Resolving urban mobility networks from individual travel graphs using massive-scale mobile phone tracking data

Jinzhou Cao¹,², Qingquan Li¹,²,³, Wei Tu¹,²*, Qili Gao¹,², Rui Cao¹,², Chen Zhong⁴

¹ Guangdong Key Laboratory of Urban Informatics, Shenzhen Key Laboratory of Spatial Smart Sensing & Research Institute of Smart Cities, Shenzhen University, Shenzhen 518060, China
² Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518060, China
³ State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China
⁴ Department of Geography, King’s College London, London WC2R 2LS, UK

Corresponding author: Dr. Wei Tu
Email: tuwei@szu.edu.cn
Address: Kejilou 1402, Shenzhen University
Nanhai road No. 3688, Shenzhen 430079, P. R. China

Acknowledgments
The authors would like to thank all colleagues and students who contributed to this study. This research was jointly supported by the Natural Science Foundation of China [42001393 and 71961137003]; the China Postdoctoral Science Foundation [2020M672803]; the Natural Science Foundation of Guangdong Provinces [2019A1515011049]; the Basic Research Program of Shenzhen Science and Technology Innovation Committee [JCJY201803053125113883], and the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation Ministry of Natural Resources[KF-2019-04-073].
Abstract

Human movements and interactions with cities are characterized by urban mobility networks. Many studies that address urban mobility are inspired by complex networks. The models of complex networks require a large amount of empirical data. However, current works relied on traditional survey data and were unable to take full advantage of the capabilities offered by complex networks; thus, the possibility of quantifying urban mobility networks by considering individual travel patterns has not yet been addressed. This study presents a data-driven approach for characterizing urban mobility networks based on massive-scale mobile phone tracking data. Individual travel motifs are first extracted using a graph-based approach. The global urban mobility network (G-UMN) and the motif-dependent urban mobility subnetworks (MD-UMNs) are then constructed. Next, network properties, including statistical measures and scaling relations between the basic measures, are proposed for characterizing mobility networks. We have conducted experiments focusing on Shenzhen, China. The results demonstrated that (1) the individual travel motifs are structurally and spatially heterogeneous, (2) the G-UMN exhibits a evolutionary hierarchical structure, and (3) the MD-UMNs show many behavioral differences in their spatial and topological properties, reflecting the impacts of the heterogeneity of the individual travel motifs. These results bridge the gap between complex network properties and urban mobility patterns and provide crucial implications and policies for data-informed urban planning.

Key Words: Spatial network; Urban mobility; Mobile phone tracking data; Complex Network analysis.
1 Introduction

Rapid urbanization has led to a great influx of residents into cities. The intra-urban movements of individuals are rapidly changing. Moreover, frequent human movements and the associated interactions with urban space pose great challenges to urban planning by demanding an urban spatial structure that is compatible with highly efficient travel for residents. Urban mobility is crucial for harmonizing urban spatial structures since it exerts significant influences on resource allocation, social equity and sustainable urban evolution (Maeda et al., 2019; Toole et al., 2015). Consequently, the ability to characterize urban mobility attracts scholarly attention in a broad range of fields, from urban planning (Ratti et al., 2006), transport (Tu et al., 2019), and urban science (Batty, 2008) to statistical physics (Bettencourt, 2013).

How to characterize urban mobility has been intensively investigated recently. However, the representation of human movements is difficult. Since an individual’s trajectory can be modeled as a graph, an innovative notion that is referred to as the ‘urban mobility network’ has been acknowledged as an effective foundation for urban mobility studies. The urban mobility network is defined as the network-oriented aggregation of individuals’ movements in urban environments (Parthasarathi, 2014). Studies that characterize urban mobility networks have been employed to reveal the properties of urban mobility (Barthélemy, 2011; Cheng et al., 2013). Recently, several studies that address urban mobility are inspired by complex networks. Complex networks theory provides models to describe the topological and spatial patterns of networks. The statistics of mobility networks can thus describe and evaluate how human mobility is distributed and developed on different scales. Therefore, these complex network-driven measures have highlighted the characteristics of urban mobility (Agryzkov et al., 2017; Zhang & Thill, 2017). The models of complex networks require a large amount of empirical data. However, current works relied on traditional survey data and were unable to take full advantage of the capabilities offered by complex networks for addressing urban mobility tasks.

With technological advances in the fields of global positioning systems (GPS) and information and communications technology (ICT), ubiquitous smart devices have become sensors that individuals carry every day (Calabrese et al., 2014). These
advances have contributed to an explosive growth of human tracking datasets, such as mobile phone positioning data (Alexander et al., 2015; Blondel et al., 2015) and GPS trajectories (Tang et al., 2015; Tu et al., 2018). These emerging datasets enable the high-precision representation of human movements (Shaw et al., 2016; Zhao et al., 2018) and create new windows for understanding human-urban interaction (Lim et al., 2018; Y. Wang et al., 2019; Xu et al., 2019). Thus, interpretable quantitative analyses of urban mobility networks are becoming possible. Some studies quantified urban mobility networks by aggregating the movements of all individuals (Hamedmoghadam et al., 2019; Louail et al., 2015; Riascos & Mateos, 2020). However, few studies simultaneously considered the heterogeneity of individual travels.

The properties of urban mobility networks are influenced by individual travel (Pinho et al., 2016; Puura et al., 2018). Because individual travel is shaped by personal characteristics and the spatial configurations of facilities, urban mobility shows various patterns (Zhang et al., 2018). Multifaceted urban mobility networks can be constructed to capture the corresponding characteristics. Therefore, this study addresses the following question: what are the heterogeneous properties of individual travels extracted from massive human tracking data? Furthermore, when aggregating individual travel into multifaceted urban mobility networks, another question is raised: what are the differences in the complex network properties of multifaceted urban mobility? These two questions highlight the necessity of a comprehensive and comparative study to investigate urban mobility networks using big human tracking data. We present a data-driven approach for resolving urban mobility networks. Individual trajectories are abstracted into standard graph-based motifs. The global urban mobility network is constructed by aggregating the travel graphs of all individuals, and multiple urban mobility subnetworks are constructed in accordance with the individual motifs; then, the resulting networks are characterized by a series of statistical measures derived from the complex network perspective. These measures allow us to reveal the patterns present in urban mobility networks. We also consider scaling relations between these measures to evaluate how urban mobility networks develop. Considering Shenzhen, China as the study area, we exploited massive-scale mobile phone tracking data to construct travel motifs of all individuals and characterized the urban mobility networks. The results of the statistical measures and
scaling relations demonstrated a multi-facet portrait of urban mobility networks, which provides crucial implications and policies for data-informed urban planning.

This study makes the following contributions. First, compared with traditional approaches, this study resolves urban mobility networks by considering the impacts of the heterogeneity of individual travels using mobile phone tracking data, which have higher penetration and a finer temporal scale. Second, the results of this study provide a deeper understanding of the structurally and spatially heterogeneous patterns of urban mobility networks. These insights thus help policy-makers to evaluate their urban development strategy, especially the urban resources allocation. Last, the findings of this study are complementary to urban studies in a different but typical urban context in the light of urban development path.

The remainder of this article is organized as follows. Section 2 reviews related works of this research. Section 3 introduces the study area and the mobile phone tracking data that are utilized. Section 4 describes the proposed methodological framework of resolving the urban mobility networks. Section 5 analyzes the results. Section 6 concludes the findings and policy suggestions and discusses future work.

2 Literature review

Urban mobility analysis is a fundamental research topic in interdisciplinary field which focuses on exploring the spatio-temporal properties as well as hidden patterns behind the intra-urban and inter-urban movements (González et al., 2008; Tu et al., 2018). The concept of urban mobility is broad in dimensions of human travels at both individual and group levels. The conceptualization of urban mobility also varies depending on the contexts of the range of applications, e.g., epidemic prevention (Gómez et al., 2018), migratory flows prediction (Huang et al., 2018), urban planning (Bokányi et al., 2019), and location-based services (Noulas et al., 2012).

The representation and characterization of urban mobility are the primary work in the urban mobility analysis (Hasan et al., 2012). In transportation planning and modeling, intra-urban human movements can be captured in the form of origin–destination (OD) matrices, where these matrices were obtained by dividing an area into a set of zones and counting the numbers of trips between two zones (Calabrese et al., 2011; Bachir et al., 2019). As a another example, inter-urban population migration can be described as a flow
network by establishing the adjacency relationships of the population flows between two cities (Pan & Lai, 2019). These studies mark underlying efforts to model the structural form of urban mobility.

Since the intra-urban or inter-urban mobility can both modeled as a graph, the notion of ‘urban mobility network’ has been viewed as an important concept for urban mobility studies. Namely, it denotes the network-structured aggregation of population’s urban travels and activities (Parthasarathi, 2014). Recent years have witnessed explosive growth of big human mobility data in urban scenarios due to the advancements in the information and communications technology and pervasive usage of smart devices. Multi-sourced and massive data provide an unprecedented opportunity for a deeper understanding of urban mobility networks. Previous studies have been employed to derive urban mobility networks from human mobility data (Belyi et al., 2017). Topics include, but are not limited to, community-based spatial structures (Gao et al., 2013; Ratti et al., 2010; Yildirimoglu & Kim, 2017), intra-urban interactions (Krings et al., 2009; Sun et al., 2015; Wu et al., 2014; Zhang et al., 2017), traffic flow dynamics (Jiang et al., 2009; Liu et al., 2012; Tang et al., 2015), scaling laws of mobility (Brockmann et al., 2006; Tachet et al., 2017; Yan et al., 2013), and inter-urban migration patterns (De Montis et al., 2005; Liu et al., 2014; Simini et al., 2012). These studies highlighted the characterizations of urban mobility networks to better understand the human behaviors and the structures of cities.

Recently, several studies that mark urban mobility networks are motivated by complex network theory (Guidotti et al., 2016). A system consisting of several non-identical elements connected by diverse interactions is considered as a complex network where the nodes are the system elements and the links are the interactions between the elements (Newman, 2010). Complex networks theory develops various quantitative measures, such as the node degree, node strength, and clustering coefficient, to characterize one network (Albert & Barabási, 2002). Important properties in complex networks, such as the small-world properties (Watts & Strogatz, 1998), scale-free properties (Barabási & Albert, 1999) and community structures (Wang et al., 2018), have also been found in the urban mobility networks, and some studies have explained the dynamic mechanism of urban mobility behind these properties (Barabási, 2005; Lera et
al., 2017). For instance, Saberi, et al. (2016) explored travel demand patterns by analyzing the measures of OD networks, including the node degree, node flux, and shortest path, using household travel survey data from Chicago and Melbourne. Zhong, et al. (2014) revealed urban spatial structures by examining the centralities of an urban network using travel survey data from Singapore. In recent years, big data have played a vital role in curving the urban mobility patterns through the complex network tools. For example, Chi, et al. (2016) and Hossmann, et al. (2011) applied complex network-driven measures to investigate mobility patterns by fusing social media check-ins, GPS trajectories, and smart card data. Louail et al. (2015) revealed the spatial structure of commuting networks extracted from mobile phone data. Although these studies revealed the overall look of urban mobility networks by aggregating the movements of large populations, there seem to be lack of the simultaneously consideration of the heterogeneity of individual travels. In other words, whether there exist any differences in properties of multifaceted urban mobility networks across various population classes remains to be better explored.

The scaling laws is seen as very effective to obtain a qualitative description of global character in urban mobility analysis. The scaling properties is proved to be widespread in urban mobility(Song, Koren, et al., 2010; X.-W. Wang et al., 2014). For example, power-law-like displacement distribution(Yan et al., 2013) and visitation frequency distribution(Zheng & Zhou, 2017) were empirically observed in many analyses of human movements. However, there is still a remarkable lack of research that would reveal the scaling relation between various complex network-based measures. With the increasing availability of human mobility datasets and the innovation in complex network methods, this paper therefore presents a complex network-based measure framework to characterize urban mobility networks from mobile phone tracking data, and further assess the development of mobility networks in a policy-oriented perspective.

3 Study area and data

3.1 Case study: Shenzhen, China

This study was conducted in Shenzhen, China, which is located in southern China and borders Hong Kong to the south, with a total area of approximately 2,050 square kilometers. The spatial map of Shenzhen is shown in Fig. 1. Shenzhen is a typical high-
density city in the world. By the end of 2015, the residential population of Shenzhen was approximately 11 million and the population density of Shenzhen had reached 5,500 per square kilometer; it ranks first in China according to national statistics (Shenzhen Municipal Statistics Bureau, 2016).

Fig. 1. Study area of Shenzhen, China.

As the first Special Economic Zone (SEZ) in China, Shenzhen has experienced rapid urbanization. Over the past forty years, Shenzhen has transformed from a small fishing village into a prominent high-tech and innovative mega-city in China. In accordance with this governmental policy, Shenzhen was divided into an SEZ and a non-SEZ during the early years. The original SEZ districts included Luohu, Futian, Nanshan and Yantian districts, which are located in southern Shenzhen. The original non-SEZ districts included Bao’an, Longhua, Longgang, Pingshan, and Dapeng districts, which are located in northern and eastern Shenzhen. The SEZ and non-SEZ districts exhibited considerable differences in their urban and transport planning, and policies, which generated enormous gaps in their economic and social development. The SEZ districts have more accessible transport systems (buses and metros), high-income job opportunities and abundant living resources, such as shopping malls, schools and universities, medical facilities, community parks. The non-SEZ districts contained many industrial parks and natural lands. In 2010, the SEZ was expanded to include the whole city; thus, an increasing number of urban resources were allocated to the original non-SEZ districts. However, these spatial differences in urban development still exist. For example, several SEZ
districts have local centers that attract huge travels from neighboring districts while the other districts lacked business and cultural centers and resulted in many cross-regional travels. This situation emphasizes the necessity of resolving urban mobility and contributing to urban planning policies, such as how to narrow the regional differences of the city in the future.

3.2 Mobile phone tracking data

The mobile phone tracking data were utilized to construct individual. The dataset employed in this study were provided by a dominant communications operator in Shenzhen collected on a working day in March 2012. Unlike the data drawn from call detail records (CDRs), which are triggered only upon receipt of communication events (such as phone calls and text messages) (Xu et al., 2016; Yang, Fang, Yin, et al., 2019), the data applied in this study were recorded every hour. The corresponding service areas were approximated by Voronoi tessellation of base towers. The locations of the mobile phone users were determined at the base tower level. Therefore, this dataset shows advantages over CDRs and other traditional travel survey datasets in terms of its higher penetration rate and temporal resolution. To protect user privacy, this dataset has been anonymized by the communication operator. No personal information, such as phone number, username, gender, or age, can be obtained from the data. Examples of records of a user are presented in Table 1. One record includes a user ID, a timestamp, and latitude and longitude coordinates. There are approximately 9.7 million mobile phone users in one day. A total of 5,926 base towers exist in the study area.

| Table 1. Examples of mobile phone tracking data |
|----------------|----------------|----------------|----------------|
| User ID         | Longitude      | Latitude       | Time (hh: mm: ss) |
| 1101032***      | 113.934***     | 22.521***      | 07: 25: 00       |
| 1101032***      | 113.882***     | 22.571***      | 08: 35: 00       |
| 1101032***      | 113.882***     | 22.571***      | 09: 26: 00       |
| 1101032***      | 113.882***     | 22.571***      | 10: 31: 00       |
| …               | …              | …              | …               |
| 1101032***      | 113.934***     | 22.521***      | 23: 28: 00       |
4 Methodology

An overarching framework has been developed to resolve urban mobility networks extracted from massive-scale mobile phone tracking data. This framework consists of three stages. The first stage abstracts individuals’ travels into motifs by processing the raw mobile phone tracking data. The second stage produces the global urban mobility network and motif-dependent urban mobility subnetworks using abstracted individual motifs. The final stage examines the statistical measures of the mobility networks and the scaling relations of these measures to characterize urban mobility. Fig. 2 illustrates the workflow of the proposed analytical framework.

4.1 Constructing individuals’ travel motifs

By leveraging mobile phone tracking data, individual trajectories were abstracted into travel motifs using a two-step method. As illustrated in Fig. 3, the raw mobile phone tracking data were first segmented into sequential stays. Each stay sequence was then utilized to abstract a graph-based travel structure, which is referred to as a motif, in which each node denotes a distinct visited place and each edge denotes the travel flow between two places.
Fig. 3. The construction of individuals’ travel motifs from mobile phone tracking records.

4.1.1 Extracting stay sequences

The records were sorted by the timestamp and clipped into time-sequential positioning records, as illustrated by the consecutive purple triquetrous points in Fig. 4. Here, sequential stays represent a set of places where users were engaged in activities (Shen & Cheng, 2016). We applied a tower-based segmentation algorithm using both spatial rules and temporal rules (Tu et al., 2017). We connected the time-sequential records with no move into candidate stays. The red vertical lines in Fig. 4 represent the five candidate stays (p2-p5, p7-p9, p10-p12, p13-p14, p15-p17). Spatial uncertainty exists because of the low spatial accuracy of cell-tower-based location technology. Consecutive records might jump between adjacent cell towers. Therefore, we calculated the spatial radius between each record that is not in any candidate stay and every candidate stay; if the distance is less than the 500 meters, the record was added to the candidate stay. As shown in Fig. 4, record p6 can be merged into candidate Stay 2. Otherwise, the point was recognized as a new candidate stay. Once all twenty-four records for a particular person were processed, the corresponding sequence of candidate stays was identified. For all candidate stays, if the temporal duration is less than 60 minutes or shorter, the stay was not considered as a true stay. The temporal duration of candidate stay p13-p14 was 50 minutes; this candidate stay was omitted. The true stay sequence was identified, as illustrated by the yellow points in Fig. 4. Note that any user with only one stay in his/her
sequence was excluded because he/she did not move throughout the whole day. These sequential stays were employed to construct a directed graph.

**Fig. 4.** Illustration of the extraction of stay sequences from mobile phone tracking data.

### 4.1.2 Constructing mobility motifs

Let $M$ be the number of users eligible for analysis. Let the sequential stays of user $u$ be denoted by $S(u) = \{Stay_1, Stay_2, \ldots, Stay_N\}$, where $N$ is the number of separate visits to locations. Accordingly, the travel graph $G(u) = \{V(u), L(u)\}$ can be constructed from $S(u)$. The vertex set $V(u) = \{v_1, v_2, \ldots, v_n\}$ contains all distinct visited locations, where $n$ is the number of distinct locations. The link set $L(u) = \{\ell_{i,j} | i, j \in V(u) \land i \neq j\}$ contains all directed trips, where $\ell_{i,j}$ is the directed flow between vertex $i$ and vertex $j$. Essentially, $G(u)$ is expressed in weighted matrix form. Each individual’s daily trips can be abstracted into a travel graph, which is referred to as a travel *motif*, where each node represents a distinct visited location and each edge represents the travel flow between a particular pair of nodes (Cao et al., 2019). Fig. 3 depicts the construction processes for motifs with three
nodes (top) and four nodes (bottom). We applied the following convention to name the motifs (Fig. 5): ID-2-1 represents the first motif with two nodes, ID-3-1 represents the first motif with three nodes, etc.

![Figure 5](image.png)

**Fig. 5.** Extracted most frequent motifs and their corresponding identities.

4.2 Generating urban mobility networks

The global urban mobility network was constructed by aggregating all individuals’ travel motifs. Considering the different topologies of the motifs, multifaceted motif-dependent urban mobility subnetworks were also constructed. These subnetworks represent the heterogeneous characteristics of the individual travel patterns, as illustrated in Fig. 6. The random and scale-free urban mobility networks that represents two extreme urban mobility patterns were also generated as references.

4.2.1 Global urban mobility network

To capture a global picture of the urban mobility network, we aggregated all individuals’ motifs to construct a weighted directed network that represented the sum of the travel flows of all individuals. We name this network the *global urban mobility network* (G-UMN), which is defined as

\[
G_{G-UMN} = (V, E, W)
\]

where \( V = Distinct(\bigcup_{u=1}^{M} V(u)) \) represents all of the distinct urban nodes (i.e., the service areas of the base towers) and \( E = \{(v_i, v_j) | (v_i, v_j) \in distinct(\bigcup_{u=1}^{M} L(u)) \}\) represents all existing flows between pairs of nodes. The edge weights \( W_{v_i,v_j} \) correspond...
to the counts of flows between two nodes. For each individual trip on edge \((v_i, v_j)\), the weight \(W_{v_i v_j}\) is incremented by 1. For each \((v_i, v_j) \in E\), we have \(W(v_i, v_j) = W_{v_i v_j}\).

4.2.2 Motif-dependent urban mobility subnetworks

Multiple urban mobility subnetworks were constructed by using these travels with one type of individual motif. The data for all individuals that exhibit the same motif were aggregated into a corresponding weighted network. We refer to these networks as motif-dependent urban mobility subnetworks (MD-UMNs). We applied the same identity convention that was applied to the motifs to express the identity of the subnetworks. The mathematical expression for an MD-UMN is

\[
G_{MD-UMN}^t = (V^t, E^t, W^t)
\]

where \(V^t = Distinct(\bigcup_{u=1}^{u=F} V(u) \in V(t) \text{ in } G_{Loc}(t))\) represents all of the urban nodes that belong to motif \(t\), \(E^t = \{(v_p, v_q) | (v_p, v_q) \in Distinct(\bigcup_{u=1}^{u=F} E(u) \in E(t) \text{ in } G_{Loc}(t))\) represents all flows that belong to motif \(t\), and \(W^t(v_i, v_j)\) represents the absolute weight of edge \((\ell_i, \ell_j)\) that belongs to motif \(t\).
Fig. 6. Illustrations of the global urban mobility network (G-UMN) and the motif-dependent urban mobility subnetworks (MD-UMNs)

4.2.3 Reference networks

Reference networks are an important baseline against which to measure the possibility of the occurrence of certain network structures, given certain properties of empirical networks. In this study, two reference networks, which represent two extreme urban conditions that characterize urban spatial heterogeneity, were generated. Specifically, a random urban mobility network and a scale-free urban mobility network were generated, as follows:

- **Random urban mobility network (RA-UMN):** The RA-UMN represents the case of entirely homogeneous neighborhoods in the city, which means that all individuals' urban travel flows are purely random, without exhibiting any preferences. It is conjectured that all resources and facilities in the urban space have a relatively uniform distribution and that individuals' trips are not restricted by the urban spatial structure. This network was simulated by means of random walks between any two nodes with the same probability $p$ and number of vertices $N$ as the G-UMN. The degree distribution shows the characteristics of a Poisson distribution, which represents the property of homogeneity (Frieze & Karoński, 2016). In addition, the clustering coefficients are very small. This network is denoted by $G_{RA-MN}$ in this paper.

- **Scale-free urban mobility network (SF-UMN):** The SF-UMN represents the case of highly heterogeneous neighborhoods in the city, which corresponds to the spatial heterogeneity derived from the relative concentrations of resources; thus, specific regions with more concentrated resources will attract a larger number of people, while other areas will experience minimal traffic. This network was generated based on preferential attachments, with the node distribution following a power law. Most nodes have only a few connections, while a few nodes possess a large number of connections. The nodes are heterogeneous, and the influence of scale disappears, which means that the network possesses the scale-free characteristic (Ferreira et al., 2018). This network is denoted by $G_{SF-MN}$ in this paper.

4.3 Characterizing the urban mobility networks
The properties of a network are essentially characterized by a set of statistical measures (Albert & Barabási, 2002; Zeng et al., 2017), such as the node degree, node strength, and clustering coefficient. Here, we employed the essential measures of a complex network analysis to characterize the urban mobility networks. Moreover, we examined two types of scaling relations among these measures and then compared these relations between the empirical urban mobility networks and the two reference networks.

### 4.3.1 Statistical measures

**Node degrees \( k_i \) and degree distribution \( P(k) \).** The node degrees \( k_i \) and the degree distribution \( P(k) \) are important quantities that reveal the spatial heterogeneities of urban mobility (Jacob et al., 2017). Nodes with larger degrees represent more highly connected areas in the city. The distribution of the node degrees captures the number of nodes with a given degree \( k \) in the mobility network. For a given network, the node degree \( k_i \) is defined as the number of nodes to which node \( i \) is connected, as shown in Equation (1) (Wu & Zhang, 2011).

\[
k_i = \sum_{j \in V} N(v_i, v_j)
\]  

Regarding the distribution of \( k \), in the SF-MN, \( P(k) \) is a fat-tailed power-law distribution, while in the RA-MN, \( P(k) \) is a Poisson distribution. In real urban mobility networks, due to the influence of physical constraints, some deviations can be observed.

**Node strengths \( s_i \) and strength distribution \( P(s) \).** The node strength \( s_i \) is employed to generalize the degree measure of weighted networks. The strength of node \( i \) is defined as the sum of the weights of the edges associated with node \( i \), as shown in Equation (2).

\[
s_i = \sum_{j \in V} W(v_i, v_j)
\]  

The strength distribution \( P(s) \) represents the number of nodes that are associated with edges (e.g., travel flows in the urban mobility network) with the strength \( s \); and a higher node strength suggests that this location attracts more travel flows from other locations and has the potential to be a hub node.

**Local and average clustering coefficients.** The local clustering coefficient of a node is a measure of the neighborhood density and captures the degree to which the neighbors of this node are linked with each other (Opsahl & Panzarasa, 2009). A high local clustering coefficient of a node indicates that individuals who visit this node will also
frequently visit its neighbors. For node $i$, its local clustering coefficient $c(i)$ is the fraction of the links that are actually present among the total possible links between its neighbors. The equation for the weighted local clustering coefficient of node $i$, as defined by Barrat, et al. (2004), is

$$c_w(i) = \frac{1}{s_i(k_i-1)} \sum_{j,k} W(\ell_i,\ell_j) + W(\ell_j,\ell_k) \frac{a_{ij}a_{jk}a_{ki}}{2}$$  \hspace{1cm} (3)$$

where $a_{ij}$ are the elements of the adjacency matrix. The average clustering coefficient of all nodes, $\langle C_w \rangle$, can be applied to quantify the density of the entire network.

$$\langle C_w \rangle = \frac{\sum_{i \in V} c_w(i)}{N}$$  \hspace{1cm} (4)$$

### 4.3.2 Scaling relations

The scaling relation examines strong trends that are observed among complex network-driven measures, such as degree, strength, and clustering coefficient. The scaling relation is a useful tool for obtaining a global trend of the mobility network of the whole city (Brú et al., 2014).

**Strength $s$ versus degree $k$.** The node strength $s^w(k)$ averaged over all nodes of degree $k$ is given by

$$s^w(k) = \frac{1}{N(k)} \sum_{i/k_i=k} s_i$$  \hspace{1cm} (5)$$

The scaling relation between $s^w(k)$ and $k$ is indicative of the statistical correlations between the weights of the network and the connectivities of the network (Barrat et al., 2004). This relation is given by

$$s^w(k) \sim Ak^\beta$$  \hspace{1cm} (6)$$

In an urban mobility network, this scaling relation quantifies the visit growth of the urban nodes of different degrees. If $s^w(k)$ grows linearly with $k$, then $\beta = 1$. If no linear increase occurs, then $\beta \neq 1$ or $\beta = 1$ with $A \neq \langle w \rangle$. Therefore, $\beta$ reflects how the travel flows per edge increase with the connectivity of the urban nodes.

**Clustering coefficient $c$ versus degree $k$.** The weighted clustering coefficient $C_w(k)$ for nodes of a given degree $k$ is calculated as

$$C_w(k) = \frac{1}{N(k)} \sum_{i/k_i=k} C(i)$$  \hspace{1cm} (7)$$
The scaling relation between $C_w(k)$ and $k$ indicates the correlations between the neighborhood density and the connectivity of the network (Liu et al., 2016). This relation can be generally expressed as

$$C_w(k) \propto k^{-B\alpha}$$  \hspace{1cm} (8)

In this case of an urban mobility network, this scaling relation quantifies how spatial neighbored clusters are organized among the nodes of different degrees. A decreasing scaling relation indicates that denser neighborhoods tend to show lower connectivity.

5 Results

5.1 Properties of individuals’ travel motifs

After processing the dataset, hundreds of motifs were identified from the raw mobile phone tracking data. A total of almost 91.7% of 9.7 million phone users could be characterized by 475 eligible unique motifs. A total of 2.5 million mobile phone users were omitted due to their one-stay sequences. We selected the top 9 motifs as the most frequent motifs for further processing. Fig. 7 depicts the chosen motif structures and their probabilities among the user population. Different colors indicate the variation in the number of nodes in a motif, which range from 2 to 5. A substantial heterogeneity exists among individuals. It can be observed that the percentage of population decreases as the number of nodes increases; the highest percentage corresponds to $n = 2$ (40.1%), followed by the motifs with $n = 3$ (25.4%) and $n = 4$ (7.5%). The most frequent motifs can be divided into two distinct motif types, i.e., the round-trip type and the multiple-trip type. The motifs of the round-trip type are ID-2-1, ID-3-1, and ID-4-1, while the other motifs belong to the multiple-trip type. Round-trip motifs have simpler structures and are a more effective way to satisfy the travel demands. Therefore, higher percentages of population do the round-trip motifs within their respective node number groups. The findings indicate that motifs with fewer nodes and round-trip structures are preferred by a larger number of individuals. These individuals show strong regularities of movements that tend to follow certain typical motifs. This observation is consistent with the results of Song, et al. (2010), who discovered that human movements are of the high regularity.
Fig. 7. Top 9 motif types extracted from the mobile phone tracking data as frequent motifs.

To uncover the spatial disparities in the top-9 motif, the entropy of the various motifs occurring at the spatial node (here, base tower) was calculated. Fig. 8 displays the distribution of the entropy values. The entropies of the motifs that occur in spatial nodes show a similar Gaussian normal distribution with a mean value of 1.15. This finding indicates that the occurrence of different type motifs is not homogeneous across the whole city. The travel motifs vary from place to place.
The probabilities of occurrences of the top 9 motifs in 10 administrative districts were calculated (Fig. 9). The suburban areas hold a higher population of simpler motifs. For example, the probabilities of two-node motifs observed in the Pingshan, Dapeng, and Longgang are 0.45, 0.44, and 0.43, respectively (Table. 2). Conversely, the corresponding value for the central areas, including Luohu, Nanshan, and Futian, are 0.36, 0.36 and 0.36, respectively. However, motifs with three or more nodes occur in higher proportions in the urban areas than in the suburban areas, with values of 0.64 versus 0.40, respectively, on average. In addition, it can be observed that for each node number group, round-trip motifs hold in higher percentages in the urban centers, while multiple-trip motifs occur in higher proportions in suburbs. For instance, ID-3-1 accounts for a proportion of 0.79 of all three-node number groups in the central areas and a proportion of 0.68 in suburban areas, while the corresponding values for ID-3-2 are 0.17 and 0.25, respectively. We conjecture that one major reason for this finding is that people who live in the suburban areas tend to have fewer activities than those who live in the urban central areas. The result is in line with some empirical studies on the human
activities in metropolitan cities, which find that residents in the suburban areas have a
simple daily activity routines (Yang, Fang, Xu, et al., 2019). This finding can be further
explained by the possible determination of the abundance level of urban resources (i.e.,
bus stations, railway stations, metros, shopping malls, hospitals parks, etc.). More
abundant urban resources and higher-level socioeconomic population may have more
efficient motifs.

![Fig. 9. Probability distributions of the 9 motif types in the 10 administrative districts of Shenzhen.](image)

<table>
<thead>
<tr>
<th>Administrative districts</th>
<th>Two-node motifs</th>
<th>Three-node motifs</th>
<th>Four-node motifs</th>
<th>Five-node motifs</th>
</tr>
</thead>
<tbody>
<tr>
<td>The SEZ districts</td>
<td>0.37</td>
<td>0.41</td>
<td>0.19</td>
<td>0.05</td>
</tr>
<tr>
<td>Nanshan</td>
<td>0.36</td>
<td>0.41</td>
<td>0.19</td>
<td>0.04</td>
</tr>
<tr>
<td>Futian</td>
<td>0.36</td>
<td>0.41</td>
<td>0.19</td>
<td>0.05</td>
</tr>
<tr>
<td>Luohu</td>
<td>0.36</td>
<td>0.41</td>
<td>0.19</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 2. Probabilities of the 4 motif groups in the 10 administrative districts of Shenzhen.
<table>
<thead>
<tr>
<th>District</th>
<th>In-degree</th>
<th>Out-degree</th>
<th>In-weight</th>
<th>Out-weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yantian</td>
<td>0.38</td>
<td>0.41</td>
<td>0.19</td>
<td>0.04</td>
</tr>
<tr>
<td>The non-SEZ districts</td>
<td>0.42</td>
<td>0.40</td>
<td>0.15</td>
<td>0.03</td>
</tr>
<tr>
<td>Bao’an</td>
<td>0.41</td>
<td>0.40</td>
<td>0.16</td>
<td>0.03</td>
</tr>
<tr>
<td>Longgang</td>
<td>0.43</td>
<td>0.40</td>
<td>0.15</td>
<td>0.03</td>
</tr>
<tr>
<td>Pingshan</td>
<td>0.45</td>
<td>0.39</td>
<td>0.14</td>
<td>0.02</td>
</tr>
<tr>
<td>Guangming</td>
<td>0.41</td>
<td>0.40</td>
<td>0.16</td>
<td>0.03</td>
</tr>
<tr>
<td>Longhua</td>
<td>0.41</td>
<td>0.40</td>
<td>0.16</td>
<td>0.03</td>
</tr>
<tr>
<td>Dapeng</td>
<td>0.44</td>
<td>0.40</td>
<td>0.15</td>
<td>0.03</td>
</tr>
</tbody>
</table>

### 5.2 Properties of the global urban mobility network

The travel motifs of all individuals were aggregated and mapped onto the geographic space, as shown in Fig. 10. After the aggregation of all individuals’ travel motifs, the G-UMN consists of 5,934 nodes, 2,725,000 edges, and 15,499,967 weights (i.e., total trips), which cover the entire study area.

![Geographical mapping of the global urban mobility network (G-UMN).](image)

Now, let us focus on the complex network-oriented measures that characterize the G-UMN. The average degree \( \langle k \rangle \) of the G-UMN is 918.4 (including in-degree and out-degree), which indicates that, on average, each base tower is connected with 460 other base towers by individuals’ movements and that the connectivity of the mobility network is relatively high. Fig. 11(a) shows the degree distribution \( P(k) \) on a log-linear plot. The
red points correspond to the empirical data that are aggregated to form the G-UMN, and the purple line corresponds to an exponential fit, which is shown by a straight line. We also show the Poisson distribution that is predicted with the same average degree $\langle k \rangle$ as the G-UMN (green line) and a similarly predicted power-law distribution (yellow line). It can be observed that the empirical $P(k)$ obeys an exponential distribution ($P(k) \propto e^{-0.001k}$). Fig. 11(b) shows the strength distribution $P(s)$ on a log-linear plot. Similarly, $P(s)$ is also fitted with an exponential distribution ($P(s) \propto e^{-0.002s}$). The exponential distribution also has an obvious long tail, which indicates a heterogeneous spatial pattern. The deviation of the empirical behavior from Poisson distribution and power-law distribution suggests that the G-UMN can be characterized neither by a completely random spatial distribution nor by a purely scale-free spatial distribution. In addition, the G-UMN shows a large average clustering coefficient $\langle C_w \rangle = 0.59$, which is significantly larger than that of the RA-UMN ($\langle C_w^{RA-MN} \rangle = 0.15$); the G-UMN is rather clustered and is far from a random distribution. These findings imply that the urban mobility in the study area is heterogeneous: some areas attract a large number of travel flows, while other areas are visited by few individuals. Residents tend to travel more frequently to nearby locations, and cross-regional travels are rare. Thus, the G-UMN forms locally clustered areas.

Fig. 11. Properties of the G-UMN. (a) Distribution of the node degrees. (b) Distribution of the node strengths.
To further investigate how the G-UMN has developed on a global scale, we analyzed the scaling relation between the number of trips (strength) and the number of nodes (degree). In Fig. 12(a), the red points and blue points represent the in-degree strength and out-degree strength, respectively, of each node in the G-UMN. The profile of strength versus degree resembles a straight line when plotted in logarithmic coordinates, as shown by the purple line, which corresponds to $s^w(k) \sim Ak^\beta$ with $\beta = 1.24$, whereas the green line corresponds to the properties of the RA-UMN with $\beta = 1$. The observation of $\beta > 1$ suggests that trips that originate or end in highly connected areas occupy more flows than they would occupy in a random network. More importantly, the volume of travel trips will have an increase at a faster rate than the increase of connectivity of urban areas. In other words, more highly connected areas in the city can attract a disproportionately larger number of travel flows. The finding suggests that the improvement of the connectivity in urban areas will accelerate population flows.

The other scaling relation of interest is that between the clustering coefficient of a spatial node and its degree. Fig. 12(b) shows the empirical behavior of the G-UMN in terms of clustering coefficients versus degrees. The blue points and red points correspond to $C_w(k)$ and $C(k)$, respectively, and the purple line and cyan dotted line represent the corresponding power-law fits, where $C_w(k) \propto k^{-0.15}$ and $C(k) \propto k^{-0.22}$, respectively. Fig. 12(b) indicates that $C(k) < C_w(k)$, which means that nodes of higher degrees accumulate a larger number of travel flows. For comparison, we also presented green points and a yellow dotted line to show the relation that characterizes the SF-UMN ($C_w(k) \propto \text{const}$). The empirical observation of a decreasing relation indicates that urban areas with denser neighborhoods do not tend to show higher connectivity; instead, the opposite tendency is observed.

As proven by the work of Dorogovtsev, et al. (2002), networks that exhibit scaling relations of the form $C_w(k) \propto k^{-B\alpha}$ are considered hierarchical networks, where a scaling exponent of $\alpha = 1$ indicates a complete hierarchy. A hierarchical structure implies that sparsely connected areas tend to be part of highly clustered areas, where the links between the different highly clustered neighborhoods are maintained by only a few hubs (Ravasz & Barabási, 2003). A few local hubs attract quantities of travel flows and form hierarchically polycentric groups within the city. Each group is internally heterogeneous.
Here, the G-UMN is empirically observed to obey this relation with $\alpha = 0.15$. This finding suggests that the city possesses an evolving hierarchically polycentric structure, which coincides with the reality of the fast-growing city in the world (Liu et al., 2016).

![Figure 12](image)

**Fig. 12.** (a) Scaling relation between strength and degree. (b) Scaling relation between clustering coefficient and degree.

**5.3 Differences in the properties of the motif-dependent urban mobility networks (MD-UMNs)**

Based on the nine extracted frequent motifs, we constructed nine MD-UMNs. We further presented a comparative quantitative analysis of the statistical measures and scaling relations of the MD-UMNs, which reflects the impacts of the heterogeneities of individual travel motifs and the differences in the urban spatial structure. Table 3 summarizes the results for the statistical properties of these nine networks. The total number of nodes for all 9 MD-UMNs was set to 5,926. There are differences among these network statistical properties. The top four MD-UMNs are the ID-2-1, the ID-3-1, the ID-3-2 and the ID-4-1 networks. The ID-5-1 and ID-5-2 networks are relatively small. Specifically, for the top 2 MD-UMNs, i.e., ID-2-1 and ID-3-1, which were constructed based on the two-node round-trip motif and the three-node round-trip motif, respectively, the average degree $\langle k \rangle$ of the ID-2-1 network ($\langle k \rangle = 450.36$) is slightly smaller than that of the ID-3-1 network ($\langle k \rangle = 464.78$), while the average strength of the ID-2-1 network ($\langle s \rangle = 2064.97$) is nearly two
times greater than that of the ID-3-1 network ($\langle s \rangle = 1406.55$). These observations indicate that these people had specific spatial dispersion patterns in terms of motifs. The ID-2-1 network has a more spatially aggregated distribution of interacted strengths, while ID-3-1 network has a more dispersed distribution of interacted strengths. This is related to the finding abovementioned in section 5.1 that people in different areas have different activity demands according to the urban abundance and socioeconomic levels.

Fig. 13 illustrates the distributions of the node strengths of the MD-UMNs on a log-linear plot. The strength values were normalized with respect to the average weights $\langle w \rangle$. The points in different colors correspond to different MD-UMNs. The distributions of all 9 MD-UMNs show similar patterns, which are well fitted by exponential distributions; however, there are differences in the rate parameters of the fitted distributions. The fitted rate parameters range from 0.006 to 1, as summarized in Table 3. The 4 round-trip MD-UMNs, i.e., ID-2-1, ID-3-1, ID-4-1, and ID-5-1 occupy the highest proportion in its respective node number group. ID-2-1 and ID-3-1 have the lowest decay rate (0.006), ID-4-1 has the median decay rate (0.01), and ID-5-1 network has the highest decay rate (0.06). The variations in parameters suggest that the spatial heterogeneities of urban mobility exist and differentiate when considering different travel motif types. ID-5-1 corresponds to more concentrated spatial patterns of hub nodes and fewer hub nodes than ID-2-1. The larger is the node number, the more complex is the individual motif, and thus, the more centralized are the spatial patterns of the urban mobility networks. The results further imply the hypothesis of complex influences of the structures of individual travel on the spatial patterns of urban mobility.
Fig. 13. Distributions of the node strengths for the nine *MD-UMNs*.
Table 3. Statistical properties of the G-UMN and the 9 MD-UMNs

<table>
<thead>
<tr>
<th></th>
<th>G-UMN</th>
<th>ID-2-1</th>
<th>ID-3-1</th>
<th>ID-3-2</th>
<th>ID-3-3</th>
<th>ID-4-1</th>
<th>ID-4-2</th>
<th>ID-4-3</th>
<th>ID-5-1</th>
<th>ID-5-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of individuals</td>
<td>5655782</td>
<td>3059250</td>
<td>1389202</td>
<td>447220</td>
<td>97395</td>
<td>406024</td>
<td>150299</td>
<td>11485</td>
<td>75937</td>
<td>18970</td>
</tr>
<tr>
<td>Number of travel flows</td>
<td>15499967</td>
<td>6118500</td>
<td>4167606</td>
<td>1788880</td>
<td>486975</td>
<td>1624096</td>
<td>751495</td>
<td>68910</td>
<td>379685</td>
<td>113820</td>
</tr>
<tr>
<td>Ave. node degree</td>
<td>919.8</td>
<td>450.36</td>
<td>464.78</td>
<td>145.59</td>
<td>37.51</td>
<td>250.14</td>
<td>98.22</td>
<td>12.88</td>
<td>83.38</td>
<td>26.41</td>
</tr>
<tr>
<td>Ave. node in-degree</td>
<td>459.8</td>
<td>225.18</td>
<td>232.39</td>
<td>72.80</td>
<td>18.75</td>
<td>125.07</td>
<td>49.11</td>
<td>6.44</td>
<td>41.69</td>
<td>13.21</td>
</tr>
<tr>
<td>Ave. strength</td>
<td>5231.17</td>
<td>2064.97</td>
<td>1406.55</td>
<td>603.74</td>
<td>164.35</td>
<td>548.13</td>
<td>253.63</td>
<td>23.26</td>
<td>128.14</td>
<td>38.41</td>
</tr>
<tr>
<td>Ave. in-strength</td>
<td>2615.59</td>
<td>1032.48</td>
<td>703.27</td>
<td>301.87</td>
<td>82.18</td>
<td>274.06</td>
<td>126.81</td>
<td>11.63</td>
<td>64.07</td>
<td>19.21</td>
</tr>
<tr>
<td>Ave. weight</td>
<td>5.69</td>
<td>4.59</td>
<td>3.03</td>
<td>4.15</td>
<td>4.38</td>
<td>2.19</td>
<td>2.58</td>
<td>1.81</td>
<td>1.54</td>
<td>1.45</td>
</tr>
<tr>
<td>Ave. undirected cc</td>
<td>0.21</td>
<td>0.23</td>
<td>0.28</td>
<td>0.17</td>
<td>0.21</td>
<td>0.22</td>
<td>0.17</td>
<td>0.17</td>
<td>0.14</td>
<td>0.13</td>
</tr>
<tr>
<td>Ave. weighted cc</td>
<td>0.60</td>
<td>0.41</td>
<td>0.54</td>
<td>0.38</td>
<td>0.58</td>
<td>0.45</td>
<td>0.43</td>
<td>0.30</td>
<td>0.33</td>
<td>0.29</td>
</tr>
<tr>
<td>Rate parameter in strength distribution</td>
<td>0.002</td>
<td>0.006</td>
<td>0.006</td>
<td>0.03</td>
<td>0.25</td>
<td>0.01</td>
<td>0.06</td>
<td>1</td>
<td>0.06</td>
<td>0.32</td>
</tr>
</tbody>
</table>
To confirm the abovementioned hypothesis, we further analyzed the spatial patterns of the node strengths of the MD-UMNs by mapping the strength values of the nodes onto a geographic space. Because the strength values vary over several orders of magnitude, we normalized them using the min-max normalization method. Each strength value is normalized by subtracting the minimum strength and dividing by the difference between the maximum strength and the minimum strength to rescale the range of the strength values to [0, 1]. The spatial distributions of the normalized strengths of the top 4 MD-UMNs are illustrated in Fig. 14. The red color indicates that the nodes have higher strengths, and thus, act as hub nodes, whereas the yellow color represents smaller values. Moreover, the larger the circle size is, the higher the strength is. The results reflect the differences in the spatial configuration of the hub nodes. In terms of ID-2-1, Fig. 14(a) reveals that hub nodes are relatively well distributed in the urban central districts. Regarding ID-3-1, Fig. 14(b) demonstrates that a cluster of hub nodes is located in the suburban districts. The central areas have a low level of strength nodes, which is quite different from those indicated by the ID-2-1 network. Regarding the ID-4-1 network, the pattern is similar to that derived from the ID-3-1 network (Fig. 14(c)). For the ID-5-1 network, the hub nodes are concentrated in the central districts (Fig. 14(d)). These observations support the hypothesis that the different spatial patterns of urban mobility are caused by the structures of individual travels. The spatial pattern of the hub nodes in ID-2-1 is relatively scattered, whereas that in ID-5-1 is relatively centralized. This finding is consistent the findings of related studies, which indicates a strong positive correlation between the number of visited locations and the scope of the spatial dispersion distribution (Xu et al., 2015).
Fig. 14. Spatial distributions of the normalized node strengths in the MD-UMNs that correspond to motifs ID-2-1, ID-3-1, ID-4-1, and ID-5-1.

The results of the scaling relations between degree and strength in the MD-UMNs are displayed in Fig. 15(a). For this analysis, the strength values were normalized with respect to the average weight $\langle w \rangle$. Points of different colors represent empirical data from different MD-UMNs, and the black dotted line corresponds to the linear scaling relation in the RA-UMN with $\beta = 1$. The scaling exponents $\beta$, which are listed in Table 4, range from 1.08 to 1.33. All of these $\beta$ are larger than 1. However, there are deviations from the scaling relation. For smaller degree values, the strength increases super-linearly with the node degree, which indicates that the strengths of the urban nodes increase at a faster rate than their degrees when the node degrees are low. This increasing trend, however, shows a linear increase for larger degree values, which suggests that an urban area of higher degree tends to proportionately attract more travel flows (of which it may be either the origin or the destination).

The results of the scaling relations of the local weighted clustering coefficients with respect to the node degrees in the MD-UMNs are shown in Fig. 15(b). Points in different
colors correspond to different MD-UMNs. As in the case of the G-UMN, this scaling relation can always be fitted with a power law for any MD-UMN. The structures of the MD-UMNs differ only in the value of the scaling exponent $\alpha$, and do not differ in the general form of the scaling relation. The scaling exponents $\alpha$, which are listed in Table 4, range from 0.13 to 0.42. All of these $\alpha$ are larger than 0. These results demonstrate that all of these networks have a decreasing scaling relation between the clustering coefficients and the node degrees. Different values of $\alpha$ imply different evolutionary states of structures of urban mobility networks. Smaller values of $\alpha$ indicate more random properties of networks, while larger values mean more hierarchical properties. The results suggest that the MD-UMNs ID-2-1 and ID-3-1 (for which $\alpha = 0.13$ and 0.21, respectively) tend to show more randomness, while the ID-4-3 and ID-3-3 (for which $\alpha = 0.40$ and 0.42, respectively) tend to be more hierarchical; the other MD-UMNs lie somewhere in between these results.

**Fig. 15.** (a) Scaling relations between degree and strength for different MD-UMNs. (b) Scaling relations of the local clustering coefficients $C_w(k)$ as functions of the degree $k$. 
Table 4. Fitting results for the scaling relations of the G-UMN and MD-UMNs

<table>
<thead>
<tr>
<th></th>
<th>G-UMN</th>
<th>ID-2-1</th>
<th>ID-3-1</th>
<th>ID-3-2</th>
<th>ID-3-3</th>
<th>ID-4-1</th>
<th>ID-4-2</th>
<th>ID-4-3</th>
<th>ID-5-1</th>
<th>ID-5-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>1.23</td>
<td>1.22</td>
<td>1.16</td>
<td>1.23</td>
<td>1.33</td>
<td>1.12</td>
<td>1.16</td>
<td>1.18</td>
<td>1.08</td>
<td>1.10</td>
</tr>
<tr>
<td>$\alpha$ (weighted)</td>
<td>0.16</td>
<td>0.13</td>
<td>0.21</td>
<td>0.22</td>
<td>0.42</td>
<td>0.24</td>
<td>0.33</td>
<td>0.40</td>
<td>0.31</td>
<td>0.38</td>
</tr>
<tr>
<td>$\alpha$ (undirected)</td>
<td>0.22</td>
<td>0.20</td>
<td>0.27</td>
<td>0.32</td>
<td>0.67</td>
<td>0.29</td>
<td>0.44</td>
<td>0.57</td>
<td>0.36</td>
<td>0.50</td>
</tr>
</tbody>
</table>
6 Discussion and conclusion

Quantitative measures for characterizing urban mobility networks have the potential to greatly advance a deeper understanding of urban mobility. By rethinking the recent science of “complex networks” and motivated by the increasing availability of big human tracking data, this paper has developed an overarching framework for characterizing urban mobility networks from the perspective of complex networks. In contrast to existing measures that focus on the aggregation of human mobility, this paper explores the impacts of the heterogeneity of individual travels and constructs multiple urban mobility networks to represent the corresponding heterogeneous characteristics. These urban mobility networks were investigated by computing statistical measures and modelling scaling relations that are based on complex network theory, which allows us to assess how the mobility networks have developed. Considering Shenzhen, China as an example, we have experimentally demonstrated the effectiveness of the proposed framework.

By investigating the properties of individuals’ travel motifs, the analysis results demonstrated that the individual travel motifs are structurally and spatially heterogeneous. This result conforms to the findings that population segregation, facility density and transport accessibility in different areas of the city are suggested as potential factors in the variation of motif distributions (Allen et al., 2012; Chen et al., 2018; Gao et al., 2018). Due to the abundance of the urban resources and accessible transport infrastructures in the central areas, residents tend to visit a larger number of places and exhibit more efficient travel motifs in these regions. Most low-socioeconomic-level population live in the suburbs of the city in China, which is different from the findings of many Western studies that these population concentrate in the urban centers (Mieszkowski & Mills, 1993). These population groups usually have less activity demands. This finding reinforces the finding that the spatial allocation of urban resources is an important factor that influence the motif choices of different population groups.

The statistical measures and scaling relations of the G-UMN enable us to better understand to what status an urban mobility network develops. The results stated that travel that originates or ends in highly connected urban areas occupies a larger number of flows, and forms some locally clustered areas. Consequently, less highly connected areas in suburban and outskirt areas attract fewer travel flows; however, areas with
denser neighborhoods show lower connectivity. The finding indicated that residents who live in the suburban areas tend to have fewer activity choices than those who live in the urban central areas. In addition, the results also implied that the G-UMN of a fast-developing city is undergoing an evolving hierarchically polycentric structure, in which it is developing from a random network into a scale-free network. Locally clustered areas may cause spatial heterogeneity in urban mobility and insufficient mobility issue, especially in the outskirts and peripheral areas. Therefore, the role of facility accessibility in these areas is central to improve the urban and transportation planning. On the one hand, it is necessary to build additional public transport and public service facilities to encourage diversified travels in suburban or outskirt areas. On the other hand, to avoid the partial congestion caused by the extreme attraction of urban hubs, when planning the establishment of urban infrastructures, policy makers should fully consider how to retain the hierarchical and polycentric structure of urban mobility. For instance, connecting central hubs with more expressways and ensuring that alleys are unblocked within each district are effective ways to maintain the polycentric structure.

Finally, the exploration of the differences in the properties among the MD-UMNs provided insights to the differences in the multifaced urban mobility networks and indicated that the spatial heterogeneities among different motif types. The results suggested that the urban network structures are influenced by individual mobility. The behavioral differences in network properties and spatial heterogeneities of urban mobility vary across the MD-UMNs. Generally, simple motifs exhibit relatively dispersed spatial patterns, while more complex motifs are associated with highly clustered and centralized structures. These findings emphasized the spatial patterns of urban mobility networks for future policymaking. The implication lies in the elucidation of the structural complexity of urban mobility networks as characterized by the diversity of individual mobility. The complex motifs easily form highly clustered network structures, such as ID-5-1 in urban central areas. In many metropolitan cities, the trend of spatial inequality in urban resources due to the urban agglomeration effect (Fang & Yu, 2017; Partridge & Rickman, 2008), which exacerbates the scarcity of urban resources in the suburbs. This will continue to impact the urban form and structure. Therefore, allocating resources in a more
dispersed manner to satisfy the complex travel demands of the citizens can effectively reduce the overload of hub nodes that are centralized in special regions.

The results not only provide a promising bridge from complex network properties to urban mobility patterns but also imply potential urban planning policies. Our primary findings are complementary to urban studies but possess a different but typical urban context in light of urban development path. Some recent studies indicated there are polycentric metropolitan form with tiers of hierarchical centers in cities of Western countries, such as the San Francisco (Cervero & Wu, 1997) and London (Roth et al., 2011). China, especially other first-tier cities, including Beijing (Deng et al., 2019), and Shanghai (Xi Liu et al., 2015), has manifested a similar pattern. Despite its similarity, China has its unique policy guidance and urban-rural gap, resulting in different urbanization processes. For example, Shenzhen is a fast-developing city and its distribution of urban resources is more affected by the policy restrictions in early years and the differences between centers and suburbs is larger than that in Western cities. Thus, these differences reinforce that the future urban-oriented policies should be more targeted among cities considering the policy restrictions and development stages. For instance, whether maintaining the hierarchical structures or reducing urban hubs in a fast-developing city or a well-developed city should be discussed city by city. Our study needs to be extended of course to other urban areas, which will complement and enrich the urban studies.

Further work is needed to fully explore the potential applications of the proposed approach. First, we understand that the lack of a longer-term dataset influences the robustness of the results and limits their validity. However, the increasing availability of suitable data sources may solve this problem. Furthermore, with the availability of time-series data, our proposed framework may be extended to achieve the dynamic monitoring of the evolution of urban systems. Last, our proposed measures are rather simple and may be enhanced to include more comprehensive measures, such as statistics that consider the social perspective and capture dynamical and topological features. Beyond these possibilities, spatial influences, such as travel distance, could also be considered.
Declaration of Competing Interest

The authors declare no conflicts of interest.

References


