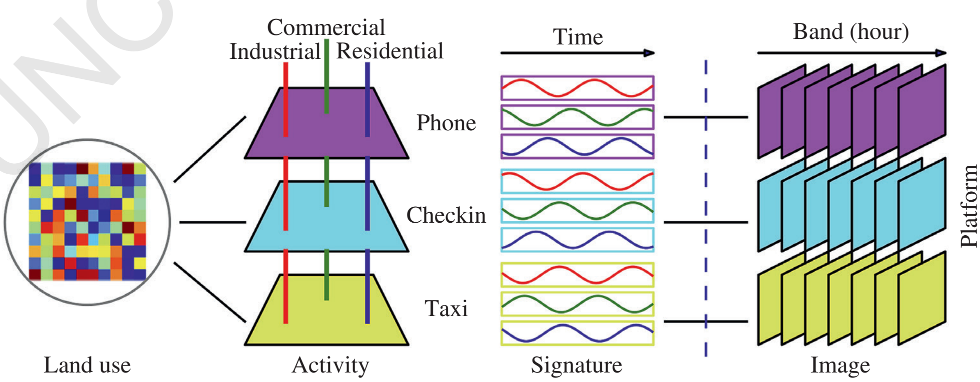

FIGURE 4.2 The generation of temporal signatures in social sensing.

In addition, one trending topic in sensing urban land uses and vibrancy is to propose a multisource or multisensor-fusion social sensing framework. By incorporating remote sensing images, POI data, and real-time social media data, Zhang et al. (2019c) proposed a cross-correlation-based urban land-use model to infer urban functions. Tu et al. (2019) combined multisource urban sensing data and proposed an entropy-based vibrancy index to investigate the heterogeneous spatial patterns of urban vibrancy in Shenzhen, China. They also investigated the demographic, socioeconomic, and environmental factors that may affect the level of urban vibrancy.

4.2.1.2 Topic Modeling

Textual data analyses based on NLP techniques have been widely used in geospatial semantics analysis (Hu, 2018). Among NLP techniques, topic modeling methods were widely adopted to discover and model the structure of latent thematic characteristics when analyzing a large number of textual documents (Papadimitriou et al., 2000; Hofmann, 1999; Blei et al., 2003). Latent Dirichlet Allocation (LDA), a popular topic modeling method, uses a bag-of-words approach to construct topics. The key idea of LDA is that documents can be represented as a joint probability distribution over latent topics, and each topic is characterized by a distribution over words (terms) (Blei, 2012).

Researchers applied a similar idea to analyze the geospatial semantics of our living environment. They view POIs (e.g. restaurants, parks, and bars) as terms; regions or neighborhoods that contain those POIs as documents; and the functions or land-use types of urban regions as topics representing the thematic characteristics of places (Gao et al., 2017; Yuan et al., 2012b). By applying the LDA topic modeling technique, we can find thematic place topics (e.g. shopping areas and art zones) and a probability distribution over place types for each topic. We can also infer the semantics of a place based on its associated topics. For example, one would assume that a *shopping area* would more likely have high colocation probabilities of *clothing stores*, *cosmetics shops*, *shoe stores*, and *cafes* (Papadakis et al., 2019). It is worth noting that human activity patterns play an essential role during the generation of the region (document)–POI type(term) frequency matrix. Therefore, researchers have incorporated location-based social network data (representing place popularity) (Gao et al., 2017) and taxi GPS trajectory data (representing human mobility) (Yuan et al., 2012b) into the LDA topic modeling process to derive urban functional regions. Previous studies also utilized sequence features and word embeddings from NLP domain with augmented spatial contexts to better infer place types (Yan et al., 2017) and urban land uses and regionalization (Yao et al., 2017; Zhai et al., 2019; Li et al., 2019b). Inspired by the idea of LDA, a street segment can be viewed as a word in topic modeling, and the street network can be viewed as a document (Zhu et al., 2017). A typical case study is to apply LDA topic modeling to identify spatial interaction patterns of vehicle movements in urban road networks (Liu et al., 2019).

4.2.2 SENSING PLACE LOCALE CHARACTERISTICS FROM STREET VIEW IMAGES

In addition to urban land uses and human activities, we can extract other characteristics of places, such as place locale, from social sensing data. Locale is one of the three basic elements of a place (i.e. location, locale, and sense of place) (Agnew, 2011). It refers to the physical settings where everyday life activities take place. Locale is composed of visible and tangible elements of a place, such as buildings, streets, and parks (Cresswell, 1992). Place locales have an impact on human activities and can reflect the socioeconomic attributes of neighborhoods. Although images include rich information about various environmental features, automating the extraction of such information used to be difficult. Recent advances in artificial intelligence provide new and efficient solutions for such tasks. A series of deep learning models have shown outstanding performances in various computer vision tasks (LeCun et al., 2015). These models can produce results comparable to or even better than human performance. The success of deep convolutional neural networks (DCNNs) is mostly attributed to its powerful ability to learn effective and interpretable image features (Karpathy et al., 2014). Compared with traditional computer vision features, DCNN is able to learn high-level cognitive information in images, such as complex visual concepts in different scenes. This helps us better capture the cultural and historical styles of places in street-level images. With the help of deep learning, researchers can look beyond the physical appearance of places to understand the culture, emotions, and semantics associated with a place.

The physical characteristics of the urban environment can be derived from street view images at two levels: object level and scene level. At the object level, a deep learning model is adopted to detect and classify each pixel in the images into an object category (Figure 4.3a). For instance, the PSPNet model (Zhao et al., 2017), trained using the ADE20K image annotation database (Zhou et al., 2017), is able to classify 150 object categories with 81% accuracy. The 150 object categories cover the most common types of objects in urban scenes, such as buildings, sky, roads, and vehicles. An urban scene can be represented as a multidimensional vector, with each dimension of the vector corresponding to the presence ratio of each visual object in a street view image (Zhang et al., 2018b). We can also train a specialized object detection model using a customized dataset to recognize a particular type of object, such as fonts on street signs. For instance, researchers found that the typefaces detected from signboards of a neighborhood are associated with the local economy (Ma et al., 2019).

At the scene level, we can apply deep learning models to recognize two types of scene characteristics from street view images: scene categories and scene attributes. As shown in Figure 4.3b, the street view image is categorized as "crosswalk" and "downtown" with a probability of 0.886 and 0.055, respectively. Regarding scene attributes, the street view image can be described as "natural light," "open area," "man-made," etc. Aside from scene category and

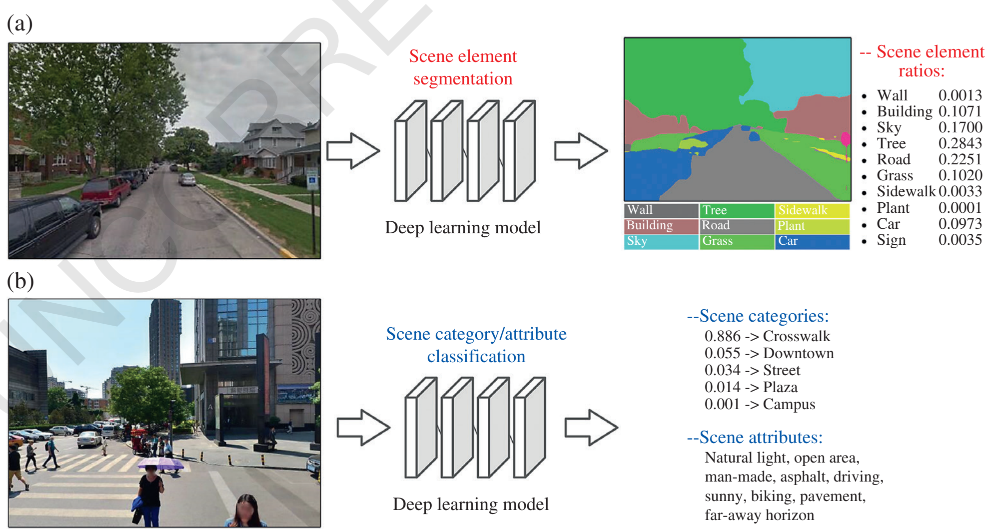

FIGURE 4.3 Sensing place locale characteristics from street view images: (a) Segmenting scene elements using image segmentation model and (b) Recognizing scene categories and attributes using scene category/attribute classification model.

scene attributes, a particular scene type, such as "street canyon" (Hu et al., 2020), can be classified using customized training datasets and DCNN models. Recently, Yan et al. (2018) found that spatial contexts and domain knowledge can improve the accuracy of state-of-the-art deep learning models such as ResNet by over 40% when classifying places.

4.2.3 SENSING HUMAN SENTIMENT AND EMOTION AT DIFFERENT PLACES

Besides physical appearances, it is also important to extract human emotions, sentiments, perceptions, and in-person experiences associated with a place (Zhang et al., 2018a,c). Emotion is a mental state stored in our neural system that is associated with human thoughts and feelings, while people may express their thoughts and feelings through (positive, negative, and neutral) sentiments. Researchers have examined the connections between the physical environment and human emotions or sentiments to enrich our understanding of places.

Two data sources, geotagged texts and images, have been widely used to extract human emotions and sentiments. Texts can be collected from social media platforms such as Twitter and Weibo. Social media users share thoughts about their daily life, nearby events, and the surrounding environments. When users post status updates to social media, their emotions, sentiments, and opinions are also revealed through these posts. Geotagged texts have been widely used to explore the geographic patterns of emotions, opinions, and sentiments with the support of NLP techniques. For instance, Mitchell et al. (2013) analyzed the distribution of emotions in the United States and explored potential socioeconomic factors that determine the emotions associated with a place. Hu et al. (2019) examined people's perceptions toward their living environment based on a sentiment analysis of online neighborhood reviews. Zheng et al. (2019) reported that high levels of air pollution may suppress users' expressions of happiness on social media platforms. Yang and Mu (2015) evaluated the degree of depression expressed by Twitter users and explored its correlation to socioeconomic and climatic factors (Yang et al., 2015).

Compared to geotagged texts, geotagged images have attracted more attention in recent years because of the rapid development of computer vision technologies and deep learning models, which can extract and analyze high-dimensional visual semantics from images. The geotagged photos uploaded to social media platforms may also contain human facial expressions. Previous studies suggested that facial expressions can reflect human emotions (Levenson et al., 1990). For example, Kang et al. (2019) extracted human emotions from facial expressions in over 6 million social media photos and explored the relationship between the physical environment and human emotions at different tourist sites. Similarly, Li et al. (2020) mapped worldwide emotion distributions based on facial expressions to better understand the associations between human emotions and the built environment. Figure 4.4 shows an example of identifying smiling and nonsmiling faces in Disney World Magic Kingdom Park based on facial expressions extracted from Flickr photos. Such facial expressions can collectively measure place sentiments, which are correlated with the socioeconomic attributes of places (Abdullah et al., 2015).

In addition to geotagged photos from social media sites, street view images can also capture real-world scenery and thereby may reflect people's emotions and perceptions of the urban environment. For instance, Zhang et al. (2018c) employed deep learning methods to measure six human perceptual and emotional indicators (i.e. safe, lively, beautiful, wealthy, depressing, and boring) to help understand "a sense of place." Furthermore, Yao et al. (2019) proposed a human-machine adversarial framework to assess users' perceptions and emotions in the urban environment. Zhang et al. (2020) uncovered "inconspicuous-nice" places in Beijing, such as beautiful but rarely visited parks or popular restaurants without an upscale ambiance, by combining street view images and Weibo check-in records. All these studies

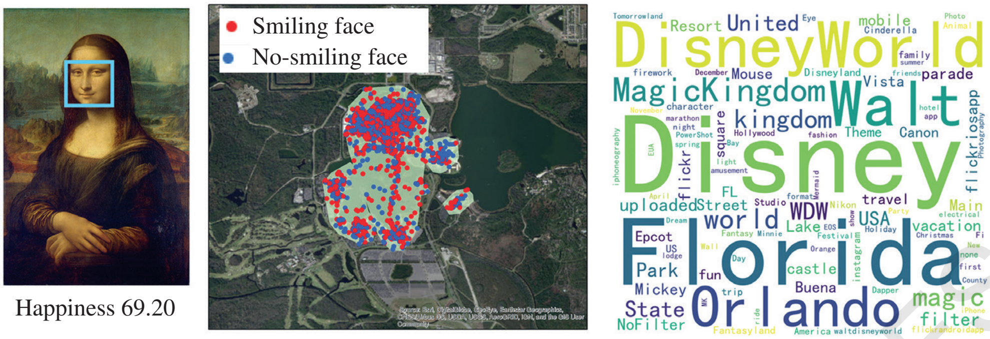

FIGURE 4.4 Sensing human emotions based on facial expressions using Flickr photos and tags: an example of identifying smiling and nonsmiling faces in the Disney World Magic Kingdom Park, Orlando, USA. *Source*: Esri, DigitalGlobe, GeoBye, Earthstar Geographics, CNES/airbus DS, USDA, USDS, AeroGRID, IGN, and other GIS user community.

show the potential of geotagged images as data sources to illustrate human emotions toward a place.

Both text- and image-based approaches have pros and cons in detecting human emotions and sentiments associated with a place. Translation and multicultural challenges may occur for textual information because there is not a universal language used by every culture, whereas facial expressions are relatively universal and consistent cross-culture (Kang et al., 2017). However, compared to abundant geotagged texts, photos on social media platforms are relatively sparse, so the low sampling rate may cause data biases (Huang et al., 2020). Moreover, there are privacy concerns when extracting facial expressions from photos. Combing text data and images may provide a more holistic view of human emotions and sentiments at different places.

4.3 SENSING SECOND-ORDER SPATIAL DEPENDENCY AND INTERACTIONS

In addition to the first-order properties of places (e.g. activity count, land-use patterns, and place semantics) (Fu et al., 2018; Marti et al., 2017; Jenkins et al., 2016), previous studies also addressed the necessity to investigate the correlation and interactions between places (Liu et al., 2016). Unlike first-order properties that measure the characteristics of individual entities, second-order properties measure the dependency and correlation between two or more entities, such as the clustering of crime incidents over space (Adepeju et al., 2016; Gelfand et al., 2010). This dependency over space is often quantified by calculating the magnitude of spatial autocorrelation, a classic measurement for understanding spatial patterns in urban systems. Meanwhile, the availability of big geo-data, such as taxi trajectories and social media check-in data, provides valuable information to reevaluate the classic question of spatial autocorrelation with the help of new multidimensional data sources (Gao et al., 2019b).

Spatial interaction, on the other hand, refers to a dynamic flow process between pairs of locations. Spatial interactions have been widely studied in many fields such as immigration, tourism, and transportation (Kerkman et al., 2017; Liu et al., 2016; Yuan et al., 2017). Although in general spatial dependency and spatial interactions are two distinct processes (Griffith et al., 2017), Getis (1991) proposed a generalized cross-product statistic to demonstrate that spatial interaction models are a special case of a general model of spatial autocorrelation. Spatial interaction focuses on a link between places *i* and *j* caused by different flows, whereas in spatial autocorrelation models, this "link" is quantified by the similarity of characteristics between *i* and *j*. In this section, we adopt the definition from Getis (1991) and view spatial dependency (autocorrelation) and spatial interactions, two related yet different cases, when sensing second-order properties of urban spaces.

4.3.1 METHODS FOR SENSING SPATIAL DEPENDENCY

Measuring spatial autocorrelation is a traditional research question in the field of geographic information sciences. Classic literature on spatial statistics proposed various global and local measures for quantifying spatial autocorrelation under different geographic scales, such as Moran's I, Getis-Ord General G, and Local indicators of spatial association (LISA) (Anselin, 1995; Fischer et al., 2010). The increasing availability of big geo-data provides abundant resources to adopt and improve these methods for investigating spatial dependency in an urban system.

First, data from social sensing are extensive in nature and can therefore be aggregated to represent patterns in different spatial units (Long and Nelson, 2013). In other words, they are particularly suitable for analyzing the clustering and dispersion of spatial phenomena in geographic units, such as the hotspots of human mobility in different urban districts. For example, Steiger et al. (2016) conducted a hotspot analysis on home- and work-related tweets in London. They found a clear indication for nonrandom behaviors of semantic similarity over space and time, suggesting that the semantics of Twitter data can potentially reflect fine-grained collective human behavior over urban spaces. Louail et al. (2014) analyzed the hotspots of mobile phone users to better understand the dynamic properties of urban structures under a fine spatial granularity. Researchers also conducted similar studies based on Bluetooth and taxi trajectories to analyze the clustering and spread of urban mobility patterns, such as the congestion of street traffic during rush hours (Silva and Moreira, 2012; Zhang et al., 2017). These studies provide new insights for understanding the dynamic functionalities of urban spaces in the big data era. Another branch of studies investigated how information diffusion on the Internet connects to spatial dependency in the geographic space. Contrary to Cairncross's statement that spatial dependency in the physical space does not impact the diffusion of information in the cyberspace (Cairncross, 2001), Yang et al. (2019) found that cities with similar scales and population tend to have similar response behaviors on social media.

Second, in addition to adopting existing methods for measuring spatial dependency, researchers also proposed new methods and models to supplement traditional methods. Because big geo-data introduce new information that was otherwise unavailable, it naturally brings new challenges and opportunities for modeling spatial dependency from a methodological perspective. For example, Gao et al. (2019b) proposed an extended Moran's *I* to measure the spatial autocorrelation of time series data. In addition, geotagged check-ins from location--based social media (LBSM) can reveal personal interests and user demographics through semantic analytics, making it possible to incorporate individual-level semantics and social-demographic information into measuring spatial dependency. For example, Steiger et al. (2016) proposed a self-organizing map (SOM)-based method to analyze the connection between geotagged tweets from a combination of spatiotemporal and semantic perspectives. Another study by Radil et al. (2010) combined a person's spatial position in a spatial contiguity matrix with this person's location in a social network to identify similar patterns. Their method was proven more effective than traditional spatial autocorrelation measures for identifying violence and gang behaviors. This branch of research was named "geosocial theories and methods" and has grown rapidly with the widespread of LBSM services (Luo and MacEachren, 2014).

4.3.2 METHODS FOR SENSING SPATIAL INTERACTIONS

Traditionally, spatial interactions result from the movement of physical objects, such as population, cargo, and vehicles. In the information age, spatial interactions can also refer to the flow of abstract concepts, such as information and data (Liu et al., 2014b). Analyzing the spatial interactions between geographic units provides an effective way for revealing the spatiotemporal characteristics of underlying urban structures. Because in the geographic space, distance plays a fundamental role in determining the connections and interactions between places, it is essential to explore the role of distance in understanding spatial interactions (Tobler, 1970). Generally, the interactions between two geographic entities decline as the distance between them increases, which is normally referred to as the "distance decay effect" (Griffith, 2009). Previous studies applied different models, such as the space–time prism (Kwan, 1998) and the gravitational model (Haynes and Fotheringham, 1984), to investigate how distance impacts spatial interactions. Many of these studies focus on deriving a distance decay function, which describes the relation between the distance and a response variable measuring spatial interactions. For instance, two commonly used distance decay functions are exponential $p\left(i,j\right)=e^{-\beta d_{ij}}$ and power-law $p\left(i,j\right)=d_{ij}^{-\beta}$ (Kang et al., 2012b), where *dij* and $p(i, j)$ represent the distance and interactions between i and j, respectively, and β is the distance decay exponent. The larger β is, the stronger impact distance has on the magnitude of interactions.

With the growing availability of individual-level human mobility data, we can examine spatial interactions from both individual and aggregated perspectives (Yue et al., 2014; Yuan et al., 2012a). At the individual level, it is feasible to investigate the distribution of travel distances based on the displacement of each individual trip. Studies show that this distribution is also impacted by the distance delay effect, meaning that most trips are relatively short, and there are very few long trips. At the aggregated level, the distance decay exponent in gravitational models can be fitted by various methods, such as linear programming (O'Kelly et al., 1995), algebraic simplification (Shen, 2004), and Monte Carlo simulation (Westerlund and Wilhelmsson, 2011). In addition, because of the spatial heterogeneity of urban spaces, the distance decay exponent obtained from fitting the distribution of individual trips can be different from the exponent obtained by fitting a gravitational model based on aggregated measurements (e.g. population count and traffic flow). Another branch of research studies the distance decay effect without calculating an explicit distance decay function; instead, researchers focus on modeling the likelihood of interactions between origins and destinations (Stouffer, 1940). Examples include the discrete-choice and utility-maximization models, with which researchers can implicitly examine the relationship between the distance and the magnitude of spatial intersections without an explicit distance decay function (Simini et al., 2012). Based on these quantitative approaches, many studies concluded that the distance decay exponents for human mobility in the geographic space are generally between 0 and 2, with 0 indicating no or very little distance decay and 2 indicating a strong distance decay of spatial interactions.

Because of the distance decay effect, geographic entities that heavily interact with others are often clustered spatially, such as urban communities (Ratti et al., 2010) and travel motifs (Schneider et al., 2013). Many quantitative methods have been developed for extracting these self-organized structural patterns as well as exploring the impact of distance decay in the formation of such patterns (Chen et al., 2015). Typical methods include graph partitioning, modularity optimization, and random walks (Fortunato, 2010). Other studies focused on examining the factors (e.g. socioeconomic and demographic attributes) that may influence the forming of communities when eliminating the impact of spatial distance (Expert et al., 2011). Based on the existing studies, various factors, such as administrational boundaries, the distance decay effect, and social-cultural characteristics, can contribute to the formation of spatial structures within a group of interacted geographic entities.

Another popular research topic that goes beyond the geographic space is to compare spatial interactions in the physical space with social interactions in the virtual space (Shaw and Yu, 2009). The rationale behind this is that the interactions in both the physical and virtual spaces can be influenced by the distance decay effect. For instance, populations in the same social group often share similar activity spaces (Wang et al., 2015), whereas the activity spaces of different social groups can be substantially different (Shi et al., 2015). Besides, human activities in the cyberspace, such as web searching and online news reports, can also provide insights on the spatial interactions between geographical entities (Yuan, 2017; Yuan et al., 2017; Grasland, 2019). As a result, many empirical studies suggest that although the distance decay effect is declining in the information era, it for sure still exists.

4.4 INTEGRATING PLACE CHARACTERISTICS WITH SPATIAL INTERACTIONS

On the one hand, the magnitude of spatial interactions between places is often determined by the characteristics of places, such as their locations and socioeconomic attributes. On the other hand, spatial interactions also add another dimension to place characteristics, one that reflects the dynamics and flows between places. Because of globalization and the increased mobility of urban residents, more and more research has focused on how increased mobility leads to stronger spatial interactions as well as the impacts of spatial interactions on redefining place characteristics (Ren et al., 2019; Zhu et al., 2020).

There are several factors to consider when analyzing the spatial interactions between places, such as the distance between places, the direction of the interaction (i.e. one- or two-way), and the intensity of the interaction. Based on Tobler's First Law of Geography, closer places are likely to have more intense interactions than faraway places (Liu et al., 2014a). In addition, the "attractiveness" of places also affects the intensity of spatial interactions (Guo et al., 2012). Here, attractiveness can be defined based on the nature of the interaction itself. For example, in tourism studies, the "attractiveness" of a place when attracting visitors is determined by its tourist resources, whereas in immigration, "attractiveness" can be influenced by a country's Gross Domestic Product (GDP) and diplomatic policies (Yang and Zhang, 2019). Therefore, spatial interactions can help us better understand place characteristics. There are many applications of using spatial interactions to enrich place characteristics in urban planning and policy making. For example, Kong et al. (2017) extracted the spatial interactions between hospitals based on taxi data and grouped hospitals into five categories. They further evaluated the characteristics of the patients treated by these five categories of hospitals. The findings can be used to optimize the spatial locations of public facilities.

Moreover, both spatial and temporal dimensions of spatial interactions can help enrich place characteristics and capture how places and urban structures evolve upon time. For example, based on the magnitude of spatial interactions between places at different times of day, we can calculate the average commute time and commute distances between places and analyze fine-scale land-use patterns (Long et al., 2012). Another study by Liu et al. (2016) investigated the similarity between places by analyzing the temporal variations of activity flows between different regions. Tao et al. (2019) constructed a four-dimensional probability tensor of time (T) × week (W) × origin (O) × destination (D) to extract spatial interactions between different areas and reexamined the partition of urban functional regions based on human movement patterns at different times. At a finer scale, Zhuo et al. (2019) analyzed the spatial interactions between buildings with different functionalities. They found that spatial interactions are more helpful than temporal signatures to improve the accuracy of inferring building functions.

From a methodological perspective, researchers have developed various models and algorithms to understand place characteristics and spatial configurations based on spatial interactions. Ratti et al. (2010) quantified and visualized the level of interactions between different regions of the United Kingdom. The results showed that the spatial interactions extracted from the locations of two parties on a phone call are highly influenced by the administrative boundaries. Similarly, Liu et al. (2015a) used a network-based method and taxi data to reveal Shanghai's two-tier urban polycentric structure. Yin et al. (2017) argued that administrative boundaries may not reflect human interactions, so they constructed a mobility network of Twitter users to better represent boundaries or urban regions as well as analyze intra- and interregional spatial interactions. However, most of these studies focused on observing spatial interactions with sufficient data to understand place characteristics. When there is insufficient data to directly observe spatial interactions, models and auxiliary attribute data can help fill in the gap. Li et al. (2018) used POI data to determine the spatial distribution of workplaces, and they applied the gravity model to estimate the number of people using different commuting modes in Shanghai. This study demonstrates an effective method to model spatial interactions with the help of auxiliary data and mathematical models.

In sum, spatial interactions can help enrich place semantics at different spatial scales, from buildings, cities, regions, countries, to the entire world. Many studies proposed a variety of algorithms to investigate place characteristics (e.g. extracting functional regions) based on the spatiotemporal patterns of spatial interactions. However, the geographical context of a place, which can be represented by the spatial interactions between the place and other places, is usually uncertain and complex (Kwan, 2012). A recently proposed graph convolutional neural network method may shed light on this issue (Zhu et al., 2020). However, understanding the complex nature of places and the interactions between places remains a challenging research question.

4.5 CONCLUSIONS

To better understand an urban system, it is crucial to integrate both physical and socioeconomic characteristics. Remotely sensed imagery already provides an effective approach to measuring physical features, so it is necessary to develop an efficient measure to sense the socioeconomic patterns of cities. Fortunately, with the development of information and communications technologies (ICTs), the availability of various big geo-data enables us to mine human behaviors and infer the socioeconomic properties of cities. The term "social sensing" was proposed to describe the capability of big geo-data in capturing human activity patterns and urban dynamics. Because most big geo-data are generated in urban areas, social sensing is especially valuable for urban studies. Note that the data quality issue of social sensing has been well recognized and widely discussed (Yuan et al., 2020). A possible solution is to integrate multisource data, including traditional survey-based small data.

This chapter summarizes the widely used big geo-data in social sensing and their applications in urban studies. Following the classic research framework in spatial statistics, we also classified social sensing methods into methods for extracting first-order place characteristics and methods for analyzing second-order spatial dependency and spatial interactions. Because social sensing effectively captures the human and socioeconomic aspects of cities, it is necessary to construct an analytical framework that integrates social sensing and remote sensing to better understand urban systems.

REFERENCES

Abdullah, S., Murnane, E. L., Costa, J. M. R. and Choudhury, T. (2015). Collective smile: Measuring societal happiness from geolocated images. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Vancouver, BC, Canada. New York, NY, USA: Association for Computing Machinery.

Adepeju, M., Rosser, G. and Cheng, T. (2016). Novel evaluation metrics for sparse spatio-temporal point process hotspot predictions - a crime case study. *International Journal of Geographical Information Science*, 30(11): 2133–2154.

Agnew, J. A. (2011). Space and place. In: *Handbook of Geographical Knowledge* (eds. Agnew, J.A. and Livingstone, D.N.), 316–331. London, UK: Sage.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2): 93–115.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4): 77–84.

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan): 993–1022.

Cairncross, F. (2001). *The Death of Distance: How the Communications Revolution is Changing our Lives.* Boston: Harvard Business School Press.

Chen, Y., Xu, J., and Xu, M. (2015). Finding community structure in spatially constrained complex networks. *International Journal of Geographical Information Science*, 29(6),889–911.

Cresswell, T. (1992). *In Place-out of Place: Geography, Ideology, and Transgression*. University of Minnesota Press.

Expert, P., Evans, T. S., Blondel, V. D. and Lambiotte, R. (2011). Uncovering space-independent communities in spatial networks. *Proceedings of the National Academy of Sciences*, 108(19): 7663–7668.

Fischer, M. M., Reismann, M. and Scherngell, T. (2010). Spatial interaction and spatial autocorrelation. In: *Perspectives on Spatial Data Analysis* (eds. Anselin, L. and Rey, S.J.), 61–79. Cham, Switzerland: Springer.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3–5): 75–174.

Fu, C., McKenzie, G., Frias-Martinez, V. and Stewart, K. (2018). Identifying spatiotemporal urban activities through linguistic signatures. *Computers, Environment and Urban Systems*, 72: 25–37.

Gao, S. (2015). Spatio-temporal analytics for exploring human mobility patterns and urban dynamics in the mobile age. *Spatial Cognition and Computation*, 15(2): 86–114.

Gao, S., Liu, Y., Wang, Y. and Ma, X. (2013). Discovering spatial interaction communities from mobile phone d ata. *Transactions in GIS*, 17(3): 463–481.

Gao, S., Janowicz, K. and Couclelis, H. (2017). Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS*, 21(3): 446–467.

Gao, Q.-L., Li, Q.-Q., Zhuang, Y., Yue, Y., Liu, Z.-Z., Li, S.-Q. and Sui, D. (2019a). Urban commuting dynamics in response to public transit upgrades: a big data approach. *PLoS One*, 14(10): e0223650.

Gao, Y., Cheng, J., Meng, H. and Liu, Y. (2019b). Measuring spatio-temporal autocorrelation in time series data of collective human mobility. *Geo-Spatial Information Science*, 22(3): 166–173.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L. and Fei-Fei, L. (2017). Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences*, 114(50): 13108–13113.

Gelfand, A. E., Diggle, P., Guttorp, P. and Fuentes, M. (2010). *Handbook of Spatial Statistics*. Boca Raton: CRC Press.

Getis, A. (1991). Spatial interaction and spatial autocorrelation - a cross-product approach. *Environment and Planning A*, 23(9): 1269–1277.

Gong, Y., Liu, Y., Lin, Y., Yang, J., Duan, Z. and Li, G. (2012). Exploring spatiotemporal characteristics of intra-urban trips using metro smartcard records. *2012 20th International Conference on Geoinformatics*, Hong Kong, China. Columbus, United States: IEEE.

Gonzalez, M. C., Hidalgo, C. A. and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196): 779–782.

Grasland, C. (2019). International news flow theory revisited through a space–time interaction model: application to a sample of 320,000 international news stories published through RSS flows by 31 daily newspapers in 2015. *International Communication Gazette*, 82(3): 231–259.

Griffith, D. A. (2009). Modeling spatial autocorrelation in spatial interaction data: empirical evidence from 2002 Germany journey-to-work flows. *Journal of Geographical Systems*, 11(2): 117–140.

Griffith, D. A., Fischer, M. M. and LeSage, J. (2017). The spatial autocorrelation problem in spatial interaction modelling: a comparison of two common solutions. *Letters in Spatial and Resource Science*, 10(1): 75–86.

Guo, D., Zhu, X., Jin, H., Gao, P. and Andris, C. (2012). Discovering spatial patterns in origin-destination mobility data. *Transactions in GIS*, 16(3): 411–429.

Haynes, K. E. and Fotheringham, A. S. (1984). *The Gravity Model and Spatial Interaction*. Sage Publication.

Hofmann, T. (1999). Probabilistic latent semantic indexing. *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Berkeley, California, USA. New York, NY, USA: Association for Computing Machinery.

Hu, Y. (2018). Geo-text data and data-driven geospatial semantics. *Geography Compass*, 12(11): e12404.

Hu, Y., Deng, C. and Zhou, Z. (2019). A semantic and sentiment analysis on online neighborhood reviews for understanding the perceptions of people toward their living environments. *Annals of the American Association of Geographers*, 109(4): 1052–1073.

Hu, C., Zhang, F., Gong, F., Ratti, C. and Li, X. (2020). Classification and mapping of urban canyon geometry using Google street view images and deep multitask learning. *Building and Environment*, 167: 106424.

Huang, Q. and Wong, D. W. S. (2016). Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us? *International Journal of Geographical Information Science*, 30(9),1873–1898.

Huang, Y., Li, J., Wu, G. and Fei, T. (2020). Quantifying the bias in place emotion extracted from photos on social networking sites: a case study on a university campus. *Cities*, 102: 102719.

Janowicz, K., McKenzie, G., Hu, Y., Zhu, R. and Gao, S. (2019). Using semantic signatures for social sensing in urban environments. In: *Mobility Patterns, Big Data and Transport Analytics* (eds. Antoniou, C., Dimitriou, L. and Pereira, F.), 31–54. Amsterdam, Netherlands: Elsevier.

Jenkins, A., Croitoru, A., Crooks, A. T. and Stefanidis, A. (2016). Crowdsourcing a collective sense of place. *PLoS One*, 11(4): e0152932.

Jia, C., Du, Y., Wang, S., Bai, T. and Fei, T. (2019). Measuring the vibrancy of urban neighborhoods using mobile phone data with an improved PageRank algorithm. *Transactions in GIS*, 23(2): 241–258.

Jiang, S., Ferreira, J. and Gonzalez, M. C. (2017). Activity-based human mobility patterns inferred from mobile phone data: a case study of Singapore. *IEEE Transactions on Big Data*, 3(2): 208–219.

Kang, C., Liu, Y., Ma, X. and Wu, L. (2012a). Towards estimating urban population distributions from mobile call data. *Journal of Urban Technology*, 19(4): 3–21.

Kang, C., Ma, X., Tong, D. and Liu, Y. (2012b). Intra-urban human mobility patterns: an urban morphology perspective. *Physica A: Statistical Mechanics and its Applications*, 391(4): 1702–1717.

Kang, Y., Wang, J., Wang, Y., Angsuesser, S. and Fei, T. (2017). Mapping the sensitivity of the public emotion to the movement of stock market value: a case study of Manhattan. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42: 1213–1221.

Kang, Y., Jia, Q., Gao, S., Zeng, X., Wang, Y., Angsuesser, S., Liu, Y., Ye, X. and Fei, T. (2019). Extracting human emotions at different places based on facial expressions and spatial clustering analysis. *Transactions in GIS*, 23(3): 450–480.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. and Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. *27th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014*, Columbus, United States: IEEE Computer Society.

Kerkman, K., Martens, K. and Meurs, H. (2017). A multilevel spatial interaction model of transit flows incorporating spatial and network autocorrelation. *Journal of Transport Geography*, 60: 155–166.

Kong, X., Liu, Y., Wang, Y., Tong, D. and Zhang, J. (2017). Investigating public facility characteristics from a spatial interaction perspective: a case study of Beijing hospitals using taxi data. *ISPRS International Journal of Geo-Information*, 6(2): 38–38.

Kwan, M.-P. (1998). Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework. *Geographic Analysis*, 30(3): 191–216.

Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102(5): 958–968.

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature*, 521(7553): 436–444.

Levenson, R. W., Ekman, P. and Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27(4): 363–384.

Li, L., Goodchild, M. F. and Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, 40(2): 61–77.

Li, M., Kwan, M.-P., Wang, F. and Wang, J. (2018). Using points-of-interest data to estimate commuting patterns in central Shanghai, China. *Journal of Transport Geography*, 72: 201–210.

Li, M., Gao, S., Lu, F. and Zhang, H. (2019a). Reconstruction of human movement trajectories from large-scale low-frequency mobile phone data. *Computers, Environment and Urban Systems*, 77: 101346.

Li, Y., Fei, T. and Zhang, F. (2019b). A regionalization method for clustering and partitioning based on trajectories from NLP perspective. *International Journal of Geographical Information Science*, 33(12): 2385–2405.

Li, Y., Fei, T., Huang, Y., Li, J., Li, X., Zhang, F., Kang, Y. and Wu, G. (2020). Emotional habitat: mapping the global geographic distribution of human emotion with physical environmental factors using a species distribution model. *International Journal of Geographical Information Science*. doi.org/10.1080/13658816.2020.1755040.

Liu, Y., Kang, C., Gao, S., Xiao, Y. and Tian, Y. (2012a). Understanding intra-urban trip patterns from taxi trajectory data. *Journal of Geographical Systems*, 14(4): 463–483.

Liu, Y., Wang, F., Xiao, Y. and Gao, S. (2012b). Urban land uses and traffic 'source-sink areas': evidence from GPS-enabled taxi data in Shanghai. *Landscape and Urban Planning*, 106(1): 73–87.

Liu, Y., Sui, Z., Kang, C. and Gao, Y. (2014a). Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data. *PLoS One*, 9(1): e86026.

Liu, Y., Wang, F. H., Kang, C. G., Gao, Y. and Lu, Y. M. (2014b). Analyzing relatedness by toponym co--occurrences on web pages. *Transactions in GIS*, 18(1): 89–107.

Liu, X., Gong, L., Gong, Y. and Liu, Y. (2015a). Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography*, 43: 78–90.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G. and Shi, L. (2015b). Social sensing: a new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 105(3): 512–530.

Liu, X., Kang, C., Gong, L. and Liu, Y. (2016). Incorporating spatial interaction patterns in classifying and understanding urban land use. *International Journal of Geographical Information Science*, 30(2): 334–350.

Liu, K., Gao, S. and Lu, F. (2019). Identifying spatial interaction patterns of vehicle movements on urban road networks by topic modelling. *Computers, Environment and Urban Systems*, 74: 50–61.

Long, J. A. and Nelson, T. A. (2013). A review of quantitative methods for movement data. *International Journal of Geographical Information Science*, 27(2): 292–318.

Long, Y. and Thill, J.-C. (2015). Combining smart card data and household travel survey to analyze jobs---housing relationships in Beijing. *Computers, Environment and Urban Systems*, 53: 19–35.

Long, Y., Zhang, Y. and Cui, C. (2012). Identifying commuting pattern of Beijing using bus smart card data. *Acta Geographica Sinica*, 67(10): 1339–1352.

Louail, T., Lenormand, M., Cantu Ros, O. G., Picornell, M., Herranz, R., Frias-Martinez, E., Ramasco, J. J. and Barthelemy, M. (2014). From mobile phone data to the spatial structure of cities. *Scientific Reports*, 4: 5276.

Luo, W. and MacEachren, A. M. (2014). Geo-social visual analytics. *Journal of Spatial Information Science*, 2014(8): 27–66.

Ma, R., Wang, W., Zhang, F., Shim, K. and Ratti, C. (2019). Typeface reveals spatial economic patterns. *Scientific Reports*, 9(1): 15946.

Marti, P., Serrano-Estrada, L. and Nolasco-Cirugeda, A. (2017). Using locative social media and urban cartographies to identify and locate successful urban plazas. *Cities*, 64: 66–78.

McKenzie, G. and Janowicz, K. (2015). Where is also about time: a location-distortion model to improve reverse geocoding using behavior-driven temporal semantic signatures. *Computers, Environment and Urban Systems*, 54: 1–13.

McKenzie, G., Janowicz, K., Gao, S. and Gong, L. (2015a). How where is when? On the regional variability and resolution of geosocial temporal signatures for points of interest. *Computers, Environment and Urban Systems*, 54: 336–346.

McKenzie, G., Janowicz, K., Gao, S., Yang, J.-A. and Hu, Y. (2015b). POI pulse: a multi-granular, semantic signature--based information observatory for the interactive visualization of big geosocial data. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 50(2): 71–85.

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S. and Danforth, C. M. (2013). The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PloS One*, 8(5): e64417.

O'Kelly, M. E., Song, W. and Shen, G. Q. (1995). New estimates of gravitational attraction by linear--programming. *Geographical Analysis*, 27(4): 271–285.

Papadakis, E., Gao, S. and Baryannis, G. (2019). Combining design patterns and topic modeling to discover regions that support particular functionality. *ISPRS International Journal of Geo-Information*, 8(9): 385–385.

Papadimitriou, C. H., Raghavan, P., Tamaki, H., Vempala, S. and Vempala, S. (2000). Latent semantic indexing: a probabilistic analysis. *Journal of Computer and System Sciences*, 61(2): 217–235.

Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T. and Zhou, C. (2014). A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science*, 28(9): 1988–2007.

Peng, H., Du, Y., Liu, Z., Yi, J., Kang, Y. and Fei, T. (2019). Uncovering patterns of ties among regions within metropolitan areas using data from mobile phones and online mass media. *GeoJournal*, 84(3): 685–701.

Radil, S. M., Flint, C. and Tita, G. E. (2010). Spatializing social networks: using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. *Annals of the Association of American Geographers*, 100(2): 307–326.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R. and Strogatz, S. H. (2010). Redrawing the map of Great Britain from a network of human interactions. *PloS One*, 5(12): e14248.

Ren, M., Lin, Y., Jin, M., Duan, Z., Gong, Y. and Liu, Y. (2019). Examining the effect of land-use function complementarity on intra-urban spatial interactions using metro smart card records. *Transportation*. doi.org/10.1007/s11116-019-09977-7.

Schneider, C. M., Belik, V., Couronné, T., Smoreda, Z. and González, M. C. (2013). Unravelling daily human mobility motifs. *Journal of the Royal Society*, 10(84): 20130246.

Shaw, S.-L. and Yu, H. (2009). A GIS-based time-geographic approach of studying individual activities and interactions in a hybrid physical–virtual space. *Journal of Transport Geography*, 17(2): 141–149.

Shen, G. (2004). Reverse-fitting the gravity model to inter-city airline passenger flows by an algebraic simplification. *Journal of Transport Geography*, 12(3): 219–234.

Shi, L., Chi, G., Liu, X. and Liu, Y. (2015). Human mobility patterns in different communities: a mobile phone data-based social network approach. *Annals of GIS*, 21(1): 15–26.

Silva, I. M. P. and Moreira, A. J. C. (2012). Evaluation of bluetooth technology as a sensor of urban mobility. *7th Iberian Conference on Information Systems and Technologies (CISTI 2012)*, Madrid, Spain. Columbus, United States: IEEE.

Simini, F., Gonzalez, M. C., Maritan, A. and Barabasi, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, 484(7392): 96–100.

Song, C., Qu, Z., Blumm, N. and Barabási, A.-L. (2010). Limits of predictability in human mobility. *Science*, 327(5968): 1018–1021.

Steiger, E., Resch, B. and Zipf, A. (2016). Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks. *International Journal of Geographical Information Science*, 30(9): 1694–1716.

Stouffer, S. A. (1940). Intervening opportunities: a theory relating to mobility and distance. *American Sociological Review*, 5(6): 845–867.

Sui, D. and Goodchild, M. (2011). The convergence of GIS and social media: challenges for GIScience. *International Journal of Geographical Information Science*, 25(11): 1737–1748.

Tao, H., Wang, K., Zhuo, L. and Li, X. (2019). Re-examining urban region and inferring regional function based on spatial temporal interaction. *International Journal of Digital Earth*, 12(3): 293–310.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46: 234–240.

Tu, W., Cao, J., Yue, Y., Shaw, S. L., Zhou, M., Wang, Z., . . ., and Li, Q. (2017). Coupling mobile phone and social media data: a new approach to understanding urban functions and diurnal patterns. *International Journal of Geographical Information Science*, 31(12),2331–2358.

Tu, W., Zhu, T., Xia, J., Zhou, Y., Yani Lai, J. J., and Li, Q. (2019). Portraying the spatial dynamics of urban vibrancy using multisource urban big data. *Computers, Environment and Urban Systems*, 80,101428.

Wang, Y., Kang, C., Bettencourt, L. M., Liu, Y. and Andris, C. (2015). Linked activity spaces: embedding social networks in urban space. In: *Computational Approaches for Urban Environments* (eds. Helbich, M., Arsanjani, J.J. and Leitner, M.), 313–336. Cham: Springer.

Westerlund, J. and Wilhelmsson, F. (2011). Estimating the gravity model without gravity using panel data. *Applied Economics*, 43(6): 641–649.

Xu, Y., Shaw, S.-L., Zhao, Z., Yin, L., Fang, Z. and Li, Q. (2015). Understanding aggregate human mobility patterns using passive mobile phone location data: a home-based approach. *Transportation*, 42(4): 625–646.

Xu, Y., Shaw, S.-L., Zhao, Z., Yin, L., Lu, F., Chen, J., Fang, Z. and Li, Q. (2016). Another tale of two cities: understanding human activity space using actively tracked cellphone location data. *Annals of the American Association of Geographers*, 106(2): 489–502.

Yan, B., Janowicz, K., Mai, G. and Gao, S. (2017). From IDTL to Place2Vec: reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Redondo Beach, CA, USA. New York, NY, USA: Association for Computing Machinery.

Yan, B., Janowicz, K., Mai, G. and Zhu, R. (2018). xnet+sc: classifying places based on images by incorporating spatial contexts. *Leibniz International Proceedings in Informatics (LIPIcs). 10th International Conference on Geographic Information Science (GIScience 2018)*, Dagstuhl, Germany: Schloss Dagstuhl--Leibniz-Zentrum fuer Informatik.

Yang, W. and Mu, L. (2015). GIS analysis of depression among Twitter users. *Applied Geography*, 60: 217–223.

Yang, Y. and Zhang, H. (2019). Spatial-temporal forecasting of tourism demand. *Annals of Tourism Research*, 75: 106–119.

Yang, W., Mu, L. and Shen, Y. (2015). Effect of climate and seasonality on depressed mood among twitter users. *Applied Geography*, 63: 184–191.

Yang, J.-A., Tsou, M.-H., Janowicz, K., Clarke, K. C. and Jankowski, P. (2019). Reshaping the urban hierarchy: patterns of information diffusion on social media. *Geo-spatial Information Science*, 22(3): 149–165.

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J. and Mai, K. (2017). Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *International Journal of Geographical Information Science*, 31(4): 825–848.

Yao, Y., Liang, Z. T., Yuan, Z. H., Liu, P. H., Bie, Y. P., Zhang, J. B., Wang, R. Y., Wang, J. L. and Guan, Q. F. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science*, 33 (12): 2363–2384.

Ye, M., Janowicz, K., Mülligann, C. and Lee, W.-C. (2011). What you are is when you are: the temporal dimension of feature types in location-based social networks. *Proceedings of the 19th ACM SIGSPATIAL International Conference on*

Advances in Geographic Information Systems, Chicago, Illinois. New York, NY, USA: Association for Computing Machinery.

Ye, X., Huang, Q. and Li, W. (2016). Integrating big social data, computing and modeling for spatial social science. *Cartography and Geographic Information Science*, 43(5): 377–378.

Yin, J., Soliman, A., Yin, D. and Wang, S. (2017). Depicting urban boundaries from a mobility network of spatial interactions: a case study of Great Britain with geo-located Twitter data. *International Journal of Geographical Information Science*, 31(7): 1293–1313.

Yuan, Y. (2017). Exploring the spatial decay effect in mass media and location-based social media: a case study of China. In: *Advances in Geocomputation: Geocomputation 2015--The 13th International Conference* (eds. Griffith, D.A., Chun, Y. and Dean, D.J.), 133–142. Cham, Switzerland: Springer.

Yuan, Y., Raubal, M. and Liu, Y. (2012a). Correlating mobile phone usage and travel behavior--a case study of Harbin, China. *Computers, Environment and Urban Systems*, 36(2): 118–130.

Yuan, J., Zheng, Y. and Xie, X. (2012b). Discovering regions of different functions in a city using human mobility and POIs. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Beijing, China. New York, NY, USA: Association for Computing Machinery.

Yuan, Y., Liu, Y. and Wei, G. (2017). Exploring inter-country connection in mass media: a case study of China. *Computers, Environment and Urban Systems*, 62: 86–96.

Yuan, Y., Lu, Y., Chow, T. E., Ye, C., Alyaqout, A. and Liu, Y. (2020). The missing parts from social media–enabled smart cities: who, where, when, and what? *Annals of the American Association of Geographers*, 110(2): 462–475.

Yue, Y., Lan, T., Yeh, A. G. O. and Li, Q.-Q. (2014). Zooming into individuals to understand the collective: a review of trajectory-based travel behaviour studies. *Travel Behaviour and Society*, 1(2): 69–78.

Zhai, W., Bai, X., Shi, Y., Han, Y., Peng, Z.-R. and Gu, C. (2019). Beyond Word2vec: an approach for urban functional region extraction and identification by combining Place2vec and POIs. *Computers, Environment and Urban Systems*, 74: 1–12.

Zhang, K. S., Sun, D., Shen, S. W. and Zhu, Y. (2017). Analyzing spatiotemporal congestion pattern on urban roads based on taxi GPS data. *Journal of Transport and Land Use*, 10(1): 675–694.

Zhang, F., Hu, M., Che, W., Lin, H. and Fang, C. (2018a). Framework for virtual cognitive experiment in virtual geographic environments. *ISPRS International Journal of Geo-Information*, 7(1): 36–36.

Zhang, F., Zhang, D., Liu, Y. and Lin, H. (2018b). Representing place locales using scene elements. *Computers, Environment and Urban Systems*, 71: 153–164.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H. and Ratti, C. (2018c). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180: 148–160.

Zhang, F., Wu, L., Zhu, D. and Liu, Y. (2019a). Social sensing from street-level imagery: a case study in learning spatio-temporal urban mobility patterns. *ISPRS Journal of Photogrammetry and Remote Sensing*, 153: 48–58.

Zhang, F., Zhou, B., Ratti, C. and Liu, Y. (2019b). Discovering place-informative scenes and objects using social media photos. *Royal Society Open Science*, 6(3): 181375.

Zhang, Y., Li, Q., Tu, W., Mai, K., Yao, Y. and Chen, Y. (2019c). Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Computers, Environment and Urban Systems*, 78: 101374–101374.

Zhang, F., Zu, J., Hu, M., Zhu, D., Kang, Y., Gao, S., Zhang, Y. and Huang, Z. (2020). Uncovering inconspicuous places using social media check-ins and street view images. *Computers, Environment and Urban Systems*, 81: 101478.

Zhao, Z., Shaw, S.-L., Xu, Y., Lu, F., Chen, J. and Yin, L. (2016). Understanding the bias of call detail records in human mobility research. *International Journal of Geographical Information Science*, 30(9): 1738–1762.

Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J. (2017). Pyramid scene parsing network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Honolulu, United States. Columbus, United States: IEEE.

Zheng, S., Wang, J., Sun, C., Zhang, X. and Kahn, M. E. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nature Human Behaviour*, 3(3): 237–243.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A. and Torralba, A. (2017). Scene parsing through ade20k dataset. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Honolulu, United States. Columbus, United States: IEEE.

Zhu, D., Wang, N., Wu, L. and Liu, Y. (2017). Street as a big geo-data assembly and analysis unit in urban studies: a case study using Beijing taxi data. *Applied Geography*, 86: 152–164.

Zhu, D., Zhang, F., Wang, S., Wang, Y., Cheng, X., Huang, Z. and Liu, Y. (2020). Understanding place characteristics in geographic contexts through graph convolutional neural networks. *Annals of the American Association of Geographers*, 110(2): 408–420.

Zhuo, L., Shi, Q., Zhang, C., Li, Q. and Tao, H. (2019). Identifying building functions from the spatio-temporal population density and the interactions of people among buildings. *ISPRS International Journal of Geo-Information*, 8(6): 247–247.