Chapter 3
Cities as Spatial and Social Networks: Towards a Spatio-Socio-Semantic Analysis Framework

Wei Luo, Yaoli Wang, Xi Liu and Song Gao

3.1 Introduction

Cities have become highly interconnected techno-social systems embedded with intricate and complicated spatial and social networks (e.g., transportation, telecommunication, and internet). A variety of flows (e.g., daily commuting, information, and disease spread) are disseminated through these complex networks (e.g., road networks, human contact networks) within cities. Batty (2013) proposed to view cities as systems consisting of points, flows, and networks in order to understand the underlying structure of cities beneath the urban form. Traditionally, spatial analysis concentrates on the identification of spatial patterns and physical movement flows (e.g., disease, crime, and travel) from the perspective of geographical locations (Bailey and Gatrell 1995). However, physical movements are also driven by social network factors that have been overlooked (Andris et al.

W. Luo (✉)
School of Geographical Sciences & Urban Planning,
Arizona State University, Tempe, AZ, USA
e-mail: wluo23@asu.edu

Y. Wang
Department of Infrastructure Engineering, University of Melbourne,
Parkville, Australia
e-mail: yaoliw@student.unimelb.edu.au

X. Liu
Department of Geography, The Pennsylvania State University,
University Park, PA, USA
e-mail: xiliu@psu.edu

S. Gao
Department of Geography, University of Wisconsin–Madison,
Madison, WI, USA
e-mail: song.gao@wisc.edu

X. Ye and X. Liu (eds.), *Cities as Spatial and Social Networks*, Human Dynamics in Smart Cities, https://doi.org/10.1007/978-3-319-95351-9_3

2018). Social network drives human interactions that induce travel movement (Crandall et al. 2010; Yuan et al. 2012b; Bapierre et al. 2015; Wang et al. 2015). For example, Cho et al. (2011) found that human short-range travel, which is periodic both geographically and temporally, is not influenced by social network structure; while long-distance travel is more impacted by social network ties. Such patterns have called on the necessity of the marriage between geographical and social network space (Geo-social space), which contributes to further insights into the interactive evolution in both spaces (Batty 2003). Both Luo and MacEachren (2014) and Andris (2016) thoroughly discussed the intimate connections between the two spaces and described the relevant techniques that allow for a combined spatial analysis and social network analysis.

Geo-social space provides new opportunities to explore the complex interactions within cities, but such integration is still short of one important component—semantic space. Semantics study the logic aspects of meanings behind behaviours and phenomena, such as motivation, sense, reference, and implications. It contains at least two levels: *conceptual* and *formal*. Conceptual semantics focus more on the cognitive aspects while formal semantics lie in the logical reasoning perspectives. In urban studies, the essential role of semantics is to develop deeper insights into the motivations of spatial behaviors (Yan et al. 2011; Ying et al. 2011). For example, similar travel patterns and activity spaces can be driven by semantics (i.e., the reasons and motivations behind such activities). Though social network can reinforce the similarity of two or multiple individuals' travel behaviors, it should also include the travel aims (i.e., what you do rather than where you go) as well. Semantics also explains the functionality structure of cities, and why different parts of a city are connected. For example, two relatively far-away regions in a city may be closely connected due to commutes of certain groups of people. In public health, semantics explains how an infectious disease is transmitted because of certain types of disease (e.g., influenza) and human activities, so it determines human transmission network over geographical space. In social segregation, semantics explains why certain activities are to be conducted collectively instead of individually, so it also overcomes the coincidence by chance. Recent research has found that social network is associated with spatial overlap (Luo et al. 2011; Radil et al. 2013; Luo et al. 2014), but it is hard to conclude that spatial/social connections are always associated with semantic similarity. The ability to explain spatial-social behaviors makes semantics an irreplaceable complement to the geo-social systems. Semantics distinguishes network analysis from abstract network analysis by mathematical studies and network analysis embedded in urban context by geographical studies.

The interconnections of spatial, social, and semantic domains are not yet well recognized, let alone the deeper theoretical integration of geography, social network, and semantics. Knowledge into the integration of geographical, social, and semantic space is far less than sufficient due to the lack of recognition on its importance, restricted access to datasets, and insufficient methods of analysis. In order to fill the gap, this work proposes a novel theoretical framework for the integration of spatial, social and semantic spaces for the urban study based on the discussion of four diverse but inherently connected research domains: urban transportation, urban structure and land use, geo-social segregation, and infectious disease transmission and control. The framework aims to improve urban studies through the spatial-socio-semantic integration. We also summarize the available datasets and discuss their advantages and disadvantages in studying the interaction among geographical, social, and semantics spaces. We conclude the vision paper by discussing potential future research challenges for advancing the integration of spatio-socio-semantic spaces.

3.2 Literature Review

The urgency of the integration of spatial, social, and semantic aspects emerges from the observations and reflections in existing works from different application domains. We will discuss four particular domains in the context of urban studies in detail, i.e., transportation, spatial and social segregation, urban structure and land use, and air-borne infectious disease transmission and control. These four domains share a common characteristic that is human movement and interactions in which a combined perspective of spatial, social and semantic analysis will be introduced.

3.2.1 Transportation

The interaction between human travel and social network has been realized in recent years. Presumably, social network contacts are possible to indicate similarity in travel behaviors in addition to random spatio-temporal co-occurrence, because they may travel together for the same activities, or have similar daily routines. However, spatial movement, i.e., the trajectories in particular, is not equivalent to travel behaviors, of which the latter conveys more the semantics behind a spatial location, e.g., the function and service provided by a place (Adams and McKenzie 2013; Tuan 2013; Liu et al. 2015). For example, the sequence of activity types (work—grocery, shopping—home) to be conducted during a day may be similar while the spatio-temporal trajectories different, which highlights the importance of activity-based travel analysis (Bhat and Koppelman 1999; Wang et al. 2016). Also, people showing up at similar places or exactly the same place may conduct different activities. Many existing works on analyzing the *stops* of movements (e.g., Miklas et al. 2007; Eagle et al. 2009; Crandall et al. 2010; Bapierre et al. 2015; Hu et al. 2015; Toole et al. 2015; Wang et al. 2015) focus on the association between social network and location similarity rather than travel behavior similarity. The spatial similarity of social network contacts has also been considered in *trajectory*-mining to infer social connections by measuring geometrically the spatio-temporal overlapping between two trajectories (Li et al. 2008; Zheng et al. 2011). However, this is yet identical to travel behavior similarity between persons beyond simply spatio-temporal coincidence. Two overlapped trajectories are not guaranteed to infer social interaction to a great extent due to random coincidence, unless semantic information is infused to attest that people have similar travel aims or activities that might indicate shared travel activity.

Locations actually are not equally effective as the indicator for social network since some locations are overall more popular and one location can be multi-functional (Cranshaw et al. 2010; Pham et al. 2013). The challenges come from many technical difficulties of a comprehensive socio-semantic trajectory analysis. Parent et al. (2013) did a thorough survey on semantic trajectory analysis, specifying the importance of spatio-temporal semantics for travel behavior analysis for stops (e.g., Liao et al. 2005; Li et al. 2008; Xie et al. 2009), movements (e.g., Zheng et al. 2010; Xiao et al. 2015), individual and collective trajectories, and many methods. It is pointed out that no precise definition of trajectory behavior is reached. Trajectory behaviors can be an individual's selection of travel mode on different travel segments, the travel schedules/activities of a person, or a group of people's convergence to a same destination, bundled travel, to name a few. A thorough understanding of travel behavior depends on the (spatio-temporal) granularity and richness (how many details) of semantics collected on that geometric segment. Some trajectory data only contains geographic location and timestamp, while others contain travel mode, the number of co-travellers, and the activities at the destination. Moreover, there is still inadequate work to find the linkage between semantics of travel and stops. For example, the association remains unknown between the frequency of travel mode choice and certain travel aims.

3.2.2 Spatial and Social Segregation

Spatial and social networks are intertwined. Although some studies on the information age have argued for "the death of distance" in the cyberspace (Cairncross 2001), scholars still find that distance plays an important role in virtual space, e.g., human telecommunications and social media platforms on the Internet (Ratti et al. 2010; Han et al. 2015). Geo-social network analysis facilitates the understanding of social behaviors that relate to both the structure of the network and the relative location-context in physical space. For example, Radil et al. (2010) investigated the network structure and geographical context of gang violence in Los Angeles, showing the effectiveness of spatialized social network analysis that allows for real-time examination on social actors' positions in spatial and social networks simultaneously. The analysis identified gangs that are similarly embedded in the territorial geography and positioned in the rivalry network. Underlying spatial structures can be revealed by social network analysis as well. Thiemann et al. (2010) analyzed a human travel network represented by the circulation of banknotes, finding that the effective boundaries partially overlap with existing administrative borders and also physical barriers like rivers and mountains. Also, by applying a network-partitioning algorithm on a large telecommunication database in Great Britain, Ratti et al. (2010) also found geographically cohesive telecommunication regions that correspond remarkably well with administrative regions. Similarly, Gao et al. (2013b) employed a modularity-based network community detection method to successfully identify urban phone-call interaction patterns in both cyberspace and physical space; they also discovered anomalies relating to urban functional regions and human activities. Network community structure in the cyberspace can also be taken as proxies for the society. Walsh and Pozdnoukhov (2011) found a clear phone-call communication divide between the south and the north in Dublin, Ireland, where there is also a social divide between these two regions separated by a physical barrier—the River Liffey. In another study in Senegal, Gao et al. (2015) discovered that both digital divide and physical divide existed in a developing country via large-scale mobile phone data analysis. Those new datasets generated from emerging information communication technologies usually have high spatiotemporal resolution but lack rich individual semantics. Traditionally, social segregation patterns are revealed based on the analysis on demographic data, household income and transportation survey data. The integration of spatio-social-semantic spaces will provide a more holistic perspective for understanding geo-social segregation and environmental factors.

3.2.3 Urban Structure and Land Use

Although cities are usually planned and designed through a top-down approach, human movement is able to show how cities are actually used by their residents. Scholars tried to infer the function of regions from the complex travel flow system in cities since 1970s (Goddard 1970). Current literature mostly focus on revealing the sub-regional structure and/or polycentric structure based on various human-mobility related data (Roth et al. 2011; Cranshaw et al. 2012; Yuan et al. 2012a; Liu et al. 2015a; Sun and Axhausen 2016), as well as temporal changes of urban structure and the influence on people's travel behaviors (Zhong et al. 2014; Sun and Axhausen 2016).

Social networks reflect urban structures on a more 'intangible' perspective. Similar to spatial networks, we can also get a sub-regional structure of cities based on social networks. This structures emphasize more on social divide instead of physical divide (Gao et al. 2015). Social networks also affect the urban space by influencing the interactions among people and social communities. For example, cross community dyads may generate new travel flows to connect places that have few interactions (Andris 2016). However, spatial and social connections alone cannot tell us 'why' certain urban structures are formed. While spatial and social networks reflect urban structure and land use based on spatial/social flows and activity intensities of people, semantics provide explanations for these underlying patterns. By enriching the semantics of urban flows, we are able to have more precise understanding on urban structure and land use, which can aid the policy making process. For example, if we know that two areas in a city are connected by the commuting behaviors of home-work separated people, we would be able to design and plan related facilities in cities to assist or reduce their commuting time and cost. Some studies are trying to incorporate semantics into social and spatial analysis: they use point of interests (POI) data to infer trip purposes (Gong et al. 2016) or area functions (Yuan et al. 2012a). It is also possible to link topics mined from Twitter to trajectories by location and time (Kling and Pozdnoukhov 2012). The activity transitions between origin and destination of trips have also been used to improve land use inference precision (Liu et al. 2016).

3.2.4 Air-Borne Infectious Disease Transmission and Control

Cities play an important role in fostering and amplifying the transmission of air-borne diseases (e.g., influenza) because of dense human contacts within cities (Meade and Earickson 2005) and both intra-city and inter-city movements (Grenfell et al. 2001). Infectious disease transmission is a mutual interaction of social and spatial relationships among individuals and locations (Bian 2004). Spatial heterogeneity in the population distribution determines the spatial layout of disease transmission while the frequent city-wide travel of individuals contribute to the temporal sequence of transmission (Mao and Bian 2010b). A travel-based vaccination strategy (i.e., prioritizing the frequent city-wide travelers) is preferable for a population with a large number of intercommunity travelers in urban areas (Mao and Bian 2010a). Guo (2007) discovered sub-regional structure within cities based on human mobility data, which can provide valuable insight for designing effective pandemic control measures. Luo et al. (2018) further found that vaccination strategies considering the sub-regional structure within an urbanized area can lead to a significant reduction of epidemic size because they can prevent spread to other regions.

Infectious disease simulation models help understand disease transmission mechanism in both geographical and social spaces, whereas internet and social media platforms such as Twitter open the new way to monitor the disease outbreak. Google Flu Trends (GFT) rely on flu-related search terms from which the estimated prevalence is calculated based on both semantic data mining and computer modelling have exactly matched the surveillance data from the Centers for Disease Control and Prevention (CDC) (Butler 2013). Compared to CDC, GFT can span much larger population and deliver faster estimates, but significantly overestimated peak flu levels. Similar research have been conducted using tweets with a mention of flu indicators and found a high correlation with CDC data (Aramaki et al. 2011; Lampos and Cristianini 2010). Thus, such new flu-tracking techniques can complement the CDC data to help build regression models to predict influenza-like illness (ILI) activity data from "sentinel" medical practices collected by CDC (Achrekar et al. 2011). The geolocation of tweets and early prediction of influenza outbreak are most important aspects to reduce impacts and help authorities plan their response in time critical situations.

The above discussion of four domains demonstrates that though semantic analysis has the potential to enrich urban studies, most research only focus on geo-social perspective with little or implicitly semantic components. Semantics can reveal human purposes and behaviors from both cognitive and logical reasoning perspectives, whereas geo-social analysis focuses more on the observed human interaction and movement patterns in the urban context. Semantics would come underneath as the latent information to explain the observed patterns in both geographical and social spaces, but the integration of both geo-social and semantics analysis is still limited. Thus, in the following section, we propose a combined framework of spatial, social and semantic analysis that is essential to future urban studies.

3.3 Conceptual Framework

Here, we propose a spatio-socio-semantic conceptual analysis framework to understand city-related topics from a comprehensive perspective. The framework is focused on the 3-*S* (spatio-socio-semantic) aspects in a layered manner that marks their different roles in the system. In the spatio-socio-semantic analysis framework (Fig. 3.1), spatial and social networks are closely intertwined in city systems as geo-social systems (Luo and MacEachren 2014). While spatial behaviors are conventionally the direct observations of geography studies, social network has been gaining attention as it is both the driver for and the result of spatial behaviors. Social connections between people have strong influence on their decisions about when and where to travel and stay, which formulates the egg-and-chicken problem on the mutual interactions between spatial and social behaviors. People travel for social purposes when they build up new social links via spatial encounter. Thus, modeling people's behaviors in a connected network instead of isolated individuals is a necessity to study cities.

In addition to the duo of spatial and social aspects, semantics come underneath as the latent information to explain their behaviors in both geographical and social spaces. Semantics provide information, such as activity types of people, to enrich the geo-social models for spatial phenomena. Semantics infer the activity behind people's spatial choices and the functions of places, transform coordinates of trajectories/spatial flows into certain types of activities, and remark locations in space with meaningful labels of functions of cities. While semantics are usually implicitly recognized, we are proposing a way to explicitly express and analyze semantics in geographical studies. For example, Gao et al. (2013a) proposed novel analysis operators of place-based GIS joins according to the semantics in order to complement traditional geometric reference systems that include coordinates, distances, topology, and directions. Specifically, these analysis operations need to take geographical background information and/or human cognition and descriptions on places to decide which point should be joined to which region. Additionally, semantics extract the general behavioral patterns in space regardless of specific

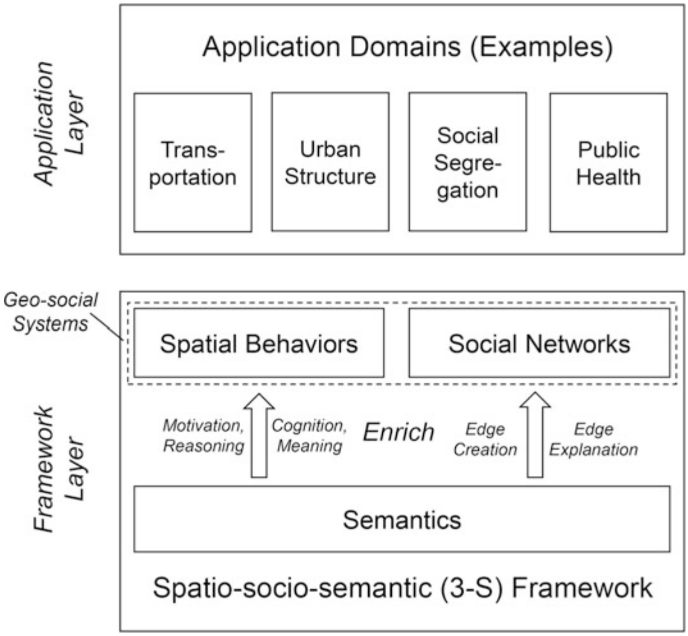

Fig. 3.1 The spatio-socio-semantics analysis framework. While spatial and social perspectives are closely intertwined as geo-social systems, semantics can enrich the systems by providing information to explain people's spatial behaviors and the aggregated patterns in cities

geographic coordinate locations. This is a recall of the importance of non-spatial attributes attached to spatial entities. For instance, the pattern of travel flow in the morning is usually inbound to city center for work, and outbound from city back home is widely observed in different cities. The semantic knowledge is not geographically constrained. Such non-spatial information helps to explain spatial behaviors from individual and collaborative (i.e., multiple persons') perspectives. The daily travel pattern induced by *activity* regularities is of a higher hierarchical level to the pattern detected from *location-visiting* regularities in space. The study of the activity regularities potentially contributes to activity-based travel analysis (e.g., Axhausen and Gärling 1992; Bhat and Koppelman 1999; Wang et al. 2016). The disease dispersion in space is likely to have similar spatial pattern regardless of particular geographical locations. For example, air-borne disease outbreak tends to spread from local growth to long distance transmission within cities (Mao and Bian 2010b) or from large 'hub' cities to smaller 'satellite' towns (Grenfell et al. 2001) because of human spatio-temporal hierarchical travelling patterns.

Rather than an independent methodology, the proposed framework works as an add-on to existing methods and models. Current methods, e.g., statistical analysis, data mining, simulations, and other computational models are feasible, but should bear the three factors of spatial, social, and semantic together. Spatial factor (including both geographical and virtual space) is quantified according to the specific study, e.g., range of an activity space, overlap of trajectories, the distribution and range of detected communities, and the affected area by a disease. Social network aspect is usually measured as the connection intensity between pairs of people, which can be represented by, but not limited to, cell phone call time duration, the number of calls or messages, and the number of shared friends. Semantic factor exists throughout any geographical phenomena, but was not explicitly analyzed in the past in geographical studies.

Semantic analysis in spatial studies is not restricted to such as text-mining or natural language processing, or semantic trajectory mining. It is more about understanding spatial processes from behavioral perspective rather than pure geometrical perspective. For example, in travel behavior analysis, knowing the type of activities can complement knowing just the locations of origin and destination; in urban structure and land use, knowing more about people's behaviors and ideas at certain places can enhance our recognition of place and space; in social and spatial segregation, knowing the type of flow (e.g., commuting flow or goods transport flow) and the underlying land use is better than simply the directions and geometric lengths of movements; in public health domain, knowing the type of disease as well as its transmission properties, and the type of activities that are more vulnerable to such transmission is better than only the spatial pattern of disease breakouts and dissemination.

Note that semantics also exist in social factors. The background information on the type of social link (e.g., friendship, phone call), the temporal stamps of link construction and maintenance, and other attributes all belong to semantics. A comprehensive framework of analysis is to bring generality into the duo-analysis of spatio-social interaction, the generality represented by semantics that seemingly diverse phenomena in particular scenarios may share common rules. The suggested way is analyzing spatio-social processes based on genres, e.g., genres of flows, of activities, of diseases, of land uses, of link types, of connection intensity. For example, infectious disease is transmitted from one person to another through direct or indirect contacts. The type of disease determines human contact network in which the type of human activities in terms of both geographical and social spaces are more vulnerable to such transmission. Chang et al. (2016) mapped HIV prevalence among people in agrarian, trading, and fishing communities in Rakai, Uganda, in which HIV prevalence shows substantial heterogeneities with the highest prevalence in fishing communities. In this way, findings in one geographical area are comparable and maybe transferrable to another area. Particular methods are contingent on specific fields.

3.4 Data Sources

The biggest challenge involved in urban related research from an integrative spatio-socio-semantics perspective would be the requirement of high-quality spatiotemporal datasets with semantics information. A combination of traditional data collection approach (e.g., households survey) with increasingly social-technological approach (e.g., location-based social media) to collect individual travel, social interaction behaviors, and semantic background information (e.g., tweet content) could be a potentially good source to tackle those challenges.

Current data collection approach to capture spatio-socio-semantic information relevant to urban studies include surveys, mobile devices, large scale human interaction simulation models, socio-technological networks (e.g., location based social media data), and sensor network (Table 3.1). Each of these approaches has advantages and disadvantages to represent geographical, social, and semantic aspects for different urban related research domains. Traditional survey approach has detailed semantic information (e.g., activities for travels) and social interactions relevant for disease transmission, but is often limited by an inherently low spatio-temporal resolution and a small number of surveyed individuals. The use of mobile devices can help to generate an ideal sample of social networks of the whole city based on people's reciprocal call patterns with high spatial resolution, but it does not provide any semantic information. POI provides information on the geographical location including details of semantic features (e.g., business, leisure). Associating POI data with individual trips generated from survey data or mobile devices, semantics of spatial flows (Alvares et al. 2007; Gong et al. 2016) and the nature of a place could be inferred (McKenzie and Janowicz 2015). However, the representativeness of such datasets is not guaranteed since it only shows activities in certain contexts. Daily activities that take a majority of time fall into the shortage of data.

While large scale human interaction simulation models with census data can provide human travel and interaction information for transportation and disease research, they cannot capture the spatio-temporal resolution for disease research and can only provide limited semantic information (e.g., demographic information) for both transportation and disease research. Socio-technological networks that can provide large and long-term datasets on social interactions and semantics information on each individual have been used to infer the function of that place and thus potential travel demand (e.g., Yang et al. 2014; McKenzie and Janowicz 2015), the latent events that triggered certain travels of people (Kling and Pozdnoukhov 2012; Coffey and Pozdnoukhov 2013), and human health behaviors in the virtual space (Lampos and Cristianini 2010; Aramaki et al. 2011). However, socio-technological networks suffer from the low temporal resolution issues for both disease transmission and transportation research, as well as low spatial resolution issues for disease transmission research. Socio-technological networks also require advanced text-mining to filter out space-related semantics (Imran et al. 2016a). Sensor network is the only approach which can capture perfect spatial resolution for human interaction that are transmitted through the close contact route, but cannot provide any semantic information for privacy concerns. There are many non-behavioral datasets, e.g., land use data, demographic dataset, which can be used to provide spatial or semantic context to complement the above datasets on human behaviors.

We can see the advantages and disadvantages among different datasets when studying urban research. An ongoing project Future Mobility Survey in Singapore

Table 3.1 Pros and cons in terms of spatio-temporal, social, and semantic aspects among six data collection methods.

Data collection	Spatial/temporal resolution	Social interaction size	Semantic richness	Example
Survey	Spatial: low temporal: low	Small	Limited	Questionnaire
Mobile device	Spatial: high temporal: high during call; low in break	Large	No	Phone call records
Socio-technological networks	Spatial: depends on domains temporal: low	Large	Enough	Twitter, Facebook check-in; POI database (e.g., Yelp)
Large-scale simulation models with census data	Spatial/temporal: depends on domains	Large	Limited	Simulation results
Sensor network	Spatial: high temporal: high	Small	No	Radiofrequency identification devices (RFID) (Cattuto et al. 2010)
Supplementary database	Spatial/temporal: depends on domains	No	Enough	Land use database; demographic census

Data resolution, size, and richness in terms of spatial, social network, and semantics are application dependent, so several descriptions (e.g., enough, limited) are ambiguous classification

can be very helpful to integrate spatio-socio-semantics aspects for transport modeling purpose. The Future Mobility Survey (FMS) is a smartphone-based prompted recall travel survey. It records the travel mode, the number of people who travel together, and activities at the destination of travel, which are all potential resources for inferring travel patterns rather than simply geographic locations. Travel patterns for different activities may be varied from solo travel to group travel. The background information with enriched semantics thus will be able to facilitate deeper insight into spatial movement flows.

3.5 Discussions and Challenges

The integrated spatio-social-semantics framework can help to explore several interesting research areas including but not limited to: (1) understanding geo-social dynamics in shaping neighborhoods with mixture of demographic characteristics (race, income, social status) and the impact of people's semantic activities;

(2) studying human mobility patterns at different spatial scales with the consideration of both geographical context and social contacts; (3) investigating the complex relationships among social segregation, urbanization, human mobility, and public health. Our analysis of the literature provides several potential research directions for further investigation of spatio-social-semantics framework. We conclude this review paper by highlighting several core challenges that will require interdisciplinary efforts to meet.

Addressing privacy issues when integrating semantic analysis into spatio-social analysis. Given that semantic analysis usually happens on a very detailed level, blending semantic analysis with spatio-social analysis puts the difficulty in collecting data and privacy issue more on the spot. Current social network information is collected either in a passive manner as by mobile phone calls and GPS trajectories, or in a voluntary manner as by location-based social networks (LBSN). The former usually yields higher spatio-temporal resolution but lacks semantic or social information, while the latter conveys ampler semantic and social information but is biased by selective occasions (e.g., people only post for certain purposes rather than continuously). Daily activities that take a majority of time fall into the shortage of data category. Dashdorj et al. (2013) enriched semantics of mobile phone dataset by integrating POIs drawn from open geographical data and given time to day. The low spatial resolution of cell phone data, nevertheless, is not resolved. That being said, it is capable of demonstrating intra-urban dynamics at an aggregated level rather than a detailed level. The voluntary data is contingent on the contributors' willingness and agreement, which brings up concerns of data quality as any type of volunteering geographic information (VGI) may do, as well as the urgency of improving privacy protection.

Addressing spatial information extraction issues from semantic information. Despite the difficulty in collecting semantics being acknowledged, attributes affiliated to spatial features such as keywords of Twitter posts, activity labels on trajectories, and texts on LSBN provide feasible foundation for the spatio-socio-semantic analysis. It, however, requires advanced text-mining to identify geographical locations (Imran et al. 2016a, b). The linguistic discernment, erroneous labelling, or inappropriate extraction of keywords lead to difficulty in location-oriented text-mining (Karimzadeh et al. 2013; Wallgrün et al. 2014). For example, though semantic data extraction from web and crowdsourced tracking systems provide a faster way to extract and monitor human health behaviors information (e.g., flu-related keywords) in the virtual space, it suffers from news bias during the excessive news period, especially after the early stage detection (Aramaki et al. 2011). Appropriate location-oriented text-mining approach to distinguish human health behaviors from news burst phenomenon is a key to provide near real-time feedback to adjust and validate disease simulation models.

Developing theory, methods, and tools to consider spatial, social network, and semantic factors simultaneously. Addressing the three aspects in the analysis at the same time is challenging. Much attention has been paid to the integration of either two components such as spatial-social (Mao and Bian 2010b; Wang et al. 2015; Luo 2016), spatial-semantic (Gao et al. 2013a; Wang et al. 2016), and social-semantic (Liu et al. 2015b). As discussed in the chapter, urban research requires the spatio-social-semantics integration. For example, both Guo (2007) and Luo et al. (2018) research shed light on the necessity and feasibility of designing sub-regional control strategies at the city scale. Current literature have already revealed various sub-regional structures according to different human mobility data with semantic information (Yuan et al. 2012a; Liu et al. 2016). We believe that integrating sub-regional structures determined by both geo-social interaction and semantic analysis into the control strategies would provide valuable insight for designing effective pandemic control measures.

To sum up, this chapter proposes a new spatio-social-semantics framework for urban relevant research. In addition to viewing cities as network systems consisting of points, flows, and interactions, this framework also emphasizes the semantic context behind individuals, locations, and connections. Spatial, social, and semantic spaces can complement each other. The increasing availability of various datasets capturing human interaction and movement and their spatial and semantic contexts make the urban research with the integrated spatio-social-semantics framework possible. Based on the data from the traditional and new sources, we are in the era of integrating heterogeneous data sources and creating innovative analytical approach to have a deeper and clearer understanding of our cities.

References

Achrekar, H., Gandhe, A., Lazarus, R., Yu, S.-H., & Liu, B. (2011) Predicting flu trends using twitter data. In *2011 IEEE Conference on Computer Communications Workshops* (*INFOCOM WKSHPS*) pp. 702–707.

Adams, B., & McKenzie, G. (2013) Inferring thematic places from spatially referenced natural language descriptions. In *Crowdsourcing geographic knowledge* (pp. 201–221). Dordrecht: Springer.

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo J. A. F., Moelans, B., & Vaisman, A. (2007) A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems*, p. 22. ACM.

Andris, C. (2016). Integrating social network data into GISystems. *International Journal of Geographical Information Science, 30*(10), 1–23.

Andris, C., Xi L., & Joseph F. Jr. (2018). Challenges for social flows. *Computers, Environment and Urban Systems, 70*, 197–207.

Aramaki, E., Maskawa, S., & Morita, M. (2011) Twitter catches the flu: Detecting influenza epidemics using Twitter. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp 1568–1576). Association for Computational Linguistics.

Axhausen, K. W., & Gärling, T. (1992). Activity-based approaches to travel analysis: Conceptual frameworks, models, and research problems. *Transport Reviews, 12*(4), 323–341.

Bailey, T. C., & Gatrell, A. C. (1995) Interactive spatial data analysis. Longman Scientific & Technical Essex.

Bapierre, H., Jesdabodi, C., & Groh, G. (2015) Mobile homophily and social location prediction. *arXiv preprint* arXiv:150607763.

Batty, M. (2003). Network geography: Relations, interactions, scaling and spatial processes in GIS. In D. Unwin (Ed.), *Re-presenting GIS* (pp. 149–170). UK: John Wiley.

Batty, M. (2013). *The new science of cities*. Cambridge, MA: MIT Press.
Bhat, C. R., & Koppelman, F. S. (1999) Activity-based modeling of travel demand. In *Handbook of transportation Science* (pp. 35–61). Boston, MA: Springer.
Bian, L. (2004). A conceptual framework for an individual-based spatially explicit epidemiological model. *Environment and Planning B, 31*(3), 381–396. https://doi.org/10.1068/b2833.
Butler, D. (2013). When Google got flu wrong. *Nature, 494*(7436), 155.
Cairncross, F. (2001). *The death of distance: How the communications revolution is changing our lives*. Cambridge, MA: Harvard Business Press.
Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.-F., & Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS ONE, 5*(7), e11596. https://doi.org/10.1371/journal.pone.0011596.
Chang, L. W., Grabowski, M. K., Ssekubugu, R., Nalugoda, F., Kigozi, G., Nantume, B., et al. (2016). Heterogeneity of the HIV epidemic in agrarian, trading, and fishing communities in Rakai, Uganda: An observational epidemiological study. *The Lancet HIV, 3*(8), e388–e396.
Cho, E., Myers, S. A., & Leskovec, J. (2011) Friendship and mobility: User movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1082–1090). New York, NY, USA: ACM. https://doi.org/10.1145/2020408.2020579.
Coffey, C., & Pozdnoukhov, A. (2013) Temporal decomposition and semantic enrichment of mobility flows. In *Proceedings of the 6th ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 34–43). ACM.
Crandall, D. J., Backstrom, L., Cosley, D., Suri, S., Huttenlocher, D., & Kleinberg, J. (2010). Inferring social ties from geographic coincidences. *Proceedings of the National Academy of Sciences, 107*(52), 22436–22441. https://doi.org/10.1073/pnas.1006155107.
Cranshaw, J., Schwartz, R, Hong, J. I., & Sadeh, N. (2012) The livehoods project: Utilizing social media to understand the dynamics of a city. In *International AAAI Conference on Weblogs and Social Media* p. 58.
Cranshaw, J., Toch, E., Hong, J., Kittur, A., Sadeh, N. (2010) Bridging the gap between physical location and online social networks. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (pp. 119–128). ACM.
Dashdorj, Z., Serafini, L., Antonelli, F., & Larcher, R. (2013) Semantic enrichment of mobile phone data records. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia* (p. 35). ACM.
Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences, 106*(36), 15274–15278.
Gao, S., Janowicz, K., MckKnzie, G., & Li, L. (2013a) Towards platial joins and buffers in place-based GIS. In *COMP@ SIGSPATIAL* (pp. 42–49).
Gao, S., Liu, Y., Wang, Y., & Ma, X. (2013b). Discovering spatial interaction communities from mobile phone data. *Transactions in GIS, 17*(3), 463–481.
Gao, S., Yan, B., Gong, L., Regalia, B., Ju, Y., & Hu, Y. (2015) Uncovering digital divide and physical divide in senegal using mobile phone data. In D. A. Griffith, Y. Chun, D. J. Dean (Eds.), Geocomputation 2015–The 13th International Conference.
Goddard, J. B. (1970) Functional regions within the city centre: A study by factor analysis of taxi flows in central London. *Transactions of the Institute of British Geographers* (pp. 161–182).
Gong, L., Liu, X., Wu, L., & Liu, Y. (2016). Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartography and Geographic Information Science, 43*(2), 103–114.
Grenfell, B., Bjørnstad, O., & Kappey, J. (2001). Travelling waves and spatial hierarchies in measles epidemics. *Nature, 414*(6865), 716–723.
Guo, D. (2007). Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographical Information Science, 21*(8), 859–877. https://doi.org/10.1080/13658810701349037.
Han, S. Y., Tsou, M.-H., & Clarke, K. C. (2015). Do global cities enable global views? Using Twitter to quantify the level of geographical awareness of US cities. *PLoS ONE, 10*(7), e0132464.

Hu, D., Huang, B., Tu, L., & Chen, S. (2015). Understanding social characteristic from spatial proximity in mobile social network. *International Journal of Computers Communications & Control, 10*(4), 539–550.

Imran, M., Mitra, P., & Castillo, C. (2016a) Twitter as a lifeline: Human-annotated Twitter corpora for NLP of crisis-related messages. arXiv preprint arXiv:160505894.

Imran, M., Mitra, P., & Srivastava, J. (2016b) Cross-language domain adaptation for classifying crisis-related short messages. arXiv preprint arXiv:160205388.

Karimzadeh, M., Huang, W., Banerjee, S., Wallgrün, J. O., Hardisty, F., Pezanowski, S., et al. (2013) GeoTxt: A web API to leverage place references in text. In *Proceedings of the 7th Workshop on Geographic Information Retrieval* (pp. 72–73). ACM.

Kling, F., Pozdnoukhov, A. (2012) When a city tells a story: Urban topic analysis. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems* (pp. 482–485). ACM.

Lampos, V., & Cristianini, N. (2010) Tracking the flu pandemic by monitoring the social web. In 2010 *2nd International Workshop on Cognitive Information Processing* (pp. 411–416). IEEE.

Li, Q., Zheng, Y., Xie, X., Chen, Y., & Liu, W., Ma, W.-Y. (2008) Mining user similarity based on location history. In *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems* (p. 34). ACM.

Liao, L., Fox, D., Kautz, H. (2005) Location-based activity recognition using relational Markov networks. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence* (pp 773–778). Edinburgh, Scotland.

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

Liu, X., Kang, C., Gong, L., & 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, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., & Shi, L. (2015b) Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, pp. 1–19. (ahead-of-print).

Luo, W. (2016). Visual analytics of geo-social interaction patterns for epidemic control. *International Journal of Health Geographics, 15*(1), 1–16. https://doi.org/10.1186/s12942-016-0059-3.

Luo, W., & MacEachren, A. M. (2014). Geo-social visual analytics. *Journal of spatial information science, 8,* 27–66. https://doi.org/10.5311/JOSIS.2014.8.139.

Luo, W., MacEachren, A. M., Yin, P., & Hardisty, F. (2011) Spatial-social network visualization for exploratory data analysis. In *SIGSPATIAL International Workshop on Location-Based Social Networks* (*LBSN 2011*) *in conjunction with the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (*GIS*) *2011*. Chicago, Illinois: ACM. https://doi.org/10.1145/2063212.2063216.

Luo, W., Yin, P., Di, Q., Hardisty, F., & MacEachren, A. M. (2014). A geovisual analytic approach to understanding Geo-Social relationships in the international trade network. *PLoS ONE, 9*(2), e88666.

Luo, W., Gao, P., & Cassels, S. (2018). A large-scale location-based social network to understanding the impact of human geo-social interaction patterns on vaccination strategies in an urbanized area. *Computers, Environment and Urban Systems Special Issue: Human Dynamics in the Mobile and Big Data Era: Towards Smart and Connected Communities.* https://doi.org/10.1016/j.compenvurbsys.2018.06.008.

Mao, L., & Bian, L. (2010a). A dynamic network with individual mobility for designing vaccination strategies. *Transactions in GIS, 14*(4), 533–545. https://doi.org/10.1111/j.1467-9671.2010.01201.x.

Mao, L., & Bian, L. (2010b). Spatial-temporal transmission of influenza and its health risks in an urbanized area. *Computers, Environment and Urban Systems, 34*(3), 204–215. https://doi.org/10.1016/j.compenvurbsys.2010.03.004.

McKenzie, G., & 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.

Meade, M., & Earickson, R. (2005) *Medical geography*. New York: The Guilford Press.

Miklas, A. G., Gollu, K. K., Chan, K. K., Saroiu, S., Gummadi, K. P., De Lara, E. (2007) Exploiting social interactions in mobile systems. In *International Conference on Ubiquitous Computing* (pp. 409–428). Berlin: Springer.

Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., et al. (2013). Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR), 45*(4), 42.

Pham, H., Shahabi, C., & Liu, Y. (2013) EBM: An entropy-based model to infer social strength from spatiotemporal data. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data* (pp. 265–276). ACM.

Radil, S. M., Flint, C., & Chi, S.-H. (2013). A relational geography of war: Actor–context interaction and the spread of World War I. *Annals of the Association of American Geographers, 103*(6), 1468–1484.

Radil, S., Flint, C., & Tita, G. (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. https://doi.org/10.1080/00045600903550428.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., et al. (2010). Redrawing the map of Great Britain from a network of human interactions. *PLoS ONE, 5*(12), e14248.

Roth, C., Kang, S. M., Batty, M., & Barthélemy, M. (2011). Structure of urban movements: polycentric activity and entangled hierarchical flows. *PLoS ONE, 6*(1), e15923.

Sun, L., & Axhausen, K. W. (2016). Understanding urban mobility patterns with a probabilistic tensor factorization framework. *Transportation Research Part B: Methodological, 91,* 511–524.

Thiemann, C., Theis, F., Grady, D., Brune, R., & Brockmann, D. (2010). The structure of borders in a small world. *PLoS ONE, 5*(11), e15422. https://doi.org/10.1371/journal.pone.0015422.

Toole, J. L., Herrera-Yaqüe, C., Schneider, C. M., & González, M. C. (2015). Coupling human mobility and social ties. *Journal of the Royal Society, Interface, 12*(105), 20141128.

Tuan, Y.-F. (2013) Topophilia: A study of environmental perceptions, attitudes, and values. New York: Columbia University Press.

Wallgrün, J. O, Hardisty, F., MacEachren, A. M, Karimzadeh, M., Ju, Y., & Pezanowski, S. (2014) Construction and first analysis of a corpus for the evaluation and training of microblog/twitter geoparsers. In *Proceedings of the 8th Workshop on Geographic Information Retrieval* (p. 4). ACM.

Walsh, F., & Pozdnoukhov, A. (2011) Spatial structure and dynamics of urban communities, pp. 1–8.

Wang, Y., Kang, C., Bettencourt, L. M., Liu, Y., & Andris, C. (2015) Linked activity spaces: Embedding social networks in urban space. In *Computational Approaches for Urban Environments* (pp. 313–336). Berlin: Springer.

Wang, Y., Kutadinata, R., & Winter, S. (2016) Activity-based ridesharing: Increasing flexibility by time geography. In *Proceedings of the 24th ACM SIGSPATIAL*. San Francisco Bay Area, CA: ACM.

Xiao, G., Juan, Z., & Zhang, C. (2015). Travel mode detection based on GPS track data and Bayesian networks. *Computers, Environment and Urban Systems, 54,* 14–22.

Xie, K., Deng, K., & Zhou, X. (2009) From trajectories to activities: A spatio-temporal join approach. In *Proceedings of the 2009 International Workshop on Location Based Social Networks* (pp. 25–32). ACM.

Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., & Aberer, K. (2011) SeMiTri: A framework for semantic annotation of heterogeneous trajectories. In *Proceedings of the 14th International Conference on Extending Database Technology* (pp. 259–270). ACM.

Yang, F., Jin, P. J., Wan, X., Li, R., & Ran, B. (2014) Dynamic origin-destination travel demand estimation using location based social networking data. In *Transportation Research Board 93rd Annual Meeting*.

Ying, J., J.-C., Lee, W.-C., & Weng, T.-C. (2011) Tseng VS Semantic trajectory mining for location prediction. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 34–43). ACM.

Yuan, J., Zheng, Y., & Xie, X. (2012a) Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 186–194). ACM.

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

Zheng, Y., Chen, Y., Li, Q., Xie, X., & Ma, W.-Y. (2010). Understanding transportation modes based on GPS data for web applications. *ACM Transactions on the Web (TWEB), 4*(1), 1.

Zheng, Y., Zhang, L., Ma, Z., Xie, X., & Ma, W.-Y. (2011). Recommending friends and locations based on individual location history. *ACM Transactions on the Web (TWEB), 5*(1), 5.

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science, 28*(11), 2178–2199.

Wei Luo [Ph.D., Pennsylvania State University], Postdoc at the Arizona State University. His research interests include geovisual analytics and high performance computing for large-scale social network analysis focusing on human dynamics and disease transmission and control (e.g., influenza, HIV), climate change impacts, international trade, social media, and criminal activity.

Yaoli Wang [Ph.D. candidate, University of Melbourne], Ph.D. candidate at the Department of Infrastructure Engineering, University of Melbourne. Her research interest includes smart mobility solutions, spatio-social behaviours, and complex network analysis. Her PhD topic is on ridesharing algorithms that integrate social networks and activity flexibility. She holds a BS in GIS from Peking University and a MS in Geography from the University of Georgia.

Xi Liu is a [Ph.D. candidate, Pennsylvania State University], Ph.D. candidate in Geography and Social Data Analytics at the Pennsylvania State University. He received his M.S. in Cartology and GIScience from Peking University, Beijing, China, and holds a B. Eng. in GIScience from Tongji University, Shanghai, China. With broad research interests in domains such as spatiotemporal data mining, spatial/social networks, and human mobility, he explores data-driven models of human activities to understand cities and solve urban issues. He currently works as a research assistant in the Friendly Cities Lab and is also a graduate affiliate of GeoVISTA Center at Penn State.

Song Gao [Ph.D., University of California, Santa Barbara], Assistant Professor at the Department of Geography, University of Wisconsin, Madison. His main research interests include Place-Based GIS, Spatiotemporal Big Data Analytics and Geospatial Semantics. He utilizes multi-data sources with spatio-temporal-network analysis methods to study human mobility patterns and urban spatial structure.