DOI: 10.1111/tgis.12849

RESEARCH ARTICLE

Transactions in GIS WILEY

Recurrent origin–destination network for exploration of human periodic collective dynamics

Xiaojian Chen[1] | Jiayi Xie[2] | Changjiang Xiao[3] | Binbin Lu[2] | Jie Shan[4]

[1]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

[2]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

[3]College of Surveying and Geo-informatics, Tongji University, Shanghai, China

[4]Lyles School of Civil Engineering, Purdue University, West Lafayette, Indiana, USA

Correspondence
Binbin Lu, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China.
Email: binbinlu@whu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 42071368, 41725006, 41871287 and U1833201

Abstract

While daily periodic movements of individuals have been widely studied, their collective dynamics are not understood. To capture periodic collective dynamics, this article represents individual daily movements as a time series of directed weighted origin–destination (OD) networks, and proposes an approach to identify a sub-network called the "recurrent OD network", which contains frequent edges appearing in each day. Taxi trajectory data over a period of 6 months in Wuhan, China are used for the case study. Here, we extracted the recurrent OD networks for each 2-h period on a given day, and compared them with the corresponding "major OD network" defined by both frequent and infrequent edges. Results show that the recurrent OD networks coincidentally exhibit spatially localized community structures and distinctive patterns of inflow and outflow for each region within a day. Overall, both methodology and findings in this study might make significant contributions in a range of fields, such as urban planning, regional economic development, and infectious disease control.

1 | INTRODUCTION

As one of the most common behaviors, people's daily intra-urban movement is the most direct manifestation of an interaction process between humans and a given city (Chen, Ma, Susilo, Liu, & Wang, 2016). Exploring people's mobility behavior is important for understanding their internal behavior mechanisms (Alessandretti, Aslak, & Lehmann, 2020), and the structural patterns within a city (Jiang, Yin, & Zhao, 2009). In recent years,

the widespread usage of location recording equipment and techniques has provided us a great opportunity to quantitatively characterize human mobility (Batabyal & Bhaumik, 2015), such as cell phone tower data (Louail et al., 2014), check-ins on social media sites (e.g., Twitter and Facebook) (Luo, Cao, Mulligan, & Li, 2016; Yang, Xiao, et al., 2019), public transportation smart card records (Hasan, Schneider, Ukkusuri, & Gonzalez, 2013), Wi-Fi connection records (Sapiezynski, Stopczynski, Gatej, & Lehmann, 2015), taxi GPS trajectory data (Zhang, Xu, Tu, & Ratti, 2018), and bike-sharing data (Shui & Szeto, 2020).

Mobility patterns are commonly studied at either the individual or the collective level (Peng, Jin, Wong, Shi, & Lio, 2012). At the individual level, people are regarded as relatively independent and the common mobility patterns are studied (Jia, Jiang, Carling, Bolin, & Ban, 2012). At the collective level, similar individual movements are aggregated and represented as directed weighted origin–destination (OD) networks or graphs (Barthelemy, 2011), where nodes indicate intra-urban spatial regions (e.g., partition regions units (Osorio-Arjona & García-Palomares, 2019; Tang, Zhang, Chen, Liu, & Zou, 2018) or points of interest (Noulas, Shaw, Lambiotte, & Mascolo, 2015), and the weight of a directed edge from node *A* to node *B* implies the number of trips with *A* as the origin and *B* as the destination regardless of the specific paths traveled. The directed weighted edges are also often called "OD flows." Thus, collective dynamics can represent human mobility from a macroscopic perspective and characterize the interaction strength between spatial regions in a straightforward manner.

Daily periodic movement is the prominent feature for individual human mobility (Huang, Cheng, & Weibel, 2019; Karamshuk, Boldrini, Conti, & Passarella, 2011). People travel in a city according to some specific schedule: leaving for work, meeting with friends, eating in preferred restaurants, returning home, and so on. These behaviors happen on most days for individuals, and numerous studies have proved that individual movements are highly predictable (Cho, Myers, & Leskovec, 2011; Song, Qu, Blumm, & Barabási, 2010). This naturally raises a fundamental question: are there any daily periodic movements or patterns of human mobility at the collective level? More specifically, is there any OD flow that occurs periodically on most days? If so, what characteristics do they share?

Previous studies have found that OD flows are similar each day (Liu, Kang, Gao, Xiao, & Tian, 2012). OD flows could be influenced by the urban environment (Liu, Kang, Gong, & Liu, 2016). Since the urban environment is relatively stable, the overall OD flows are recurrent to some extent each day, despite the fact that each individual OD flow between two regions may be somewhat random. For example, round trips between residential and working regions are the most important activities in people's daily lives (Ekman, Keränen, Karvo, & Ott, 2008), and numerous interactions within or among popular working zones form part of the most important economic activities (Yang, Sun, Shang, Wang, & Zhu, 2019) happening every day. These daily regular mobilities also lead to the high similarity of the 24-h fluctuation of the volume of OD flows in each day (Barroso, Albuquerque-Oliveira, & Oliveira-Neto, 2020). Also, daily periodicity is often used as an important feature in the field of deep learning for OD-related problems (Wang, Fu, Zhang, Li, & Li, 2018). Therefore, the daily periodic phenomenon of OD flows has been directly or indirectly proven from many perspectives. However, how to directly extract the periodic OD flows from massive individual movement data and characterize the patterns are rarely studied.

In this study, we propose a concept of "recurrent OD network." The directed weighted edges of this network are called "recurrent OD flows" to capture the day-to-day periodic OD flows from massive individual mobilities. To identify these, we need to construct three types of networks in sequence. Firstly, we construct an OD network with mobility data for each day, known as the *dynamic OD network* (Holme & Saramaki, 2012). This is fundamental for capturing day-to-day periodic OD flows. Secondly, to filter out random movements and capture the major OD flows, we apply a head/tail breaks rule (Jiang, 2013) to retain edges with relatively large weights in the dynamic OD network, forming a *dynamic major OD sub-network*. Finally, to capture the periodic OD flows, frequent edges in the dynamic major OD sub-network are further preserved to form a so-called *recurrent OD network*.

To highlight the patterns of a recurrent OD network, we compare it with the corresponding *major OD network*, which contains all the edges of the dynamic major OD sub-networks. As such, the major OD network contains both frequent and infrequent OD flows. The frequent ones are identified as recurrent OD flows, while the infrequent ones are called "non-recurrent OD flows." A recurrent OD network can be regarded as the periodic part

of the corresponding major OD network. Thus, comparing the recurrent OD network based on the major OD network leads to a deeper understanding of the periodic pattern.

With recurrent OD networks identified, the perspectives of community structure and weight flux of each region are analyzed. On the one hand, the community structure of regions provides the mesoscopic properties (Yan, Jeub, Flammini, Radicchi, & Fortunato, 2018). It represents close interactions among city regions, and can be used to uncover the spatial structure of recurrent OD flows. On the other hand, the flux of each region gives insight into the microscopic characteristics. It indicates the number of total movements in a region. We use the net-flux ratio to analyze the pattern of recurrent OD flows at a node level. A case study is conducted with taxi trajectory data over a period of 6 months in Wuhan, China.

All in all, the main contributions of this work can be preliminarily summarized as follows. Firstly, to capture the periodic OD flows, we propose a novel framework to detect a network called the recurrent OD network. This framework incorporates daily dynamic OD networks, the head/tail breaks method and frequent edges. The identified recurrent OD network provides novel insights into the OD flows from both spatial and temporal perspectives. Secondly, we compare the community structures between the major and recurrent OD networks, and reveal the patterns of periodic and non-periodic interactions in regional interactions. Thirdly, this study introduces a novel measurement, the net-flux ratio, to analyze the patterns of regional flux, that is, the patterns of periodic and non-periodic OD at the node level.

The structure of this article is organized as follows. Section 2 introduces the related literature. Section 3 describes the concepts and methods proposed here. Section 4 describes the test data and analyzes the patterns. Section 5 discusses the potential contributions in some practical applications. Section 6 discusses future work and concludes.

2 | RELATED WORK

In recent years a large number of high spatial-temporal resolution location-based data sets have recorded the movement behaviors of people in the city, providing an objective description for urban dynamics (Pappalardo et al., 2015). To reveal significant patterns of urban dynamics, individual movements are often mapped into spatially embedded OD networks. There are generally two perspectives for understanding OD networks: static and dynamic (Li, Cao, Li, & Wu, 2020).

From the static perspective, all the movements are checked together without viewing the flows on each day. This concept of OD networks has been studied in geography for decades and has been applied to aspects ranging from economic research (Bronzini, Herendeen, Miller, & Womer, 1974) and population migration (Tobler, 1970) to transportation (Magnanti & Mirchandani, 1993). The long history of the spatial OD network is reviewed by Barthelemy (2011). One interesting work on OD networks within a city is by Chowell, Hyman, Eubank, and Castillo-Chavez (2003). Due to the limited observations at that time, they employed a pseudo-agent-based approach to mimic human mobilities in Portland, Oregon. In their work, the urban OD network was constructed by representing physical locations in the city as nodes, and simulating movements of agents to construct directed weighted edges. With the emergence of empirical location-based data, more work on urban OD networks has been done. For example, the properties of node degree, weight, travel distance, and centrality have been verified by following the power law (Guidotti, Monreale, Rinzivillo, Pedreschi, & Giannotti, 2016; Shao, Sui, Yu, & Sun, 2019). Moreover, strong correlations between the degree and weight (Saberi, Ghamami, Gu, Shojaei, & Fishman, 2018; Saberi, Mahmassani, Brockmann, & Hosseini, 2017; Tang et al., 2016) and community structures (Hossmann, Spyropoulos, & Legendre, 2011; Yang, He, Song, Fu, & Wang, 2018; Yildirimoglu & Kim, 2018) were also observed in real data sets. Besides mining patterns, various models to generate static OD networks have also been proposed, such as gravity (Roy & Thill, 2004), maximum entropy (Xie, Kockelman, & Waller, 2011), radiation (Simini, Gonzalez, Maritan, & Barabási, 2012), and matrix decomposition (Hamedmoghadam, Ramezani, &

Saberi, 2019; Louail et al., 2015; Qi et al., 2021). Therefore, the analysis of and models on static OD networks give insight into the laws of urban dynamics as a whole, and complex network theories are used extensively to reveal the patterns.

Different from static OD networks, dynamic OD networks are able to organize movements on a daily basis. These dynamic networks depict urban dynamics in a realistic way and have attracted the attention of many researchers. For example, Ferreira, Poco, Vo, Freire, and Silva (2013) presented a new system for the exploration of dynamic OD trips by using a visual query model allowing quick selection data slices on each day. Kim and Chung (2018) used a Gaussian mixture model to model the travel time distribution within a given day and analyze the OD-based day-to-day travel time variability. Bimpou and Ferguson (2020) incorporated day-to-day travel time reliability to measure regions' accessibility. Recently, there has been work focusing on predicting dynamic urban OD flows, primarily by adopting deep learning methods to capture spatiotemporal features automatically. Toqué, Côme, El Mahrsi, and Oukhellou (2016) were among the first to use the long short-term memory (LSTM) neural network model to estimate future OD flows by using the historic OD flows as input. Since their work, more complex deep learning models have been proposed to estimate dynamic future flows, among them convolutional LSTM (Duan et al., 2019), contextualized spatial-temporal networks (Liu et al., 2019), dual-stage graph convolutional recurrent neural networks (Hu, Yang, Guo, Jensen, & Xiong, 2020), spatial-temporal LSTM (Li et al., 2020), spatial-temporal encoder-decoder residual multi-graph convolutional networks (Ke et al., 2021), and dynamic node–edge attention networks (Zhang, Xiao, Shen, & Zhong, 2021). Therefore, dynamic OD networks provide a more comprehensive and detailed day-to-day description for urban dynamics and have become a powerful tool for understanding collective dynamics in recent years.

While many studies have been conducted using dynamic OD networks, less attention has been paid to periodic flows. The work of Andrienko, Andrienko, Fuchs, and Wood (2017) is the closest to our study. They proposed a spatial-temporal clustering approach for identifying significant OD flows. Rather than analyzing the patterns, their work mainly focused on detecting periodic and long-term temporal patterns visually. However, our work proposes a simple and intuitive approach to directly detect frequent sub-networks (called “recurrent OD networks”) from dynamic OD networks. Also, complex network theories are introduced to characterize the patterns.

3 | METHODOLOGY

This section first introduces the framework for identifying the recurrent OD network from individual movement data. Then complex network analysis, including community detection and flux of each region, are applied to explore the network structural pattern and specific recurrent OD flows in each region.

3.1 | Algorithm for identifying the recurrent OD network

The key problem of our work is to identify OD flows happening frequently from daily movement data. The technical framework for identifying recurrent OD network is shown in Figure 1. Its detailed procedure is defined as the following three steps.

1. Construct a dynamic OD network with daily movement data.
2. Identify large flows (i.e., edges with large weights) in each OD network from the dynamic OD network to produce a dynamic major OD sub-network.
3. Extract edges frequently appearing in the dynamic major OD sub-network and assign average weight through days to construct the recurrent OD network.

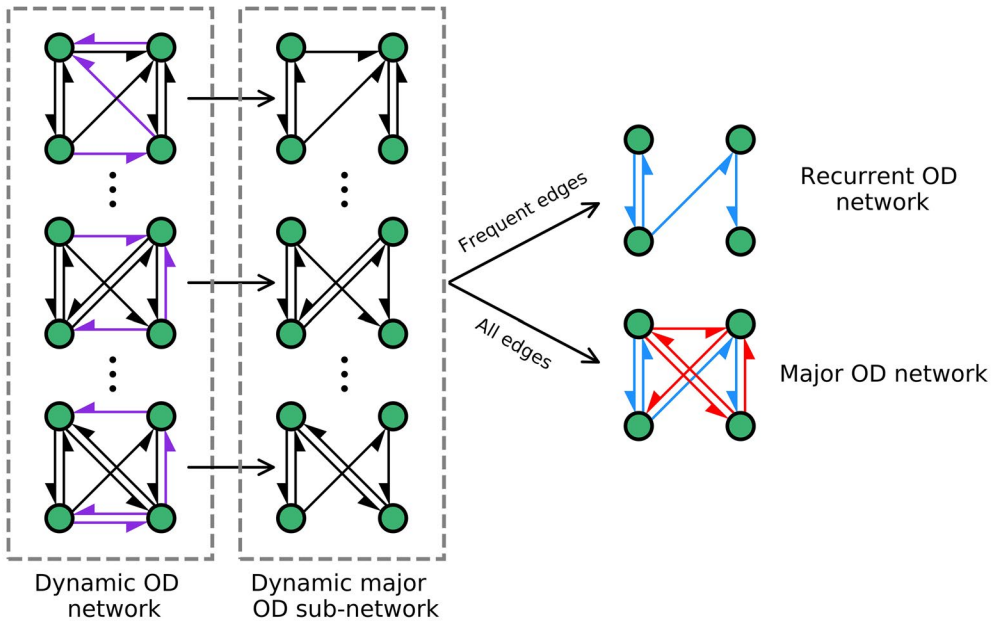

FIGURE 1 The technical framework of defining recurrent OD networks and major OD networks. Black and purple edges in the dynamic OD network indicate edges with strong and weak connections, respectively. Blue and red edges in the recurrent OD network and major OD network denote the edges appearing frequently and infrequently in the dynamic major OD sub-network

Here, the frequent edges indicate that the edges exist in most networks of the dynamic major OD sub-network. The concept of the major OD network is also introduced, which contains all edges of the dynamic major OD sub-network. The recurrent OD network and major OD network will be used to highlight the characteristics of periodic OD flows.

3.1.1 | Dynamic OD network

In the first place, we intend to construct a time series of OD networks which describes OD movements on each day from the movement of individuals. This is fundamental for exploring day-to-day periodic OD flows in the subsequent research and analysis. A common way to obtain this time-dependent network is to construct the corresponding network over each period of time (Holme & Saramaki, 2012). As such, a time series of OD networks is extracted, which is also called "dynamic OD flows" (Duan et al., 2019) or "dynamic OD matrix" (Liu et al., 2019) in some of literature. In this study, we construct an OD network on each day for daily periodic pattern analysis.

Also, the geographical meaning of a node is required for the construction of the OD network. We use the most common uniform grid measuring 1 km × 1 km as the unit region (Nanni, Tortosa, Vicent, & Yeghikyan, 2020), while the scale size can be adapted according to the applications. Also, a non-uniform partition such as traffic analysis zones can be applied if the information is available (Yu, 2019).

Finally, we have the following definition.

Definition 1 (Dynamic OD Network) A dynamic OD network $N = \{N_t, 1 \le t \le n, t \in \mathbb{N}^+\}$ is a time series of spatial directed weighted networks, where the OD network $N_t = (V_{N_t}, E_{N_t}, W_{N_t})$ is a snapshot of day t, V_{N_t} is the set of nodes that represent 1 km × 1 km partition grids, E_{N_t} is the set of directed edges which represent the OD flows between each pair of spatial nodes, and W_{N_t} is the set of weights that indicate the number of trips occurring to each edge.

3.1.2 | Dynamic major OD sub-network

With the dynamic OD network constructed, we aim to produce the corresponding dynamic sub-network that only retains edges with large weights. For an OD network, the larger the weight of an edge, the stronger the connections between the corresponding origin–destination pair. In this sense, edges with large weights can largely express the dominant structure of interactions in a specific OD network, while low weights are considered to be potentially caused by random behaviors (Serrano, Boguna, & Vespignani, 2009). Therefore, we try to keep the high-weighted edges to ensure that the subsequently extracted recurrent OD flows are core interactions.

This is done by setting a global weight threshold for a network, which is one of the most popular approaches in many practical applications, such as the brain (Lynall et al., 2010), airlines (Sawai, 2012), food webs (Bellingeri & Bodini, 2013), and biological networks (Allesina, Bodini, & Bondavalli, 2006). However, in practice, the selection of the threshold value is difficult (Yan et al., 2018). It often needs to be set by the user with the actual meaning and characteristics of the network itself. For the urban OD network on each day, it is still unknown how to select the threshold to capture the core interactions.

In this step, a data-driven head/tail breaks rule is introduced to guide the selection of the weight threshold for an OD network. This is because the weight distribution of an OD network is usually heavy-tailed, with the majority having smaller values and the minority larger values (Barthelemy, 2011), and the head/tail breaks rule is widely applied to deal with this distribution of the data (Jiang, 2013). It divides data into head (above the mean) and tail (below the mean) parts by the geometric mean, and then the head part can be further partitioned in this way. Under this partition scheme, the underlying hierarchy of the data can be captured. We iteratively run this partition until the ratio of the total evolved weight in the head is less than .4, where this threshold is recommended by some works to distinguish the minority from the majority (Ma, Osaragi, Oki, & Jiang, 2020; Ma, Sandberg, & Jiang, 2015). Therefore, this corresponds to repeated partitioning by the geometric mean of the remaining edges' weights until the upper 40th percentile set of weights for inclusion in the major sub-network is reached.

Note that the above method is a little different from finding the upper 40% quantile of data directly. The head/tail breaks method splits the data into multiple levels according to the above iterative procedure, thus representing a hierarchy of heavy-tailed distributed data (Li et al., 2016). Therefore, the above threshold selection method based on the head/tail breaks rule is equivalent to keeping the data exceeding the upper 40% quantile according to the level as the basic unit rather than the specific value of the data itself. See the specific example in Section 4.2.

To be specific, we define and construct a dynamic major OD sub-network as follows.

Definition 2 (Dynamic major OD sub-network) A dynamic major OD sub-network $M = \{M_t, 1 \le t \le n, t \in \mathbb{N}^+\}$ is a time series of spatial directed weighted networks extracted from the dynamic OD network $N = \{N_t, 1 \le t \le n, t \in \mathbb{N}^+\}$. $M_t = (V_{M_t}, E_{M_t}, W_{M_t}) \subseteq N_t = (V_{N_t}, E_{N_t}, W_{N_t})$, where $V_{N_t}, E_{N_t}, W_{N_t}$ are the corresponding sets of nodes, edges, and weights of N_t, $E_{M_t} = \{e_i | w_i \ge w_{\min}(t), (e_i, w_i) \in (E_{N_t}, W_{N_t})\}$, V_{M_t}, W_{M_t} are sets of corresponding involved nodes and weights. $M_t \subseteq N_t$ indicates that for a weighted edge $(e_i, w_i) \in (E_{M_t}, W_{M_t})$, it has $(e_i, w_i) \in (E_{N_t}, W_{N_t})$, and M_t is called the sub-network of N_t. The threshold $w_{\min}(t)$ is calculated by iteratively applying the head/tail breaks procedure until the first time when $\sum_{w_i \in W_{N_t}} w_i \mathbb{I}_{\{w_i \ge w_{\min}(t)\}} / \sum_{w_i \in W_{N_t}} w_i < .4$, where $\mathbb{I}$ is the indicator function.

3.1.3 | Recurrent OD network

The purpose of this step is to identify the edges appearing frequently in the time series of major OD sub-networks $M = \{M_t, 1 \le t \le n, t \in \mathbb{N}^+\}$. As such, "frequent" means the edges in most networks of the dynamic major OD sub-network. For an edge v representing the interaction from node a to node b, the number of occurrences of v in each M_t is at most 1 (with 0 meaning not present), although the weight may vary. Therefore, the number of an edge v in $M = \{M_t, 1 \le t \le n, t \in \mathbb{N}^+\}$ represents the frequency of its appearance as a core

interaction on most days. After normalization according to the total number of days n, we keep the edge v if the corresponding ratio exceeds the parameter $\beta \in [0, 1]$ and assign v with average weight in the study period. These retained weighted edges represent frequently occurring OD interactions on most days and form a spatial directed weighted network (called a recurrent OD network), so they are suitable for describing periodic OD interactions.

In detail, a recurrent OD network could be defined as follows:

Definition 3 (Recurrent OD network) Given a dynamic OD network $N = \{N_t = (V_{N_t}, E_{N_t}, W_{N_t}), 1 \le t \le n, t \in \mathbb{N}^+\}$, the corresponding dynamic major OD sub-network $M = \{M_t = (V_{M_t}, E_{M_t}, W_{M_t}), 1 \le t \le n, t \in \mathbb{N}^+\}$ and the frequency parameter β, the recurrent OD network $\overline{R} = \{V_{\overline{R}}, E_{\overline{R}}, W_{\overline{R}}\}$ is a spatial directed network with weighted edges, where $E_{\overline{R}} = \{e_i | \sum_{1 \le t \le n} \mathbb{I}_{e_i \in E_{M_t}}(e_i) \ge n\beta\}$, $V_{\overline{R}}$ is the evolved nodes, and $w_{i,\overline{R}} \in W_{\overline{R}}$ is calculated as:

$$w_{i,\overline{R}} = \frac{1}{n} \sum_{w_{i,t} \in W_{N_t}, 1 \le t \le n} w_{i,t} \tag{1}$$

3.1.4 | Major OD network

In this section we introduce the motivation to construct a spatial directed weighted network called a "major OD network". With the above definitions, the recurrent OD network is defined by those frequent edges in the dynamic major OD network. This also implies that the dynamic major OD network contains both frequent and infrequent edges. Therefore, to highlight the characteristics of those frequent OD flows, we collect all edges from the dynamic major OD sub-network and compare the major OD network as a benchmark with the OD recurrent network.

In detail, the major OD network can be defined as follows.

Definition 4 (Major OD network) Given a dynamic OD network $N = \{N_t = (V_{N_t}, E_{N_t}, W_{N_t}), 1 \le t \le n, t \in \mathbb{N}^+\}$ and the corresponding dynamic major OD sub-network $M = \{M_t = (V_{M_t}, E_{M_t}, W_{M_t}), 1 \le t \le n, t \in \mathbb{N}^+\}$, the major OD network $\overline{M} = \{V_{\overline{M}}, E_{\overline{M}}, W_{\overline{M}}\}$ is a spatial directed network with weighted edges, where $V_{\overline{M}} = \cup_{1 \le t \le n} V_{M_t}$, $E_{\overline{M}} = \cup_{1 \le t \le n} E_{M_t}$ and $w_{i,\overline{M}} \in W_{\overline{M}}$ is calculated as:

$$w_{i,\overline{M}} = \frac{1}{n} \sum_{w_{i,t} \in W_{N_t}, 1 \le t \le n} w_{i,t} \tag{2}$$

From the definitions above, mathematically we have $E_{\overline{R}} \subseteq E_{\overline{M}}$. In addition, for $e_i \in E_{\overline{R}}$ with weight $\overline{w}_{i,\overline{R}}$, we have $\overline{w}_{i,\overline{R}} = \overline{w}_{i,\overline{M}}$, since they are all calculated by the average weights from W_{N_t} of the dynamic OD network $N = \{N_t = (V_{N_t}, E_{N_t}, W_{N_t}), 1 \le t \le n, t \in \mathbb{N}^+\}$. Therefore, $\overline{R}$ is the subset of $\overline{M}$, that is all the weighted edges in the recurrent OD network are all in the major OD network. We call the weighted edges of $\overline{R}$ "recurrent OD flows" to capture the periodic OD flows, and the weighted edges in $\overline{M}/\overline{R}$ (i.e., in $\overline{M}$ but not in $\overline{R}$), are called "non-recurrent OD flows" to describe the non-periodic OD flows.

3.2 | Complex network analysis

This section first introduces the network community detection method used to analyze the network structure of the major OD network and the recurrent OD network. Then indices related to flux in each region are adopted to describe the detailed recurrent and non-recurrent OD flows at the node level.

3.2.1 | Community detection

In a network, a community refers to a cluster of nodes in which most of the nodes are connected internally while the connections between clusters are sparse. Different from a spatial cluster [e.g., max-*p*-regions (Duque, Anselin, & Rey, 2012) and REDCAP (Guo, 2008)] which mainly focus on aggregating close spatial objects, the network community is detected based on the connections of weighted edges between nodes. As such, the network community structure directly characterizes the regions densely connected by recurrent OD flows as well as non-recurrent OD flows, which reveals the periodic and non-periodic interaction structures between regions.

The major OD network $\overline{M}$ and recurrent OD network $\overline{R}$ are directed weighted networks, thus, the popular Infomap (Rosvall & Bergstrom, 2008) method considering both the direction and weight is adopted to identify communities. It captures the structure of a network by depicting a random walker on the network in a two-level representation: between clusters and within a cluster. If it finds an efficient two-level representation, clusters are then naturally identified. Infomap proposes to give a unique Huffman codeword to each cluster but reuse Huffman codewords within each cluster, which is similar to the fact that different cities could have the same street name. Under this two-level representation, the efficiency of a cluster partition C (i.e., dividing n nodes to m clusters) is given by:

$$L(C) = q_{\curvearrowright} H(\mathcal{Q}) + \sum_{i=1}^{m} p^i_{\circlearrowright} H\left(\mathcal{P}^i\right) \tag{3}$$

The first term describes the entropy of a random walk between clusters, and the second term is the summation of entropy of random walks within each cluster. Specifically, $q_{\curvearrowright}$ is the probability that a random walk moves among clusters, $H(\mathcal{Q})$ is the entropy of Huffman codeword of clusters, $p^i_{\circlearrowright}$ is the probability that movements happen in the cluster i, $H\left(\mathcal{P}^i\right)$ is the entropy of the Huffman codeword within cluster i. Finally, the most efficient partition C is computed by minimizing $L(C)$.

3.2.2 | Regional flux

The flux of a node (Saberi et al., 2018), also called "strength" (Zhong, Arisona, Huang, Batty, & Schmitt, 2014), refers to the total weights of a region. The edge weight represents the number of movements between two regions. From the perspective of a region (or node), there are inflow- and outflow-weighted OD flows, which means the flows ending or starting from the region. As such, the total inflow or outflow weights (called "influx" or "outflux") measure the regional attractiveness, which can be used to capture the daily regular movement of urban dynamics (Pan, Qi, Wu, Zhang, & Li, 2013; Qian, Liu, Tao, & Zhou, 2020; Zhou, Liu, Qian, Chen, & Tao, 2020). Therefore, analyzing the influx and outflux of recurrent and non-recurrent OD flows could provide the patterns of periodic and non-periodic urban dynamics at the node level.

In the following, we introduce the flux-related statistics with their forms in the major OD network $\overline{M}$ and the recurrent OD network $\overline{R}$.

Influx and outflux

Mathematically, for a region v_i, we have influx x (F_{in}) and outflux (F_{out}) defined *as* (Saberi et al., 2018):

$$F_{\text{in}}\left(v_i\right) = \sum_{v_j \in v_{\text{in}}(v_i)} w\left(v_j, v_i\right) \tag{4}$$

$$F_{\text{out}}\left(v_i\right) = \sum_{v_j \in v_{\text{out}}(v_i)} w\left(v_i, v_j\right) \tag{5}$$

where $v_{in}(v_i)$ ($v_{out}(v_i)$) is a set of nodes that have directed edges pointing to (starting from) node i, and $w(v_j, v_i)$ is the weight from node v_j to v_i.

For our analysis here, since the major OD network $\overline{M}$ contains recurrent and non-recurrent OD flows, for a region $v_i \in V_{\overline{R}} \subset V_{\overline{M}}$, the influx and outflux of the major OD network ($F_{in}^{\overline{M}}(v_i)$ and $F_{out}^{\overline{M}}(v_i)$) can be expressed as the form of flux resulting from recurrent and non-recurrent OD flows:

$$F_{in}^{\overline{M}}(v_i) = \sum_{v_j \in v_{in}^{\overline{M}}(v_i)} w(v_j, v_i) = \sum_{v_j \in v_{in}^{\overline{R}}(v_i)} w(v_j, v_i) + \sum_{v_j \in v_{in}^{\overline{M/R}}(v_i)} w(v_j, v_i) \triangleq F_{in}^{\overline{R}}(v_i) + F_{in}^{\overline{NR}}(v_i) \tag{6}$$

where $v_{in}^{\overline{M}}(v_i)$, $v_{in}^{\overline{R}}(v_i)$ and $v_{in}^{\overline{M/R}}(v_i)$ are the sets of nodes that have directed edges pointing to node i in $\overline{M}$, $\overline{R}$ (i.e., the recurrent OD flows) and $\overline{M/R}$ (i.e., the non-recurrent OD flows), and $\sum_{v_j \in v_{in}^{\overline{M/R}}(v_i)} w(v_j, v_i)$ is recorded as $F_{in}^{\overline{NR}}(v_i)$ for simplification. Similarly, we have:

$$F_{out}^{\overline{M}}(v_i) = F_{out}^{\overline{R}}(v_i) + F_{out}^{\overline{NR}}(v_i) \tag{7}$$

Therefore, the decomposition form of flux allows us to explain the periodic and non-periodic patterns more clearly in the following research.

Net-flux ratio

To comprehensively measure the dominated type and degree of influx and outflux in a region, we introduce the net-flux ratio (NFR) as:

$$\mathrm{NFR}(v_i) = \frac{F_{in}(v_i) - F_{out}(v_i)}{F_{in}(v_i) + F_{out}(v_i)} \tag{8}$$

Since $F_{in}(v_i)$ and $F_{out}(v_i)$ are both greater than 0, the value of $\mathrm{NFR}(v_i) \in [-1, 1]$. $\mathrm{NFR}(v_i) > 0$ indicates the region v_i is dominated by influx and that the stronger dominated degree is larger. Specifically, $\mathrm{NFR}(v_i) = 1$ means there is no outflux; and vice versa.

By applying NFR to the major OD network $\overline{M}$ ($NFR_{\overline{M}}$) and the recurrent OD network $\overline{R}$ ($NFR_{\overline{R}}$), we have the following form:

$$NFR_{\overline{M}}(v_i) = \frac{F_{in}^{\overline{M}}(v_i) - F_{out}^{\overline{M}}(v_i)}{F_{in}^{\overline{M}}(v_i) + F_{out}^{\overline{M}}(v_i)} = \frac{F_{in}^{\overline{R}}(v_i) + F_{in}^{\overline{NR}}(v_i) - F_{out}^{\overline{R}}(v_i) - F_{out}^{\overline{NR}}(v_i)}{F_{in}^{\overline{R}}(v_i) + F_{in}^{\overline{NR}}(v_i) + F_{out}^{\overline{R}}(v_i) + F_{out}^{\overline{NR}}(v_i)} \tag{9}$$

$$NFR_{\overline{R}}(v_i) = \frac{F_{in}^{\overline{R}}(v_i) - F_{out}^{\overline{R}}(v_i)}{F_{in}^{\overline{R}}(v_i) + F_{out}^{\overline{R}}(v_i)} \tag{10}$$

Ratio of three types of regions based on NFR

Furthermore, based on $NFR(v_i)$, we can classify the v_i into three classes: NFR(v_i) belongs to [–1, –1/3) when v_i is in the outflow region, to [–1/3, 1/3] when v_i is in the balance region, and to (1/3, 1] when v_i is in the outflow region. And we calculate the ratio of the three classes as:

$$r_{\mathrm{outflow}} = \frac{N_{\mathrm{outflow}}}{N_{\mathrm{outflow}} + N_{\mathrm{balance}} + N_{\mathrm{inflow}}} \tag{11}$$

$$r_{\mathrm{balance}} = \frac{N_{\mathrm{balance}}}{N_{\mathrm{outflow}} + N_{\mathrm{balance}} + N_{\mathrm{inflow}}} \tag{12}$$

$$r_{\text{inflow}} = \frac{N_{\text{inflow}}}{N_{\text{outflow}} + N_{\text{balance}} + N_{\text{inflow}}} \tag{13}$$

where N_{outflow}, N_{balance}, and N_{inflow} denote the number of v_i classifying as outflow, balance, and inflow. Therefore, these three ratios describe the flux types of regions based on OD flows within a city, which can be applied to the major OD network and recurrent OD network to capture periodic and non-periodic urban dynamics.

4 | RESULTS

This section conducts a case study with the taxi trajectory data of Wuhan. Section 4.1 introduces the study area and data information in this study. In Section 4.2 we divide the day into 12 segments (2 h per slot), and identify the major OD network and recurrent OD network for each time-slot. Section 4.3 compares the two networks quantitatively from the perspectives of community structure and flux of each region.

4.1 | Study area and data

The study area is the central urban part of Wuhan city. It lies between 30.462°N and 30.659°N latitude (around 22 km) and between 114.180°E and 114.409°E longitude (around 22 km); see Figure 2. Our study uses 1 km as the grid scale, which partitions the study area into $22 \times 22 = 484$ squares. Note that the study area is divided into east and west parts by the Yangtze River. Figure 2 also shows how the city area is partitioned into districts. The central business district (CBD) is located at the intersection of Qiaokou, Jianghan, and Jiangan near the riverbank; this is the political and economic center with the densest population in Wuhan.

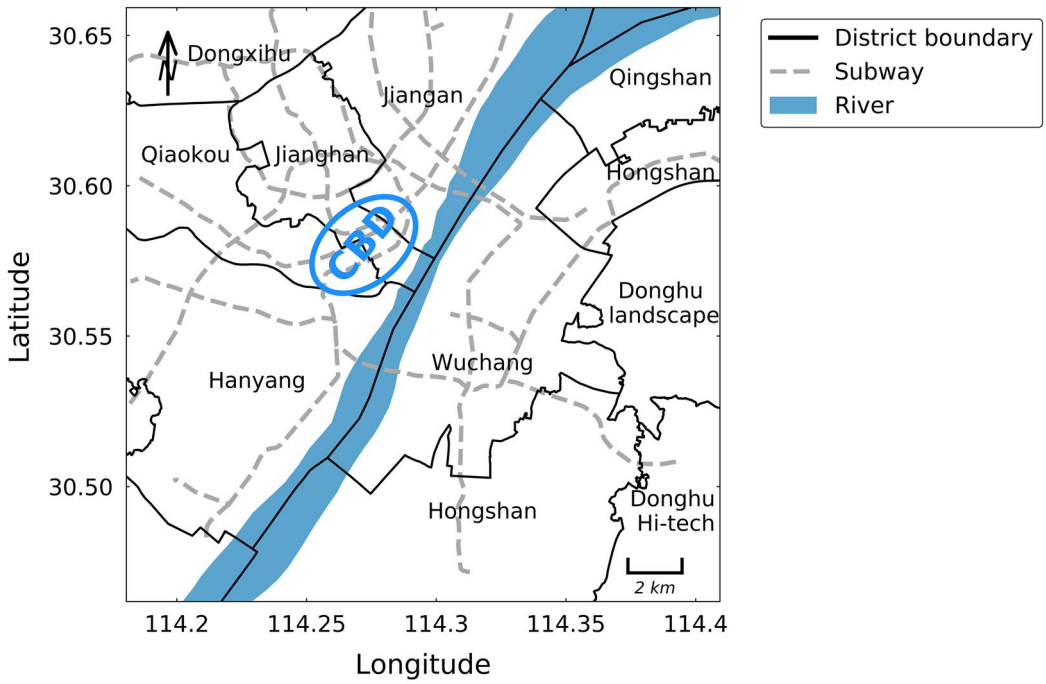

FIGURE 2 Map of Wuhan central urban area

TABLE 1 An example of taxi record when carrying passengers

Taxi ID	100
Time	February 12, 2015 14:03:38
Longitude	114.3234
Latitude	30.5521
Occupation status	Occupied

The data used in this case study contains trajectories collected from 7,983 taxis in Wuhan from February 1 to August 10, 2015. Each record consists of the taxi ID, time, longitude, latitude and occupation status ("empty" and "occupied"). A sample record is shown in Table 1. The data are provided by the Wuhan Municipal Transportation Administration, and only audited research institutions can access the data. Specific information on the driver corresponding to the taxi ID is not available to users due to privacy considerations. Moreover, the data do not contain any passenger information, recording only the passenger occupation status of the taxi.

Generally, an origin–destination pair for a taxi trip is meaningful for representing an individual movement only if it carries one or more passengers. We therefore chose the effective flows by taking the start and end of each trajectory segment when the status was "occupied." Note that the Chinese New Year (February 18–25), also known as the Spring Festival, is included in this period. On considering the irregular population movements during this specific period, which might distort our conclusions, we thus removed all the records collected in this period. In addition, we excluded days with only a few trajectories based on suggestions from the data provider. This is because bad weather or other unknown reasons make the number of trajectories very small in these days. As such, we retain a data set of trajectories of up to 29,557,561 records in 173 days. Finally, since our study mainly focuses on the movements between different regions, the records with pick-up and drop-off points in the same grid square are removed, which leads to 28,625,700 records in 173 days.

4.2 | Identifying major and recurrent OD network

Since people's movement behaviors are different at different times of the day, the OD flows also show changing patterns. To capture the dynamic patterns of the recurrent OD network, we divide they day into 12 time periods on average.

We first show the results of using the head/tail partition method to extract the major OD sub-network with relatively higher weights. We take the OD network $N_t = (V_{N_t}, E_{N_t}, W_{N_t})$ for 12:00–14:00 on March 5 as an example. In Figure 3, the ratio of weight w_0 is calculated as $\sum_{w_i \in W_{N_t}} w_i \mathbb{I}_{\{w_i = w_0\}} / \sum_{w_i \in W_{N_t}} w_i$, which describes the proportion of the number of trips with weight w_0. The linear decreasing trend at the log-log scale shows that most of the edges have very low weights and only a small fraction of the edges have relatively higher weights. The geometric mean of all the weights $w_i \in W_{N_t}$ is 1.27, and the first partition filters out the edges with a weight smaller than 1.27, so that a weight ratio of 0.63 remains (Figure 3a). We use this partition rule again in the remaining head part, which leads to a threshold (i.e., the geometric mean of the remaining weights) of 2.92 with a remaining ratio of 0.44 (Figure 3b). Then after the third partition, the geometric mean is 4.26. In this case, the data with weights 3 and 4 are classified as being at the same level, and after filtering them out, the remaining ratio is $0.24 < 0.4$, satisfying the stopping condition (Figure 3c). Although the remaining ratio is $0.32 < 0.4$, satisfying the stopping condition by filtering out only the data with a weight of 3, the idea of the head/tail breaks method is to maintain the hierarchical structure by level, so in the end, we retain all the edges with a weight larger than or equal to 5 as part of the major OD sub-network.

In this 2-h network example, there are a total of 10,799 edges (or interactions) with a total weight of 19,398. This shows that a total of 19,398 movements have occurred, but the number of different regional pairs involved is also very large (10,799). Therefore, at the scale of a 1 km grid, people's taxi travel behavior for 2 h a day is not so consistent. This may have a certain relationship with the high theoretical upper limit (233,772, see Appendix A in the online supporting Information) of interaction pairs in different regions at this scale. In addition, although the weight range is wide, from 1 to 19, all edges with a weight less than 4 account for 76% of the movement, which is close to the "80–20 rule" (Newman, 2005).

Figure 4 counts the weight thresholds of edges identified by the head/tail breaks rule on each day of the study. We find that the largest remaining edge threshold is 5 during 10:00–14:00 (when people mainly commute between business and recreation centers) and 20:00–22:00 (the peak time for people to return home from entertainment venues or workplaces). On the other hand, the weight threshold is mainly 2 at 2:00–6:00. The possible reason for this relatively lower threshold is that there are fewer taxi trips in the city during this time period. As a result, the interaction of most OD flows is 1, and its weight ratio exceeds 0.4. Except for the above two situations, the threshold in other time periods is mainly 3, which is relatively stable (see more details in Appendix B).

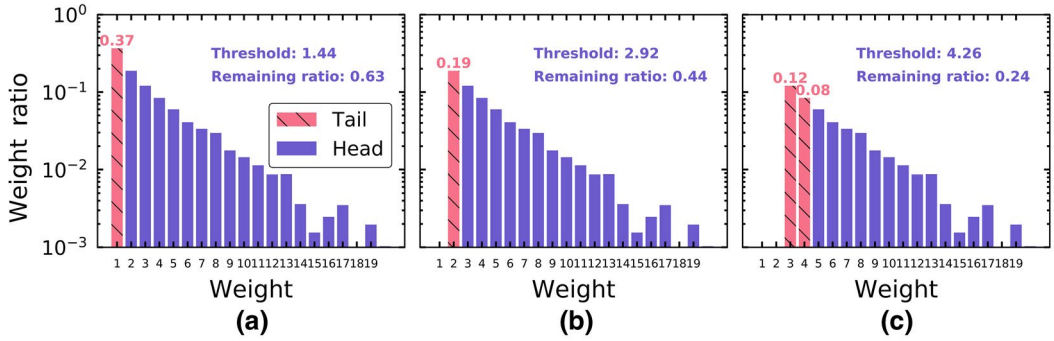

FIGURE 3 The head/tail partition result of the OD network for 12:00–14:00 on March 5. (a) Weight ratio for all weighted edges, and its first partition result. (b) Weight ratio for the remaining weighted edges after the first partition (with weight greater than 1.4), and its second partition result. (c) Weight ratio or the remaining weighted edges after the first partition (with weight greater than 2.92), and its third partition result

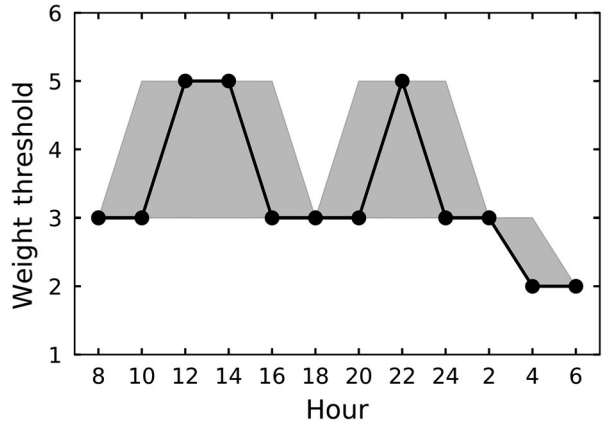

FIGURE 4 The weight threshold determined by head/tail partitioning in the different time periods. The solid line indicates the median value, and the shaded area corresponds to the 5%–95% quantile range across the 173 days. The value H on the x-axis indicates the time period between $H - 2$ and H

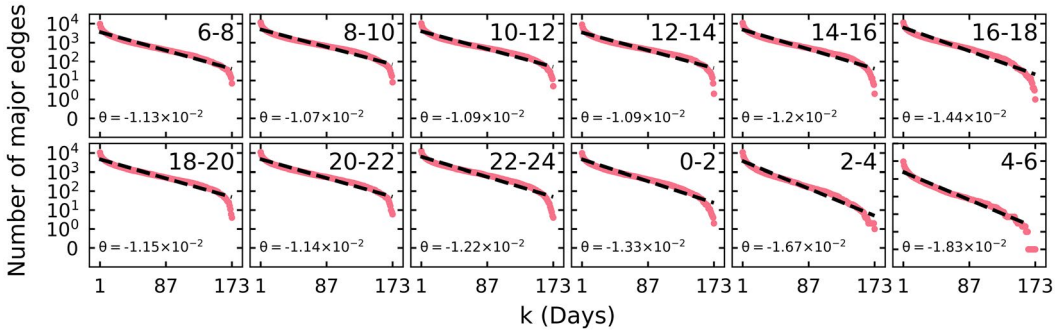

FIGURE 5 The number of major edges that appear for at least k days. The value k on the x-axis indicates the number of days is k. The value $y(k)$ in the y-axis indicates the number of major edges appearing on at least k days over the study period. The dashed line is fitted by $y(k) \propto 10^{\theta k}$. Graphs are plotted on the linear-log scale

After constructing the dynamic major OD sub-network, the parameter β (in Definition 3) is required to be pre-defined for identifying the recurrent OD network. It determines the minimum frequency at which a recurrent major edge should appear. Here, the major edges refer to the edges in the dynamic major OD network. We count the number of major edges (y) that appear for at least k days over the study period, which shows that y decays with k as $y(k) \propto 10^{\theta k}$ ($\theta < 0$) (Figure 5). This exponential decay indicates that most of the major edges do not appear frequently, but are accidental on some days. Moreover, these curves decay at different speeds over different time-slots. This shows that the frequency of occurrence of a major edge is not consistent over time. In this study, we uniformly set the parameter $\beta = 0.5$ by experience, indicating that a major edge appeared in more than half of the days in the study period, and then we consider it as being recurrent.

Figure 6 shows the results of the major OD network $\overline{M}$ and the recurrent OD network $\overline{R}$ in different time-slots. It can be seen that the major OD network forms dense interconnections between most regions of Wuhan, showing that people's periodic and non-periodic movements are widespread among Wuhan regions. However, the recurrent OD network only involves a small number of regions, and the physical length between OD flows is relatively shorter (Appendix A), which results in some regions being separated and disconnected from each other. This indicates that the periodic movements in the Wuhan area mainly exist between a small number of regions close in space. In addition, we have observed major and recurrent OD network changes at different times. In particular, during the during 2:00–6:00 when most people are resting at home, the number of recurrent OD flows is significantly reduced compared to other times. Therefore, periodic movement as a part of the hybrid movement (combination of periodic and non-periodic) shows some unique characteristics.

4.3 | Characterizing the recurrent OD network

In this section we compare the structure of OD interactions by identifying the community structure between the major OD network and the recurrent OD network. The evolution of flux within a day is then analyzed.

4.3.1 | Community structure

Community structure deconstructs the OD network by identifying the community which represents the densely connected regions. Previous studies have found that the community structure of an OD network in a city can be represented as multiple spatially connected blocks (Rinzivillo et al., 2012), which is largely due to the city's

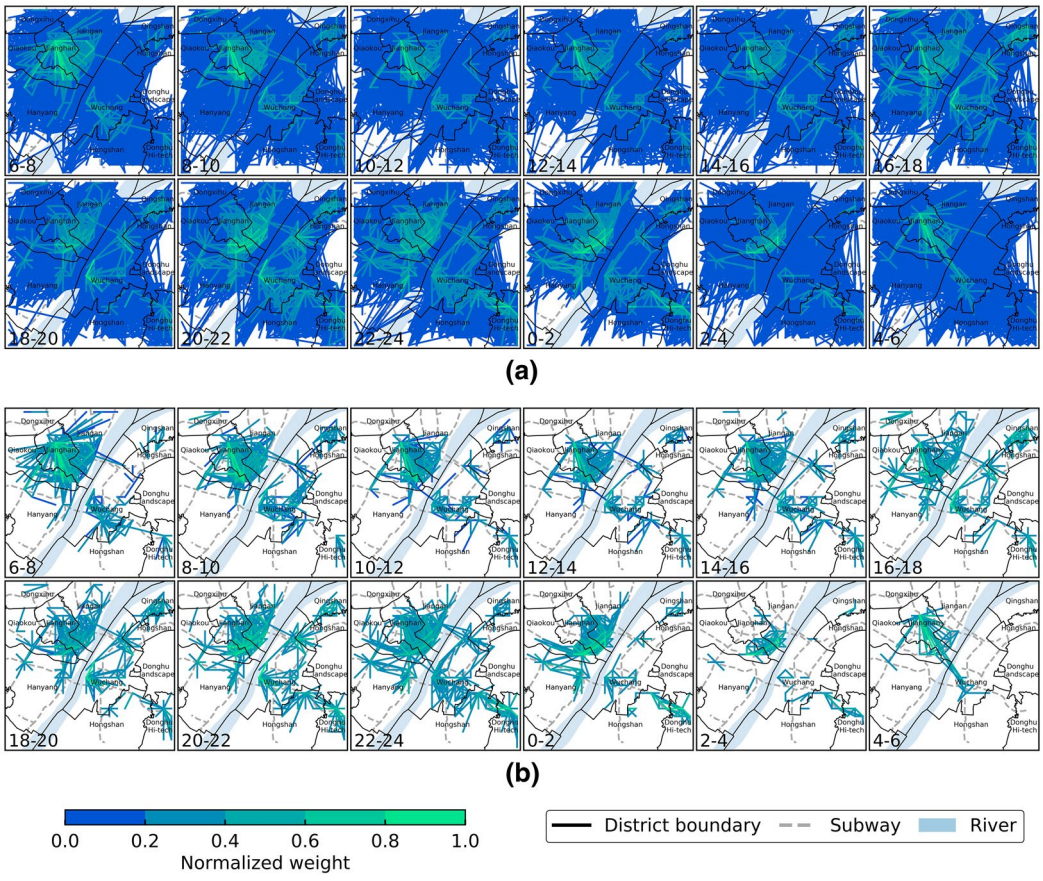

FIGURE 6 Spatial distributions of (a) major OD network $\overline{M}$ and (b) recurrent OD network $\overline{R}$. For visualization, the weight (w) of each edge is normalized by $w_{\text{nor}} = (\max_{w_i \in W_{\overline{M}}} w_i - w)/(\max_{w_i \in W_{\overline{M}}} w_i - \min_{w_i \in W_{\overline{M}}} w_i)$ within each time period

polycentral structure pattern and the tendency of people to move short distances. Furthermore, the observed spatially connected blocks sometimes showed a certain degree of expansion across official district boundaries (Zhou, Yue, Li, & Wang, 2016). While it has been observed that the recurrent OD network is a more spatially localized network at the base of the major OD network, it is natural to ask what is the relationship between the community structure of the two.

The Infomap approach is applied to identify the community of the network. Figure 7 displays the spatial distribution of the community in the major OD network and the recurrent OD network. It finds that the community structure of the major OD network is a large-scale spatial connection block restricted by the river (Figure 7a). The communities are distributed on both sides of the river, due to the fact that the river creates a natural barrier to people's movement, and people can only move between the east and west banks via a small number of bridges across the river. This makes the regional OD interactions between the two banks less intense than within each. Within each bank, the spatial connection block is formed. On the one hand, it may be due to the polycentric spatial structure of urban spatial planning (Gordon, Richardson, & Wong, 1986; McMillen & McDonald, 1997). There are many popular areas in the city, surrounding their respective regional centers, forming a large space including large shopping malls, entertainment venues, office buildings, and residential buildings. On the other hand, people tend to travel as close as possible in the city, which leads to local interaction clusters between nearby regions (Barrat, Barthelemy, & Vespignani, 2005). These two reasons make the communities with densely interactive regions in

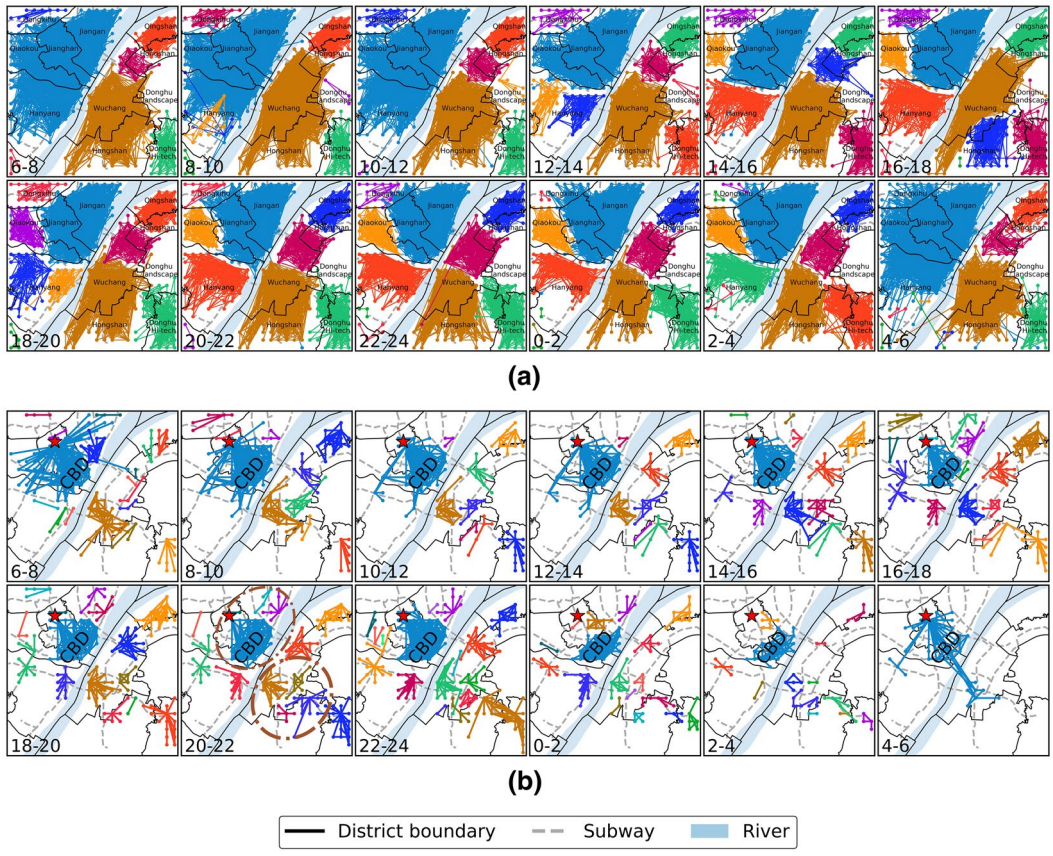

FIGURE 7 Spatial distributions of the community structure of: (a) the major OD network $\overline{M}$; and (b) the recurrent OD network $\overline{R}$. The red star indicates the largest railway station, Hankou. The brown circles in (b) 20:00–22:00 show some small communities of the $\overline{R}$ which belong to the same big community in $\overline{M}$ in the corresponding time period

the city appear as spatially connected blocks. However, these communities are relatively large-scale in space, and some communities cover or span several districts during most time periods of the day, such as the communities composed of OD interactions in Qiaokou, Jianghan, and Jiangan districts (above the west bank of the river), and the communities cross Wuchang and Hongshan districts (below the east bank of the river). This phenomenon of cross-district expansion of the OD network community is consistent with that reported previously (Zhou et al., 2016). This may be due to the convenient transportation that allows people to move freely in a larger area, thus forming a densely interactive space block between districts.

On the other hand, we observe that the community of the recurrent OD network is more spatially localized at the base of the community of the major OD network $\overline{M}$ (Figure 7b). A community in the major OD network $\overline{M}$ may decompose into multiple smaller communities (e.g., those circled at 20:00–22:00) or shrink into a spatially localized community (e.g., the community around Qingshan district above the east bank of the river at 20:00–22:00) of the recurrent OD network $\overline{R}$. Figure 7b shows Hankou, the largest railway station, and the CBD area in Wuhan, which contains a large number of offices, entertainment, and high-end residences. We find that apart from the larger cross-district communities formed by the railway station and the CBD, other communities are affected by districts to a certain extent. Most of the connections in the community occur within one district, or a small amount of recurrent OD flows across two districts. Since urban spatial configurations are usually arranged by district, the spatial settings of the city may have a relatively strong influence on the periodic OD flows. Citizens may tend to

periodically move between typical regions of interest within the districts (see Appendix C in the online Supporting Information). Specifically, at 4:00–6:00, there is only one community across the river, mainly connecting multiple railway stations and the CBD (see Appendix C). This may be due to the fact that there are fewer OD interactions during this time period.

To summarize, for most cases, we observe that the community structure of the major OD network is larger in space. Although it is restricted by the river, many cross-district connections are still formed in the communities on both sides of the respective rivers. On the other hand, the communities of the recurrent OD network are spatially shrunken or split from the communities of the major OD network, and are relatively strongly affected by the scheme of districts with relatively fewer OD interactions across multiple districts. This community structure relationship between the two OD networks illustrates that the dense connectivity pattern of regions within a community can be further represented as one or more spatially localized periodic OD flows, supplemented by many spatially global non-periodic connections. Thus, the periodic and non-periodic OD movements are the local and global characterizations of the interaction structure of the urban area.

4.3.2 | Pattern of regional flux

In this section to deepen understanding of the periodic and non-periodic OD movements' patterns, the regional flux (see Section 3.2.2) is further compared between the major OD network $\overline{M}$ and the recurrent OD network $\overline{R}$. While the evolution of the relationship between influx and outflux within a day is strongly related to citizens' activities (Zhou et al., 2020), is the pattern of the flux formed by the recurrent OD network $\overline{R}$ the same as the flux formed by the major OD network $\overline{M}$?

To answer this question, the net-flux ratio is introduced to comprehensively measure the flow type and dominated degree of a region. Figure 8 displays the spatial distributions of $NFR_{\overline{M}}(v_i)$ and $NFR_{\overline{R}}(v_i)$, for $v_i \in V_{\overline{R}} \subset V_{\overline{M}}$. Red (blue) color indicates that the inflow (outflow) dominates the regions, and the darker the color the higher the degree of dominance. While the flux is dynamic throughout the day, we surprisingly find that the regional flux of the recurrent OD network $\overline{R}$ has colors similar to but darker than that of the major OD network $\overline{M}$ in the corresponding time period (see Appendix D in the online Supporting Information). This implies that regions $v_i \in V_{\overline{R}}$ are dominated by similar types of recurrent OD flows and major OD flows (containing both recurrent and non-recurrent OD flows), but also the recurrent OD flows amplify the dominated degrees.

Under this amplification process of the recurrent OD network $\overline{R}$, the dynamic travel patterns of people are highlighted. This is illustrated by checking the ratios of influx, balance, and outflux regions (defined in Section 3.2.2) according to the NFR throughout the day. If we consider the flux of the major OD network $\mathrm{NFR}_{\overline{M}}(v_i)$, the balance region dominates nearly all the time-slots (Figure 9a). This is because $\mathrm{NFR}_{\overline{M}}(v_i)$ is too close to 0 to be classified into the influx or outflux region. However, after amplification in the recurrent OD network $\overline{R}$, the three kinds of ratios dominate in turn (Figure 9b), which is consistent with the travel activities of citizens. There are typically three periods.

1. 4:00–10:00. The ratio of the outflow region is relatively high. People mainly go to work, starting the day's economic activity in the city. They tend to go to work via bus or train stations, which implies that people live in relatively dispersed areas but places for working activities are more concentrated within Wuhan.

2. 10:00–20:00. The ratio of the balance region is the highest. This period reflects the relatively stable and balanced inflows and outflows of people in various regions, when most people are already located where they will need to be for the next 8–10 h. There is no obvious tendency for people to go or leave for each region.

3. 20:00–4:00. The ratio of the inflow region gradually increases and becomes the highest. This is because people mainly go home during this period.

Actually, the amplification phenomenon of the $\overline{R}$ net-flux ratio is mainly due to the relatively different intensity of the dominant flow exhibited by each of the recurrent and non-recurrent flows themselves. More specifically, for

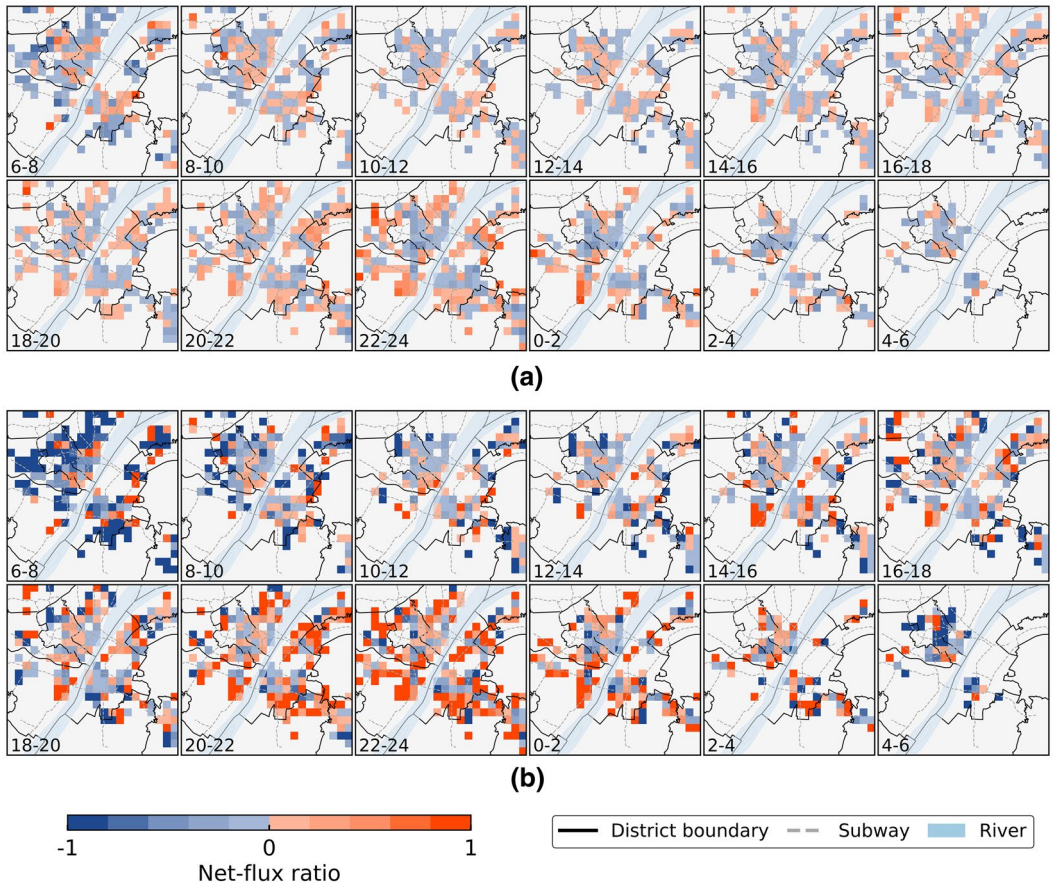

FIGURE 8 Spatial distributions of net-flux ratio for nodes of: (a) the major OD network $\overline{M}$; and (b) the recurrent OD network $\overline{R}$. We only present the nodes $v_i \in V_{\overline{R}} \subset V_{\overline{M}}$ for comparison

a region v_i dominated by inflows, the amplification phenomenon can be deduced as (see Appendix E in the online Supporting Information):

$$\frac{F_{in}^{\overline{R}}(v_i)}{F_{out}^{\overline{R}}(v_i)} > \frac{F_{in}^{\overline{NR}}(v_i)}{F_{out}^{\overline{NR}}(v_i)} \tag{14}$$

Therefore, the difference between in- and out- flux of the recurrent OD flows deserves more attention than that of non-recurrent OD flows. There are both recurrent and non-recurrent OD flows in the major OD network $\overline{M}$, the recurrent dominant inflows are largely diluted by the non-recurrent OD flows. The situation is similar for a region v_i dominated by outflows, and then we have:

$$\frac{F_{out}^{\overline{R}}(v_i)}{F_{in}^{\overline{R}}(v_i)} > \frac{F_{out}^{\overline{NR}}(v_i)}{F_{in}^{\overline{NR}}(v_i)} \tag{15}$$

To summarize, we use the net-flux ratio to measure the regional dominated flow type and degree. We observe the recurrent OD network having a similar dominated flow type but a stronger degree than the major OD network. Furthermore, by checking the ratios of the three types of regions defined by NFR, we found that daily dynamics

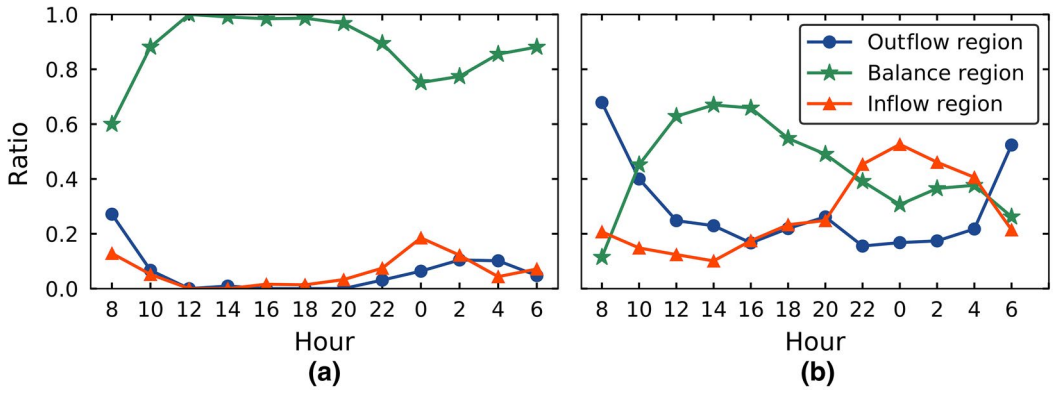

FIGURE 9 Ratios of outflow, balance, and inflow region of $v_i \in \overline{R} \subset \overline{M}$ in (a) $\overline{M}$ and (b) $\overline{R}$. The value H on the x-axis indicates the time period between $H-2$ to H

described by the major OD network $\overline{M}$ are apparently more unordered than those by the corresponding recurrent OD network $\overline{R}$. The dynamic changes in the ratios of $\overline{R}$ are more consistent with people's daily travel regularities in common sense. Furthermore, the dynamic travel patterns are highlighted with this amplification process. Therefore, the results illustrate that periodic OD flows can better characterize whether an area is dominated by inflow or outflow without being diluted by non-periodic OD flows. This periodic OD flow with more prominent regional in-/out- flux characteristics can more clearly describe the attractiveness of a region at different time periods and better capture the temporally stable urban dynamic pattern from the whole.

5 | DISCUSSION

Our study proposes a method of identifying the recurrent OD network to capture the periodic OD flows between regions within a city. Compared with the major OD network containing both recurrent and non-recurrent OD networks, it finds that the recurrent OD network has a more spatially localized community structure and amplifies the regional dominated degree of in or out flows. These results show several potential contributions in practical applications related to spatial OD interactions.

One benefit is that they can provide city managers with better guidance on planning administrative districts. OD network community structure describes a space block with close interaction between regions objectively, which has been applied to help the design of the district boundary by many researchers (e.g., Rinzivillo et al., 2012; Zhou et al., 2016). In our study, the spatial localized community structure of the recurrent OD network describes those time-stable interactive structures. City managers can consider their spatial distribution to design district boundaries, thereby capturing the core structure from the perspective of time to better manage people's daily travels.

Second, the results may be conductive to developing some more targeted regional economic development strategies. The daily stable population flow between urban regions is one of the important driving factors to ensure the vitality of a regional economy (He, Zhang, & Xiu, 2019). The recurrent OD flows precisely describe this kind of strong interaction phenomenon that frequently occurs in daily life, and can provide some guidance for planning the economic development in the city. First, for those regions that form a community of the recurrent OD network, more roads and more rationally designed traffic lights can be built between the regions to ensure smooth movements. It can also support more urban facilities around their surrounding regions to drive the development of the surrounding economy. Moreover, the prominent characteristics of regional inflows or outflows can better define the attractiveness of a region at the different time periods of the day, so as to understand the time-stable

role of a region in the regional economic interaction and make corresponding economic development strategy in different time periods of a day.

Third, they help to develop more effective infectious disease control strategies. The movements of the citizens are one of the important reasons for the spread of epidemics (Peixoto, Marcondes, Peixoto, & Oliva, 2020). The identified recurrent OD flows represent those day-to-day repeated interactions between different regions. In addition, recurrent OD flows show more prominent inflow and outflow characteristics in a region, which can better identify a region that is likely to spread epidemics to other regions or be infected. Therefore, in the early stages of the epidemic, when people's movements are still the same as before, focus on monitoring and interventions for movement between these regions may better interrupt the rapid spread of epidemics.

6 | CONCLUSIONS AND FUTURE WORK

To capture the daily periodic OD flows, our study proposed the extraction of recurrent OD networks based on the idea of frequent sub-networks. The algorithm first constructed the dynamic OD network by organizing the individual movements as an OD network on each day, then edges with large weights were retained by the head/tail breaks method. Finally, the edges appearing on most days were identified to construct the recurrent OD network. To highlight the patterns of the recurrent OD network, the major OD network containing both recurrent and non-recurrent OD flows was compared at different times of the day.

We compared the community structure and regional flux to characterize the network from both mesoscopic and microscopic pespectives. Firstly, we adopted the Infomap approach to analyze the community structure. Results confirmed that, in comparison with a major OD network $\overline{M}$, the community structure of the corresponding recurrent OD network $\overline{R}$ better illustrated both the periodic and non-periodic OD flows. Then the NFR was introduced to measure the dominant flow and degree of a region. By comparing the NFR heatmap, it was observed that the recurrent OD network was of a color similar to but darker than the major OD network. This showed that, the periodic OD flows of a region had similar but more significant in- and out- flow characteristics. Under this amplification process, the dynamic travel patterns of citizens measured by the net-flux ratio were clearer within the day. Finally, the potential benefits in terms of urban planning, regional economic development, and infectious disease control were discussed.

However, there remain some limitations and work to be done in the future. Firstly, the factors influencing the recurrent OD network need to be further studied. Through preliminary network map matching with semantic information on points of interest, we found that the basic setting of the city was a possible factor that affects the formation of recurrent OD flows between the two regions. A more comprehensive socioeconomic and human demographic information can be considered to quantitatively explore specific influencing factors.

Secondly, the analysis of spatio-temporal patterns of the community structure of the recurrent OD network is an important future direction. Future work can combine the degree distribution, centrality, connectivity, clustering and other indicators of the network to further characterize the features of the community and examine its daily evolution. Moreover, regional socioeconomic and human demographic information can be incorporated to deepen the understanding of spatial patterns of evolution.

Thirdly, an interesting topic is to compare the recurrent OD flows to the subway network to determine whether taxis serve as a complement or as competition to the subway system in Wuhan. These day-to-day repeated movements describe the stable demand for public transportation movement between regions. How many of the taxi trips are likely to start or end at a subway stop? How many taxi trips could have been replaced by passengers taking the subway instead? Future research can answer these questions and reveal the role of taxis and subways in the stable demand of cities by comparing taxi and subway data.

Finally, the influence of different thresholds on recurrent OD flows needs further analysis and discussion. The weight threshold and the frequency threshold determine the extracted recurrent and non-recurrent OD flows.

Comparisons and discussions between different thresholds are important for further understanding the periodic OD movements.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (NSFC: 42071368, 41725006, 41871287, U1833201).

CONFLICT OF INTEREST

None.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

Xiaojian Chen https://orcid.org/0000-0002-3945-1621
Binbin Lu https://orcid.org/0000-0001-7847-7560

REFERENCES

Alessandretti, L., Aslak, U., & Lehmann, S. (2020). The scales of human mobility. *Nature*, *587*(7834), 402–407. https://doi.org/10.1038/s41586-020-2909-1

Allesina, S., Bodini, A., & Bondavalli, C. (2006). Secondary extinctions in ecological networks: Bottlenecks unveiled. *Ecological Modelling*, *194*(1–3), 150–161. https://doi.org/10.1016/j.ecolmodel.2005.10.016

Andrienko, G., Andrienko, N., Fuchs, G., & Wood, J. (2017). Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE Transactions on Visualization and Computer Graphics*, *23*(9), 2120–2136. https://doi.org/10.1109/Tvcg.2016.2616404

Barrat, A., Barthelemy, M., & Vespignani, A. (2005). The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2005*, P05003. https://doi.org/10.1088/1742-5468/2005/05/P05003

Barroso, J. M. F., Albuquerque-Oliveira, J. L., & Oliveira-Neto, F. M. (2020). Correlation analysis of day-to-day origin-destination flows and traffic volumes in urban networks. *Journal of Transport Geography*, *89*, 102899. https://doi.org/10.1016/j.jtrangeo.2020.102899

Barthelemy, M. (2011). Spatial networks. *Physics Reports: Review Section of Physics Letters*, *499*(1–3), 1–101. https://doi.org/10.1016/j.physrep.2010.11.002

Batabyal, S., & Bhaumik, P. (2015). Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey. *IEEE Communications Surveys and Tutorials*, *17*(3), 1679–1707. https://doi.org/10.1109/Comst.2015.2419819

Bellingeri, M., & Bodini, A. (2013). Threshold extinction in food webs. *Theoretical Ecology*, *6*(2), 143–152. https://doi.org/10.1007/s12080-012-0166-0

Bimpou, K., & Ferguson, N. S. (2020). Dynamic accessibility: Incorporating day-to-day travel time reliability into accessibility measurement. *Journal of Transport Geography*, *89*, 102892. https://doi.org/10.1016/j.jtrangeo.2020.102892

Bronzini, M., Herendeen, J., Jr., Miller, J., & Womer, N. (1974). A transportation-sensitive model of a regional economy. *Transportation Research*, *8*(1), 45–62. https://doi.org/10.1016/0041-1647(74)90017-3

Chen, C., Ma, J. T., Susilo, Y., Liu, Y., & Wang, M. L. (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

Cho, E., Myers, S. A., & Leskovec, J. (2011). Friendship and mobility: User movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA (pp. 1082–1090). New York, NY: ACM.

Chowell, G., Hyman, J. M., Eubank, S., & Castillo-Chavez, C. (2003). Scaling laws for the movement of people between locations in a large city. *Physical Review E*, *68*, 066102. https://doi.org/10.1103/PhysRevE.68.066102

Duan, Z. T., Zhang, K., Chen, Z., Liu, Z. Y., Tang, L., Yang, Y., & Ni, Y. Y. (2019). Prediction of city-scale dynamic taxi origin-destination flows using a hybrid deep neural network combined with travel time. *IEEE Access*, *7*, 127816–127832. https://doi.org/10.1109/Access.2019.2939902

Duque, J. C., Anselin, L., & Rey, S. J. (2012). The max-*p*-regions problem. *Journal of Regional Science*, *52*(3), 397–419. https://doi.org/10.1111/j.1467-9787.2011.00743.x

Ekman, F., Keränen, A., Karvo, J., & Ott, J. (2008). Working day movement model. In *Proceedings of the First ACM SIGMOBILE Workshop on Mobility Models*, Hong Kong (pp. 33–40). New York, NY: ACM.

Ferreira, N., Poco, J., Vo, H. T., Freire, J., & Silva, C. T. (2013). Visual exploration of big spatio-temporal urban data: A study of New York City taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2149–2158. https://doi.org/10.1109/TVCG.2013.226

Gordon, P., Richardson, H. W., & Wong, H. L. (1986). The distribution of population and employment in a polycentric city: The case of Los Angeles. *Environment and Planning A*, *18*(2), 161–173. https://doi.org/10.1068/a180161

Guidotti, R., Monreale, A., Rinzivillo, S., Pedreschi, D., & Giannotti, F. (2016). Unveiling mobility complexity through complex network analysis. *Social Network Analysis and Mining*, *6*(1), 59. https://doi.org/10.1007/s13278-016-0369-2

Guo, D. (2008). Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographical Information Science*, *22*(7), 801–823. https://doi.org/10.1080/13658810701674970

Hamedmoghadam, H., Ramezani, M., & Saberi, M. (2019). Revealing latent characteristics of mobility networks with coarse-graining. *Scientific Reports*, *9*, 7545. https://doi.org/10.1038/s41598-019-44005-9

Hasan, S., Schneider, C. M., Ukkusuri, S. V., & Gonzalez, M. C. (2013). Spatiotemporal patterns of urban human mobility. *Journal of Statistical Physics*, *151*(1–2), 304–318. https://doi.org/10.1007/s10955-012-0645-0

He, H., Zhang, J., & Xiu, D. (2019). China's migrant population and health. *China Population and Development Studies*, *3*(1), 53–66. https://doi.org/10.1007/s42379-019-00032-7

Holme, P., & Saramaki, J. (2012). Temporal networks. *Physics Reports: Review Section of Physics Letters*, *519*(3), 97–125. https://doi.org/10.1016/j.physrep.2012.03.001

Hossmann, T., Spyropoulos, T., & Legendre, F. (2011). A complex network analysis of human mobility. In *Proceedings of the 2011 IEEE Conference on Computer Communications Workshops*, Shanghai, China (pp. 876–881). Piscataway, NJ: IEEE.

Hu, J., Yang, B., Guo, C., Jensen, C. S., & Xiong, H. (2020). Stochastic origin-destination matrix forecasting using dual-stage graph convolutional, recurrent neural networks. In *Proceedings of the 36th IEEE International Conference on Data Engineering*, Dallas, TX (pp. 1417–1428). Piscataway, NJ: IEEE.

Huang, H. S., Cheng, Y., & Weibel, R. (2019). Transport mode detection based on mobile phone network data: A systematic review. *Transportation Research Part C: Emerging Technologies*, *101*, 297–312. https://doi.org/10.1016/j.trc.2019.02.008

Jia, T., Jiang, B., Carling, K., Bolin, M., & Ban, Y. F. (2012). An empirical study on human mobility and its agent-based modeling. *Journal of Statistical Mechanics: Theory and Experiment*, *2012*, P11024. https://doi.org/10.1088/1742-5468/2012/11/P11024

Jiang, B. (2013). Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *The Professional Geographer*, *65*(3), 482–494. https://doi.org/10.1080/00330124.2012.700499

Jiang, B., Yin, J. J., & Zhao, S. J. (2009). Characterizing the human mobility pattern in a large street network. *Physical Review E*, *80*, 021136. https://doi.org/10.1103/PhysRevE.80.021136

Karamshuk, D., Boldrini, C., Conti, M., & Passarella, A. (2011). Human mobility models for opportunistic networks. *IEEE Communications Magazine*, *49*(12), 157–165. https://doi.org/10.1109/Mcom.2011.6094021

Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., & Ye, J. (2021). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network. *Transportation Research Part C: Emerging Technologies*, *122*, 102858. https://doi.org/10.1016/j.trc.2020.102858

Kim, S. H., & Chung, J. H. (2018). Exploration on origin-destination-based travel time variability: Insights from Seoul metropolitan area. *Journal of Transport Geography*, *70*, 104–113. https://doi.org/10.1016/j.jtrangeo.2018.05.021

Li, D. W., Cao, J. M., Li, R. Y., & Wu, L. F. (2020). A spatio-temporal structured LSTM model for short-term prediction of origin-destination matrix in rail transit with multisource data. *IEEE Access*, *8*, 84000–84019. https://doi.org/10.1109/Access.2020.2991982

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., … Cheng, T. (2016). Geospatial big data handling theory and methods: A review and research challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 119–133. https://doi.org/10.1016/j.isprsjprs.2015.10.012

Liu, L. B., Qiu, Z. L., Li, G. B., Wang, Q., Ouyang, W. L., & Lin, L. (2019). Contextualized spatial-temporal network for taxi origin-destination demand prediction. *IEEE Transactions on Intelligent Transportation Systems*, *20*(10), 3875–3887. https://doi.org/10.1109/Tits.2019.2915525

Liu, X., Kang, C. G., Gong, L., & Liu, Y. (2016). Incorporating spatial interaction patterns in classifying and understanding urban land use. *International Journal of Geographical Information Science*, *30*(2), 334–350. https://doi.org/10.1080/13658816.2015.1086923

Liu, Y., Kang, C. G., Gao, S., Xiao, Y., & Tian, Y. (2012). Understanding intra-urban trip patterns from taxi trajectory data. *Journal of Geographical Systems*, *14*(4), 463–483. https://doi.org/10.1007/s10109-012-0166-z

Louail, T., Lenormand, M., Cantu Ros, O. G., Picornell, M., Herranz, R., Frias-Martinez, E., … Barthelemy, M. (2014). From mobile phone data to the spatial structure of cities. *Scientific Reports*, *4*, 5276. https://doi.org/10.1038/srep05276

Louail, T., Lenormand, M., Picornell, M., García Cantú, O., Herranz, R., Frias-Martinez, E., … Barthelemy, M. (2015). Uncovering the spatial structure of mobility networks. *Nature Communications*, *6*, 6007. https://doi.org/10.1038/ncomms7007

Luo, F. X., Cao, G. F., 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

Lynall, M. E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., & Bullmore, E. T. (2010). Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, *30*(28), 9477–9487. https://doi.org/10.1523/Jneurosci.0333-10.2010

Ma, D., Osaragi, T., Oki, T., & Jiang, B. (2020). Exploring the heterogeneity of human urban movements using geo-tagged tweets. *International Journal of Geographical Information Science*, *34*(12), 2475–2496. https://doi.org/10.1080/13658816.2020.1718153

Ma, D., Sandberg, M., & Jiang, B. (2015). Characterizing the heterogeneity of the OpenStreetMap data and community. *ISPRS International Journal of Geo-Information*, *4*(2), 535–550. https://doi.org/10.3390/ijgi4020535

Magnanti, T. L., & Mirchandani, P. (1993). Shortest paths, single origin-destination network design, and associated polyhedra. *Networks*, *23*(2), 103–121. https://doi.org/10.1002/net.3230230205

McMillen, D. P., & McDonald, J. F. (1997). A nonparametric analysis of employment density in a polycentric city. *Journal of Regional Science*, *37*(4), 591–612. https://doi.org/10.1111/0022-4146.00071

Nanni, M., Tortosa, L., Vicent, J. F., & Yeghikyan, G. (2020). Ranking places in attributed temporal urban mobility networks. *PLoS ONE*, *15*(10), e0239319. https://doi.org/10.1371/journal.pone.0239319

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323–351. https://doi.org/10.1080/00107510500052444

Noulas, A., Shaw, B., Lambiotte, R., & Mascolo, C. (2015). Topological properties and temporal dynamics of place networks in urban environments. In *Proceedings of the 24th International Conference on World Wide Web*, Florence, Italy (pp. 431–441). New York, NY: ACM. https://doi.org/10.1145/2740908.2745402

Osorio-Arjona, J., & García-Palomares, J. C. (2019). Social media and urban mobility: Using Twitter to calculate home-work travel matrices. *Cities*, *89*, 268–280. https://doi.org/10.1016/j.cities.2019.03.006

Pan, G., Qi, G. D., Wu, Z. H., Zhang, D. Q., & Li, S. J. (2013). Land-use classification using taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems*, *14*(1), 113–123. https://doi.org/10.1109/Tits.2012.2209201

Pappalardo, L., Simini, F., Rinzivillo, S., Pedreschi, D., Giannotti, F., & Barabási, A. L. (2015). Returners and explorers dichotomy in human mobility. *Nature Communications*, *6*, 8166. https://doi.org/10.1038/ncomms9166

Peixoto, P. S., Marcondes, D., Peixoto, C., & Oliva, S. M. (2020). Modeling future spread of infections via mobile geolocation data and population dynamics. An application to COVID-19 in Brazil. *PLoS ONE*, *15*(7), e0235732. https://doi.org/10.1371/journal.pone.0235732

Peng, C. B., Jin, X. G., Wong, K. C., Shi, M. X., & Lio, P. (2012). Collective human mobility pattern from taxi trips in urban area. *PLoS ONE*, *7*(4), e34487. https://doi.org/10.1371/journal.pone.0034487

Qi, G., Ceder, A., Huang, A., & Guan, W. (2021). A methodology to attain public transit origin-destination mobility patterns using multi-layered mesoscopic analysis. *IEEE Transactions on Intelligent Transportation Systems*, 1–19. Online (Early Access). https://doi.org/10.1109/TITS.2020.2990719

Qian, Z., Liu, X. T., Tao, F., & Zhou, T. (2020). Identification of urban functional areas by coupling satellite images and taxi GPS trajectories. *Remote Sensing*, *12*(15), 2449. https://doi.org/10.3390/rs12152449

Rinzivillo, S., Mainardi, S., Pezzoni, F., Coscia, M., Pedreschi, D., & Giannotti, F. (2012). Discovering the geographical borders of human mobility. *KI - Künstliche Intelligenz*, *26*(3), 253–260. https://doi.org/10.1007/s13218-012-0181-8

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105

Roy, J. R., & Thill, J. C. (2004). Spatial interaction modelling. *Papers in Regional Science*, *83*(1), 339–361. https://doi.org/10.1007/s10110-003-0189-4

Saberi, M., Ghamami, M., Gu, Y., Shojaei, M. H., & Fishman, E. (2018). Understanding the impacts of a public transit disruption on bicycle sharing mobility patterns: A case of Tube strike in London. *Journal of Transport Geography*, *66*, 154–166. https://doi.org/10.1016/j.jtrangeo.2017.11.018

Saberi, M., Mahmassani, H. S., Brockmann, D., & Hosseini, A. (2017). A complex network perspective for characterizing urban travel demand patterns: Graph theoretical analysis of large-scale origin destination demand networks. *Transportation*, *44*(6), 1383–1402. https://doi.org/10.1007/s11116-016-9706-6

Sapiezynski, P., Stopczynski, A., Gatej, R., & Lehmann, S. (2015). Tracking human mobility using WiFi signals. *PLoS ONE*, *10*(7), e0130824. https://doi.org/10.1371/journal.pone.0130824

Sawai, H. (2012). Reorganizing a new generation airline network based on an ant-colony optimization inspired small-world network. In *Proceedings of the 2012 IEEE Congress on Evolutionary Computation*, Brisbane, QLD, Australia (pp. 1–8). Piscataway, NJ: IEEE.

Serrano, M. A., Boguna, M., & Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(16), 6483–6488. https://doi.org/10.1073/pnas.0808904106

Shao, F. J., Sui, Y., Yu, X., & Sun, R. C. (2019). Spatio-temporal travel patterns of elderly people: A comparative study based on buses usage in Qingdao, China. *Journal of Transport Geography*, *76*, 178–190. https://doi.org/10.1016/j.jtrangeo.2019.04.001

Shui, C. S., & Szeto, W. Y. (2020). A review of bicycle-sharing service planning problems. *Transportation Research Part C: Emerging Technologies*, *117*, 102648. https://doi.org/10.1016/j.trc.2020.102648

Simini, F., Gonzalez, M. C., Maritan, A., & Barabási, A. L. (2012). A universal model for mobility and migration patterns. *Nature*, *484*(7392), 96–100. https://doi.org/10.1038/nature10856

Song, C. M., Qu, Z. H., Blumm, N., & Barabási, A. L. (2010). Limits of predictability in human mobility. *Science*, *327*(5968), 1018–1021. https://doi.org/10.1126/science.1177170

Tang, J. J., Zhang, S., Chen, X. Q., Liu, F., & Zou, Y. J. (2018). Taxi trips distribution modeling based on entropy-maximizing theory: A case study in Harbin city—China. *Physica A: Statistical Mechanics and its Applications*, *493*, 430–443. https://doi.org/10.1016/j.physa.2017.11.114

Tang, J. J., Zhang, S., Zhang, W. H., Liu, F., Zhang, W. B., & Wang, Y. H. (2016). Statistical properties of urban mobility from location-based travel networks. *Physica A: Statistical Mechanics and its Applications*, *461*, 694–707. https://doi.org/10.1016/j.physa.2016.06.031

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*(Suppl. 1), 234–240. https://doi.org/10.2307/143141

Toqué, F., Côme, E., El Mahrsi, M. K., & Oukhellou, L. (2016). Forecasting dynamic public transport origin-destination matrices with long-short term memory recurrent neural networks. *In Proceedings of the 19th IEEE International Conference on Intelligent Transportation Systems*, Rio de Janeiro, Brazil (pp. 1071–1076). Piscataway, NJ: IEEE.

Wang, P. Y., Fu, Y. J., Zhang, J. W., Li, X. L., & Li, D. (2018). Learning urban community structures: A collective embedding perspective with periodic spatial-temporal mobility graphs. *ACM Transactions on Intelligent Systems and Technology*, *9*(6), 63. https://doi.org/10.1145/3209686

Xie, C., Kockelman, K. M., & Waller, S. T. (2011). A maximum entropy-least squares estimator for elastic origin-destination trip matrix estimation. *Transportation Research Part B: Methodological*, *45*(9), 1465–1482. https://doi.org/10.1016/j.trb.2011.05.018

Yan, X. R., Jeub, L. G. S., Flammini, A., Radicchi, F., & Fortunato, S. (2018). Weight thresholding on complex networks. *Physical Review E*, *98*(4), 042304. https://doi.org/10.1103/PhysRevE.98.042304

Yang, C., Xiao, M., Ding, X., Tian, W., Zhai, Y., Chen, J., … Ye, X. (2019). Exploring human mobility patterns using geo-tagged social media data at the group level. *Journal of Spatial Science*, *64*(2), 221–238. https://doi.org/10.1080/14498596.2017.1421487

Yang, J., Sun, Y. Z., Shang, B. W., Wang, L., & Zhu, J. (2019). Understanding collective human mobility spatiotemporal patterns on weekdays from taxi origin-destination point data. *Sensors*, *19*(12), 2812. https://doi.org/10.3390/s19122812

Yang, Y., He, Z., Song, Z. Y., Fu, X., & Wang, J. W. (2018). Investigation on structural and spatial characteristics of taxi trip trajectory network in Xi'an, China. *Physica A: Statistical Mechanics and its Applications*, *506*, 755–766. https://doi.org/10.1016/j.physa.2018.04.096

Yildirimoglu, M., & Kim, J. (2018). Identification of communities in urban mobility networks using multi-layer graphs of network traffic. *Transportation Research Part C: Emerging Technologies*, *89*, 254–267. https://doi.org/10.1016/j.trc.2018.02.015

Yu, W. H. (2019). Discovering frequent movement paths from taxi trajectory data using spatially embedded networks and association rules. *IEEE Transactions on Intelligent Transportation Systems*, *20*(3), 855–866. https://doi.org/10.1109/Tits.2018.2834573

Zhang, D., Xiao, F., Shen, M., & Zhong, S. (2021). DNEAT: A novel dynamic node-edge attention network for origin-destination demand prediction. *Transportation Research Part C: Emerging Technologies*, *122*, 102851. https://doi.org/10.1016/j.trc.2020.102851

Zhang, X. H., Xu, Y., Tu, W., & Ratti, C. (2018). Do different datasets tell the same story about urban mobility: A comparative study of public transit and taxi usage. *Journal of Transport Geography*, *70*, 78–90. https://doi.org/10.1016/j.jtrangeo.2018.05.002

Zhong, C., Arisona, S. M., Huang, X. F., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, *28*(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521

Zhou, M., Yue, Y., Li, Q. Q., & Wang, D. G. (2016). Portraying temporal dynamics of urban spatial divisions with mobile phone positioning data: A complex network approach. *ISPRS International Journal of Geo-Information*, *5*(12), 240. https://doi.org/10.3390/ijgi5120240

Zhou, T., Liu, X. T., Qian, Z., Chen, H. X., & Tao, F. (2020). Automatic identification of the social functions of areas of interest (AOIs) using the standard hour-day-spectrum approach. *ISPRS International Journal of Geo-Information*, *9*(1), 7. https://doi.org/10.3390/ijgi9010007

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of the article at the publisher's website.

How to cite this article: Chen, X.J., Xie, J., Xiao, C., Lu, B., & Shan, J. (2021). Recurrent origin–destination network for exploration of human periodic collective dynamics. *Transactions in GIS*, 00, 1–24. https://doi.org/10.1111/tgis.12849