Entangled Footprints: Understanding Urban Neighbourhoods by Measuring Distance, Diversity, and Direction of Flows in Singapore

Qingqing Chen¹, I-Ting Chuang², Ate Poorthuis³

¹Department of Geography, College of Arts and Sciences, University at Buffalo, 105 Wilkeson, Buffalo, New York, United States, 14261

²School of Architecture and Planning, University of Auckland, New Zealand 1010

³Department of Earth and Environmental Sciences, KU Leuven, Celestijnenlaan 200E, Leuven, Belgium, 3001

Published as: Chen, Q., Chuang, I.-T. and Poorthuis, A. (2021) 'Entangled footprints: Understanding urban neighbourhoods by measuring distance, diversity, and direction of flows in Singapore', *Computers, Environment and Urban Systems*, 90, p. 101708.
doi:10.1016/j.compenvurbsys.2021.101708

Abstract

Traditional approaches to human mobility analysis in Geography often rely on census or survey data that is resource-intensive to collect and often has a limited spatio-temporal scope. The advent of new technologies (e.g. geosocial media platforms) provides opportunities to overcome these limitations and, if properly leveraged, can yield more granular insights about human mobility. In this paper, we use an anonymized Twitter dataset collected in Singapore from 2012 to 2016 to investigate this potential to help understand the footprints of urban neighbourhoods from both a spatial and a relational perspective.

We construct home-to-destination networks of individual users based on their inferred home locations. In aggregated form, these networks allow us to analyze three specific mobility indicators at the neighbourhood level, namely the distance, diversity, and direction of urban interactions. By mapping these three indicators of the spatial footprint of each neighbourhood, we can capture the nuances in the position of individual neighbourhoods within the larger urban network. An exploratory spatial regression reveals that socio-economic characteristics (e.g. share of rental housing) and the built environment (i.e. land use) only partially explain these three indicators and a residual analysis points to the need to explicitly include each neighbourhood's position within the transportation network in future work.

Keywords: Human Mobility · Location-Based-Services (LBS) · Activity Space · Twitter · Spatial Networks · Urban Neighbourhoods · Singapore

¹ Email address: qchen47@buffalo.edu

² Email address: iting.chuang@auckland.ac.nz

³ Email address: ate.poorthuis@kuleuven.be

1. Introduction

Human mobility has been a long-standing interest in Geography. This interest has intensified in the last two decades with the advent of new technologies (e.g. GPS, geosocial media) allowing increasingly finely-grained analyses of mobility patterns, with applications ranging from urban planning to epidemiology (Kraemer et al., 2020; Shi et al., 2020; B. Zhang et al., 2020). Of particular interest to urban researchers has been the relationship between human mobility and the underlying urban structure that can be revealed by mobility patterns (Liu et al., 2020; Saxon, 2020; Sun et al., 2016; Zhong et al., 2017). The perennial discussion about the degree of centrality or polycentricity that urban systems exhibit is a clear example of this interest.

We take this promising intersection of urban structure and human mobility analysis as a starting point to further unpack the constituent parts of the urban system. Concretely, instead of taking a city as the unit of analysis and determining the degree of its centrality, our analysis focuses on a finer-grained spatial scale: individual neighbourhoods. To operationalize these neighbourhoods, we tessellate geographical space in a regular grid of hexagonal cells that we refer to as ‘neighbourhoods’ in the remainder of the paper. This allows us to address questions regarding specific neighbourhoods and their role within the system. In other words, we are trying to understand where and how people move within and between neighbourhoods. Which specific neighbourhoods operate as functional centres? Are particular neighbourhoods dependent on other ‘central’ locations, or are they principally self-contained?

These questions – and others alike – have direct urban policy implications, perhaps best exemplified in the recent uptake of ‘20-minute’ planning goals pursued by cities all over the world (e.g. Capasso Da Silva et al., 2020; Moreno et al., 2021; Weng et al., 2019). We argue that a human mobility lens on urban neighbourhoods can generate valuable insights for these

planning questions, especially when combined with the analytical latitude offered by new types of mobility data that provide more flexible definitions of the size and shape of neighbourhoods.

To illustrate this potential, we put the proposed method to the test in an applied case study focused on Singapore. Singapore has a specific history of master planning which has explicitly focused on decreasing the reliance on the central downtown core through the formation of a polycentric structure based on satellite town centres. Previous studies have determined this masterplan to be more or less successful (and the city thus to be either more monocentric or more polycentric), depending on the analytical perspective and datasets used (Dale, 1999; Han, 2005; Zhong et al., 2017).

For our analysis, we utilize a de-identified dataset that consists of all tweets sent between 2012 and 2016 from Singapore (Chen and Poorthuis, 2021). Although we use Twitter data in our case, the methodology proposed here can be applied to data from other Location-Based Services (LBS), such as Instagram, cell phone data, Weibo, Facebook, etc. As such, we consistently refer to the used dataset as ‘LBS data’ in the remainder of the paper. Crucially for the analysis at hand, such datasets allow for the potential to draw connections between each user’s ‘home’ location and all other locations they have visited across the city.

In concrete terms, we calculate three core metrics for each neighbourhood for both incoming and outgoing flows: the distance of flows (i.e. the distance between the source and sink of a flow or interaction); the diversity of flows (i.e. for incoming flows: the diversity of sources; for outgoing flows: diversity of sinks); and the cardinal direction of flows (i.e. the average orientation of incoming or outgoing flows). We will present these three metrics through a series of visualizations and use them to reveal specific patterns in the spatial footprints of neighbourhoods in Singapore.

In doing so, we illustrate four specific methodological affordances that are created by the application of human mobility data on the structure of urban neighbourhoods. First, it gives

us the flexibility of creating neighbourhood boundaries and scales tailored to the research questions at hand, allowing analyses at a very granular level. Second, it provides an innovative way to capture aspects of the diversity of a neighbourhood by considering the origin of its visitors, instead of only the characteristics of its residents or neighbourhood facilities. Third, mobility data complements understanding neighbourhoods based on a residential census or survey research with a view based on actual flows of people. Fourth and related, we can perform analyses for both inflow and outflow of people, each offering a different perspective on the same neighbourhood.

In the next section we will discuss the current state-of-the-art of the use of human mobility data in urban analysis in more detail. After that we will describe the details of the dataset and the methodology used to construct the three mobility metrics (i.e. *Distance*, *Direction*, and *Diversity*), and the regression model for understanding the relationship between these metrics and socio-economic and built-environment characteristics. We discuss the results through a series of visualizations and specific neighbourhood examples before discussing the core findings and potential future work.

2. Related Work

The mechanisms behind people's movement patterns are important in a wide range of applications, from predicting the spread of diseases, and urban planning to transportation networks and health geography (Campbell et al., 2021; Kraemer et al., 2020; Xia et al., 2020). Conventional approaches often use census or survey data to analyze individual travel patterns (Ewing and Cervero, 2001; Vilhelmson, 1999). Although such approaches can obtain detailed demographic information and explicit itineraries of travel movements, they have a limited spatio-temporal scope and are resource-intensive and time-consuming to collect.

To overcome these limitations, alternative data sets, such as bank notes (Brockmann et al., 2006), public transport transactions (Gong et al., 2017; Xia et al., 2018), WiFi probe data

(Poucin et al., 2018; Traunmueller et al., 2018), GPS logs (Kong et al., 2018; Wu et al., 2019), and mobile phone data (Li et al., 2019; Zhao et al., 2019) have made their entry in the field due to their increased spatio-temporal resolution and scale. Studies using such data often focus on finding mathematical patterns and regularities and use system-level metrics such as mobility rate, radius of gyration, diversity of destinations, potential power-laws, and Zipf's law (Alessandretti et al., 2020; Calabrese et al., 2013; Shi et al., 2020; Traunmueller et al., 2018).

With the advent of geosocial media platforms, such as Twitter, Facebook, and Weibo, phone-based location data is increasingly used for analysing human mobility, as well as investigating its relation to urban space from temporal, spatial and social perspectives (Li et al., 2020; Martí et al., 2019; Shaw and Sui, 2020; Wang et al., 2019). For example, the mobility patterns that we collectively create through our daily activities can help to classify urban space, as Lee et al. (2013) did based on behavioural patterns inferred from geotagged-tweets. Similarly, Sun et al. (2016) used a Location-Based Social Network (LBSN) 'check-in' dataset to detect the 'true' centre of a city with different cluster detection algorithms. Zhang et al. (2020) combined social media 'check-in' data and street view images to analyze human activity patterns and perceptual features of places in order to detect unassuming-yet-popular places, which is helpful for understanding and improving urban vibrancy (Kang et al., 2020). These new types of datasets and approaches, while often initially applied to a single case study, have the potential to be scaled up to a wide variety of countries and cities. This is an enticing promise of big data since mapping human mobility patterns not just for a single city, but at the scale of a system of cities or even an entire country is a powerful, if still underused, analytical lens (Nelson and Rae, 2016; Rae, 2009).

Performing urban analysis at this scale is not a new academic pursuit in and of itself. For example, urban geography and regional science have a long history of studying the polycentric nature of urban systems from both theoretical and empirical perspectives (Burger

and Meijers, 2012; Kloosterman and Musterd, 2001; Liu and Liu, 2018; van Meeteren et al., 2016). As van Meeteren et al.'s (2016) bibliometric analysis of this discourse shows, the polycentricity concept has been approached from many different angles (from intra- and inter-urban to regional perspectives). What many of these perspectives share is a reliance on data that is ultimately derived from employment or business data (e.g. commuting, employment data, business networks etc.).

Trying to move beyond the limitations of these more conventional datasets, researchers have started to evaluate the potential of new data sources for studying polycentricity. For instance, Liu et al. (2015) analyzed the city structure of Shanghai using taxi trip data, while Fang et al. (2017) used mobile phone operator data to identify polycentric clusters in Shenzhen. Other examples of using both mobile phone and transportation data abound (Roth et al., 2011; Zhong et al., 2017). Although some of these studies explicitly concentrate on a search for universal laws (cf. Schwanen, 2016), they have also enabled an increased focus on smaller, intra-urban patterns and a move away from looking at spatial structures solely based on an economic or employment (i.e. commuting data) perspective.

Within the polycentricity literature, Singapore has received ample attention as a case study. For example, Han (2005) used property sales transactions to map global and local spatial autocorrelation indices to identify spatial clusters. The result showed a redistribution of condominium projects towards the outer planning regions. Similarly, Hou (2019) found some evidence for a polycentric structure when looking at non-work trips of Singapore's elderly population based on a travel survey. By using the same travel survey over a long period, Zhong et al. (2017) also identified a pattern of increased polycentricity. However, they simultaneously observed a parallel process in which the downtown core is growing in importance as well.

Much of the aforementioned work regards a city, or even a country, as the unit of analysis and attempts to determine whether the system as a whole is polycentric by studying

commuting flows, job dispersion, the land use mix, the density or popularity of a location and so on. In contrast, we disaggregate the whole system into individual neighbourhoods and study the human footprints at this more granular scale with the mobility data.

To operationalize our research question – where and how people move within and between neighbourhoods – we design three specific metrics - *Distance*, *Diversity* and *Direction* (discussed in detail in the methodology section) - that allow us to understand how one neighbourhood relates to or fits within the larger urban system. Importantly, because of the LBS data used in this study, we can flexibly choose the neighbourhood size to conduct this analysis at a relatively granular level.

Changing the spatial scale to the neighbourhood also enables us to compare metrics about the urban structure with socio-economic characteristics of each neighbourhood. In this paper, we use this affordance to explore potential socio-economic covariates that help explain specific patterns in the distance and diversity of mobility flows. The impact of individual characteristics on a person's mobility patterns has been well-studied. For example, Kwan (1999) explored the gendered nature of social life and found that women experience more spatial constraints in their daily mobility patterns than men. Xu et al. (2018) compared mobility patterns derived from mobile phone data with a user's socio-economic status, proxied through real estate prices, and found that residents in Singapore neighbourhoods with high real estate prices tend to travel shorter distances. In their comprehensive analysis, Limtanakool et al. (2006) analyzed the impact of a range of socio-economic variables (gender, age, income, educational attainment etc.) and land use on transportation mode choice for different types of trips using a conventional travel survey. Using Twitter data instead, Huang and Wong (2016) also found a relationship between land use and socio-economic variables on the one hand and Twitter user's activity spaces on the other hand. Against this backdrop, we employ a (spatial) regression model in this paper to formalize the link between our mobility metrics (e.g. distance

and diversity) and various socio-economic characteristics of the neighbourhood, as is discussed in the next section.

3. Methodology

3.1 Case Study & Data

To illustrate our proposed approach, we take Singapore (see Fig. 1 for geographic context) as a case study. Singapore has a specific urban planning history, in which it has focused on decreasing the reliance on the central downtown core and transforming itself into a polycentric city. Singapore has experienced tremendous urban development and growth since its independence in 1965, in large part due to the specific master planning and intervention of its government. After an initial focus on improving the living environment and meeting the housing demand of a booming population, the city's Urban Redevelopment Authority (URA), together with other agencies such as the Housing and Development Board (HDB), started working towards an increased spatial decentralization from the 1980s onwards (Han, 2005). This is highlighted in the 1991 Concept Plan⁴ that outlined a blueprint for a hierarchy of decentralized business hubs. The downtown CBD would be accompanied by four new regional centres, several sub-regional centres, and fringe centres. The four new regional centres were to be directly integrated into the transit system and included local industrial and business parks to bring employment closer to home and thus reduce congestion.

⁴ [The 1991 Concept Plan of Singapore](#) - Living the Next Lap: Towards a Tropical City of Excellence.

Figure 1. Map of study area, Singapore. Labelled places are reference points to orient the reader. 1) City Hall, photography by the author; 2) East Coast, photography by Tamaki Hayashi; 3) National University of Singapore, photography by Melvin Yap; 4) Sentosa, photography by William Cho; 5) Changi Airport, photography by Geoff Whalan; 6) Tampines, photography by Dickson Phua; 7) Central Water Catchment, photography by CW Gan; 8) Woodlands, photography by L-W; 9) Serangoon Centre, photography by Ellen Forsyth; 10) Jurong Gateway, photography by Edsel Little.

As stated earlier, instead of focusing on the city as a whole, we investigate Singapore's individual neighbourhoods and their relationship within the larger urban system. To do so, we use an LBS dataset to construct a network of connections between users' home locations and all other locations visited in the city. Considering the varying shapes and sizes of administrative regions, and balancing privacy and ethics with keeping sufficient observations in individual neighbourhoods for subsequent analyses, we have chosen to use a grid cell with 750m resolution (cf. Chen and Poorthuis (2021) for a discussion of this scale) as the unit of analysis. As such, we divide Singapore into a total number of 1942 hexagonal grid cells and refer to each cell as a neighbourhood (see also Appendix A.1). Each LBS data point, originally consisting of a relatively precise geotag, is aggregated to the constructed neighbourhoods.

The de-identified LBS dataset used for this analysis consists of around 20.9 million tweets created by approximately 130.3k users sent from Singapore between 2012 and 2016 (Chen and Poorthuis, 2021). Each data point includes three attributes - a unique identifier for the user; a unique identifier for the spatial neighbourhood for the data point; and a timestamp that reflects the time the data point was created. The overall distribution of the number of users and the number of tweets is presented in Appendix A.2. It should be noted that the dataset used here is likely to be biased and not representative for the population at large. Nonetheless, it does open possibilities beyond the scalar limitations of traditional surveys and census data and can still be used to uncover insights into human mobility patterns.

To account for bots or spam users, we remove the top 0.1% of the most active users based on the frequency of tweets. Conversely, to eliminate sporadic users, we remove users with less than 10 tweets during the study period, or with only tweets sent from a single location, as they do not provide sufficient data for mobility analysis. Although these filtering steps might remove some ‘real’ users, this is not necessarily an issue for the subsequent analysis when we exercise caution with claims that generalize to the entire population. After the final filter step, the dataset consists of around 20.07 million tweets sent by 128,554 users.

3.2 Distinction between Locals and Visitors

To analyze the spatial footprints in each neighbourhood, we first distinguish between ‘locals’ and ‘visitors’. Locals refer to users with homes within that specific neighbourhood or hexagonal grid cell and visitors refer to all other users with homes outside of the grid cell. As such, an essential step is to identify a user’s home location based on the data available from the social media platform.

To operationalize, we use the same ensemble approach proposed by Chen and Poorthuis (2021) which is based on the consensus between four different embedded algorithms. As such, we can determine 21,863 (17%) users’ home locations (see Appendix A.3). Users whose home

location cannot be inferred are removed from the dataset. It is important to point out that although this is a small subset of users, the ensemble approach is meant to prevent false positives to avoid additional bias (by users ‘wrongly’ assigned to a location) from entering our analysis. With the identified home locations, we can distinguish visitors and locals in each neighbourhood, and build the home-to-destination networks of each user. Specifically, for each neighbourhood, we can pinpoint where the visitors come from (i.e. inflow network) and where the locals go out to (i.e. outflow network), which yields a total of 8,551,017 trajectories. It should be noted that these trajectories do not correspond to the exact trip a user might have made to get to their destination as intermediate points from which no tweets are sent are excluded. All subsequent analyses are applied to both inflow and outflow networks.

In this aggregated form, the ‘home-to-destination’ network allows us to identify the characteristics of each neighbourhood’s aggregated footprints. This network forms the basis for the three mobility metrics and the subsequent regression analysis, which we will further detail in these sections. Figure 2 illustrates the overall research design.

Figure 2. Research design overview

3.3 Distance of Flows

Based on the home-to-destination networks, we measure both incoming and outgoing travel distances for each neighbourhood. Considering issues of privacy and ethics, we solely keep neighbourhoods with more than five visitors or locals in the subsequent analysis. As such, we are left with 1,196 neighbourhoods in the inflow network and 482 neighbourhoods in the outflow network. For each neighbourhood, the incoming distance is estimated based on trajectories of visitors between the centroids of their home locations to the centroid of that neighbourhood, and the outgoing travel distance is estimated based on the trajectories of locals between the home neighbourhood centroid and the centroids of all other neighbourhoods they've visited. We utilize the mean incoming and outgoing travel distances per neighbourhood in the visualizations and spatial regression modelling in the following sections.

3.4 Diversity of Flows

To operationalize our metric of diversity, we take inspiration from ecology (Morris et al., 2014). Concretely, we measure the inflow and outflow diversity according to the variety and intensity of the origin neighbourhoods (for inflow diversity) and destination neighbourhoods (for outflow diversity). We consider users from a single neighbourhood to be similar to how ecologists use the concept of a 'species'. However, the 750-meter grid cells would yield too many 'species' with too few observations for each neighbourhood. To address this issue, we create unique and dynamic concentric sectors for each neighbourhood (see Fig. 3). This allows us to increase the size of the sectors as they further move from the neighbourhood. Sectors further away from the neighbourhood generally attract fewer visitors so their larger size compensates for this. The underlying idea here is that the difference between a visitor from 1km or 5km away can be notable, but the difference between a visitor from 20km or 25km away is much less significant. We consider visitors with home locations within the same sector or locals visiting the same sector as the same 'species'. To create sectors around the centroid

of a grid, we first draw eight buffers, clipped to the land boundary of Singapore, with radii of 1km, 3km, 5km, 7km, 10km, 20km, 30km, and 60km around the geographical centre of the grid (see Fig. 3a). Each grid cell is assigned to a buffer based on which buffer overlaps with its centroid.

Apart from distance alone, users coming from different directions (e.g. north, south, east, west, or centre) in the same distance band may represent distinct demographic or neighbourhood backgrounds. As such, we cut each buffer into four directions ('*top-left*', '*top-right*', '*bottom-left*', and '*bottom-right*'). We consider users within the first, walking-distance, buffer (1km) to all belong to the same neighbourhood, so refrain from cutting the first inner buffer (see Fig.3b). As such, the total number of created sectors for each grid cell can be calculated as follows:

$$4 * (n - 1) + 1$$

where n is the number of buffers used. In our case, we create a total of 29 ($n = 8$) circular sectors for each grid cell. These sectors create the foundation for the measurement of diversity in the subsequent section.

Figure 3. Concentric sectors of a grid cell (a) eight buffers around the geographical centre of a grid cell; (b) concentric sectors; (c) a selected grid cell; (d) examples of the innermost sector and a selected sector.

To measure diversity, we use Shannon’s diversity index (H) (Nagendra, 2002) as it is commonly used for species diversity and accounts for both richness and evenness of the species present. The Shannon’s diversity index (H) is calculated as follows:

$$H = - \sum_{i=1}^S p_i \ln p_i$$

where S is the number of “species” (i.e. number of sectors in our case) and p_i is the proportion of individuals that belong to species i . The higher the Shannon’s index (H) value is, the more diverse the origins of visitors are for inflow diversity. Conversely, for outflow diversity a higher H means that locals visit a more diverse set of neighbourhoods throughout the city.

As the larger, outer sectors have a higher likelihood of users originating there, we use the ‘user density’ (i.e. the ratio of users over the sector area) to replace the ‘number of users’ in p_i . We apply the Shannon’s index (H) to all neighbourhoods that have more than 100 tweets sent by at least 5 users, which yields a total of 1051 grid cells for this analysis.

3.5 Direction of Flows

Apart from the distance and diversity of flows, we add one final mobility metric – direction – to capture another aspect of mobility flows. Flows into or out of a neighbourhood can happen in all cardinal directions and with different intensities and distances. Visualizing the entire network of these connections is not very insightful for making neighbourhood-level inferences. This is why we create a single metric that summarizes the overall direction and strength of a flow to or from a neighbourhood, inspired by Tobler’s (1981) work on vector fields to display flow patterns.

We first aggregate the sectors of a grid into four zones, with each zone corresponding to a specific angle and vector, relative to ‘East’ (0°): the North East (NE) zone (45°), North West (NW) zone (135°), South West (SW) zone (225°), and South East (SE) zone (315°) (see

Fig. 4a). This simplification alleviates computational complexity while maintaining the overall flow orientation of a neighbourhood. For each vector, its direction represents where visitors come from or where locals go out to, and its magnitude indicates the number of people in that flow. The resultant vector - the sum of the four vectors - is used to represent the overall flow direction of each grid (see Fig. 4b). To visualize all resulting vectors uniformly on a single map, we log-transform the length of each vector (to prevent some arrows from being much longer than others). Furthermore, we also represent the number of users active in the grid with the width of the resultant vector arrows.

Figure 4. Analysis of footprints flow direction

3.6 Spatial Regression Analysis

To contextualize the mobility patterns in our analysis, we compare two of the aforementioned metrics – distance and diversity – with the socio-economic and built environment characteristics of each neighbourhood through a spatial regression analysis. We do so by

drawing on a set of publicly available datasets, including the HDB Resale Flat Prices data, HDB Property Information data, SLA Street Directory data, and LTA Train Station data. These datasets allow us to create six independent variables shown as follows:

- 1) Mean resale price per square meter;
- 2) Percentage of 1-Room and 2-Room rental flats;
- 3) Percentage of residential building area;
- 4) Percentage of industrial building area;
- 5) Distance to the nearest MRT station;
- 6) Distance to the central grid cell.

We aggregate the variables to individual neighbourhoods (i.e. with the same spatial resolution as neighbourhoods) and only keep neighbourhoods that have public housing (see Appendix A.5 for more details). It is important to stress that we employ this regression model not to predict travel distance or mobility diversity based on these independent variables but rather to provide an initial quantitative exploration of the potential underlying covariates that influence mobility patterns in Singapore.

4. Results and Discussion

4.1 Distance of Flows

Figure 5 visualizes the mean and standard deviation of the travel distance in each grid cell for both inflow and outflow networks. To provide consistency, we use an orange colour scheme for the inflow network and a purple colour scheme for the outflow network in this and all subsequent figures. The mean travel distance for the inflow network is about 10.3km (S.D. 6.1km) and its spatial pattern presents a clear core-periphery structure. Neighbourhoods with a large travel distance ($>10\text{km}$) generally appear around the edges of the city, notably in the west, east, and southern regions. Part of this can be explained due to Singapore's island geography, but on closer inspection neighbourhoods with such a high distance tend to be fairly specific.

In the east, the cluster with high inflow distance is home to Changi Airport (indicated with (5) in Fig. 1), while the area in the west is home to Singapore's harbour, a large industrial zone, as well as a university. Although the average distance travelled to these areas is high, so is its standard deviation. It indicates that, while people on average are willing to travel quite far to these areas, they also attract people from much closer by. This is different for the cluster in the South that is home to Sentosa (indicated with (4) in Fig. 1), the country's beach and resort destination, directly adjacent to the downtown core. Similarly, we see a high incoming travel distance but the standard deviation here is low, indicating that people consistently travel from far away to this area. These areas perform specific urban functions that are not (easily) found elsewhere within the city and do not have a high density of residential housing. As such, their relatively high incoming travel distance is indeed expected. Conversely, the area on the northern edge of the city – Woodlands (indicated with (8) in Fig. 1) – is also home to a large industrial estate but does have significant residential estates nearby. As such, its incoming travel distance is far lower than the areas previously discussed.

For the outflow network (see Fig. 5c/5d), the average travel distance is around 6.6 km (S.D. 2.4km), which is lower than that of the inflow network. This is because the outflow is based on people's home location and thus captures their overall mobility footprint, while the inflow includes all types of locations (residential and non-residential) thus also covers special urban functions that are travelled to over relatively large differences, but far less often than (nearby) daily errands.

Here we observe a pattern emerging in the peripheral areas where people travel longer distances than those living in the central areas. It is noteworthy that both incoming and outgoing flows around Serangoon Central (indicated with (9) in Fig. 1) have relatively short distances, indicating that the area might form a local centre that serves many needs for its own residents, but does not attract visitors from the outside.

(a) Distance of Flows: average incoming distance

(b) Distance of Flows: S.D. of incoming distance

(c) Distance of Flows: average outgoing distance

(d) Distance of Flows: S.D. of outgoing distance

Figure 5. Distance of flows (a) Spatial distribution of average inflow distance; (b) Spatial distribution of standard deviation of inflow distance; (c) Spatial distribution of average outflow distance; (d) Spatial distribution of standard deviation of outflow distance.

4. 2 Diversity of Flows

Figure 6 visualizes the diversity of flows for both the inflow and the outflow network. We observe a distinct difference in the spatial patterns between the inflow diversity (see Fig. 6a) and the outflow diversity (see Fig. 6b). Concretely, the city centre (broadly defined areas around label (9) in Fig.1) displays a high inflow diversity, which corresponds with its central urban function within Singapore. It also benefits from the road and public transportation networks that are built to converge on this area. In contrast, the outflow diversity in the area is relatively low, which indicates residents living near the centre might not have to travel widely to meet their daily essentials. This is similar to the phenomenon we observed with the travel distance in the previous section.

We also see high inflow diversity for Changi Airport (indicated with (5) in Fig. 1) and Sentosa (indicated with (4) in Fig. 1), which perform as leisure destinations in the city. This is in line with current urban policy as Singapore is actively developing the local urban function of Changi Airport and has recently opened Jewel Changi, a multi-use complex (e.g. gardens, attractions, retail, restaurants, hotels), signalling a new perspective on airports as hubs for urban planning beyond international travel.

Looking beyond central urban functions, we can zoom in on some regional centres in the city, as designated in the Concept Plan 1991. Tampines (indicated with (6) in Fig. 1) is one such centre and is also one of the most populous residential areas in Singapore. Tampines has modest scores of inflow diversity, indicating that it generally services residents from a limited set of (likely nearby) neighbourhoods. Its outflow diversity is relatively low as well, which indicates its residents can obtain their daily needs locally, corroborated by the relatively low travel distance in Figure 5c. It is a testament to the relative success of Tampines as a regional centre, with its provision of shopping malls, hawker centres, other retail amenities, as well as a sizable number of jobs.

For Woodlands, another regional centre (indicated with (8) in Fig. 1), we also observe low diversity in both inflow and outflow networks. This is especially pertinent in comparison with the outgoing travel distance in Figure 5c/d, where we find residents of this area travel comparatively far with marked deviations. Despite their high travel distance, the low diversity in Figure 6a/b indicates that they travel to a relatively limited number of neighbourhoods and that these neighbourhoods are likely to be in the same cardinal direction. In practice, this is a first indication that, beyond their local neighbourhood, residents in Woodlands primarily focus on the city centre. It is this notion that we will explore further in the next section.

(a) Diversity of Flows: normalized inflow diversity

(b) Diversity of Flows: normalized outflow diversity

Figure 6. Diversity of flows (a) Spatial distribution of normalized inflow diversity; (b) Spatial distribution of normalized outflow diversity.

4.3 Direction of Flows

To more clearly understand the directional orientation of each neighbourhood, we visualize the direction for incoming and outgoing flows using the approach discussed in section 3.5. In Figures 7 and 8, we visualize these directional summaries for each grid cell as a vector flow map. To provide a context for these flows, we display a neighbourhood ratio in the underlying choropleth map. The neighbourhood ratio is measured based on the weighted number of tweets sent within 3 km from the geographical centre of the visited neighbourhood for the inflow network or from the geographical centre of the home for the outflow network. To make comparison across maps easier, the mean travel distance, the standard deviation of travel distance, and the diversity of flows are also presented in sub-figures.

In light of our earlier discussion on polycentricity, these vector flow maps are useful visual summaries of directional orientation patterns within an urban system. From the inflow network shown in Figure 7, it seems that the directional orientation of neighbourhoods in Singapore does not exhibit a monocentric pattern, but a rather polycentric one where footprint flows anchor onto a series of sub-centres. It should be noted that this pattern is likely to shift when we would only focus on, for example, commuting interactions but the LBS dataset here does not differentiate between different types of interactions.

In the outflow network displayed in Figure 8, we observe a starkly different spatial pattern, where the direction of outgoing interactions does focus heavily on the centre of Singapore (indicated with (1) in Fig.1). This does not necessarily mean that all residents, or even a majority, travel all the way to the centre, as this may solely reflect flows in the general direction of the centre to locations before or after the centre itself.

To further help understand the nuances in this pattern, the underlying neighbourhood ratio choropleth provides valuable context. For example, the neighbourhood ratio in the central areas is often well over 50%, indicating most outgoing interactions happen within a 3km radius for residents here. In contrast, the neighbourhood ratio in Woodlands (indicated with (8) in Fig.1) is much lower and hovers somewhere between 30-40%. As a regional centre, Woodlands might not perform as well as planned. Tampines (indicated with (6) in Fig.1) scores much better on this neighbourhood ratio, corroborating the findings on diversity and distance for this area in the previous section. The regional scale on which these centres function is also reflected in Figure 7c, showing that well over half of all incoming interactions come from within a 3km radius. In contrast, this ratio is less than 20% for the city centre itself. Based on this perspective, Tampines again seems to be the most successful regional centre in Singapore.

(a) Direction of Flows: background is neighbourhood ratio

(b) Direction of Flows: background is average incoming distance

(c) Direction of Flows: background is S.D. of incoming distance

(d) Direction of Flows: background is normalized inflow diversity

Figure 7. Flow maps with different metric backgrounds for the inflow network (a) Background with neighbourhood (within 3km) ratio; (b) Background with average incoming distance; (c) Background with standard deviation of incoming distance; (d) Background with normalized inflow diversity. Note: there is a reference map in Figure 1.

(a) Direction of Flows: background is neighbourhood ratio

(b) Direction of Flows: background is average outgoing distance

(c) Direction of Flows: background is S.D. of outgoing distance

(d) Direction of Flows: background is normalized outflow diversity

Figure 8. Flow maps with different metric backgrounds for the outflow network (a) Background with neighbourhood (within 3km) ratio; (b) Background with average outgoing distance; (c) Background with standard deviation of outgoing distance; (d) Background with normalized outflow diversity. Note: there is a reference map in Figure 1.

4.4 Socio-economic and built-environment covariates of spatial footprints

Although the patterns in Figure 5-8 can be explained in large part through a qualitative interpretation and analysis, in this section we conduct an additional quantitative analysis of the underlying processes that might explain these flow patterns. To do that, we conduct a regression analysis with distance and diversity as dependent variables and six socio-economic and built-environment characteristics discussed in Section 3.6 as independent variables.

As a standard OLS regression model exhibited strong spatial autocorrelation in the residuals (Moran's $I > 0.7$, see Appendix A.6.1 for more details), we adopt a Spatial Error Model (SEM) for its theoretical fit with the two modelled spatial processes. In other words, we assume that the spatial correlation in the error term results from some unmeasured spatial covariate that we cannot capture in the independent variables, partly due to limited public data available at this spatial scale. Table 1 summarizes the model results for both the inflow and the outflow networks.

For the distance of flows, we observe that the percentage of industrial building area in a grid cell has a negative effect on the mean distance of interactions for both inflow and outflow networks. In the Singapore context, with its relatively strict zoning in residential and industrial estates, this means that industrial estates generally attract people from nearby neighbourhoods for work-related travel. Concrete examples of this are the industrial estates of Ubi (in the south-eastern corner of the central region), Woodlands (indicated with (8) in Fig. 1), and Jurong (indicated with (10) in Fig. 1), which indeed seem to attract mostly people from within a 3km radius (see Fig. 7c). In contrast, we find a positive relationship between the percentage of residential area and distance. In other words, (dense) residential grids generally have both incoming and outgoing flows that travel further for different purposes, such as for shopping, recreation, etc. These opposite effects highlight a possibility to distinguish travel purposes and urban functions based on the sign of these coefficients.

The distance to the most central grid cell, another independent model term, has a positive effect on the flow distance. This reflects that neighbourhoods near the centre generally attract people from further away, which is also born out of the qualitative analysis of Figures 5-7. However, for outflow, this positive effect is somewhat surprising and contradicts the previous findings that the central areas have relatively short outgoing flows on average. A potential explanation can be found in the residuals of the SEM (see Fig. 9b), which show positive residuals throughout the city centre. This indicates that the effect of distance from the centre might indeed be present and positive but only apply for the difference between, for example, Woodlands (indicated with (8) in Fig.1) and Serangoon (indicated with (9) in Fig. 1) but the effect does not apply once we get into the centre itself.

For the diversity of flows, we observe a slightly different picture that is more dependent on the position of the grid cell in the built environment. Both the distance to the nearest MRT station and the distance to the centre have a negative effect on the diversity of flows. In other words, the centrality of a location, both in the urban system as a whole and in its public transport network, has a positive effect on both incoming and outgoing diversity.

Overall, what the SEM reveals is that potential proxies for socio-economic class (average resale price and the percentage of rentals) do not have a clear effect on either the distance and diversity of mobility flows. This is somewhat inconsistent with prior research that shows socio-economic characteristics are closely associated with human travel behaviour (Czepkiewicz et al., 2020; Gadepalli et al., 2020; Jang and Yi, 2021; Xu et al., 2018). One reason for this might be Singapore's unique city-island-state character and its well-designed public transportation system, which plays a crucial role in residents' day-to-day commuting. These specifics might mediate the impact of socio-economic characteristics on travel behaviour.

When we zoom in to specific locations, the residuals of the SEM provide additional clues to help explain the pattern of mobility flows in this analysis. For example, the model strongly underestimates the distances travelled to and from the residential area around Jurong West (indicated with (10) in Fig. 1). This could partly be explained by the connection to the public transportation network, which is only via a single MRT line that goes directly to the centre. As such, it might be easier for residents in Jurong West to travel further to central locations rather than nearby neighbourhoods that are not accessible via MRT. Similarly, the SEM overestimates the diversity of the area around Punggol. Although Punggol connects to the MRT system on the North-East Line, this resembles the situation in Jurong West, as this line travels directly to the city centre, while it remains comparatively difficult to travel to Punggol from geographically nearby neighbourhoods. Furthermore, there are several geographic features (water) and two military airports that make Punggol more isolated in the urban system than the model currently captures. These two examples clearly indicate that for a more foundational explanation of flow patterns, one would, at the minimum, need to account for the exact position in the urban transportation network and not merely use a crude proxy (distance to the nearest station) as we have done here.

The SEM employed here only provides an initial exploratory analysis of these mobility flows, limited by the relative lack of relevant, granular, available data. A comparison of the SEM with the Spatial Lag Model (SLM) (see Appendix A.6.2/3) shows that although the overall trends remain the same, the effect sizes and p-values of the independent variables vary. This potential lack of robustness is another indication that subsequent research on understanding the causal processes behind the distance and diversity of flows would be required.

(a) Residuals of average incoming distance (Moran I statistic: -0.09) (b) Residuals of average outgoing distance (Moran I statistic: -0.09)

(c) Residuals of inflow diversity (Moran I statistic: -0.05)

(d) Residuals of outflow diversity (Moran I statistic: -0.08)

Figure 9. Spatial distribution of residuals of Spatial Error Model (a) residuals of mean incoming distance (b) residuals of mean outgoing distance (c) residuals of inflow diversity (d) residuals of outflow diversity.

		Average Incoming distance	
Independent variables	Estimate	z value	Pr(> z)
(Intercept)	7.681274404	8.8870	< 0.00000000000000022
Mean resale price per square meter	0.000049152	0.4654	0.641649
Percentage of 1-Room and 2-Room rental flats	0.519593765	1.2634	0.206444
Percentage of industrial building area	-3.230181140	-2.5433	0.010982
Percentage of residential building area	2.443680855	2.7616	0.005752
Distance to the nearest MRT station	-0.235250179	-1.5634	0.117956
Distance to the central grid cell	0.079052859	1.7674	0.077161
Lambda: 0.80697, LR test value: 318.77, p-value: < 0.00000000000000022		ML residual variance (sigma squared): 0.56934, (sigma: 0.75455)	
Asymptotic standard error: 0.026271		Number of observations: 353	
z-value: 30.717, p-value: < 0.00000000000000022		Number of parameters estimated: 9	
Wald statistic: 943.52, p-value: < 0.00000000000000022		AIC: 910.76, (AIC for lm: 1227.5)	
Log likelihood: -446.3775 for error model			
Average Outgoing distance			
Independent variables	Estimate	z value	Pr(> z)
(Intercept)	7.12260427	8.1881	0.00000000000000022
Mean resale price per square meter	0.00010133	0.9773	0.32841
Percentage of 1-Room and 2-Room rental flats	0.59249114	1.4105	0.15841
Percentage of industrial building area	-3.33092235	-2.5402	0.01108
Percentage of residential building area	1.71113222	1.9402	0.05236
Distance to the nearest MRT station	-0.08804982	-0.5913	0.55432
Distance to the central grid cell	0.11128362	2.4223	0.01542
Lambda: 0.81384, LR test value: 317.35, p-value: < 0.00000000000000022		ML residual variance (sigma squared): 0.53831, (sigma: 0.7337)	
Asymptotic standard error: 0.025714		Number of observations: 340	
z-value: 31.65, p-value: < 0.00000000000000022		Number of parameters estimated: 9	
Wald statistic: 1001.7, p-value: < 0.00000000000000022		AIC: 862.26, (AIC for lm: 1177.6)	
Log likelihood: -422.1293 for error model			
Normalized inflow diversity			
Independent variables	Estimate	z value	Pr(> z)
(Intercept)	0.8774461556	13.7522	< 0.00000000000000022
Mean resale price per square meter	0.0000130204	1.4840	0.137807
Percentage of 1-Room and 2-Room rental flats	-0.0106980279	-0.2918	0.770460
Percentage of industrial building area	-0.1218683156	-1.1025	0.270251
Percentage of residential building area	0.0548738401	0.7087	0.478486
Distance to the nearest MRT station	-0.0376369288	-3.1287	0.001756
Distance to the central grid cell	-0.0156651221	-7.0694	0.00000000001556
Lambda: 0.57378, LR test value: 104.21, p-value: < 0.00000000000000022		ML residual variance (sigma squared): 0.0044611, (sigma: 0.066792)	
Asymptotic standard error: 0.046713		Number of observations: 353	
z-value: 12.283, p-value: < 0.00000000000000022		Number of parameters estimated: 9	
Wald statistic: 150.87, p-value: < 0.00000000000000022		AIC: -854.4, (AIC for lm: -752.19)	
Log likelihood: 436.2011 for error model			
Normalized outflow diversity			
Independent variables	Estimate	z value	Pr(> z)
(Intercept)	1.040150898	12.7917	< 0.00000000000000022
Mean resale price per square meter	-0.000010795	-1.0743	0.282670
Percentage of 1-Room and 2-Room rental flats	-0.027652761	-0.6758	0.499145
Percentage of industrial building area	0.053329555	0.4189	0.675324
Percentage of residential building area	0.154867576	1.8040	0.071232
Distance to the nearest MRT station	-0.045330767	-3.1562	0.001598
Distance to the central grid cell	-0.026936375	-6.7224	0.00000000001788
Lambda: 0.78611, LR test value: 276.51, p-value: < 0.00000000000000022		ML residual variance (sigma squared): 0.0050956, (sigma: 0.071384)	
Asymptotic standard error: 0.028634		Number of observations: 340	
z-value: 27.454, p-value: < 0.00000000000000022		Number of parameters estimated: 9	
Wald statistic: 753.7, p-value: < 0.00000000000000022		AIC: -731.23, (AIC for lm: -456.72)	
Log likelihood: 374.6151 for error model			

Table 1. Statistical summary of Spatial Error Model (SEM) results. Variables with negative effects and low p-values are highlighted in orange colour, variables with positive effects and low p-values are highlighted in blue colour.

5. Conclusion

In this paper we propose a method to analyze individual neighbourhoods within a larger urban system based on the mobility flows of people. We do so by operationalizing three specific mobility indicators, namely the average distance, diversity, and direction of flows, in both the inflow and outflow home-to-destination networks. We derive these indicators from a de-identified LBS dataset in Singapore. Each of these metrics offers a different but complementary perspective on the same neighbourhood.

As such, we argue for moving beyond analysis of urban networks as a whole, but rather looking at specific neighbourhoods as their constituent parts. This helps better understand and evaluate the role and performance of each neighbourhood within the larger urban system. In addition, the analytical latitude provided by new types of mobility data, such as the LBS dataset used here, gives more flexibility to define and use neighbourhoods with the size and shape appropriate for the research question at hand.

The presented study of Singapore serves as an illustrative case for the affordance of big data for smaller-scale urban research with mobility data. The three mobility metrics supplement each other in explaining flow patterns, and the patterns found in their spatial distributions allow us to reflect and examine urban policies and their implications. In Singapore, the development of regional centres, as designated in the Concept Plan 1991, is a particularly compelling example. By comparing the mobility metrics between different regional centres, our results show that beyond the city centre itself, Tampines can be considered as the most successful regional centres in Singapore, with the other regional centres retaining a relatively large focus on the city centre. The results also demonstrate that the urban interaction in Tampines generally comes from its nearby neighbourhoods. Such findings can be useful for local planning as they, for example, provide a way to evaluate and build better connections between neighbourhoods.

This approach can also be extended to investigate the determinants of a successful regional centre by including other types of data, such as local employment, the variety of retail services, etc. Apart from a qualitative interpretation of individual neighbourhood results, the spatial regression analysis provides an initial quantitative exploration of covariates behind these patterns. Here we find, among other things, specific zoning policies (e.g. industrial estates) to have measurable effects on the subsequent distance and diversity of flow networks in and out of neighbourhoods. However, socio-economic characteristics seem to have a relatively small impact on the three indicators, perhaps pointing towards the equalizing role of

Singapore's public transportation. The residuals of the regression also point towards the need to explicitly consider a neighbourhood's position within the urban transportation system to better understand its scores on each indicator.

The work presented here provides only an illustrative case study of this approach and can be extended in multiple ways. First, the LBS data has clear limitations, in terms of generalizability, spatial accuracy and precision, sampling bias and a lack of demographic information (Chen et al., 2016; Yuan et al., 2020). It should also be noted that geotagged data is becoming less common on Twitter, especially after 2015, due to changes in the platform and both privacy and commercial concerns make future access to LBS data uncertain. This could be partially addressed by using, for example, cell phone data and incorporating other public datasets (e.g. census data). More effort is needed to generalize these types of data before they can be used in urban policy, such as calibrating or evaluating its bias. (Longley et al., 2015; Luo et al., 2016). Nonetheless, as people with different demographic characteristics (e.g. age, gender, education, etc) or with different identities (e.g. locals or visitors) can have distinct movement patterns (Dong et al., 2020; F. Zhang et al., 2020) a further disaggregation of population sub-groups within such data might also provide useful insights. Second, we currently aggregate the mobility network over the entire study period. However, some places serve different urban functions during the day and night, or from season to season or year to year – with subsequent shifts in human mobility patterns. As such, subsequent work could disaggregate the temporal scale.

Nonetheless, analyzing and visualizing the three specific mobility indicators for both inflow and outflow networks provides an insightful set of perspectives for understanding the position of neighbourhoods within an urban system. With LBS and similar data allowing for analysis at a finer grained spatial scale, we hope that mobility-based perspectives on

neighbourhood and urban structure will become increasingly useful for urban researchers and practitioners working and planning at the neighbourhood level.

Research Data and Code Availability Statement

All code and data used to reproduce the analysis and figures within this paper are available at: <https://doi.org/10.6084/m9.figshare.13514746.v1>.

Acknowledgement

This research, led together with the Housing and Development Board, is supported by the Singapore Ministry of National Development and the National Research Foundation, Prime Minister's Office under the Land and Liveability National Innovation Challenge (L2 NIC) Research Programme (L2 NIC Award No. L2NICTDF1-2017-4). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Housing and Development Board, Urban Redevelopment Authority, Singapore Ministry of National Development and National Research Foundation, Prime Minister's Office, Singapore.

References

- Alessandretti, L., Aslak, U., Lehmann, S., 2020. The scales of human mobility. *Nature* 587, 402–407. <https://doi.org/10.1038/s41586-020-2909-1>
- Brockmann, D., Hufnagel, L., Geisel, T., 2006. The scaling laws of human travel. *Nature* 439, 462–465. <https://doi.org/10.1038/nature04292>
- Burger, M., Meijers, E., 2012. Form Follows Function? Linking Morphological and Functional Polycentricity. *Urban Studies* 49, 1127–1149. <https://doi.org/10.1177/0042098011407095>
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., Ratti, C., 2013. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies* 26, 301–313. <https://doi.org/10.1016/j.trc.2012.09.009>
- Campbell, M., Marek, L., Hobbs, M., n.d. Reconsidering movement and exposure: Towards a more dynamic health geography. *Geography Compass* n/a. <https://doi.org/10.1111/gec3.12566>

- Capasso Da Silva, D., King, D.A., Lemar, S., 2020. Accessibility in Practice: 20-Minute City as a Sustainability Planning Goal. *Sustainability* 12, 129. <https://doi.org/10.3390/su12010129>
- Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M., 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies* 68, 285–299. <https://doi.org/10.1016/j.trc.2016.04.005>
- Chen, Q., Poorthuis, A., 2021. Identifying home locations in human mobility data: an open-source R package for comparison and reproducibility. *International Journal of Geographical Information Science* 0, 1–24. <https://doi.org/10.1080/13658816.2021.1887489>
- Czepkiewicz, M., Heinonen, J., Næss, P., Stefansdóttir, H., 2020. Who travels more, and why? A mixed-method study of urban dwellers' leisure travel. *Travel Behaviour and Society* 19, 67–81. <https://doi.org/10.1016/j.tbs.2019.12.001>
- Dale, 1999. Urban planning in Singapore :the transformation of a city /Ole Johan Dale. – National Library [WWW Document]. URL <https://www.nlb.gov.sg/biblio/9146945> (accessed 11.26.20).
- Dong, Z., Chen, Y.-C., Dobra, A., 2020. A statistical framework for measuring the temporal stability of human mobility patterns. *Journal of Applied Statistics*.
- Ewing, R., Cervero, R., 2001. Travel and the Built Environment: A Synthesis. *Transportation Research Record* 1780, 87–114. <https://doi.org/10.3141/1780-10>
- Fang, Z., Liu, L., Shaw, S.-L., Yin, L., 2017. An Extended Community Detection Algorithm to Compare Human Mobility Flow Based on Urban Polycentric Cluster Boundaries: A Case Study of Shenzhen City, in: Zhou, C., Su, F., Harvey, F., Xu, J. (Eds.), *Spatial Data Handling in Big Data Era: Select Papers from the 17th IGU Spatial Data Handling Symposium 2016*, *Advances in Geographic Information Science*. Springer, Singapore, pp. 111–124. https://doi.org/10.1007/978-981-10-4424-3_8
- Gadepalli, R., Tiwari, G., Bolia, N., 2020. Role of user's socio-economic and travel characteristics in mode choice between city bus and informal transit services: Lessons from household surveys in Visakhapatnam, India. *Journal of Transport Geography* 88, 102307. <https://doi.org/10.1016/j.jtrangeo.2018.08.017>
- Gong, Y., Lin, Y., Duan, Z., 2017. Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. *Computers, Environment and Urban Systems* 64, 169–183. <https://doi.org/10.1016/j.compenvurbsys.2017.02.003>
- Han, S.S., 2005. Polycentric Urban Development and Spatial Clustering of Condominium Property Values: Singapore in the 1990s. *Environ Plan A* 37, 463–481. <https://doi.org/10.1068/a3746>
- Hou, Y., 2019. Polycentric urban form and non-work travel in Singapore: A focus on seniors. *Transportation Research Part D: Transport and Environment* 73, 245–275. <https://doi.org/10.1016/j.trd.2019.07.003>
- Huang, Q., Wong, D.W.S., 2016. Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us? *International Journal of Geographical Information Science*.

- Jang, S., Yi, C., 2021. Imbalance between local commuting accessibility and residential locations of households by income class in the Seoul Metropolitan Area. *Cities* 109, 103011. <https://doi.org/10.1016/j.cities.2020.103011>
- Kang, C., Fan, D., Jiao, H., 2020. Validating activity, time, and space diversity as essential components of urban vitality. *Environment and Planning B: Urban Analytics and City Science* 2399808320919771. <https://doi.org/10.1177/2399808320919771>
- Kloosterman, R.C., Musterd, S., 2001. The Polycentric Urban Region: Towards a Research Agenda. *Urban Studies* 38, 623–633. <https://doi.org/10.1080/00420980120035259>
- Kong, X., Song, X., Xia, F., Guo, H., Wang, J., Tolba, A., 2018. LoTAD: long-term traffic anomaly detection based on crowdsourced bus trajectory data. *World Wide Web* 21, 825–847. <https://doi.org/10.1007/s11280-017-0487-4>
- Kraemer, M.U.G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., Group†, O.C.-19 D.W., Plessis, L. du, Faria, N.R., Li, R., Hanage, W.P., Brownstein, J.S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O.G., Scarpino, S.V., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368, 493–497. <https://doi.org/10.1126/science.abb4218>
- Kwan, M.-P., 1999. Gender, the Home-Work Link, and Space-Time Patterns of Nonemployment Activities. *Economic Geography* 75, 370–394. <https://doi.org/10.1111/j.1944-8287.1999.tb00126.x>
- Lee, R., Wakamiya, S., Sumiya, K., 2013. Urban area characterization based on crowd behavioral lifelogs over Twitter. *Pers Ubiquit Comput* 17, 605–620. <https://doi.org/10.1007/s00779-012-0510-9>
- Li, M., Gao, S., Lu, F., Zhang, H., 2019. Reconstruction of human movement trajectories from large-scale low-frequency mobile phone data. *Computers, Environment and Urban Systems* 77, 101346. <https://doi.org/10.1016/j.compenvurbsys.2019.101346>
- Li, W., Wang, S., Zhang, X., Jia, Q., Tian, Y., 2020. Understanding intra-urban human mobility through an exploratory spatiotemporal analysis of bike-sharing trajectories. *International Journal of Geographical Information Science* 34, 2451–2474. <https://doi.org/10.1080/13658816.2020.1712401>
- Limtanakool, N., Dijst, M., Schwanen, T., 2006. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography* 14, 327–341. <https://doi.org/10.1016/j.jtrangeo.2005.06.004>
- Liu, K., Murayama, Y., Ichinose, T., 2020. Using a new approach for revealing the spatiotemporal patterns of functional urban polycentricity: A case study in the Tokyo metropolitan area. *Sustainable Cities and Society* 59, 102176. <https://doi.org/10.1016/j.scs.2020.102176>
- Liu, X., Gong, L., Gong, Y., Liu, Y., 2015. Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography* 43, 78–90. <https://doi.org/10.1016/j.jtrangeo.2015.01.016>
- Liu, Z., Liu, S., 2018. Polycentric Development and the Role of Urban Polycentric Planning in China's Mega Cities: An Examination of Beijing's Metropolitan Area. *Sustainability* 10, 1588. <https://doi.org/10.3390/su10051588>

- Longley, P.A., Adnan, M., Lansley, G., 2015. The Geotemporal Demographics of Twitter Usage. *Environ Plan A* 47, 465–484. <https://doi.org/10.1068/a130122p>
- Luo, F., Cao, G., Mulligan, K., Li, X., 2016. Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago. *Applied Geography* 70, 11–25. <https://doi.org/10.1016/j.apgeog.2016.03.001>
- Martí, P., Serrano-Estrada, L., Nolasco-Cirugeda, A., 2019. Social Media data: Challenges, opportunities and limitations in urban studies. *Computers, Environment and Urban Systems* 74, 161–174. <https://doi.org/10.1016/j.compenvurbsys.2018.11.001>
- Moreno, C., Allam, Z., Chabaud, D., Gall, C., Pratlong, F., 2021. Introducing the “15-Minute City”: Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities* 4, 93–111. <https://doi.org/10.3390/smartcities4010006>
- Morris, E.K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T.S., Meiners, T., Müller, C., Obermaier, E., Prati, D., Socher, S.A., Sonnemann, I., Wäschke, N., Wubet, T., Wurst, S., Rillig, M.C., 2014. Choosing and using diversity indices: insights for ecological applications from the German Biodiversity Exploratories. *Ecology and Evolution* 4, 3514–3524. <https://doi.org/10.1002/ece3.1155>
- Nagendra, H., 2002. Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied Geography* 22, 175–186. [https://doi.org/10.1016/S0143-6228\(02\)00002-4](https://doi.org/10.1016/S0143-6228(02)00002-4)
- Nelson, G.D., Rae, A., 2016. An Economic Geography of the United States: From Commutes to Megaregions. *PLOS ONE* 11, e0166083. <https://doi.org/10.1371/journal.pone.0166083>
- Poucin, G., Farooq, B., Patterson, Z., 2018. Activity patterns mining in Wi-Fi access point logs. *Computers, Environment and Urban Systems* 67, 55–67. <https://doi.org/10.1016/j.compenvurbsys.2017.09.004>
- Rae, A., 2009. From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK census. *Computers, Environment and Urban Systems* 33, 161–178. <https://doi.org/10.1016/j.compenvurbsys.2009.01.007>
- Roth, C., Kang, S.M., Batty, M., Barthélemy, M., 2011. Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows. *PLOS ONE* 6, e15923. <https://doi.org/10.1371/journal.pone.0015923>
- Saxon, J., 2020. The local structures of human mobility in Chicago. *Environment and Planning B: Urban Analytics and City Science* 2399808320949539. <https://doi.org/10.1177/2399808320949539>
- Schwanen, T., 2016. Geographies of transport II: Reconciling the general and the particular. *Progress in Human Geography*. <https://doi.org/10.1177/0309132516628259>
- Shaw, S.-L., Sui, D., 2020. Understanding the New Human Dynamics in Smart Spaces and Places: Toward a Splatial Framework. *Annals of the American Association of Geographers* 110, 339–348. <https://doi.org/10.1080/24694452.2019.1631145>
- Shi, S., Wang, L., Xu, S., Wang, X., 2020. Prediction of Intra-urban Human Mobility by Integrating Regional Functions and Trip Intentions. *IEEE Transactions on Knowledge and Data Engineering* 1–1. <https://doi.org/10.1109/TKDE.2020.3047406>

- Sun, Y., Fan, H., Li, M., Zipf, A., 2016. Identifying the city center using human travel flows generated from location-based social networking data. *Environ Plann B Plann Des* 43, 480–498. <https://doi.org/10.1177/0265813515617642>
- Tobler, W.R., 1981. A Model of Geographical Movement. *Geographical Analysis* 13, 1–20. <https://doi.org/10.1111/j.1538-4632.1981.tb00711.x>
- Traunmueller, M.W., Johnson, N., Malik, A., Kontokosta, C.E., 2018. Digital footprints: Using WiFi probe and locational data to analyze human mobility trajectories in cities. *Computers, Environment and Urban Systems* 72, 4–12. <https://doi.org/10.1016/j.compenvurbsys.2018.07.006>
- van Meeteren, M., Poorthuis, A., Derudder, B., Witlox, F., 2016. Pacifying Babel’s Tower: A scientometric analysis of polycentricity in urban research. *Urban Studies* 53, 1278–1298. <https://doi.org/10.1177/0042098015573455>
- Vilhelmson, B., 1999. Daily mobility and the use of time for different activities. The case of Sweden. *GeoJournal* 48, 177–185.
- Wang, J., Kong, X., Xia, F., Sun, L., 2019. Urban Human Mobility: Data-Driven Modeling and Prediction. *SIGKDD Explor. Newsl.* 21, 1–19. <https://doi.org/10.1145/3331651.3331653>
- Weng, M., Ding, N., Li, J., Jin, X., Xiao, H., He, Z., Su, S., 2019. The 15-minute walkable neighborhoods: Measurement, social inequalities and implications for building healthy communities in urban China. *Journal of Transport & Health* 13, 259–273. <https://doi.org/10.1016/j.jth.2019.05.005>
- Wu, L., Yang, L., Huang, Z., Wang, Y., Chai, Y., Peng, X., Liu, Y., 2019. Inferring demographics from human trajectories and geographical context. *Computers, Environment and Urban Systems* 77, 101368. <https://doi.org/10.1016/j.compenvurbsys.2019.101368>
- Xia, F., Wang, J., Kong, X., Wang, Z., Li, J., Liu, C., 2018. Exploring Human Mobility Patterns in Urban Scenarios: A Trajectory Data Perspective. *IEEE Communications Magazine* 56, 142–149. <https://doi.org/10.1109/MCOM.2018.1700242>
- Xia, F., Wang, J., Kong, X., Zhang, D., Wang, Z., 2020. Ranking Station Importance With Human Mobility Patterns Using Subway Network Datasets. *IEEE Transactions on Intelligent Transportation Systems* 21, 2840–2852. <https://doi.org/10.1109/TITS.2019.2920962>
- Xu, Y., Belyi, A., Bojic, I., Ratti, C., 2018. Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Computers, Environment and Urban Systems* 72, 51–67. <https://doi.org/10.1016/j.compenvurbsys.2018.04.001>
- Yuan, Y., Lu, Y., Chow, T.E., Ye, C., Alyaqout, A., Liu, Y., 2020. The Missing Parts from Social Media–Enabled Smart Cities: Who, Where, When, and What? *Annals of the American Association of Geographers* 110, 462–475. <https://doi.org/10.1080/24694452.2019.1631144>
- Zhang, B., Chen, S., Ma, Y., Li, T., Tang, K., 2020. Analysis on spatiotemporal urban mobility based on online car-hailing data. *Journal of Transport Geography* 82, 102568. <https://doi.org/10.1016/j.jtrangeo.2019.102568>
- Zhang, F., Zu, J., Hu, M., Zhu, D., Kang, Y., Gao, S., Zhang, Y., Huang, Z., 2020. Uncovering inconspicuous places using social media check-ins and street view images. *Computers,*

Environment and Urban Systems 81, 101478.
<https://doi.org/10.1016/j.compenvurbsys.2020.101478>

- Zhao, Z., Shaw, S.-L., Yin, L., Fang, Z., Yang, X., Zhang, F., Wu, S., 2019. The effect of temporal sampling intervals on typical human mobility indicators obtained from mobile phone location data. *International Journal of Geographical Information Science* 33, 1471–1495. <https://doi.org/10.1080/13658816.2019.1584805>
- Zhong, C., Schlöpfer, M., Müller Arisona, S., Batty, M., Ratti, C., Schmitt, G., 2017. Revealing centrality in the spatial structure of cities from human activity patterns. *Urban Studies* 54, 437–455. <https://doi.org/10.1177/0042098015601599>