# DAILY LIFE PATTERN OF A CITY: DELINEATING ACTIVITY SPACE AND TIME USING SOCIAL MEDIA DATA

WORKING PAPER

**Wei Chien Benny Chin***
Department of Geography
National Taiwan University
Taipei, 10617, Taiwan
wcchin.88@gmail.com

## ABSTRACT

People live in cities. However, different groups of people are active at varying times at distinct places in their daily lives, forming different activity spaces and times within a city. Previous studies focused on the distribution of the activity or the partition of regions based on similarity or strength of the interaction, but not the collective activity spaces and times of the people who live within the area. Using only the geotagged/timestamped social media as a proxy, this study intended to delineate activity spaces and times of eight selected cities, including Tokyo, Osaka, Hongkong, Singapore, Bangkok, Jakarta, Manila, and Penang. Activity space is defined as a geographic extent where people undertake their daily life. This study generated two co-occurrence networks (spatial and temporal) for each city and delineated the activity spaces and times using a network community detection method. In summary, the results showed a clear pattern for both activity spaces and times in the eight cities. The activity spaces results showed spatially continuous communities with clear borders, indicated the boundaries of human movements, which may be affected by political or natural separation. The activity time results existed a cyclic pattern on a daily and weekly basis, indicating the habits of people in each city, and which pattern is slightly different between cities. In conclusion, this study demonstrated a framework for delineating activity spaces and provided a novel perspective for representing the space and time patterns of daily life in a city.

***Keywords*** Spatial community · temporal community · urban structure · daily activity pattern · East Asia · Southeast Asia

## 1 Introduction

A city can be viewed as an organism (Batty, 2012). People live in cities. But, different groups of people are active on varying time at distinct places in their daily lives. A regular daily life of a person is usually a combination of where the person stays, where the person works or study, and where the person is getting his/her daily needs (foods, transportation facilities, etc.). The daily travel experience of a person is known as the individual activity spaces (Golledge and Stimson, 1997; Schönfelder and Axhausen, 2003; Lee et al., 2016). A similar group of people, i.e. the people who live within a same neighborhood and their offices or schools are close to each other, forms a community (Yuan and Nara, 2015; Sobolevsky et al., 2018). Thus, groups of similar people are forming the collective activity spaces within a city (Shelton et al., 2015).

Activity space is defined as the places where a person is accessing or visiting for his/her daily life activity (Golledge and Stimson, 1997; Schönfelder and Axhausen, 2003). The concept of activity spaces is related to several concepts and terms in previous studies. The time-geography concept focuses on the time-space trajectory of a person (or a group of people), and aims to understand the potential movements of people within the space and time dimensions through space-time paths or space-time prisms (Hägerstraand, 1970; Kwan, 2000; Kwan and Lee, 2004; Cao et al., 2015; Song

*Currently, a Research Fellow with SGP Cities, Singapore University of Technology and Design. benny_chin@sutd.edu.sg

and Miller, 2015; Yuan and Nara, 2015). Functional region is a concept describing the networked regions where people travel from one place to another for some specific activities or purposes (Brown and Holmes, 1971; Farmer and Fotheringham, 2011; Lloyd and Cheshire, 2017), e.g. travel to work/school, catchment areas of retail centers. From the perspective of ecological studies, the concept of activity spaces resembles the concept of home-range. Home-range is a concept that focuses on the spatial extent or boundary where an animal visited or moved during its daily activity (Burt, 1943; Ciucci et al., 1997; Šálek et al., 2015). While these concepts and related studies focused on different issues, scales, or objects, they were studied about movements in time and space.

The data of human movement is a key to uncover the activity spaces. However, accessing the movement data of the population is not feasible due to the cost of data collection and the necessity of privacy protection. Human movement data is often not available for most developing countries. With adequate privacy protection and data anonymization, previous studies used various types of data as alternative data for human movements, including mobile phones data (Blondel et al., 2008; Järv et al., 2014; Ge and Fukuda, 2016), Taxi GPS records (Zhang et al., 2017; Zhu et al., 2017), and social media data (Blanford et al., 2015; Nguyen et al., 2016; Comito et al., 2016; Shen and Karimi, 2016; Jendryke et al., 2017).

In the age of internet of thing (IoT) and Internet of People (IoP), the so-called ubiquitous computing technologies and products are embedded in everyone's life, especially for the people who live in cities (Miranda et al., 2015; Liu et al., 2015). Social media is a platform where people shares their stories, ideas, emotions, and news with each other. The externality of the posting records might also show the location and time when they are posting to their social media timeline. In other words, human activities are recorded as a side product on the social media platform, and which can be used as a proxy for revealing the human activities and urban geographical structures, including activity spaces and times (Jiang and Miao, 2015; Shelton et al., 2015).

Previous studies using social media data as a proxy of human activity mainly focused on the intensity of activities, including the distribution patterns (Jiang et al., 2016; Ma et al., 2018) and the intensity of movements (Liu et al., 2014; Wu et al., 2014; Hawelka et al., 2014). While the relationships of places sharing a large visiting group of people indicated a strong interaction, and which interaction could help for delineating the regions of activity spaces (Sobolevsky et al., 2018). However, the collective activity space or time that could help on revealing the activity pattern or lifestyle of a city was neglected in previous studies.

On the other hand, although the concept of activity times is defined in decades ago (Golledge and Stimson, 1997), the time dimension of activities were not focused in previous studies, i.e. time dimension is often treated as an extension in spatial analysis. Similar to the activity spaces, the activity times can be defined as the temporal ranges that were mainly dominated by some groups of people. This information can be useful in the understanding of the patterns of life of a city or region (Yuan and Nara, 2015).

This study intended to delineate the activity spaces and times, using only the location and time of the social media records with the concept of activity spaces. The working definition of the activity space in this study is defined as the places which is frequently visited by a similar group of people. The non-western regions are often neglected or received less attentions in previous studies, especially for the developing countries in Southeast Asia. Therefore, this study used eight cities located in East Asia and Southeast Asia as case studies.

# 2 Methods

## 2.1 Study sites

Concentrated to East Asia and Southeast Asia, eight cities were selected to be analyzed in this study, including Tokyo (Japan), Osaka (Japan), Hongkong (People's Republic of China, PRC), Singapore, Bangkok (Thailand), Jakarta (Indonesia), Manila (Philippines), and Penang (Malaysia). The first four cities are developed cities, whereas the last four are developing cities. While the aim is to analyze the activity spaces, which is the boundaries of daily human movements, the study areas of the eight cities were designed to includes the urbanized regions. In addition, the results were intended to be compare with the political boundaries and natural boundaries, hence some of the neighboring regions were included, such as the Shenzhen for Hongkong case study, and a part of Johor for the Singapore case study.

Tokyo and Osaka (Keihanshin) are two major Metropolitan Areas in Japan; only the core areas of the two were included in this study, which is the Tokyo 23 special wards ($632km^2$), and the combinations of Osaka City, Sakai City with 21 neighboring cities or towns ($1002km^2$) were used as the Tokyo and Osaka cases, respectively. Hongkong and Singapore are important harbour cities in Asia. Both Hongkong and Singapore are located at the edge of the Asia continent and are separated by some natural borders from their neighboring cities or countries: Hongkong is separated from mainland China by the Sham Chun River, whereas Singapore is separated from the Peninsular Malaysia by the Johore Strait. For these two cases, the neighboring areas are also included in the studying area, which is the Shenzhen City (PRC) for Hongkong case (a total of $1397km^2$) and Johor (Malaysia) for Singapore cases (a total of $2389km^2$).

Although the last four cities are not as developed as the first four cities, they are also highly urbanized and populated cities. Bangkok, the capital city of the Kingdom of Thailand, is an inland city surrounded by several adjacent provinces. The Bangkok case in this study contains the Bangkok Metropolitan Region, which is the Bangkok City and some surrounding provinces ($7652km^2$). Jakarta, the capital city of Indonesia, is the most populous city in the Java Island. Together with the surrounding urbanized municipalities, the major area of the Greater Jakarta Metropolitan Area ($3524km^2$) in included as the Jakarta case in this study. While Manila is the capital city of the Philippines, Metro Manila contains the Manila City and its surrounding areas that were also highly urbanized. This study includes the Metro Manila (or the National Capital Region) area as the Manila case ($1801km^2$). Penang is an urbanized state in Malaysia, which geographical area is separated into two parts by the Penang Strait: the Penang island and Seberang Perai, which is on the Peninsular Malaysia ($1056km^2$).

### 2.2 Data collection

To collect data for each city, this study used the Twitter API through a Python package (python-twitter, an open-source package provided by The Python-Twitter Developers) to search for the geocoded tweets within the study area. The cities were divided into $1km \times 1km$ grids, and each cell was divided into four equal size quadrants and which centroids were used for calling the tweet searching API (with radius equal to 250 meter). The data collection procedure contained two steps: first, this study searched for the users who had been twitting within the study area (using the "GetSearch" function); second, this study acquired the Twitter timeline of these users (by using the "user_timeline" function). In simple words, these procedures were designed to identify the persons who had tweeted within the study area, and then collect the tweeting history of these people. After the user timeline was collected and the non-geocoded tweets were filtered out, a list of tweets that recorded the time and location of each tweet were prepared for each user. Then, the location of each tweet is remapped into the cell of the $1km \times 1km$ grid.

### 2.3 Analysis framework

An activity space is defined as a geographical extent where people undertake their daily life. In other words, the activity space is a spatial boundary that a similar group of people live inside and have their daily life activities, including sleeping, working, eating, socializing, etc. Therefore, this study adopted a network community detection method to identify the activity space based on the location co-users network (hereafter referred to as locations network), which is described as follow.

For each city, a two-mode network was constructed based on the aforementioned user "timeline" data, with the users on one side (mode), and the grid cells (locations) where each user had visited as the other side (mode) (the top left network on Figure 1). This two-mode network is then being converted into the locations network, which nodes are the locations, and the weights of the links between nodes represent the number of users that appeared (tweetted) on both locations (cells) (the bottom left network on Figure 1 shows the network for two users, the top right network on Figure 1 shows the complete network for all users). If a link is weighted high, this indicates that the two locations share a high number of users (the top right network on Figure 1).

After a location network is generated for a city, it is used to perform a community detection analysis to identify the groups of locations by the weights of sharing users (the bottom right network on Figure 1). In other words, the nodes in a same community means that they have strong ties between them, i.e., a large group of users who are visiting the locations within the community. The communities are then used as the working definition of activity spaces in this study.

The community detection algorithm in this study was often referred to as the Louvain method (Blondel et al., 2008). This community detection method attempts to maximize the network modularity (Q) through a heuristic approach. For a given community structure for a weighted network (with weights on links), the corresponding modularity (Q) can be calculated using Equation 1 (Newman, 2004).

$$Q = \frac{1}{2m} \sum_{i,j} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \tag{1}$$

where, $A$ is the weighted matrix representing the weight of links between nodes, thus $A_{ij}$ is the weight of the link between node i and j; $k_i$ is the the total weight of links moving out from node i ($k_i = \sum_j A_{ij}$); $c_i$ represent the community which the node i is assigned, and the $\delta(c_i, c_j)$ function is equal to one if $c_i = c_j$, or else equal to zero; $m$ is half of the sum of all links ($m = 0.5 \times \sum_{i,j} A_{ij}$).

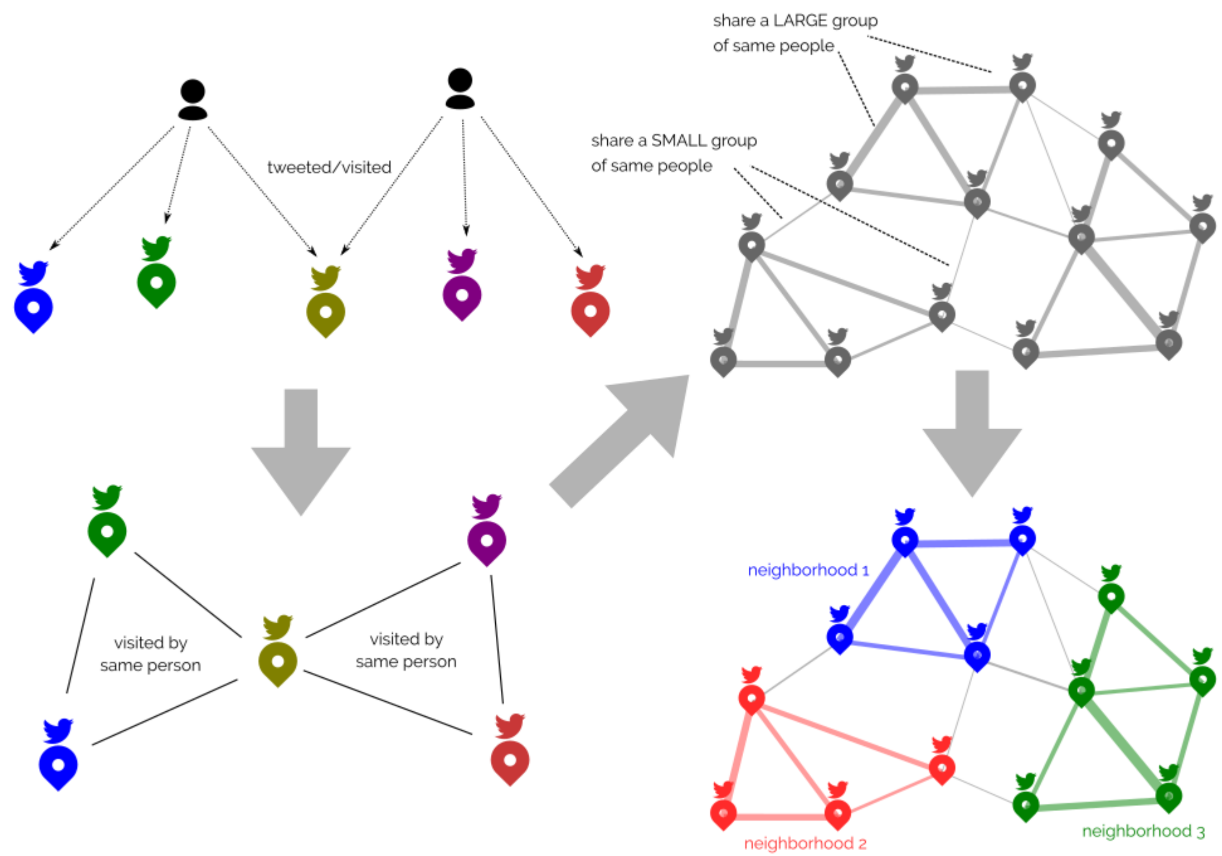


Figure 1: The procedures of delineating spatial communities from the user-tweets network. The first network is the users-tweets two-mode network; the second network is a network which nodes are locations (grid cells) and the links are added between nodes if the two nodes share a user, so the three locations on the left (and also the three on the right) were connected with each other; after processing all user-tweets relationships into the locations share-users relationships, the weight of links indicated the number of the same users visited both locations; the activity spaces can be found by performing the community detection method on the network with weighted links.

The range of modularity is between -1 and 1. A higher modularity value indicates that the interactions within each community are strong (high weights within each community) and the interactions between communities are weak. The Louvain method used a heuristic approach to maximize the modularity and to detect community partitions. This study used a Python package for performing the Louvain method (python-louvain, an open-source package provided by Thomas Aynaud).

Similarly, a time co-users network (hereafter, time network) is also constructed with the same framework for each city, by replacing location cells with hours in a week (a total of 24*7 units). Then, the community detection algorithm is applied to the time network, and the results are used as the working definition of activity time in this study, which indicated the temporal range that is occupied by the same group of users.

# 3 Results

## 3.1 Descriptive statistics

Table 1 shows the descriptive statistics about the collected data and the optimized modularity (Q) results, including the date range of the tweets within each city, the number of users and tweets, the number of occupied cells, the spatial modularity and temporal modularity. The percentage in brackets in the last column indicates the ratio of occupied cells to the total number of cells in the grid for each city. The Tokyo and Osaka cases had a higher occupied ratio because only the core areas of the cities were included in the study. The area of some cities (e.g. Hongkong and Bangkok) covered places without daily human activity, which leads to a lower occupied ratio (and more blank areas in the result of activity spaces).

Table 1: Descriptive statistics of the Twitter data collected for the eight cities and the spatial and temporal modularity (Q). The percentage in brackets represents the ratio of cells that are occupied by at least one tweet. The spatial and temporal modularity indicated the modularity of the detected communities distribution.

| city | date range | #users | #tweets | #cells | spatial Q | temporal Q |
|---|---|---|---|---|---|---|
| Tokyo | '10/09 - '17/07 | 15865 | 1076552 | 2674 (96%) | 0.457 | 0.101 |
| Osaka | '10/10 - '17/09 | 6132 | 346026 | 3184 (74%) | 0.610 | 0.122 |
| Hongkong | '10/09 - '17/09 | 1752 | 138005 | 2190 (33%) | 0.378 | 0.066 |
| Singapore | '10/09 - '17/09 | 7118 | 636184 | 4269 (41%) | 0.391 | 0.054 |
| Bangkok | '10/09 - '17/09 | 9986 | 867866 | 9998 (32%) | 0.328 | 0.055 |
| Jakarta | '10/09 - '17/09 | 16560 | 2568086 | 9165 (62%) | 0.367 | 0.043 |
| Manila | '10/09 - '17/09 | 12345 | 1180583 | 4216 (55%) | 0.325 | 0.046 |
| Penang | '11/03 - '17/09 | 3051 | 146651 | 2067 (45%) | 0.372 | 0.048 |

The number of users, tweets, and cells provide some information on the study scale of each city. The cases of Bangkok and Jakarta had the largest study area; Penang and Hongkong had the smallest study area. Although the study area of Tokyo case was small, it had the highest density of active users in this study (about six users/cell), following by Manila case (about three users/cell); the Osaka case and Jakarta case had about two users/cell. In terms of the density of tweets, the Tokyo case had the highest density (400 tweets/cell), following by Jakarta and Manila (both were 280 tweets/cell). These comparisons served only for providing some ideas about the data which would be used in the following analyses.

The results of modularity shows the highest modularity during the heuristic process of community detection. The spatial modularity results are about 0.35, except for Tokyo and Osaka are slightly higher (0.45 and 0.61, respectively). Although 0.35 is not a low value of modularity value, this indicates that the interaction between communities exists, i.e. the locations between activity spaces also share some users. The temporal modularity results are lower than spatial modularity, which is about 0.05 (Tokyo and Osaka are about 0.10). This suggested that the activity times were dominated by some significant groups of people, but there exists some random people who normally tweets on different hours also tweeted within the activity period. In other words, the groups of people are not completely distinct.

### 3.2 Activity spaces

Figure 2 shows the distribution of activity spaces of the eight cities. While the analysis did not consider the spatial continuous constraints (the spatial neighboring relationships between cells), most of the communities did show clear spatial boundaries. For instance, Osaka was separated into two major communities, the north and south parts, which resemble the Osaka City and Sakai City administrative boundaries; both Jakarta and Manila were separated into six major communities with clean polygon shapes; Bangkok was separated into five major communities that were scattered as a result of terrain constraints. This indicated that large and similar groups of people visited similar parts of the city. Although the transportation facilities are ready to connect the locations within these cities, people tend to go somewhere similar for their daily life activities. In other words, this result shows the boundary of human movements for daily activity.

While the transportation system is the key to connect places, it is also a key to location partitioning. The lines in figure 2 show the major street network and the railway/subway network. The density of the transportation network indicates the density of human activity, e.g. in the Hongkong and Bangkok cases, the cells that were identified as part of the major communities were scattered around the transportation system. The transportation system in the eight cities shows a radiation shape that concentrates at a city core area, and the density decreases with the distance to the city core. The shape of the major activity spaces in some cities, including Tokyo, Singapore, Bangkok, and Manila, sliced the city core into pieces and extended each piece to remote or sprawling areas. In other words, each of the communities has a part in the city core area, and a larger area of the sprawling area. While the core area is mixed by different groups of users, the spatial structure is more complex and thus has less effect on the formation of community; the structure of communities is then being dominated by the structure of the non-core areas, which have a clear separation of users. This situation leads to the results that the shapes of communities were separated into parts starting from the remote areas and the boundaries meet each other near the city core area.

In addition, the results from the three cities including Hongkong, Singapore, and Penang, showed clear separations of activity spaces caused by political boundaries, including the administrative boundary between Hongkong and Shenzhen, and the country boundary between Singapore and Johor; natural boundaries, including the Penang Strait between Penang Island and Seberang Perai, and the Victoria Harbour between Hongkong Island and Kowloon Peninsula. Although the

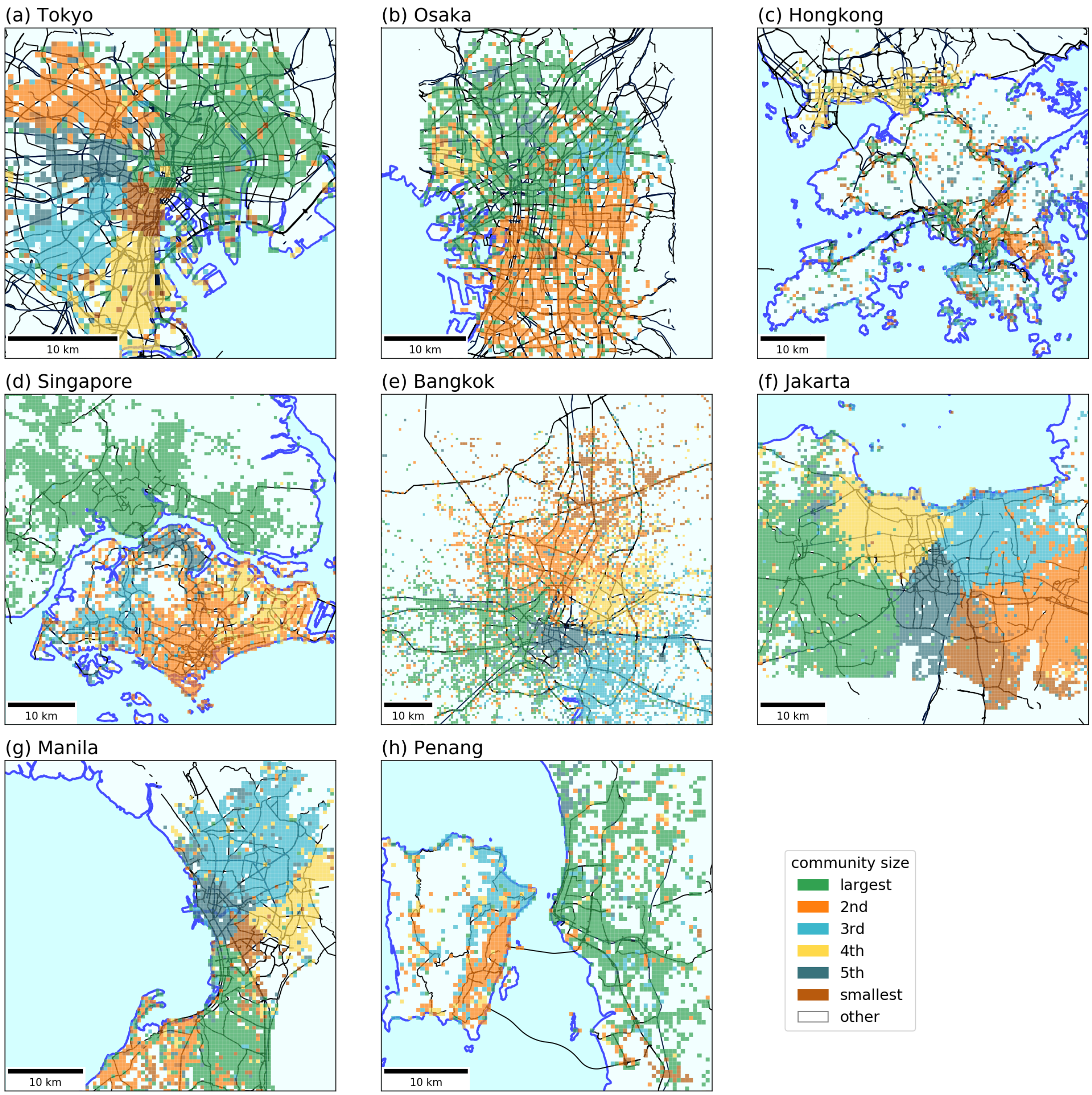


Figure 2: The activity spaces in the eight cities.

spatial interactions between these areas were dense because of the daily movements of cross-regions or cross-border commuters, the formation of activity spaces was strongly affected by political or natural boundaries.

### 3.3 Activity times

Figure 3 shows the distribution of the activity times of the eight cities. In general, the result shows a common cyclic pattern for each city in a weekly and daily basis. The weekdays had a pattern while the Saturday and Sunday had a different pattern. Each of the cities had its own daily cyclic pattern. For the example of Tokyo, during a weekday, a group of people started to be active from 05:00 to 09:00 in the morning; another group of people became active from 10:00 to 16:00, and were replaced by another group of people during 17:00-18:00; the activity on Twitter platform is then be dominated by another group of people starting from 19:00 to 04:00 of the next weekday. This result suggests that the major group of active users on a social media platform is different from hours to hours, and is different between weekends and weekdays.

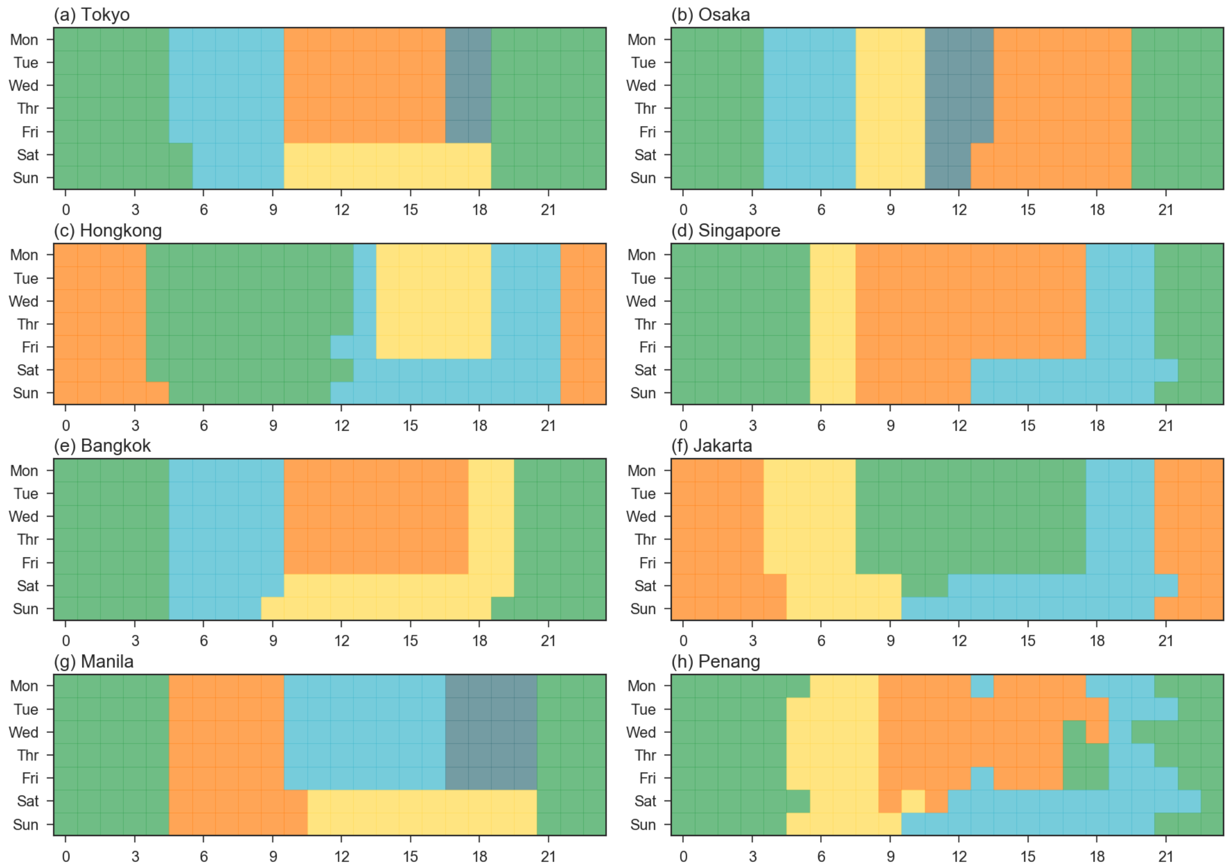


Figure 3: The activity times in the eight cities.

Based on Figure 3, we can observed that the activity times were slightly different for the varying cities. While they were different, some similar pattern could be observed. For instance, a group of people was active during night hours until the next early morning, which people were probably night owls; a group of people was active on social media in the morning for everyday (until about 9 a.m., with exception for Hongkong that the temporal range is extended to about 12:00), which might be the people who may Tweets during their commuting (weekdays) or breakfast time (weekends); during weekdays, a group of people was active from about 09:00 to 17:00, which time range is the working hours; for most of the cities (except for Osaka) a short period after the aforementioned period was occupied by another group of people who tweet during their dinner hours or on their way back home; for weekends, most of the day time (about 10:00 to 19:00) was occupied by a group of people, who might also be active during the dinner hours of weekdays.

Although Twitter data can only reflect the activity of users on social media, the time of tweets as an external information of being active also captures the daily habits of people in a city.

## 4 Discussion and conclusion

Based on the user-tweets records, this study presented a framework for delineating activity space and time. The activity spaces can be extracted based on the intensive interactions between locations, whereas the externality of the time of tweets can be used to reveal the daily activity pattern of cities (activity times). This study shows the activity spaces and activity times of eight cities located in East Asia and Southeast Asia. This study presented a framework for delineating activity spaces and time, and provided a novel perspective for representing the spatial and temporal patterns of the daily life in a city.

The spatially continuous community with clear separation indicates that a city can be divided into separated parts by the person who lives within it. Places can be delineated based on the relationships of frequently visited by the same group of people. Previous study by (Jiang and Miao, 2015) delineated the natural cities from a continent scale using the TIN and HT-index framework based on social media dataset. This study presented a way to delineate the activity space, which is a series of places within a city. In the results of the eight cases, the segments of areas showed clear boundaries between the activity spaces, while the spatial continuous properties were not considered in the analysis. These activity spaces indicated that some groups of people were dominating the activity within some distinctive and segmented regions. In most of the cases, the activity spaces split the city core into pieces and sprawl to the outer ring of the city. In other words, the core of a city is often mixed with a different group of people, and thus it is a emerging area that different activity spaces interacted with each other. Transportation system plays an important role in the daily activity of people who live within a city. It connects places and provides ways of commuting from where people stay to where they work. In the case of this study, although the transportation infrastructures were ready for connecting places, people tended to work and stay at somewhere nearby that was affordable to reduce the cost of movement (Hickman and Banister, 2015; Hamiduddin, 2017) or to avoid congestion. This situation might be reflected in the result that clear boundaries within the cities.

The boundaries of activity spaces resemble the boundary of natural divisions and administrative borders. The natural boundaries, such as straits (Hongkong, Singapore and Penang cases), and mountains (Hongkong) separated the activity of people. This may be because the natural boundaries will raised the cost of movement naturally. On the other hand, the within-city administrative border (for the cases of Tokyo, Osaka, Singapore, Manila, and Jakarta), and the between-countries/cities (e.g. Hongkong and Singapore) also become a key in delineating the activity space. This may because the places were planned for different land uses and managed by different local governments; the cost of cross-national border travels in terms of money and time are usually higher than within nation travels.

The delineated activity spaces showed the distribution of people and the boundary of the movement of most people. This information is useful for various types of managements, such as infectious disease controls (Huang et al., 2013; Wen and Chin, 2015), marketing and advertisement targeting (Goss, 1995; Hofstede et al., 2002), and resources allocations for maximizing the accessibility (Katsaros et al., 2003; Tsai et al., 2012). The framework in this study used only the records of social media that included timestamps and geotags from Twitter platform. In addition, the analysis procedures used open-source packages. This means that the dataset and framework are ready for applying to other cities and to be used for the aforementioned purposes.

The results of activity times show the collective lifestyle pattern of people in each city. Although the records of tweets represent only the activities on social media, the externality of the time of tweeting, e.g., the time when a group of people started to tweet may represent their free time, and the time when this group of people stopped tweeting indicated that they are working or sleeping. This study make use of the externality of the tweet records to capture the lifestyle of a city, and identified the activity times , i.e., the groups of hours that were dominated by the same groups of people, respectively. Each city has its own lifestyle, and which is slightly different from other cities. E.g., the commuters of some cities wake up earlier than others; people in some of the cities leave their offices later than others; the weekend activity pattern of each city are different. This study demonstrated a novel perspective of presenting the lifestyle of people in a city.

This study has several limitations and which requires further investigation. First, this study only includes the dataset of the activity on the Twitter platform, which may not be the major social media for some of the cities. Using the geolocated datasets from the major social media for each city may get a more accurate result, but which is not feasible due to the privacy policy for most social media. Second, people tended to spend their time skimming on the tweets they followed, rather than posting tweets all time. This leads to the situation that it is difficult to generate a full daily movement trajectory for each user. Therefore, this study used a relatively longer period of dataset to analyze the activity space as a collective result of the people who live within the study area. Third, the underlying information, such as land-use pattern, socio-demographic distribution, infrastructure, and natural environments were not considered in this study. These underlying information can provide in-depth understanding of the distribution of human activity in time and space. Fourth, this study combined all data within the study area in all years. The movement of people may be a result of commuting (daily) or house moving (permanent), thus the analysis results might be diluted. This study assumed that most people do not move their houses within the period of dataset, thus the effect of the house moving is not significant. Future study is suggested to differentiate commuting and house moving through addition models before the community detection analysis.

## References

M. Batty. Building a science of cities. *Cities*, 29:S9–S16, 2012.

J. I. Blanford, Z. Huang, A. Savelyev, and A. M. MacEachren. Geo-located tweets. enhancing mobility maps and capturing cross-border movement. *PloS one*, 10(6):e0129202, 2015.

V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

L. A. Brown and J. Holmes. The delimitation of functional regions, nodal regions, and hierarchies by functional distance approaches. *Journal of regional science*, 11(1):57–72, 1971.

W. H. Burt. Territoriality and home range concepts as applied to mammals. *Journal of mammalogy*, 24(3):346–352, 1943.

G. Cao, S. Wang, M. Hwang, A. Padmanabhan, Z. Zhang, and K. Soltani. A scalable framework for spatiotemporal analysis of location-based social media data. *Computers, Environment and Urban Systems*, 51:70–82, 2015.

P. Ciucci, L. Boitani, F. Francisci, and G. Andreoli. Home range, activity and movements of a wolf pack in central Italy. *Journal of Zoology*, 243(4):803–819, 1997.

C. Comito, D. Falcone, and D. Talia. Mining human mobility patterns from social geo-tagged data. *Pervasive and Mobile Computing*, 33:91–107, 2016.

C. J. Farmer and A. S. Fotheringham. Network-based functional regions. *Environment and Planning A*, 43(11): 2723–2741, 2011.

Q. Ge and D. Fukuda. Updating origin–destination matrices with aggregated data of GPS traces. *Transportation Research Part C: Emerging Technologies*, 69:291–312, 2016.

R. G. Golledge and R. J. Stimson. *Spatial behavior: A geographic perspective*. Guilford Press, 1997.

J. Goss. "We know who you are and we know where you live": The instrumental rationality of geodemographic systems. *Economic Geography*, 71(2):171–198, 1995.

T. Hägerstraand. What about people in regional science? *Papers in regional science*, 24(1):7–24, 1970.

I. Hamiduddin. Journey to Work Travel Outcomes from 'City of Short Distances' Compact City Planning in Tübingen, Germany. *Planning Practice & Research*, pages 1–20, 2017.

B. Hawelka, I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti. Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3):260–271, 2014.

R. Hickman and D. Banister. New household location and the commute to work: Changes over time. *R. Hickman, D. Bonilla, M. Givoni, e D. Banister, International Handbook on Transport and Development*, pages 73–86, 2015.

F. T. Hofstede, M. Wedel, and J.-B. E. Steenkamp. Identifying spatial segments in international markets. *Marketing Science*, 21(2):160–177, 2002.

C.-Y. Huang, T.-H. Wen, and Y.-S. Tsai. FLUed: A novel four-layer model for simulating epidemic dynamics and assessing intervention policies. *Journal of Applied Mathematics*, 2013, 2013.

O. Järv, R. Ahas, and F. Witlox. Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*, 38:122–135, 2014.

M. Jendryke, T. Balz, S. C. McClure, and M. Liao. Putting people in the picture: Combining big location-based social media data and remote sensing imagery for enhanced contextual urban information in Shanghai. *Computers, Environment and Urban Systems*, 62:99–112, 2017.

B. Jiang and Y. Miao. The evolution of natural cities from the perspective of location-based social media. *The Professional Geographer*, 67(2):295–306, 2015.

B. Jiang, D. Ma, J. Yin, and M. Sandberg. Spatial distribution of city tweets and their densities. *Geographical Analysis*, 48(3):337–351, 2016.

D. Katsaros, A. Nanopoulos, M. Karakaya, G. Yavas, Ö. Ulusoy, and Y. Manolopoulos. Clustering mobile trajectories for resource allocation in mobile environments. In *International Symposium on Intelligent Data Analysis*, pages 319–329. Springer, 2003.

M.-P. Kwan. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. *Transportation Research Part C: Emerging Technologies*, 8(1-6):185–203, 2000.

M.-P. Kwan and J. Lee. Geovisualization of human activity patterns using 3D GIS: a time-geographic approach. *Spatially integrated social science*, 27, 2004.

J. H. Lee, A. W. Davis, S. Y. Yoon, and K. G. Goulias. Activity space estimation with longitudinal observations of social media data. *Transportation*, 43(6):955–977, 2016.

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

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

A. Lloyd and J. Cheshire. Deriving retail centre locations and catchments from geo-tagged Twitter data. *Computers, Environment and Urban Systems*, 61:108–118, 2017.

D. Ma, I. Omer, T. Osaragi, M. Sandberg, and B. Jiang. Why topology matters in predicting human activities. *Environment and Planning B: Urban Analytics and City Science*, 2018. doi:10.1177/2399808318792268.

J. Miranda, N. Mäkitalo, J. Garcia-Alonso, J. Berrocal, T. Mikkonen, C. Canal, and J. M. Murillo. From the Internet of Things to the Internet of People. *IEEE Internet Computing*, 19(2):40–47, 2015.

M. E. Newman. Analysis of weighted networks. *Physical review E*, 70(5):056131, 2004.

Q. C. Nguyen, S. Kath, H.-W. Meng, D. Li, K. R. Smith, J. A. VanDerslice, M. Wen, and F. Li. Leveraging geotagged Twitter data to examine neighborhood happiness, diet, and physical activity. *Applied Geography*, 73:77–88, 2016.

S. Schönfelder and K. W. Axhausen. Activity spaces: measures of social exclusion? *Transport policy*, 10(4):273–286, 2003.

T. Shelton, A. Poorthuis, and M. Zook. Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information. *Landscape and Urban Planning*, 142:198–211, 2015.

Y. Shen and K. Karimi. Urban function connectivity: characterisation of functional urban streets with social media check-in data. *Cities*, 55:9–21, 2016.

S. Sobolevsky, P. Kats, S. Malinchik, M. Hoffman, B. Kettler, and C. Kontokosta. Twitter Connections Shaping New York City. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

Y. Song and H. J. Miller. Beyond the boundary: New insights from inside the space-time prism. In *Space-Time Integration in Geography and GIScience*, pages 189–209. Springer, 2015.

Y.-S. Tsai, P. C.-I. Ko, C.-Y. Huang, and T.-H. Wen. Optimizing locations for the installation of automated external defibrillators (AEDs) in urban public streets through the use of spatial and temporal weighting schemes. *Applied Geography*, 35(1-2):394–404, 2012.

M. Šálek, L. Drahníková, and E. Tkadlec. Changes in home range sizes and population densities of carnivore species along the natural to urban habitat gradient. *Mammal Review*, 45(1):1–14, 2015.

T.-H. Wen and W. C. B. Chin. Incorporation of Spatial Interactions in Location Networks to Identify Critical Geo-Referenced Routes for Assessing Disease Control Measures on a Large-Scale Campus. *International journal of environmental research and public health*, 12(4):4170–4184, 2015.

L. Wu, Y. Zhi, Z. Sui, and Y. Liu. Intra-urban human mobility and activity transition: Evidence from social media check-in data. *PloS one*, 9(5):e97010, 2014.

M. Yuan and A. Nara. Space-time analytics of tracks for the understanding of patterns of life. In *Space-Time Integration in Geography and GIScience*, pages 373–398. Springer, 2015.

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

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

## Notes on contributor

**Wei Chien Benny Chin** (B.Sc., M.Sc., PhD) is currently a Postdoctoral Researcher with the Singapore University of Technology and Design under SGP Cities. He is a geographer and a Geographical Information Scientist. His research interest includes GIScience, analysis of spatial structures and complex network analysis. He is participating in projects associated with complex human movement and spatial structures' horizontal and vertical dimensions.