Transactions in GIS

Research Article Transactions in GIS, 2014, 18(S1): 126–142

Uncovering the Directional Heterogeneity of an Aggregated Mobile Phone Network

Tao Pei,*,‡ Stanislav Sobolevsky,† Carlo Ratti,† Alexander Amini† and Chenghu Zhou*

*State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences
†MIT Senseable City Lab
‡Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application

Abstract
The aggregated mobile phone network (AMPN) (i.e. the calling time or numbers are aggregated at every vertex), which records the call volume between different places over time, has been studied extensively to reveal the mobility patterns of residents, etc. Nevertheless, most previous works were implemented based on the non-directionality of the network model. This simplification may overlook some important characteristics of AMPN. To explore the AMPN as a directional network model, we introduce the concept of directional heterogeneity in the study of AMPN data. The heterogeneity is twofold: (1) the imbalance of vertex (difference between outgoing and incoming calls of the vertex); and (2) the reciprocity of each edge (difference between the directed weights of the same edge). Taking the data of Singapore as an example, we systematically analyze the directional heterogeneity of AMPN. Our findings include three aspects. First, the AMPN shows as more unbalanced in the night-time than in the daytime, and its imbalance decreases as vertex granularity increases. Second, the directional heterogeneity varied with locations. Specifically, the residential area is dominated by deficits and others by surpluses. Third, the trajectories of incoming and outgoing calls follow a similar geographical pattern (i.e. southeast-north-south-north-southeast), indicating the calling behavior and routine mobility of users over time and space.

1 Introduction

Aggregated mobile phone networks (AMPNs) (i.e. the calling time or the number are aggregated at every vertex), which are seen as the outcome of the communication between different places over time, have been extensively studied to reveal the underlying patterns of human mobility and interactions on difference scales (Ratti et al. 2006; Pulselli et al. 2006; Jacobs-Crisioni and Koomen 2012; Loibl and Peters-Anders 2012; Yu and Pei 2013; Amini et al. 2014). Nevertheless, most previous studies relating to the AMPN are based on the non-directional or dyadic network model. That is, in the model, edges between vertexes are considered as non-directional, or the interaction between vertexes is viewed as binary. In fact, the

Address for correspondence: Tao Pei, State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A Datun Road, Beijing, China 100101. E-mail: peit@lreis.ac.cn
Acknowledgments: The authors wish to thank SingTel for providing the dataset. This study was funded through support from the National Natural Science Foundation of China (Project Number: 41171345 and 41231171), a grant from the National Key Technology R&D Program of China (No. 2012AA12A403). We further thank Ericsson, the MIT SMART Program, the Center for Complex Engineering Systems (CCES) at KACST and MIT CCES program, the National Science Foundation, the MIT Portugal Program, the AT&T Foundation, Volkswagen ERL, BBVA, The Coca Cola Company, Expo 2015, Ferrovial, the Regional Municipality of Wood Buffalo and all the members of the MIT Senseable City Lab Consortium for supporting the research.

 doi: 10.1111/tgis.12128

AMPN is a directional network model, and the interaction between vertexes is not only directional but also unbalanced. The simplification in previous works may cancel the directional interactions and overlook the flows (say, the flows of information and people) between vertexes, which can be indicated by the directional heterogeneity of AMPN and seen as a very useful tool for observing the city (Batty 2013). To explore the AMPN as a directional network model, we introduce the concept of directional heterogeneity in the study of AMPN data. Here, the directional heterogeneity includes two aspects.

1. One is the imbalance of a vertex, which is defined as the difference between incoming and outgoing calls from the said vertex. For example, if a Base Transceiver Station (BTS, a piece of equipment that facilitates wireless communication between user equipment and a network, so the number, or the time, of calls within its service area can be obtained) has more outgoing calls than incoming calls, it generates surplus; if, on the other hand, it has more incoming calls than outgoing calls, it generates a deficit.
2. The other is the reciprocity of an edge, which measures the difference between the reciprocated weights of the same edge. The analysis of the directional heterogeneity may help in better understanding the moving urban landscape and dynamics of a city. Research on this topic, although important, is not sufficient. To the best of our knowledge, our article is the first approach to the directional heterogeneity of the AMPN.

The reciprocity and imbalance of binary networks have been explored in different disciplines to reveal patterns of growth of networks, understand dynamic processes (e.g. diffusion or percolation processes), and describe the onset of higher-order structures (e.g. triadic motifs) (Meyers et al. 2006; Zamora-Lopez et al. 2008; Perra et al. 2009; Zlatic and Stefancic 2009; Squartini et al. 2013). Nevertheless, few studies have been conducted in describing the directional heterogeneity of AMPN. Below is a brief review of the literature of AMPN.

Studies referring to AMPN can be grouped into two categories. The first is based on the calling data aggregated at the level of BTSs. Since the calling activities in a certain period are highly related to the number of users, they were used to estimate population in cities (Vieira et al. 2010; Manfredini et al. 2011; Rubioa et al. 2013). The spatiotemporal variation of calling activities was also employed to describe the urban landscape, essentially, the residents' mobility pattern (Sevtsuk and Ratti 2010; Sun et al. 2011; Grauwin et al. 2014; Kung et al. 2014). Apart from this, by combining POIs, specific social groups can be distinguished from the aggregated mobile phone data (Vaccari et al. 2009), and specific social events can be identified by calculating the temporal deviation from the average calling activity at a place (Traag et al. 2011; Laura et al. 2014). In addition to retrieving human mobility, the time sequences aggregated hourly at the level of BTS were also analyzed to classify the urban land use types (Soto and Frias-Martinez 2011; Toole et al. 2012; Pei et al. 2014).

The second is focused on the edges between vertexes, which are further divided into two aspects: (1) analyzing the routine mobility patterns derived from individual trajectories, which are constructed from the mobile phone usages transferred between antennas (Gonzalez et al. 2008; Song et al. 2010; Calabrese et al. 2011); and (2) utilizing the information about the connections between vertexes (either through communication or human mobility) to find the community structure in a network. Here are some representative approaches. Ratti et al. (2010) delineated the community structure based on the telecommunications network of the UK and found that regions generated by the community detection method correspond remarkably well with administrative regions. Based on the analysis of mobile phone data from countries of various scales across Europe, Asia, and Africa, Sobolevsky et al. (2013) found that the cohesiveness and matching of official regions still exists at the country scale; moreover, this phenomenon can also be observed on a finer level. Kang et al. (2013) compared community structure generated by taxicab trips and that by mobile phone data, and revealed that mobile phone data produces more fine-grained spatially cohesive communities than those from taxicab trips.

Although significant progress has been made regarding the AMPN, most of them were limited to the non-directional or dyadic network model. The directional heterogeneity of the AMPN, which may indicate the urban flows and the relationship between each other, still does not attract enough attention. In order to enhance the research on the directional heterogeneity of AMBN and apply it to urban study, based on the network theory, we use several indices to measure the directional heterogeneity of AMPN. Then, taking the mobile data of Singapore as an example, we systematically analyze the imbalance and reciprocity of AMBN in terms of time and space.

The remainder of the article is structured as follows. Section 2 introduces some basic concepts for measuring the imbalance of vertexes and the reciprocity of edges in an AMPN. In Section 3, the mobile phone data used in this article are described. Section 4 shows the temporal characteristics of directional heterogeneity in the AMPN. Section 5 describes the spatial characteristics of the heterogeneity. Section 6 analyzes the micro spatiotemporal heterogeneity of the AMPN from the trajectories of the centers of calls. Section 7 summarizes the characteristics of the AMPN and presents our plans for future work.

2 Some Concepts of Imbalance and Reciprocity

Before the directional heterogeneity of an AMPN is analyzed, some basic concepts should be defined. The directional heterogeneity of an AMPN is twofold. The first is the imbalance of a vertex; the other is the reciprocity between vertexes. Next, we introduce six concepts, which were proposed by Squartini et al. (2013).

Definition 1: The reciprocated weight between i and j is defined as $\overleftrightarrow{w}_{ij} = \min(w_{ij}, w_{ji})$, where w_{ij} is the weight from vertex i to j, and w_{ji} is the weight from j to i.

Definition 2: The non-reciprocated weights from i to j is defined as $\Delta w_{ij}(i \neq j)$, which can be calculated as $w_{ij} - w_{ji}$. The normalized non-reciprocated weight from vertex i to j can be calculated by $\dfrac{\Delta \mathrm{w}_{ij}}{\sum_{j,i\neq j} w_{ij} + \sum_{j,j\neq i} w_{ji}}$.

Definition 3: The imbalance of vertex i can be measured by $s_i^{out} - s_i^{in}$. $s_i^{out} = \sum_j w_{ij}$ is the outflow and $s_i^{in} = \sum_j w_{ji}$ is the inflow (Koylu and Guo 2014). The outflow and the inflow are also referred to as the out-strength and in-strength, respectively in the network literature (Squartini et al. 2013). If the imbalance is positive, then it is called surplus; otherwise, deficit. The normalized imbalance of vertex i can be measured by $\dfrac{s_i^{out} - s_i^{in}}{s_i^{out} + s_i^{in}}$.

Definition 4: The imbalance of a network can be measured by $B = \dfrac{\Delta S}{S}$, where $\Delta S = \sum_i |s_i^{out} - s_i^{in}|$ and $S = \sum_i (s_i^{in} + s_i^{out})$. ΔS is called the sum of non-reciprocated strengths. Note that value of B is between 0 and 1. If B is higher, the network is more unbalanced; otherwise, less unbalanced.

Definition 5: The non-reciprocity of vertex i can be measured by $\sum_{j,i\neq j}|w_{ij}-w_{ji}|$. The normalized non-reciprocity of vertex i can be measured by $r_i = \frac{\sum_{j,i\neq j}|w_{ij}-w_{ji}|}{\sum_{j,i\neq j}w_{ij}+\sum_{j,j\neq i}w_{ji}}$.

Definition 6: The reciprocity of a network can be measured by $R = \frac{\vec{W}}{W}$, where $W = \sum_i \sum_{j\neq i} w_{ij}$ and $\vec{W} = \sum_i \sum_{j\neq i} \vec{w}_{ij}$. $\vec{W}$ is the sum of reciprocated weights. Note that values of R are between 0 and 1. If R is higher, the network is more reciprocated; otherwise, less reciprocated.

3 Mobile Phone Data

The mobile phone data we used in this article are the hourly aggregated number of calls between BTSs in Singapore. There are over 5,000 BTSs in Singapore and the average distance between two neighboring BTSs is less than 500 m over the entire city. To study the directional heterogeneity of the AMPN of Singapore, we collected a week of mobile phone data (i.e. from 28th March (Monday) to 3rd April (Sunday) of 2011). In order to reveal the relationship between the network heterogeneity and different land use types, we also collected the urban plan map of the city and combined the land use types to form the ultimate map. In this article, the city is divided into five land use types: Residential, Business, Commercial, Open space, and Other (including school, hospital, government office or civic and community institution) (Figure 1).

The hourly number of outgoing calls is normalized by the total number of outgoing calls for the whole week. Figure 2 displays the normalized number of outgoing calls for the whole week. From the figure, we find that the number of calls shows a very good periodic pattern. Specifically, the number of calls is very low in the early mornings and increases in the daytime. The curve reaches its maximum at 6 p.m. and then decreases after that time. Interestingly, Monday to Friday generate a high peak at 6 p.m. whereas the weekends exhibit a more flat platform in the same period.

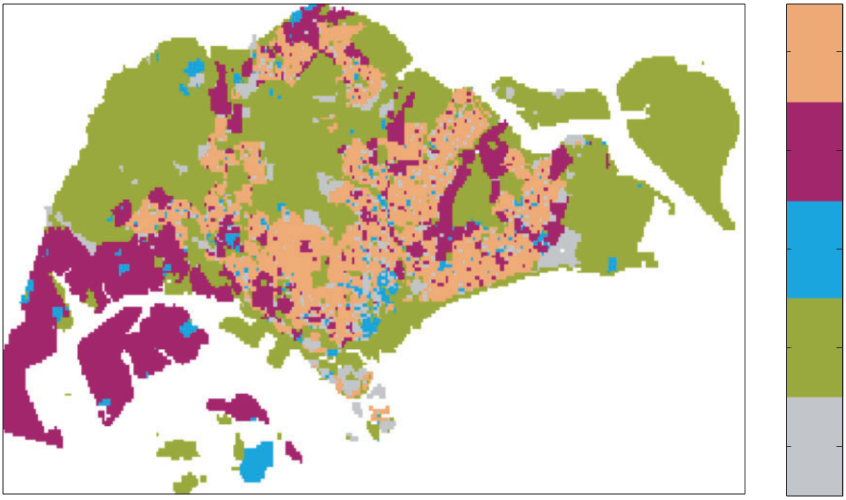


Figure 1 Land use in Singapore

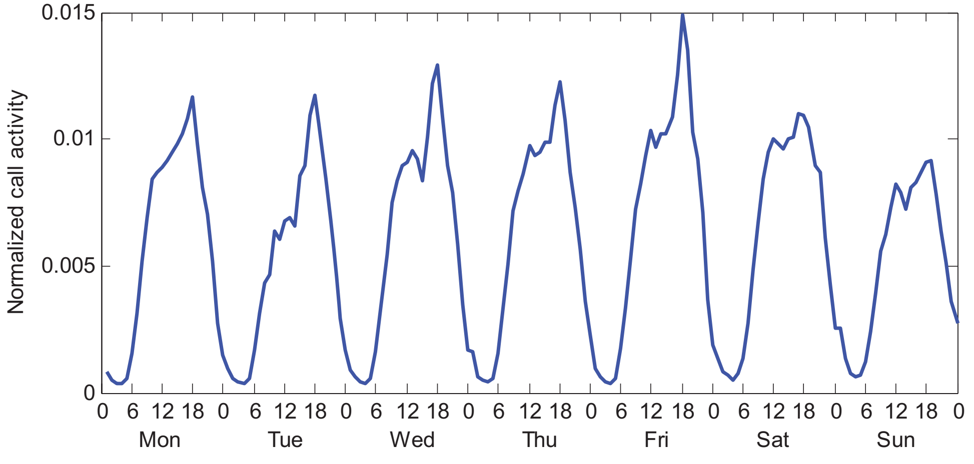


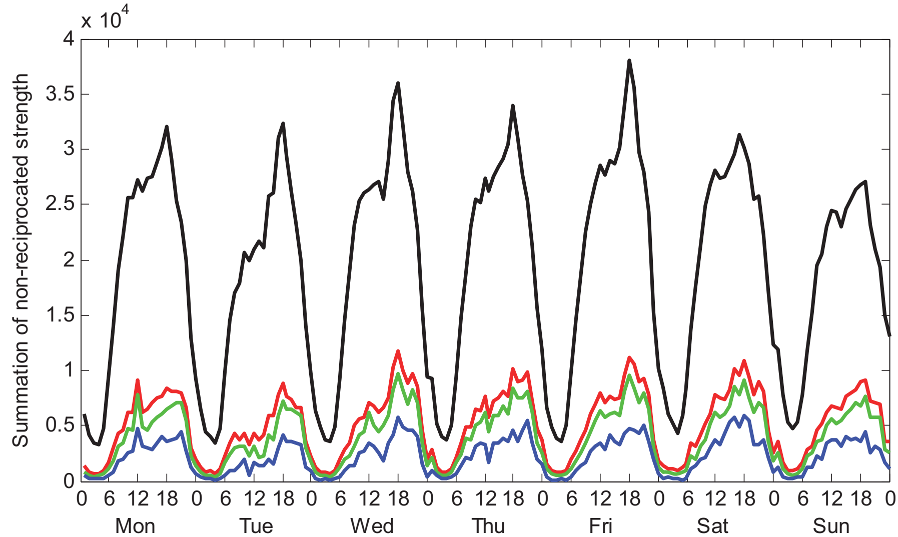
Figure 2 Normalized call activity in Singapore

4 Temporal Imbalance and Reciprocity

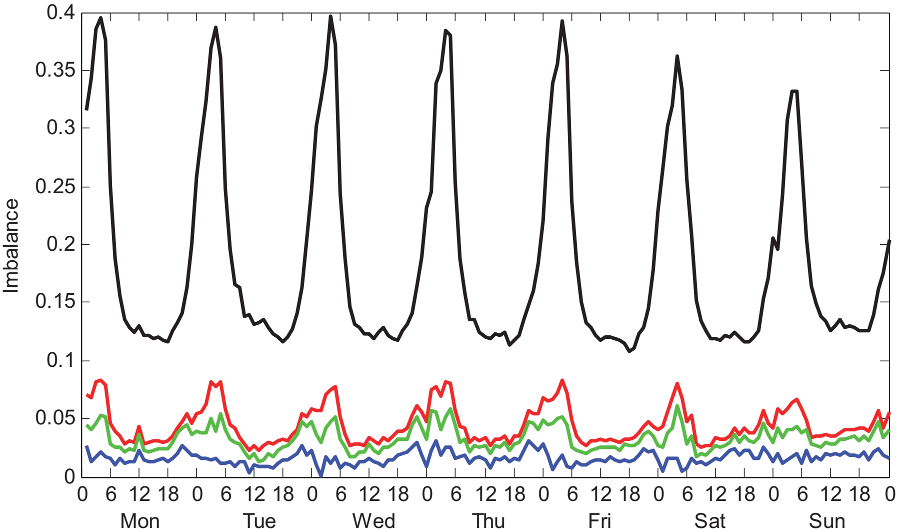
4.1 Entire Imbalance of AMPN at Different Scales

We analyzed the temporal characteristics of entire imbalance of the AMPN for the whole week. The hourly sum of non-reciprocated strengths (i.e. ΔS in Definition 4) and the hourly imbalance (i.e. B in Definition 4) are displayed in Figure 3, each of which is calculated at four different scales, specifically, from small to large, the BTS, the 2 km * 2 km grid cell, the 54-district (i.e. the city is divided into 54 districts) and 5-district level (i.e. the city is divided into five districts). Note the vertex size changes at different scales (say, at the 54-district level, the call volume is aggregated for every district, and the directional calls between any two districts form an AMPN). In Figure 3, each time sequence contains 168 points representing every hour in the whole week.

Figure 3a shows the sum of non-reciprocated strength (ΔS) at four different scales, i.e. from top to bottom, the BTS, the 2 km * 2 km grid cell, the 54-district and the 5-district level. These four levels all show a similar periodic pattern in terms of trend and local maxima. The difference is as the granularity of vertex increases the sum of non-reciprocated strength decreases. The curve of the BTS level shows a similar pattern to that of normalized call activities (Figure 2). Overall, the network, at all scales, generates larger sums of non-reciprocated strength in the daytime (the maximum is found at 6 p.m.) and smaller sums in the nighttime (the minimum is located at 4 a.m.). Figure 3b shows the imbalance of the network (B) at different scales. Being the direct opposite of ΔS, B reaches its maximum at 4 a.m. for each day and minimum at 6 p.m. (Figure 3b). A possible underlying reason for why ΔS's daily minimum but B's maximum occurs during the early morning could be since phone calls initiated at this time are regarding unusual circumstances (i.e. emergency calls), which are most likely one way. Therefore, although these calls are a relatively rare occurrence and thus few in number, they will still generate a large imbalance. Conversely, ΔS's daily maximum but B's minimum at 6 p.m. is probably the result of the larger number of calls initiated at that time which therefore reduce the possible impact of imbalance between out-flow and in-flow at every vertex. In addition to the similar pattern existing between different days, the dissimilarity is also obvious. Specifically, the Friday and Saturday nights show smaller imbalances than other days',



(a)



(b)

Figure 3 Temporal variation of AMPN imbalance: (a) Sum of non-reciprocated strength; (b) entire network imbalance (*B* in Definition 4). The black lines indicate the BTS level; the red lines the 2 km * 2 km grid cell level; the green lines the 54-district level and the blue lines the 5-district level

especially at the small granularities. The reason might be that more calls are generated on weekend nights and the calls are more dispersive than those on weekday nights. The further explanation is that on weekend nights more people are likely to stay late and are more widely distributed across the city.

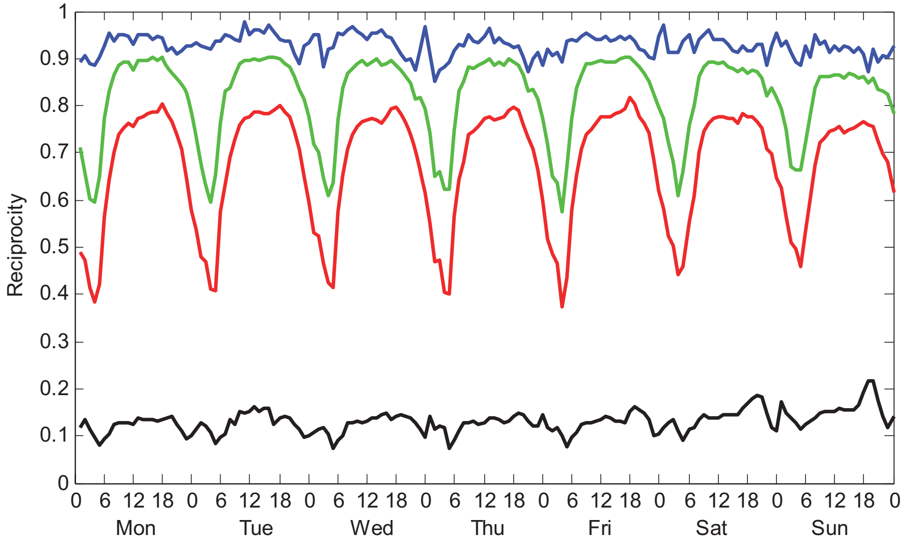


Figure 4 Temporal variation of AMPN reciprocity (R in Definition 6). The black line indicates the BTS level; the red line the 2 km * 2 km grid cell level; the green line the 54-district level and the blue line the 5-district level

4.2 The Reciprocity at Different Scales

To determine the reciprocity of the city, we calculated the value of R at different scales in light of Definition 6. Figure 4 displays the reciprocity at four different scales, which, from top to bottom, are the BTS, 2 km * 2 km grid cell, 54-district, and 5-district level. As the scale increases, the reciprocity also increases; specifically, the reciprocity is larger than 0.8 at the 5-district level and smaller than 0.2 at the BTS level. The former is totally reciprocated and the latter is almost entirely random. No significant pattern is found in either of these two curves. However, those curves at median scales demonstrate very distinctive and clear patterns. Specifically, high reciprocity is generated in the daytime for every day (the maximum is found at 6 p.m.) and low reciprocity is produced in the night-time (the minimum is found at 4 a.m.). The pattern is similar to that of the sum of non-reciprocated strength and opposite to the imbalance. The phenomenon in these curves indicates that the periodic pattern can only be found in the medium scales. The small scale shows the entire heterogeneity and randomness, whereas the large scale displays total homogeneity and reciprocity. The phenomenon in these curves indicates that the scale effect also exists in the imbalance and reciprocity of the AMPN. The reason for generating the reciprocity is similar to that for the imbalance.

5 Spatial Imbalance and Reciprocity

5.1 The Constant Temporal Deficit and Surplus Areas in Singapore

To identify the spatial heterogeneity of AMPN in Singapore, we calculate the hourly normalized imbalance (Definition 3) of cells at a level of a 500 m by 500 m grid. Figures 5a–d display the normalized imbalance of weekday morning (5–11 a.m.), the weekday noon (11 a.m.–2

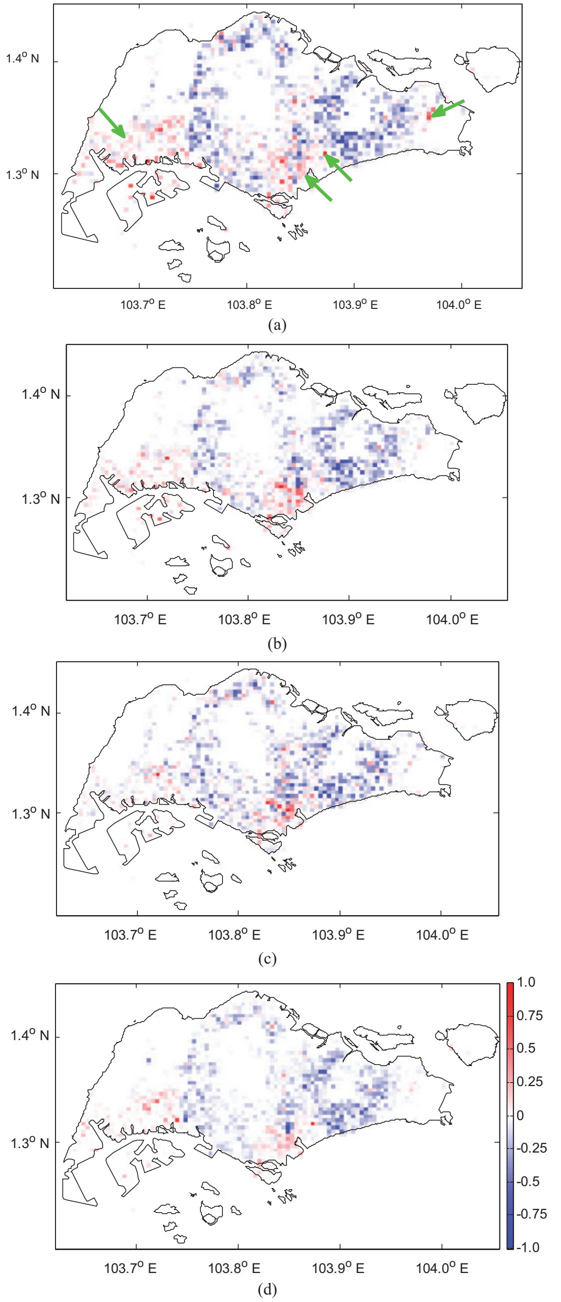


Figure 5 The normalized surplus and deficit areas for weekdays: (a) weekday morning (5–11 a.m.); (b) weekday noon (11 a.m.–2 p.m.); (c) weekday afternoon (2–7 p.m.); and (d) weekday night (7 p.m.–5 a.m.). In Figure 5a, the business area (indicated by the left green arrow) is located in the west; the downtown area (indicated by the second left green arrow) is located in the south; the sport center (indicated by the second right green arrow) is located to the northeast of the downtown area; the airport (indicated by the right green arrow) is located in the east

p.m.), the weekday afternoon (2–7 p.m.), and the weekday night (7 p.m.–5 a.m.), respectively. Figure 6 displays the normalized imbalance for the same periods as those for weekdays. Note that the normalized surplus is between 0 and 1 whereas the normalized deficit is between 0 and −1. Figure 5 indicates that the area generating a deficit is larger than that generating a surplus. The spatial variation of normalized imbalance decreases from morning to night. The same characteristics are also found for weekends.

The constant surpluses and deficits may indicate areas with special social functions. After comparison, we can locate areas constantly generating a surplus and those constantly generating a deficit. From Figures 5 and 6, we can interestingly see that the downtown area (indicated by the second left arrow in Figure 5a, most of it belongs to Commercial) and the southwestern area (indicated by the left arrow in Figure 5a, most of it belongs to Business) are those constantly generating a surplus while the cells constantly generating a deficit are mainly located in the middle, the north and the east (most of them belong to Residential). It is also interesting to see that the airport constantly produces a surplus (indicated by the right arrow in Figure 5a). The reason for the airport generating a surplus is that passengers tend to initiate calls rather than receive calls regardless of whether they leave or arrive. Another constant surplus, indicated by the second right arrow, is a sports center. The reason for it generating a constant surplus is that during physical exercise people are unable to answer calls, and more outgoing calls are call-backs after missed calls are found. All analysis indicates that the surplus or deficit, with other information, might help us identify locations with special social functions.

5.2 The Imbalance and Reciprocity at Different Land Use Types

In order to determine the difference of imbalance between various land use types, we calculate four indices (Figures 7 and 8). In detail, the imbalance of land use is showed in Figure 7a, the normalized imbalance in Figure 7b, the non-reciprocity of land use in Figure 8a and the normalized non-reciprocity in Figure 8b. We found that different land use types demonstrate distinctively different imbalance and non-reciprocity. Figure 7a displays that a deficit dominates Residential the most of time, which indicates more incoming calls. The surplus dominates other land use types, which indicates more outgoing calls. This characteristic is consistent with previous results we have found in Figures 5 and 6. Interestingly, in addition to the high deficit value appearing in the nights, we also find one local minimum repeat periodically for Residential in weekdays. Specifically, it is constantly located at 12 a.m. The inverse pattern is found for Business and Commercial. The high deficit of Residential and surplus of Business and Commercial found at 12 a.m. only on weekdays might be due to the fact that working individuals during their noon break may initiate more calls to their family members at home (say, their retired parents) to check their safety, etc. Figure 7b shows that the normalized imbalance is very low (between –0.05 and 0.05) for all land use types. Among them, Commercial shows the strongest normalized imbalance and the most intensive variation, whereas Residential displays the lowest and smoothest normalized imbalance; however, not all of them demonstrate distinct periodic-patterns.

Figure 8a displays that Residential generates the highest non-reciprocity. We also find two local maxima in weekdays. The first is located at 12 a.m. (except for the relatively small value on Tuesday) and the second is located between 6 and 8 p.m., indicating an opposite numerical pattern to that of imbalance (Figure 7b). Meanwhile, the weekends are slightly more non-reciprocated than weekdays. Figure 8b shows a very weak normalized non-reciprocity for all land use types, most of which are lower than 0.1. Commercial demonstrates more intensive variation than other land use types, which corresponds to that of normalized imbalance (Figure 7b).

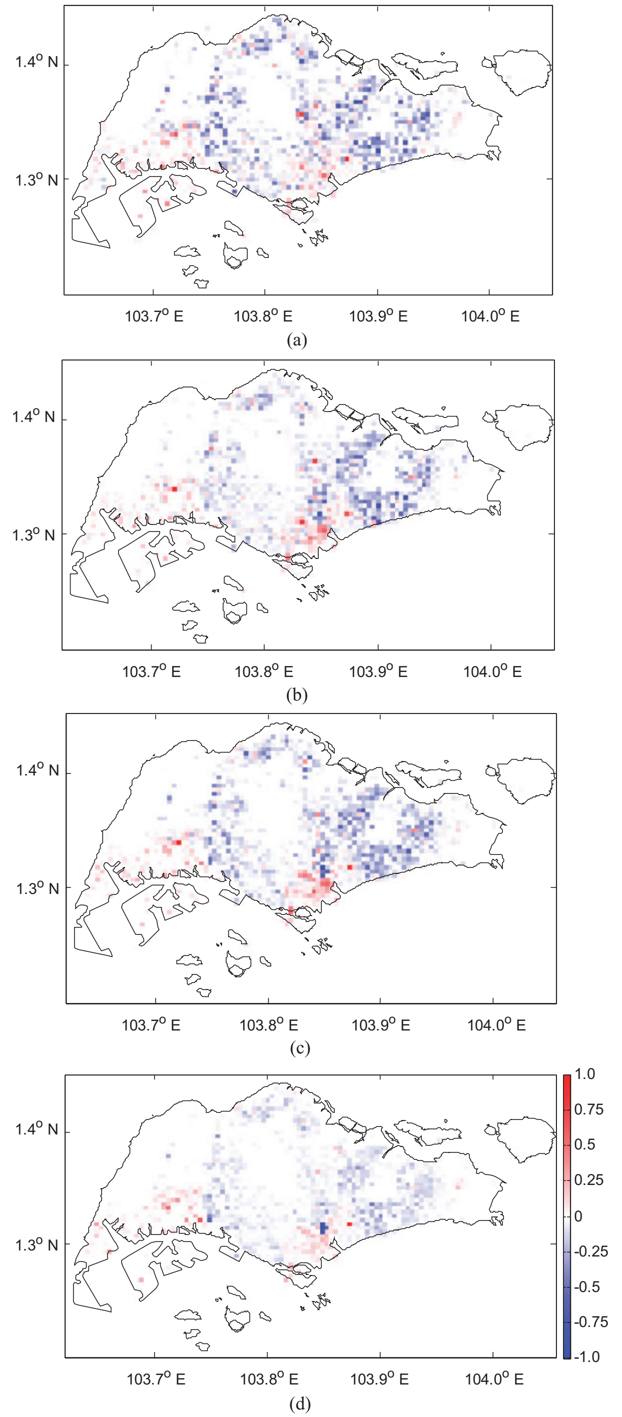


Figure 6 The normalized surplus and deficit areas for weekends: (a) weekend morning (5–11 a.m.); (b) weekend noon (11 a.m.–2 p.m.); (c) weekend afternoon (2–7 p.m.); and (d) weekend night (7 p.m.–5 a.m.)

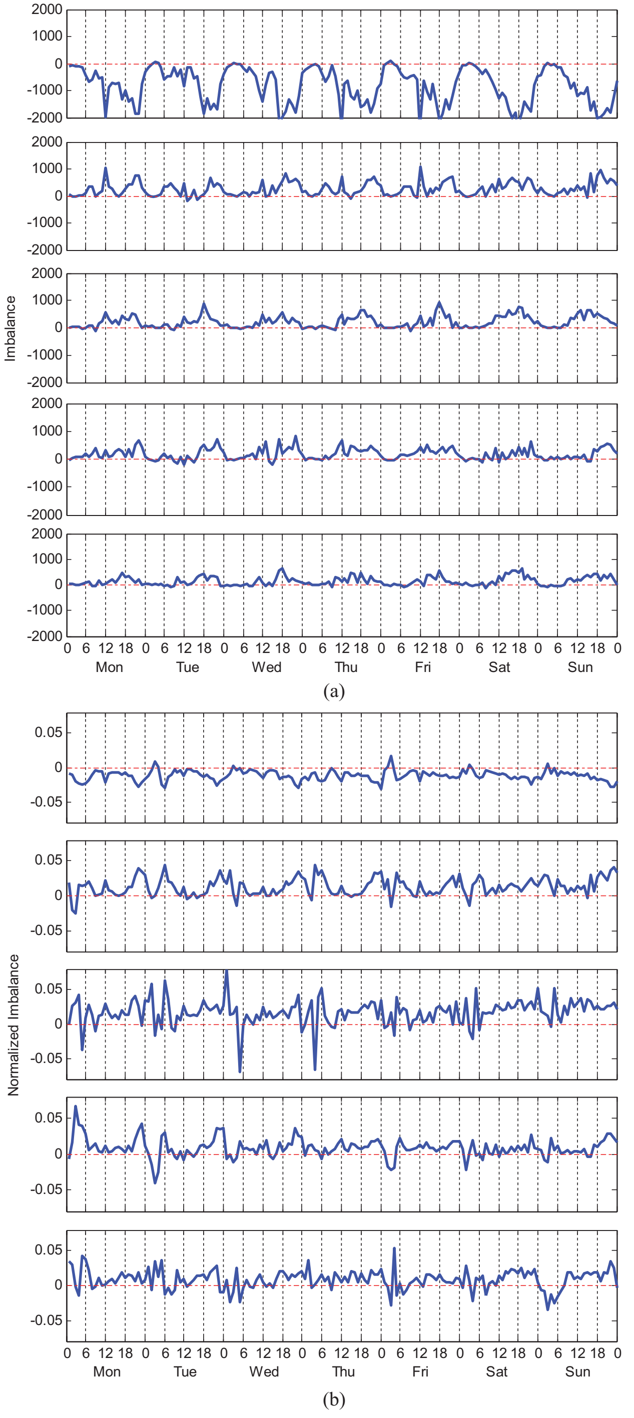


Figure 7 Imbalance of different land use types: (a) Imbalance of land use; and (b) Normalized imbalance of land use. From top to bottom are Residential, Business, Commercial, Open space and Other

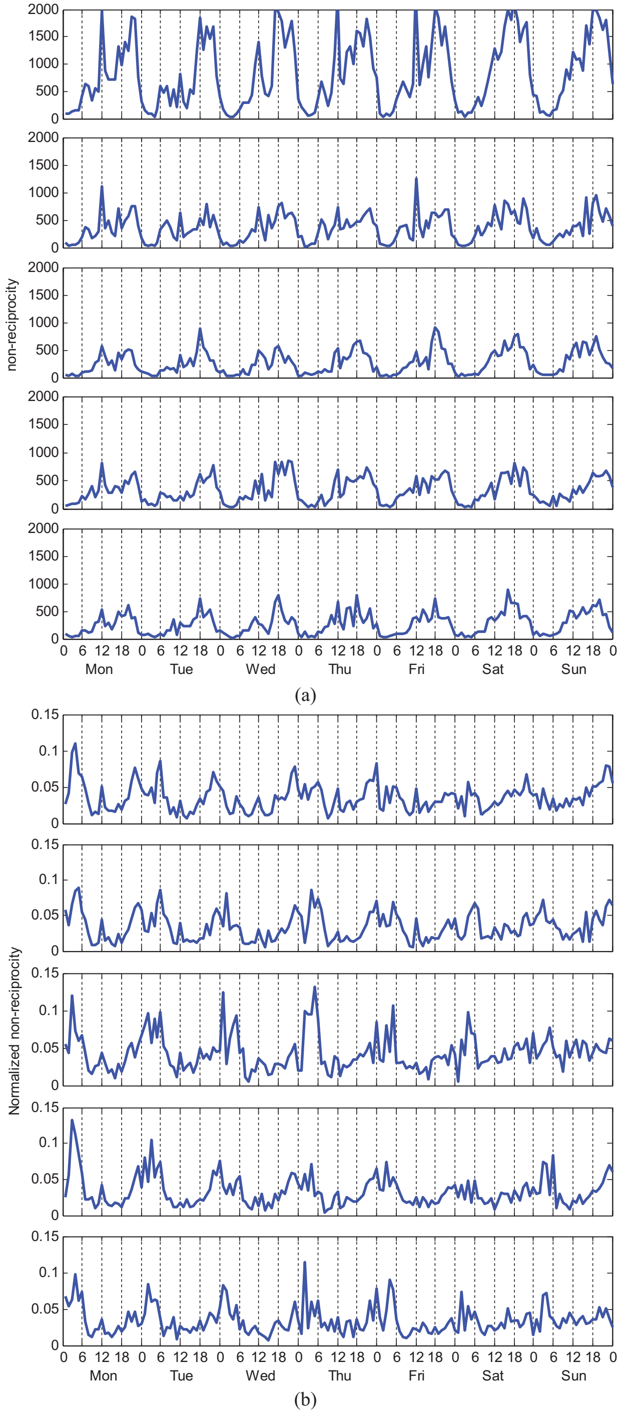


Figure 8 Non-reciprocity of different land use types: (a) Non-reciprocity of land use; and (b) Normalized non-reciprocity of land use. From top to bottom are Residential, Business, Commercial, Open space and Other

Table 1 Normalized non-reciprocated weights of different land use types for weekdays

	Residential	Business	Commercial	Open Space	Other
Residential	0	–0.0086	–0.0061	–0.0084	–0.0060
Business	0.0147	0	–0.0010	0.0021	–0.0001
Commercial	0.0194	0.0018	0	0.0001	0.0018
Open Space	0.0161	–0.0024	–0.0001	0	0.0009
Other	0.0155	0.0001	–0.0015	–0.0012	0

Table 2 Normalized non-reciprocated weights of different land use types for weekends

	Residential	Business	Commercial	Open Space	Other
Residential	0	–0.0110	–0.0096	–0.0084	–0.0090
Business	0.0218	0	–0.0013	0.0049	0.0001
Commercial	0.0323	0.0022	0	0.0022	0.0019
Open Space	0.0168	–0.0050	–0.0013	0	0.0011
Other	0.0229	–0.0001	–0.0014	–0.0014	0

In order to explain the characteristics of the imbalance and reciprocity mentioned above, we calculate the normalized non-reciprocated weights between land use types for weekdays and weekends in Tables 1 and 2. From these tables we find that Residential is always negative to the other land use types on both weekdays and weekends whereas Commercial always exhibits a positive pattern. Although other land use types do not exhibit the same pattern as that of Residential or Commercial, they are all dominated by the positive weights in Commercial. The reasons for Commercial being more positive than the other land use types are still unclear. However, we guess that it may be related to profit-making and the high mobility of residents in that area. The reason for Residential generating negative weights might be the same as that for the local minimum of imbalance, which we referred to earlier. That is, due to separation from family members during the day, more calls are targeted to home for their information. In addition, all the normalized non-reciprocated weights for weekends are not smaller than those for weekdays (this is also consistent with the phenomenon found in Figure 8a). The reason might be that at weekends the city is more heterogeneous, because people in Business and Commercial are fewer than usual. The heterogeneity of population distribution increases the non-reciprocity of the AMPN.

As indicated in the book of "the new science of cities" (Batty 2013), the flows (e.g. flows of people, information, and energy) are a very important aspect of studying cities. The information flow, being one of the important flows, may have a close relationship with those of people and energy. If the mobile phone communication can be seen as an important component of the information flow, then all analysis of imbalance and reciprocity above might indicate some characteristics of information flow between different land use types. Specifically, the negative weights for the residential area indicate that more calls are probably directed home, which implies that the residential area attracts more attention (if we think incoming calls indicate more attention is attracted).

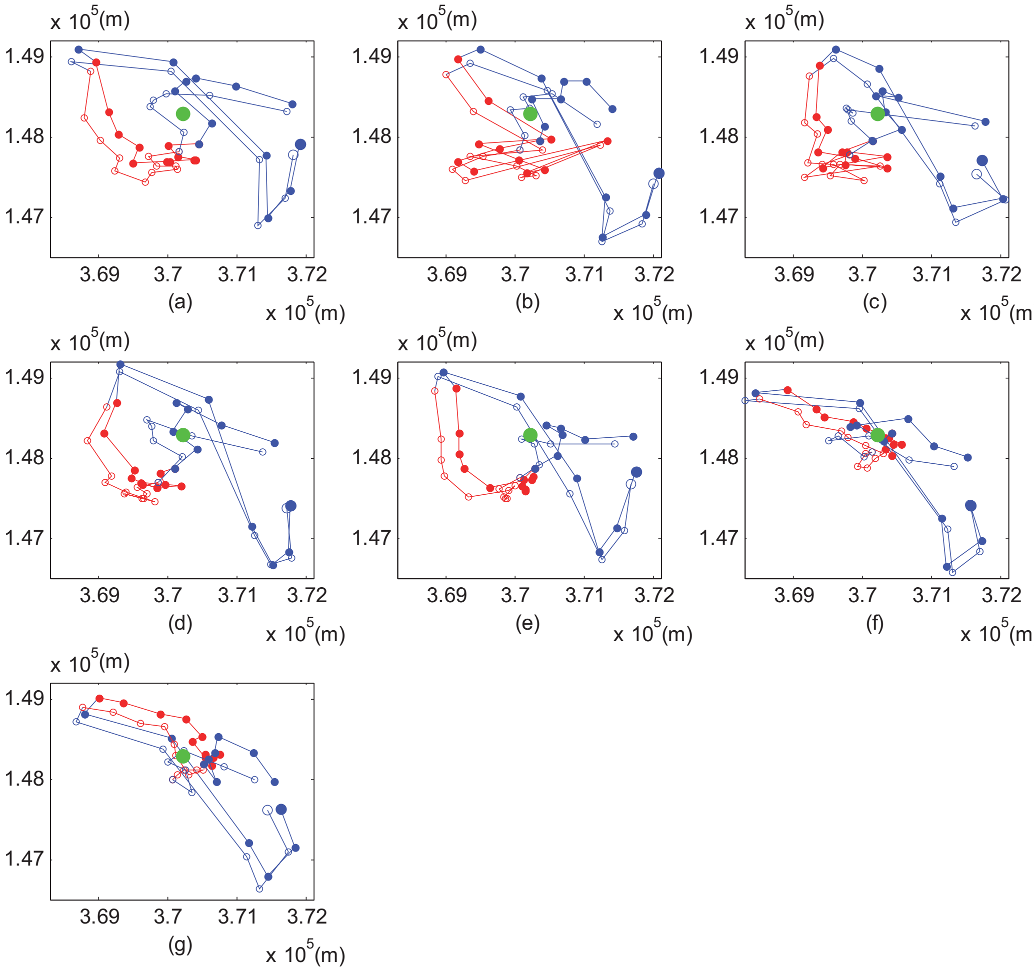


Figure 9 Trajectories of gravity centers of incoming and outgoing calls: (a) Monday; (b) Tuesday; (c) Wednesday; (d) Thursday; (e) Friday; (f) Saturday; (g) Sunday (Blue lines indicate the trajectories at nighttime and red lines indicate those in the daytime. Dots indicate the centers of incoming calls and circles indicate those of outgoing calls, where the larger blue dots indicate the start point of trajectories of incoming calls; the larger blue circles indicate the start point of trajectories of outgoing calls; the green dots indicate the geometric center of the BTS network). Note that axes x and y are in geodetical coordinates

6 Spatiotemporal Heterogeneity

Figures 9a–g display the trajectories of the gravity centers of outgoing and incoming calls for each day. The centers of outgoing calls were calculated according to the following formulas: $X_t = \sum_i \frac{x_i n_{i,t}}{\sum_i n_{i,t}}$ and $Y_t = \sum_i \frac{y_i n_{i,t}}{\sum_i n_{i,t}}$, where X_t is the x-coordinates of the center for the tth hour, Y_t is the y-coordinates of the center for the tth hour, x_i is the x-coordinate of the ith BTS, y_i is the y-coordinate of the ith BTS, $n_{i,t}$ is the number of outgoing calls initiated from the ith

BTS within the *t*th hour. The centers of incoming calls can be calculated in the same way. Therefore, if more calls are initiated (received) in the north, the centers of outgoing (incoming) calls will be located in the north, vice versa. The move of the centers can be seen as the overall flow of people in the whole city, and so it can be used to analyze the direction and the rhythm of the flow. The trajectories of outgoing and incoming calls show a similar pattern with small distances departed. Specifically, for most of the trajectories, the centers of outgoing calls are located to the west or the south of those of incoming calls. The reason could be that the commercial and the business areas (generating major surpluses) are mainly located to the south and the southwest of most of the residential area (generating major deficit). This can also be confirmed by Figures 5 and 6.

In addition to the difference between incoming and outgoing calls, we find an interesting spatiotemporal pattern from both trajectories. That is, all trajectories follow a similar general geographic sequencing (i.e. southeast-north-south-north-southeast), which indicates the mobility pattern and calling behavior of users. In the middle of the night (from 11 p.m. to 4 a.m.), the trajectories go from the southeast to the north. The reason for calling centers moving this way is that most calls are initiated or received in the downtown area and airport during this period, so the centers are located in the southeast. The center then goes north, because users in the residential areas (north of the downtown area) wake up as time passes. After that (i.e. from 5 to 8 a.m.), the trajectories turn to the south, indicating the users' mobility pattern (from home to work places). The mobility pattern coincides with the fact that the commercial and business areas are also located in the south. In working time (9 a.m.–4 p.m.), the centers are constrained in the south, and the distances between nodes are smaller than those generated in the morning and late night, indicating less mobility of users. This also corresponds to the fact that most users are in work places and do not move in this period. In the evening (i.e. from 5 to 8 p.m.) the trajectories go back to the north, indicating users' mobility after work (from work places to home). In the night (from 9 to 11 p.m.), the trajectories go from the north to the southeast, indicating the centers of call activity return to the downtown area and to places like airports. Unlike weekdays, during the daytime the centers of outgoing and incoming calls move to the place near the geometric center of the network on the weekends. The reason for the difference between workdays and weekends could be that at the weekends some people may stay at home and do not go out (especially to the south).

7 Conclusions

This article systematically analyzes the spatiotemporal imbalance and reciprocity of the AMPN in Singapore. Our findings include the following aspects. First, regarding the temporal pattern, we find that the AMPN is more unbalanced at night-time than in the daytime while the reciprocity also shows the corresponding (numerically inverse) pattern. Meanwhile, the maxima of imbalance are found in the early morning (around 4 a.m.), and the minima in the evening (around 6 p.m.) on both weekdays and weekends, whereas the reciprocity shows the inverse. Both imbalance and reciprocity demonstrate the scale effect as the granularity size of vertex changes. Second, regarding the spatial pattern over different land use types, we discover that the Residential area is dominated by a deficit on both weekdays and weekends whereas the other land-use types show surpluses. The reason could be that more calls are initiated from other land-use types than from Residential due to the separation between family members. As a result, the strong and constant surplus may provide useful information to delineate the commercial area (or even more specific places, say the airport terminal and the sports center) whereas the deficit is confined to the residential area. Third, the spatiotemporal analysis reveals that the trajectories of incoming and outgoing calls follow the same pattern, and the relative position between them is decided by the configuration of land-use types and their call activities. The centers of outgoing and incoming calls move from the southwest to the north in the late night and early morning, indicating the discrepancy of human activity between different land-use types. Then, the centers go to the south in the morning, fluctuate at noon, and go back to the north, shedding light on the residents' mobility pattern. All of the information may help urban planners understand the true dynamics of the city and thereby facilitate them to plan more effectively.

In this article, we mainly focus on the issue of how to describe and analyze the directional heterogeneity of an AMPN. More topics deserve further studying, such as the reason for Commercial constantly generating a surplus and Residential generating a deficit, the relationship between the heterogeneity of a city and its communication network, and that between the human mobility (the flow of people) and communication patterns (the flow of information). Our future work may be extended to study these issues and finally reveal the unban dynamics based on tracking information like mobile phone data.

References

Amini A, Kung K, Kang C G, Sobolevsky S, and Ratti C 2014 The impact of social segregation on human mobility in developing and industrialized regions. *EPJ Data Science* 3: 6

Batty M 2013 *The New Science of Cities*. Cambridge, MA, MIT Press

Calabrese F, Lorenzo G D, Liu L, and Ratti C 2011 Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing* 10(4): 36–44

Gonzalez M, Hidalgo C, and Barabasi A 2008 Understanding individual human mobility patterns. *Nature* 453: 779–82

Grauwin S, Sobolevsky S, Moritz S, Gódor I, and Ratti C 2014 Towards a comparative science of cities: Using mobile traffic records in New York, London and Hong Kong. Unpublished manuscript (available at arXiv:1406.4400)

Jacobs-Crisioni C G W and Koomen E 2012 Linking urban structure and activity dynamics using cell phone usage data. In *Proceedings of the Workshop on Complexity Modeling for Urban Structure and Dynamics, Fifteenth AGILE International Conference on Geographic Information Science*, Avignon, France

Kang C G, Sobolevsky S, Liu Y, and Ratti C 2013 Exploring human movements in Singapore: A comparative analysis based on mobile phone and taxicab usages. *In Proceedings of the Second ACM SIGKDD International Workshop on Urban Computing*, Chicago, Illinois

Koylu C and Guo D S 2014 Smoothing locational measures in spatial interaction networks. *Computers, Environment and Urban Systems* 41: 12–25

Kung K S, Greco K, Sobolevsky S, and Ratti C 2014 Exploring universal patterns in human home-work commuting from mobile phone data. *PLoS ONE* 9(6): e96180

Laura F, Marco M, and Massimo C 2014 Discovering events in the city via mobile network analysis. *Journal of Ambient Intelligence and Humanized Computing* 5: 265–77

Loibl W and Peters-Anders J 2012 Mobile phone data as source to discover spatial activity and motion patterns. In Jekel T, Car A, Strobl J, and Griesebner G (eds) *GI_Forum 2012: Geovizualisation, Society and Learning*, Berlin, Herbert Wichmann Verlag: 524–32

Manfredini F, Tagliolato P, and Rosa C D 2011 Monitoring temporary populations through cellular core network data. In Murgante B, Gervasi O, Iglesias A, Taniar D, and Apduhan B O (eds) *Computational Science and Its Applications: ICCSA 2011 International Conference Proceedings, Part II*. Berlin, Springer Lecture Notes in Computer Science Vol. 6783: 151–61

Meyers L A, Newman M E J, and Pourbohloul B 2006 Predicting epidemics on directed contact networks. *Journal of Theoretical Biology* 240: 400–18

Pei T, Sobolevsky S, Ratti C, Shaw S L, and Zhou C H 2013 A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science* 28: 1988–2007

Perra N, Zlatic V, Chessa A, Conti C, Donato D, and Caldarelli G 2009 PageRank equation and localization in the WWW. *Europhysics Letters* 88: 48002

Pulselli R M, Ratti C, and Tiezzi E 2006 City out of chaos: Social patterns and organization in urban systems. *International Journal of Ecodynamics* 1: 125–34
Ratti C, Pulselli R M, Williams S, and Frenchman D 2006 Mobile landscapes: Using location data from cell phones for urban analysis. *Environment and Planning B* 33: 727–48
Ratti C, Sobolevsky S, Calabrese F, Andris C, Reades J, Martino M, Claxton R, and Strogatz H S 2010 Redrawing the map of Great Britain from a network of human interactions. *PLoS ONE* 5(12): e14248
Rubioa A, Sanchezb A, and Frias-Martineza E 2013 Adaptive non-parametric identification of dense areas using cell phone records for urban analysis. *Engineering Applications of Artificial Intelligence* 26: 551–63
Sevtsuk A and Ratti C 2010 Does urban mobility have a daily routine?: Learning from the aggregate data of mobile networks. *Journal of Urban Technology* 17: 41–60
Sobolevsky S, Szell M, Campari R, Couronné T, Smoreda Z, and Ratti C 2013 Delineating geographical regions with networks of human interactions in an extensive set of countries. *PLoS ONE* 8(12): e81707
Song C, Qu Z, Blumm N, and Barabasi A-L 2010 Limits of predictability in human mobility. *Science* 327: 1018–21
Soto V and Frias-Martinez E 2011 Automated land use identification using cell-phone records. In *Proceedings of the Third ACM International Workshop on MobiArch*, Bethesda, Maryland: 17–22
Squartini T, Picciolo F, Ruzzenenti F, and Garlaschelli D 2013 Reciprocity of weighted networks. *Scientific Reports* 3: 2729
Sun J B, Yuan J, Wang Y, Si H B, and Shan X M 2011 Exploring space–time structure of human mobility in urban space. *Physica A* 390: 929–42
Toole J L, Ulm M, González M C, and Bauer D 2012 Inferring land use from mobile phone activity. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, Beijing, China
Traag V A, Browet A, Calabrese F, and Morlot F 2011 Social event detection in massive mobile phone data using probabilistic location inference. In *Proceedings of the Third IEEE Conference on Social Computing*, Boston, Massachusetts: 1–4
Vaccari A, Gerber A, Biderman A, and Ratti C 2009 Towards estimating the presence of visitors from the aggregate mobile phone network activity they generate. In *Proceedings of the Eleventh International Conference on Computers in Urban Planning and Urban Management*, Hong Kong: 1–11
Vieira M R, Frias-Martinez V, Nuria O, and Frias-Martinez E 2010 Characterizing dense urban areas from mobile phone-call data: Discovery and social dynamics. In *Proceedings of the IEEE International Conference on Social Computing/IEEE International Conference on Privacy, Security, Risk and Trust*, Minneapolis, Minnesota: 241–8
Yu X P and Pei T 2013 Analysis on degree characteristics of mobile call network. *Acta Physica Sinica* 62: 208901
Zamora-Lopez G, Zlatic V, Zhou C, Stefancic H, and Kurths J 2008 Reciprocity of networks with degree correlations and arbitrary degree sequences. *Physical Review E* 77: 016106
Zlatic V and Stefancic H 2009 Influence of reciprocal edges on degree distribution and degree correlations. *Physical Review E* 80: 016117

