

## Exploring patterns of socio-spatial interaction in the public spaces of city through Big Data

**Aminreza Iranmanesh, Resmiye A. Atun**

Department of architecture, Eastern Mediterranean University, Famagusta, North Cyprus  
E-mail: aminreza.iranmanesh@emu.edu.tr, resmiye.alpar@emu.edu.tr

**Abstract.** *Research on the socio-spatial aspect of cities has never been so vibrant and exciting. The form of urban life is changing and evolving with new advancements in communication and technology. Digital communication and social media have reshaped the way people as the actors of society interact with each other and with the network of the city. New social networks and widespread of mobile devices can be used to create and reinforce existing social ties. Mobile devices also change the role of citizens from consumers into producers of data; they are the new reporters, photographers, videographers of everyday life. This production creates large quantities of data known as the “Big Data.” Big data has created many new opportunities for researchers to investigate new aspects of cities. This paper aims to explore how people access urban public spaces through social media by taking the parameter of distance and physical proximity into account. We tried to investigate if different levels of accessibility effects the way people interact with space through social media. Through this process, the study explored different socio-spatial patterns in the city that are being affected by social media. The data was collected in two layers in the city of Lefkoşa located in Northern Cyprus: first, the geo-tagged social media data was collected from the target group, and it was located on the map. Twitter as a microblogging medium was selected for data collection due to its public nature, geo-tagged abilities, and manageable short content. Second, degrees of accessibility in local and global scale was calculated using Space Syntax. The result shows the importance of local accessibility even when social media is the effective variable.*

*Keywords: Lefkoşa, Socio-spatial Interaction, Big data, mobile devices, Twitter, Urban Network, Accessibility*

### Introduction

Cities and their socio-spatial structure have become more complex due to new developments in digital communication and social media. Exploring these structure through the window of scientific and analytical techniques has become the norm. New data and high computational power enable the researchers to approach more complex phenomenon (Batty, 2012). Cities can be defined as the container of interactions, the spatial structure of communication and exchange, but the tangibility of these words

has been changing dramatically over the past few decades. Therefore the very definition of the city has become a new. The majority of interactions, communications, and exchanges that are taking place in urban space are invisible and not bound with proximal space (Willis, 2007). Human activities in public space or the production of urban space is a complex network of interaction. Anytime somebody mentions a space or engage in an activity inside an urban space a tie forms which is a socio-spatial interaction. The beauty of social media is that it enables us to explore

and read these interactions in the space as they happen. Furthermore, the formation of new interactions on social media might work as a new attraction that leads to more interactions. Therefore, we need new ways of understanding and reading urban space that includes this new phenomenon.

Accordingly, this research tries to contribute to the existing literature by exploring the possibilities of using social media as a research tool on top of spatial analysis to read the effects of distance free mobile communication on the socio-spatial form of urban life. The study uses geo-tagged Twitter data to read a layer of interaction with space and Space syntax to explore the accessibility to urban space in different distances.

### **The mobile-networked society**

Social media has affected human interactions in three main ways, offering more diversity and opportunities, increasing access to information, and immediate access to network through mobile devices that transforms peoples' way of interaction. This revolution has shifted the social structure from groups to network individuals (Rainie & Wellman, 2012). The rate of access to communication through various mediums is growing. Looking at this growing lifestyle together with the rising population of cities indicates the necessity of conducting research on the effects and consequences of this relatively new phenomenon in city life. The access to social media and free communication is not a luxury that a particular group of people could access; it becomes more available and accessible to more people around the world even in underdeveloped and developing countries (ITU, 2013).

The constant changes in means and way of communication in the city reidentify the social and socio-spatial ties. Consequently, people's relationships have become more individual and more connected at the same time. Digital communication has affected the interpretation of network in the contemporary society (Chambers, 2006). Accordingly, there is a critical need for more research on people's interaction with space and with each other

within the spatial content of the city. Castells (1989) argue that the changing dynamic of information technology reshapes the meaning of time and space. He describes this new hybrid space as the 'space of flow' which is a space between physical and digital space. This space makes the physical proximity a less critical player in the everyday life of the city. However, this analysis was done before the spread of mobile devices. Therefore, the dynamic of the socio-spatial form of urban life needs reevaluation.

### **Reading urban space through big data**

Studies on big data cover a wide variety of methods, techniques, approaches (Klauser & Albrecht, 2014), in this research, we primarily focus on peoples' interaction with space through social media. Mobile devices carry many embodied data collection tools such as cameras, global positioning system, and instant access to data sharing facilities. As Ang and Seng (2016) argue the recording capacity of these devices provide potential data collection which was impractical through traditional methods. The geo-tagged data, in particular, shows strong promises in providing additional layers to our approach in exploring urban space (Dunkel, 2015). The geo-tagged data creates a link between an interaction and physical space in real time. Arribas-Bel (2014) argues that the availability of these new data sources can be used as a complementary alternative to the census or survey data; this does not mean ignoring the existing data, but rather using the full potential of new sources for carrying out new research and strategies regarding urban space.

According to Boyd & Crawford (2012), Big Data affects the scholarly activities in its' technological, analytical, and methodological aspects. Rabari and Storper (2014) describe an emerging 'digital skin' which is being created by human interaction recorded through digital communication. The becoming of this skin opens up new questions and theoretical issues in urban policy making. Anderson (2004) argues that this new data can be more natural, accurate, and objective. As the quantity of data increases

rapidly, the results are becoming more accurate, or using his words: “With enough data, the numbers speak for themselves”. Availability and accessibility of digital communication transforms the role of citizens as producers and reproducers of content through mobile devices (Bontcheva & Rout, 2014). Therefore, this production is an inherited interactive layer of the contemporary city and not merely a tool for research. Thus, new methodological and theoretical frameworks must be explored to understand this data. Big data is beneficial for citizens as well as policymakers as it provides alternative visions and decision-making patterns in everyday life of the city (Kitchin, 2014). The big data can be used to increase people’s indirect participation in urban planning, but it requires numerous further studies as the field of study is still very young and emerging (Crampton et al., 2013; Hao, Zhu, & Zhong, 2015). The interactions through social media should not be merely considered as a way of reading the network but an inherited part of its reality. To understand the reality of urban space and how people interact with it the nature of big data which is produced by people should be carefully explored. Mobile devices with their internal global positioning system can tie any interaction to a specific location on earth which provides endless opportunities for researchers. Furthermore, the big data together with advanced computational power and statistical analysis could provide insight into urban studies.

### **Distance and accessibility in mobile networked society**

In the early frameworks of social network analysis distance is treated as a fundamental principle. However, digital communication has decreased the physical proximity as a necessity of forming interactions. The nature of digital communication is less bounded with distance and more with the structure of the information network (Castells, 1996). Being present in space is not essential for participating in socio-spatial dynamics of the contemporary city (Zachary, 2012), but not being necessary

does not imply the obsolescence of distance (see Batty, 2012). Cairncross (2001) argue that the wireless technology kills the location and puts it into our pockets. People are becoming more connected to each other and at the same time more individual centered. Yuan, Raubal, and Liu (2012) show a positive correlation between the usage of digital mobile networking and the distances people travel inside the city.

On the contrary, it could be argued that social media and mobile devices be tools that enhance people’s interaction with space. Digital communication has the potential to add an extra layer of interaction to the urban space, causing the rise of diversity (Kotzias, Lappas, & Gunopulos, 2014). It seems that there is a misconception that people’s online interactions are detached from their everyday life (Rainie & Wellman, 2012).

### **Case study, exploring the socio-spatial interactions in the city of Lefkoşa**

This study examines the concept of accessibility through social media in the North Nicosia (Lefkoşa) located in the Island of Cyprus. The city is physically divided into two parts as the result of 1974 conflict by the United Nations buffer zone (Figure 1). In this article, we focused on the northern part of the border. The data was collected in two parts: first, the social media data was gathered through Twitter public records, and second, the space syntax analysis was conducted in different levels of physical accessibility.

### **Twitter data**

Twitter was used as the data collection method for this research due to its short and manageable content; its geotag capability and its’ public availability for researchers. One of the challenges in using Twitter data is the process of filtering irrelevant entries (see Bollier & Firestone, 2010). For instance advertisements and people sending continues tweets from a single residential location. This research addressed this issue by selecting a particular target group. Throughout a process



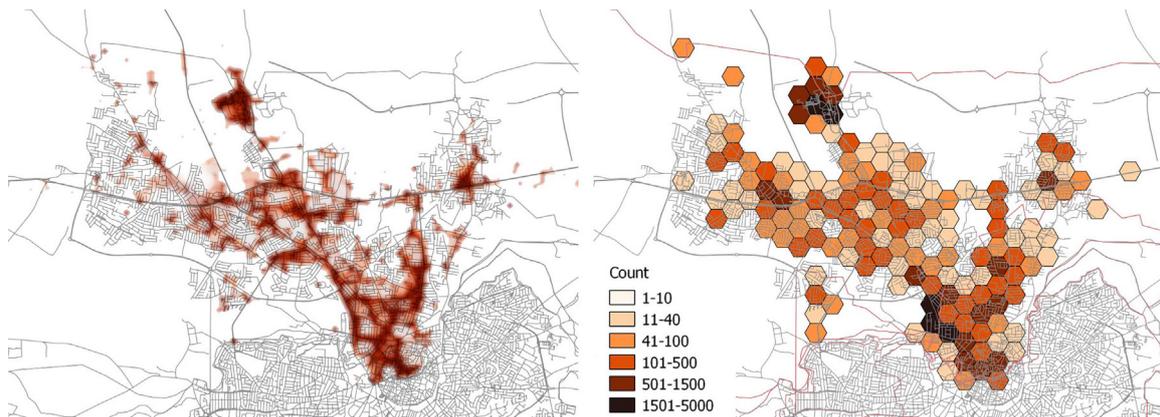
**Figure 1. The city of Nicosia and UN buffer zone (by author).**

of screening the interactions in public space, 1189 users have been chosen. The selection was based on the regularity of their activity and being engaged with urban space. Accordingly, the Twitter record of the target group in a one year span was collected through NodeXL platform. More than 1.2 million were collected, out of this entries, 53430 were geo-tagged and within the borders of the case study. Figure 2 (left) shows the heat-map of the Twitter data. The university campus on the north, walled city on the south and the vibrant urban spaces on the west side are visible just by looking at the raw data. The data was further categorized into a hexagonal grid (see Shelton, Poorthuis, & Zook, 2015) with a diameter of 500m to facilitate an easier way of making sense of the data and overlapping it with space syntax outcome (Figure 2 right). There are 270 cells with significant data. The same grid is used to measure accessibility through space syntax.

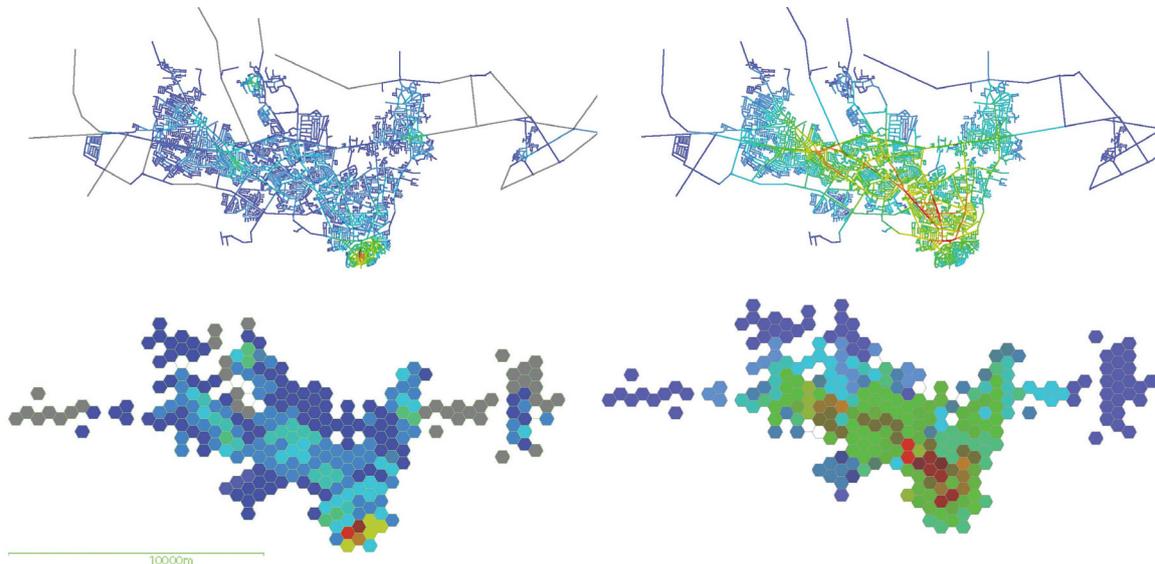
### **The space syntax**

Space syntax uses mathematical properties of the spatial network to estimate the movement is a network of spaces. It can be used for understanding pedestrian and vehicular movements in the city (Hillier, 1996; Hillier & Hanson, 1984). Space syntax measures a variety of variables for the urban network. To address the primary objective of this research, we used To-movement (integration) and Through-movement (choice) in ten radii from local to global. 3000 meters was considered a global accessibility radius and 600 meters as local (Bill Hillier & Ing, 2012).

To-movement measures destination. It evaluates the accessibility of each part of the network from all other parts in a given physical distance. This measurement helps the research to understand which parts of the spatial network have more potential of being a destination for



**Figure 2. Left: The heat map of collected Twitter data, right, counting the twitter data in the hexagonal grid**



**Figure 3. Left: To-movement in local level (600m), Right: To-movement in global level (3000m)**

public activities. Through-movement (choice or betweenness) explores the space as a passage, a move through space, or in-between spaces; It can be used to analyze the possibility of people choosing a particular path to reach all possible destinations in a given distance. The space syntax technique provides flexibility in measuring various levels of distance for each variable

The Space syntax analysis was conducted using UCL DepthmapX. Accordingly, the Axial map was prepared N=7266; then it was broken down into segment map. Both measurements were performed with 200 meters of interval, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, and 3000 as the suggested global reach. The data was put into the hexagonal grid, the

average value of all lines inside a cell was calculated. Figure 3 shows the To-movement value and its average in the hexagonal grid in 600 meters (Local) and 3000 meters (Global). The raw data shows the strong local accessibility in the historic core of the city close to the border and high global accessibility in contemporary developments of the city.

### Data Analysis

The final table consists of 270 cells with three main variables, twitter count, average to-movement, and average through-movement. Pearson correlation was performed to explore the correlation between variables.

The first stage of analysis looks into the relationships between Twitter data and integration. The result shows statistical significance at all levels of integration, but it demonstrates stronger correlations with local measures (Table 1). The correlation between Twitter data and integration reaches its pick at 600m. The result demonstrates the importance of local access to destination spaces in the urban spatial network. It could be argued that the local accessibility plays a more significant role in the formation of socio-spatial interactions within the urban public space.

The second stage of analysis looks into the correlations between Twitter data and Choice (Table 2). The data starts to show statistical significance in 1400m and substantial significant ( $p < 0.01$ ) at 1000 meters. It seems that the local measurements are more dominant

in Through-movement analysis as it does not show any correlations with global measures. It appears that even in in-between spaces which are facilitating peoples' movement to their destination higher local measurements is correlated with higher probability of socio-spatial interaction.

The study used the regression analysis to detect the outliers in the data sets. Unlike the usual approach of identifying and removing the outliers, we considered them a valuable source of information. These are spaces that are not fitting within the expected statistical patterns socio-spatial interactions. The cumulative distribution function of chi-square for Mahalanobis distance with 10 degrees of freedom on the integration variable was conducted to find the outliers. Figure 4 shows the detected outlier cells on the map.

Table 1. Correlations between Twitter data and different levels of To-movement (integration)

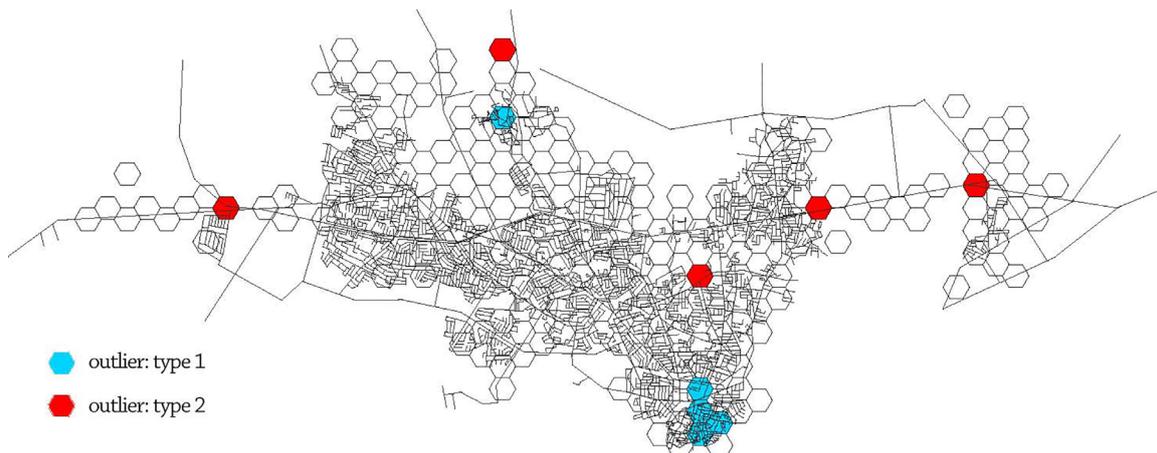
		TWITTER DA	TI024 Integration R3000 metric	TI024 Integration R2000 metric	TI024 Integration R1800 metric	TI024 Integration R1600 metric	TI024 Integration R1400 metric	TI024 Integration R1200 metric	TI024 Integration R1000 metric	TI024 Integration R800 metric	TI024 Integration R600 metric	TI024 Integration R400 metric
TWITTER DA	Pearson Correlation	1	.146*	.170**	.176**	.185**	.194**	.207**	.224**	.236**	.226**	.191**
	Sig. (2-tailed)		.021	.007	.005	.004	.002	.001	.000	.000	.000	.002
	N	270	248	248	248	248	248	248	248	248	248	248

\*. Correlation is significant at the 0.05 level (2-tailed).  
 \*\*. Correlation is significant at the 0.01 level (2-tailed).

Table 2. Correlations between Twitter data and different levels of Through-movement (choice)

		TWITTER DA	TI024 Choice R3000 metric	TI024 Choice R2000 metric	TI024 Choice R1800 metric	TI024 Choice R1600 metric	TI024 Choice R1400 metric	TI024 Choice R1200 metric	TI024 Choice R1000 metric	TI024 Choice R800 metric	TI024 Choice R600 metric	TI024 Choice R400 metric
TWITTER DA	Pearson Correlation	1	.084	.101	.108	.118	.133*	.152*	.168**	.180**	.182**	.183**
	Sig. (2-tailed)		.187	.114	.089	.063	.037	.017	.008	.004	.004	.004
	N	270	248	248	248	248	248	248	248	248	248	248

\*. Correlation is significant at the 0.05 level (2-tailed).  
 \*\*. Correlation is significant at the 0.01 level (2-tailed).



**Figure 5. The detected outliers**

Two main types of outliers could be described in this case: type 1 represents parts of the spatial network that show high local integration levels, shallow global integration, and extremely high density through social media. In this case, the historic core of the city and the university campus are fitting this behavioral pattern. These two spaces are different in nature, but both have their internal life within the wider network of the city, and they offer high-quality public spaces for interaction. Type 2 represents spaces with low local and global integration and somewhat high visibility through social media. These parts of the network are the activity driven spaces. Interesting functionalities that attract people regardless of travel distance in . Therefore, in addition to our previous findings, it could be argued that accessibility and distance can be both critically important and obsolete in some cases.

## Conclusion

Digital communication and social media have changed the way people interact with the city and with each other. Widespread and availability of mobile devices have brought the new type of interaction into the spatial network of the city in real time. The side product of this transformation is 'big data,' which is being recorded by the people. 'Big data' opens up new horizons for researchers, and it provides unprecedented data collection

and evaluation techniques in comparison to traditional methods. The field of 'big data' is still in its infancy. Theoretical, technological, and methodological frameworks are being created and tested every day. The 'big data' in its raw form cannot be self explanatory; it must be supported and evaluated together with other types of data. In this research, we tried to develop a methodological framework to read urban space through Twitter data (big data) and space syntax. The research objective was to explore the relationship between accessibility and the formation of socio-spatial ties through social media. The Twitter data has been monitored and processed into a fine-grained collection of socio-spatial interactions and free of spams and unwanted tweets. On the other hand, accessibility was measured using space syntax. To-movement (integration) and through-movement (choice) were measured to explore the urban space as both destinations and move through parts of the network.

The findings show the critical importance of both global and local accessibility in the formation of socio-spatial interactions. To-movement seems to be more correlating with Twitter data at the local level. Therefore, it could be argued that spaces that are more locally accessible as urban destinations have the potential to contain more interactions within. Through-movement shows significant only at local level. It could be assumed that when people are moving from one point to another if pedestrian (local), they are more likely to interact with space. On the other hand,

global movements in the network are probably vehicular movement. Thus there are fewer chances for socio-spatial interactions. This paper argues that when the data is properly collected and sorted, exploring the outliers play a critical role in undersetting socio-spatial interactions. In this case, we found two types of spaces that are not fitting into the statistical analysis. First, spaces with high local and low global accessibility that are heavily reported through social media, and second, spaces with attractive functions which are low in both local and global accessibility of spatial network.

There are a variety of opportunities for research in the urban socio-spatial network through the window of social media. Perhaps in further studies, including more variables like land use, economic drivers, or domestic destinations might shine more light on the subject.

## References

- Anderson, C. (2004). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Geografiska Annaler: Series B, Human Geography*, 86(1), 33-34.
- Ang, L.-M., & Seng, K. P. (2016). Big Sensor Data Applications in Urban Environments. *Big Data Research*.
- Arribas-Bel, D. (2014). Accidental, open and everywhere: Emerging data sources for the understanding of cities. *Applied Geography*, 49, 45-53. doi:10.1016/j.apgeog.2013.09.012
- Batty, M. (2012). Building a science of cities. *Cities*, 29, S9-S16.
- Bill Hillier, M., & Ing, C. S. D. (2012). Compositional and urban form effects on centres in Greater London.
- Bollier, D., & Firestone, C. M. (2010). The promise and peril of big data: Aspen Institute, Communications and Society Program Washington, DC.
- Bontcheva, K., & Rout, D. (2014). Making sense of social media streams through semantics: a survey. *Semantic Web*, 5(5), 373-403.
- Cairncross, F. (2001). The Death of Distance: How the Communications Revolution Is Changing Our Lives. Harvard Business School Press.
- Castells, M. (1989). The informational city: Information technology, economic restructuring, and the urban-regional process: Blackwell Oxford.
- Castells, M. (1996). The Information Age: Economy, Society and Culture Volume 1: The Rise of the Network Society. 2nd ed. Oxford: Wiley Blackwell.
- Chambers, D. (2006). *New Social Ties, Contemporary Connections in a Fragmented Society*. London: PALGRAVE MACMILLAN.
- Crampton, J. W., Graham, M., Poorthuis, A., Shelton, T., Stephens, M., Wilson, M. W., & Zook, M. (2013). Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb. *Cartography and Geographic Information Science*, 40(2), 130-139. doi:10.1080/15230406.2013.777137
- Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, 142, 173-186.
- Hao, J., Zhu, J., & Zhong, R. (2015). The rise of big data on urban studies and planning practices in China: Review and open research issues. *Journal of Urban Management*, 4(2), 92-124.
- Hillier, B. (1996). *Space is the Machine: A Configurational Theory of Architecture*. Cambridge: Press Syndicate.
- Hillier, B., & Hanson, J. (1984). *The Social Logic of Space*. Cambridge Cambridge University Press.
- ITU. (2013). World in 2013, ICT facts and figures Retrieved from Geneva, Switzerland:
- Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1-14.
- Klauser, F. R., & Albrechtslund, A. (2014). From self-tracking to smart urban infrastructures: Towards an interdisciplinary research agenda on Big Data. *Surveillance & Society*, 12(2), 273.
- Kotzias, D., Lappas, T., & Gunopulos, D. (2014). Addressing the Sparsity of Location Information on Twitter. Paper presented at the EDBT/ICDT Workshops.

- Rabari, C., & Storper, M. (2014). The digital skin of cities: urban theory and research in the age of the sensed and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions, Economy and Society*, rsu021.
- Rainie, L., & Wellman, B. (2012). *Networked : the new social operating system*. Cambridge: MIT Press.
- Shelton, T., Poorthuis, A., & Zook, M. (2015). Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information. *Landscape and Urban Planning*, 142, 198-211.
- Willis, K. S. (2007). Sensing Place: mobile and wireless technologies in urban space. *Encountering Urban Places: Visual and Material Performances in the City*, 155-169.
- Yuan, Y., Raubal, M., & Liu, Y. (2012). Correlating mobile phone usage and travel behavior—A case study of Harbin, China. *Computers, Environment and Urban Systems*, 36(2), 118-130.
- Zachary, N. P. (2012). *The Connected City: How Networks are Shaping the Modern Metropolis*. Oxon: Routledge.