Uncovering and modeling the hierarchical organization of urban heavy truck flows

Yang, Yitao; Jia, Bin; Yan, Xiao Yong; Zhi, Danyue; Song, Dongdong; Chen, Yan; de Bok, Michiel; Tavasszy, Lóránt A.; Gao, Ziyou

DOI
10.1016/j.tre.2023.103318

Publication date
2023

Document Version
Final published version

Published in
Transportation Research Part E: Logistics and Transportation Review

Transportation Research Part E 179 (2023) 103318

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

Uncovering and modeling the hierarchical organization of urban heavy truck flows

Yitao Yang [a,c], Bin Jia [a,b,*], Xiao-Yong Yan [a,*], Danyue Zhi [a,d], Dongdong Song [a], Yan Chen [e], Michiel de Bok [c], Lóránt A. Tavasszy [c], Ziyou Gao [a]

[a] *School of Systems Science, Beijing Jiaotong University, Beijing 100044, China*
[b] *School of Economics and Management, Xi'an Technological University, Xi'an 710021, China*
[c] *Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg1, Delft 2628 CN, the Netherlands*
[d] *TUM School of Engineering and Design, Technical University of Munich, Munich 80333, Germany*
[e] *School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China*

ARTICLE INFO

Keywords:
Complex networks
Heavy truck flows
Community structure
Interaction dynamics

ABSTRACT

Knowledge of the hierarchical organization of urban heavy truck flows is important for understanding the structure of urban freight system and underlying interactions dynamics, providing insights to assess and develop freight policies. The complexity and dynamic nature of urban freight system pose significant challenges in comprehensively capturing structured arrangement of heavy truck movements. In this paper, we uncover the hierarchical organization of urban heavy truck flows by using complex network theory. We use large-scale heavy truck GPS data and urban freight location point-of-interest (POI) data to construct urban heavy truck mobility networks, and detect their community structure. The empirical results suggest different sets of locations are closely linked to each other to form multiple clusters. By integrating the categories of locations, we reveal the cluster-specific industry concentration and industry-specific location roles, informing evidence-based policy formulation. To capture the interaction dynamics of locations, we develop a spatial network growth model that considers the spatial agglomeration of industrial clusters and interaction pattern of locations. The model provides a mathematical tool to simulate the formation process of real-world networks for logistics planning and management.

1. Introduction

Modern cities are supported by freight transport system, which guarantees the supply of household goods, industrial raw materials and construction materials (Li et al., 2017). Understanding the urban freight transport system provides insights for policymakers and business economists to assess and develop freight policies, which are of vital importance for improving the livability and sustainability of cities (Tavasszy and De Jong, 2013). The structure of urban freight transport system can be interpreted as the arrangement of functional freight locations (e.g., companies, supermarkets and logistics facilities) and their underlying spatial interactions represented

* Corresponding authors at: School of Systems Science, Beijing Jiaotong University, Beijing 100044, China (B. Jia and X. Y. Yan).
E-mail addresses: yitao-yang@bjtu.edu.cn (Y. Yang), bjia@bjtu.edu.cn (B. Jia), yanxy@bjtu.edu.cn (X.-Y. Yan), danyue-zhi@bjtu.edu.cn (D. Zhi), dongdongsong@bjtu.edu.cn (D. Song), 18125794@bjtu.edu.cn (Y. Chen), M.A.deBok@tudelft.nl (M. de Bok), L.A.Tavasszy@tudelft.nl (L.A. Tavasszy), zygao@bjtu.edu.cn (Z. Gao).

https://doi.org/10.1016/j.tre.2023.103318
Received 17 March 2023; Received in revised form 5 September 2023; Accepted 4 October 2023
Available online 11 October 2023
1366-5545/

by freight vehicle flows (Rodrigue, 2020). Revealing the spatial layouts of freight locations and explaining the underlying interaction dynamics help us understand the structure of urban freight transport system.

The spatial layouts of freight locations refer to the physical arrangement and distribution of facilities and areas involved in freight-related activities (Aljohani and Thompson, 2020). This includes the placement and organization of warehouses, distribution centers, terminals, ports, manufacturing plants and other nodes in the supply chain. In some cases, freight-related facilities exhibit clustering and concentration patterns (Cidell, 2010). This means that certain areas or regions tend to have a higher density of logistics hubs. In other cases, facilities may be strategically distributed across different regions or areas to cater to specific markets or reduce transportation distances. The spatial layouts of freight locations reflect the physical structure of urban freight transport system, and this can be uncovered by analyzing the movements of freight vehicles. Freight vehicle flows indicate the interactions between freight facilities. By examining the hierarchical organization of freight vehicle flows, we can gain insights into the spatial relationships between different levels of facilities across global and regional scales, helping understand the connectivity and dependency of locations in urban freight transport system (Guo et al., 2023; Wang and Yan, 2023; Wang et al., 2023).

The underlying interaction dynamics between locations refers to the dynamic processes and mechanisms through which locations interact with each other (Barthelemy, 2011). It encompasses the underlying factors and forces that drive the interactions between locations, such as supply and demand, transportation infrastructure and economic factors. Location interactions refer to the direct connections that occur between individual locations, and they are crucial for understanding and managing urban freight systems. The interaction dynamics of locations consider the broader system-level processes and relationships that influence the flow of goods, services and information between different locations within a freight network. Modeling the hierarchical organization of freight vehicle flows allows for a structured and systematic analysis of the interaction dynamics between locations. The understanding of the underlying interaction dynamics helps optimize the freight network, improve coordination and enhance the overall efficiency of freight-related activities (Ferrari, 2014, 2015; Nassar et al., 2023).

Heavy trucks are an important component of urban freight vehicles, undertaking high-volume transport tasks between functional freight locations to establish their spatial interactions (Yang et al., 2022b; Zhao et al., 2020). In the big data era, massive heavy truck mobility data become available (Demissie and Kattan, 2022), providing the possibility to understand the structure of urban freight transport system from the perspective of complex networks (Bombelli et al., 2020; Cheng et al., 2022; Cheung et al., 2020; Ghanei et al., 2023). Uncovering and modeling the hierarchical organization of heavy truck flows can shed light on the distinguishing features, e.g., spatial agglomeration, of freight locations and the underlying interactions dynamics, providing supports for regulating urban freight transport system.

In the last decades, the use of complex networks to reveal the urban structure by analyzing the hierarchical organization of individual movement flows has become widespread (Anda et al., 2021; Henry et al., 2022; Louail et al., 2015; Murali et al., 2016; Yildirimoglu and Kim, 2018). Most previous studies (Bassolas et al., 2019; Huang et al., 2018; Yildirimoglu and Kim, 2018) have devoted a great deal of efforts to reveal the urban population-related structure by using massive individual movement data. In terms of urban freight-related structure, previous studies are scarce and mostly focused on the spatial layouts of freight locations, while ignoring the more important aspect, i.e., underlying interactions between locations. With regard to modeling the hierarchical organization of individual movement flows, previous studies (Chen et al., 2022; Lancichinetti et al., 2008; Watts, 2004) have proposed various community-based evolving network models. These models aim to capture the dynamic nature of real-world networks and the evolving patterns of interactions between nodes. However, these models do not consider the spatial agglomeration of industrial clusters, which are characterized by the co-location of firms, suppliers, service providers and other supporting institutions within a particular region or locality (Cong and Zou, 2017; Mori and Smith, 2015). The spatial agglomeration of industrial clusters brings together related industries, promotes collaboration and facilitates economies of scale, and are widely observed in urban freight transport systems. Therefore, to gain a deeper understanding of the interaction dynamics of freight locations, a more realistic network model that considers the spatial agglomeration of industrial clusters needs to be developed.

To fill in the previous gaps, in the paper, we use massive heavy truck GPS data to capture urban heavy truck flows between freight locations, and to construct empirical urban heavy truck mobility networks. We uncover the hierarchical organization of urban heavy truck flows by characterizing the community structure of networks. According to the empirical results, we analyze the spatial distribution of freight location clusters and the roles of individual freight locations, and further explore the cluster-specific industry concentration and industry-specific location roles. To explain the interaction dynamics of freight locations underlying the hierarchical organization of heavy truck flows, we develop a spatial network growth model considering both spatial agglomeration of industrial clusters and spatial interaction pattern of freight locations. Finally, we analyze the practical implications inspired by our model for regulating urban freight transport system, and discuss the potential applications in practice.

Our study contributes to the literature are threefold. (1) We construct urban heavy truck mobility networks by massive movement flows and characterize the community structure of empirical networks. (2) We uncover the hierarchical organization of urban heavy truck flows across communities and nodes, and the cluster-specific industry concentration and industry-specific location roles. (3) We develop a community-based spatial network growth model to capture the interaction dynamics of freight locations, providing policy supports for regulating urban freight transport system.

The remainder of this paper is organized as follows: **Section 2** gives the literature review. **Section 3** presents the collection and details of heavy truck GPS data and urban freight location POI data. **Section 4** provides the methods of constructing urban heavy truck mobility networks, detecting and characterizing community structure, developing and validating the spatial network growth model. **Section 5** describes the empirical and model results, and the policy implications inspired by the model. **Section 6** at the end, offers concluding insights.

2. Literature review

The past decade has witnessed a great deal of efforts into understanding the urban structure and underlying interaction dynamics by exploring the hierarchical organization of individual movement flows. The research provide insights into the functioning and organization of cities, enables the development of efficient transportation systems, helps address societal challenges and capitalizes on technological advancements (Chen et al., 2020). Multi-source individual movement datasets, such as mobile phone data (Bachir et al., 2019; Yan et al., 2017), GPS data (Siripirote et al., 2020; Yang et al., 2022c), social network data (Sala et al., 2021) and smart card data (Xia et al., 2020), have been exploded due to the development of information technology.

With regard to understanding the urban structure, previous studies (Bassolas et al., 2019; Chi et al., 2016; Huang et al., 2018; Liu et al., 2015; Saberi et al., 2017; Xia et al., 2020; Yildirimoglu and Kim, 2018; Zhang and Thill, 2017) have revealed various urban population-related structure at different scales from the perspective of complex network theory (Barthelemy, 2011). In terms of urban freight-related structure, previous studies were scarce and mostly focused on the relationships between transport nodes and networks and the regions in which they are situated (Cui et al., 2015). For example, Grobar (2008) reported that urban freight infrastructure nodes in the USA, such as ports, are often surrounded by low-income and minority-ethnic communities, with residents that are disproportionately exposed to health risks and noise. Similarly, some studies (Dablanc et al., 2013) found that ports, airports and inter-modal terminals that host trade-related activities tend to be concentrated within metropolitan areas. Warehousing and distribution facilities are often clustered around these terminals, or are located near each other at the urban periphery. Some studies (Jacobs et al., 2011; Mullen and Marsden, 2015) also explored the links between urban transport nodes and regional development. In addition to the exploration of the spatial relationships between transport nodes and networks, previous studies have also focused on the distribution of location clusters, i.e., freight community structure. For example, Zheng et al. (2018) investigated the spatial clustering of ports and explored the optimal location of hub ports for various communities. Bai et al. (2023) identified overlapping community structures and key nodes in global liner shipping network to assess network resilience. Nguyen et al. (2020) found Chinese dry ports have strong community structure in the Belt and Road Initiatives. Besides, there are many previous studies (Kale et al., 2007; Krutein and Goodchild, 2022; Mesa-Arango and Ukkusuri, 2015; Ouyang et al., 2022; Zhang et al., 2019) that have explored freight-related community structures.

These previous studies have improved our understanding of urban freight systems, however, most of them focused only on the spatial layouts of freight locations, while ignoring the more important aspect, i.e., underlying interactions between locations. In-depth studies on understanding urban freight transport system by integrating the interactions characterized by freight vehicle (especially heavy truck) flows between freight locations are still lacking.

With regard to understanding the interaction dynamics of locations, the research aims to develop a community-based evolving network model to reproduce the hierarchical organization of individual movement flows, and to explain how urban structure is formed. A community-based evolving model refers to a computational framework that incorporates both community dynamics and spatial patterns to understand and predict the evolution of a system. This model combines elements of community detection, spatial analysis, and evolutionary dynamics to capture the interactions and transformations within a complex system over time (Fortunato, 2010; Fortunato and Hric, 2016). Early studies mainly concentrated on the unweighted network models, and the most representative one is the Barabasi-Albert (BA) model (Barabasi and Albert, 1999). Subsequently, many network models (Barrat et al., 2004, 2005; Louf et al., 2013) have been proposed to reproduce and explain a variety of phenomena in real-world systems. However, these network models cannot capture the hierarchical organization (also well-known as communities) widely observed in many real-world systems (Fortunato, 2010). To this end, previous studies proposed various network models with community structure. A special class of models, so-called planted ℓ-partition model (Condon and Karp, 2001), is quite popular in the past decades. The model partitions a graph in ℓ communities with equal number of nodes each. The nodes of the same community are linked with a given probability, whereas the nodes of different communities are linked with another probability. Inspired by this, many other network models with community structure, such as Gaussian random model (Brandes et al., 2003), LFR model (Lancichinetti et al., 2008), relaxed caveman model (Watts, 2004) and embedded hierarchy model (Arenas et al., 2006), were proposed. Most of these models are used for generating benchmark graphs of community detection algorithms. however, they need to be refined to provide a better description of real-world networks with community structure. Accordingly, previous studies proposed a variety of community structured evolving network models (Kossinets and Watts, 2006; Li and Chen, 2006). These models were used to capture the community structure of many real-world networks, such as social networks (Hanaki et al., 2007; Kossinets and Watts, 2006), biological networks (Rives and Galitski, 2003) and citation networks (Rosvall and Bergstrom, 2008), but were not specifically designed for transportation systems. Because these models did not consider the spatial relationships between nodes when characterizing the community structure, they may not fully capture the spatial dependencies and patterns that are inherent in transportation systems (Rodrigue, 2020). Recently, Chen et al. (2022) proposed two evolving network models with distance preferences, named MoncSid-N and MoncSid-E, which provide insights for explaining the effects of space on the formation of community structure.

These previous models help us understand the dynamical mechanisms of many real-world systems, but they cannot be applied to capture the interaction dynamics of freight locations, because they do not consider the spatial agglomeration of industrial clusters observed in urban freight transport systems (Wu et al., 2022). companies of specialized industries tend to concentrate within small geographical areas, forming industry clusters based on factors like raw materials, markets, and transport costs (Shakib, 2020; Wu et al., 2022). Companies within the same industry cluster play similar roles in the urban freight transport system, leading to shared business partnerships and specific interaction patterns. However, previous network models fail to explain the interaction dynamics of freight locations associated with this phenomenon. Therefore, it is crucial to develop a community-based spatial evolving model that considers the spatial agglomeration of industrial clusters to better understand and explain the interaction dynamics of freight locations.

This study explores the urban freight-related structure by integrating the interactions between locations characterized by large-scale heavy truck flows. We reveal the spatial layout and industry-related interactions between locations by uncovering the hierarchical organization of urban heavy truck flows. We develop a spatial network growth model that considers the spatial agglomeration of industrial clusters to explain the interaction dynamics underlying the urban freight-related structure.

3. Data

3.1. Heavy truck GPS data

We capture urban heavy truck flows between freight locations by using massive heavy truck GPS data. Our heavy truck GPS data are from the China Road Freight Supervision and Service Platform (https://www.gghypt.net/). This platform is used to record the real-time geographic locations of all heavy trucks with a load exceeding 12 tons in China and monitor their traffic violations (speeding, fatigue driving, etc.). We obtain the GPS trajectories of 2.6 million heavy trucks over the period from May 18, 2018 to May 31, 2018. The attributes of the GPS data include truck ID, timestamp, longitude, latitude, speed and direction angle. The number of records is greater than 41 billion.

Raw GPS data often contains erroneous and redundant information, such as data jumps and drifts caused by factors like GPS signal reflection or obstruction in urban areas (Demissie and Kattan, 2022) These issues can compromise the quality and reliability of the data, particularly in scenarios like tunnels where GPS signal loss is prevalent. To ensure the integrity of the data, we employ data preprocessing techniques that specifically address three main types of abnormal data: duplicate or missing data, unreasonable data, and data jumps. In dealing with duplicate or missing data, we adopt a straightforward approach of removing corresponding GPS records. This step helps maintain the authenticity and validity of the data by eliminating any redundant or incomplete observations. Additionally, we identify and eliminate unreasonable data points, such as GPS points located outside national borders, as they are unlikely to represent valid movement patterns or locations. To detect and handle data jumps, we calculate the average speed and acceleration between successive GPS points and compare them against predefined maximum threshold values (e.g., 120 km/h for speed and 5 m/s^2 for acceleration). If the calculated average speed or acceleration exceeds these thresholds, indicating a significant deviation from normal movement, we consider it a data jump and remove the corresponding GPS records.

3.2. Freight location POI data

We use freight location POI data to construct urban heavy truck mobility networks. We use the application programming interface (API) of Amap (https://lbs.amap.com) to crawl freight location POIs in each city. In the Amap application, developers store POIs in a hierarchical format by industry categories. According to the correlation between POIs and heavy truck freight activities (Amer and Chow, 2017; Dernir et al., 2014), we choose four base categories of POIs, including enterprises, shopping, daily life service and transportation service. First, the category of enterprises involves manufacturing, distribution and other industries that are significant generators and recipients of freight flows. They often serve as origins or destinations for goods, and their spatial distribution has a direct impact on urban freight patterns (Pamucar et al., 2022). Second, the category of shopping involves markets and stores that attract a substantial amount of freight traffic due to their role in the distribution and sale of goods. These locations serve as key nodes to connect suppliers and customers (Lim et al., 2019). Third, the category of daily life service involves logistics facilities that play a crucial role in the handling, consolidation, and distribution of goods within urban areas. These facilities act as intermediate points within the freight transport systems, facilitating the transfer of goods (de Oliveira et al., 2022). Fourth, the category of transportation service involves transshipment terminals that facilitate the consolidation, sorting and redistribution of freight (Mohammed et al., 2023). The four selected categories involve different stakeholders, including manufacturers, distributors, retailers and transportation service providers, across the urban supply chains, and thus have sufficient representativeness of urban freight networks.

The subcategories under the base category of enterprises include advertisement & decoration company, construction company, medical company, machinery & electronics company, chemical & metallurgy company, network science & technology company,

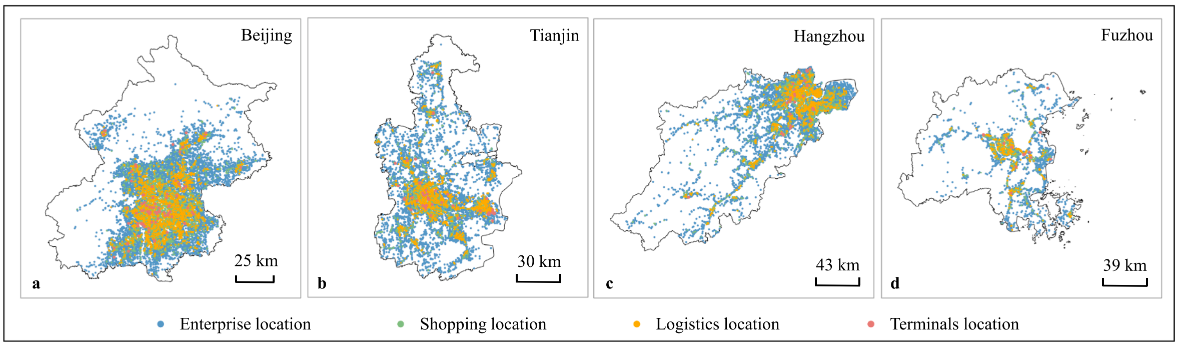

Fig. 1. Geographic distributions of four categories of POIs in four cities.

commercial trade company, telecommunication company, mining company and factory. The subcategories under the base category of shopping include shopping plaza, convenience store, sports store, clothing store, franchise store, home electronics hypermarket, personal care items shop, supermarket, plants & pet market, home building materials market, comprehensive market, stationary store and special trade house. The subcategories under the base category of daily life service include logistics service, logistics warehouse space. The subcategories under the base category of transportation service include airport related location, railway station, port & marina and border crossing location. The record of each POI provides its name, geographic location and category. The geographic distributions of four base categories of POIs in four case cities are shown in Fig. 1. The results show that freight-related companies tend to be located on the suburbs of cities, while shopping and logistics nodes tend to be located in city centers. Freight terminals are the least numerous and are more evenly distributed across the city.

4. Method

4.1. Construction of urban heavy truck mobility networks

In the paper, we first construct urban heavy truck mobility networks, and then uncover the hierarchical organization of urban heavy truck flows from the perspectives of complex networks (Barthelemy, 2011). We construct urban heavy truck mobility networks by using heavy truck GPS data and freight location POI data, as shown in Fig. 2. To begin with, we use a recent trip origin–destination (OD) identification algorithm (Yang et al., 2022b) to extract freight trips of each heavy truck from its GPS trajectory (see Fig. 2**a-b**). In this OD identification algorithm, a speed threshold is first determined by analyzing the truck speed distribution characteristics to identify truck stops from GPS data, and then the multilevel time thresholds are determined by using a nonparametric iterative method to dynamically identify trip ODs from all stops of a truck. For each identified trip origin or destination, freight-related POIs and urban road networks are used to determine whether it is a real trip end to ensure algorithmic accuracy. Next, we can obtain the heavy truck flows, i.e., the integration of heavy truck trips, between each pair of freight locations (see Fig. 2**c**). The volume of heavy truck flows, i.e., the number of heavy truck trips, between each pair of freight locations indicates the spatial interaction strength between them, as shown in Fig. 2**d**. Finally, we construct urban heavy truck mobility networks by integrating the heavy truck flows between all pairs of freight locations (see Fig. 2**e**).

Urban heavy truck mobility networks are weighted undirected networks $G(\boldsymbol{N}, \boldsymbol{E}, \boldsymbol{W})$, where $\boldsymbol{N}$ is the set of nodes represented by freight locations, and their geographical coordinates are given by latitude and longitude. $\boldsymbol{E}$ is the set of links represented by spatial interactions between freight locations and $\boldsymbol{W}$ is the set of link weights represented by interaction strengths, i.e., truck flows, between freight locations.

4.2. Community detection and characterization

4.2.1. Community detection method

We use the Infomap (Rosvall and Bergstrom, 2008) method to detect the communities of urban heavy truck mobility networks, and uncover the hierarchical organization of urban heavy truck flows by characterizing the community structure. Infomap is one of the most popular community detection method: it uses the probability flow of random walks on a network as a proxy for information flows

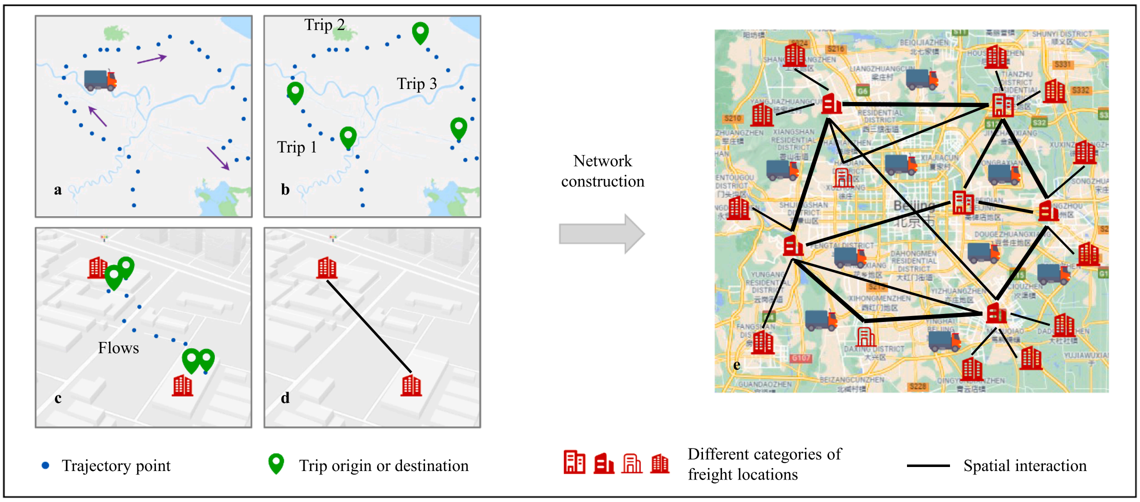

Fig. 2. Network construction by using heavy truck GPS trajectory data and urban freight location point-of-interest data. We first extract heavy truck trips (panel **b**) from GPS trajectories (panel **a**) by using trip ends identification method, and then obtain the heavy truck flows (panel **c**) and spatial interactions (panel **d**) between freight locations. **e** Constructing network by integrating spatial interactions of all freight location pairs. The network nodes and links are represented by freight locations and spatial interactions between them. Line width in the panel indicates interaction strength.

in the real system and decompose the network into communities by compressing a description of the probability flow. In the previous empirical comparisons of algorithms to find communities (Lancichinetti and Fortunato, 2011), Infomap showcased remarkable advantages in accuracy and computational performance, and is applicable to large-scale networks. The Infomap method, originally designed for unweighted networks, has also been extended to handle weighted networks (Rosvall and Bergstrom, 2008). In the case of weighted networks, the algorithm takes into account both the topology of the network (connections between nodes) and the weights associated with those connections. To detect the community structure of urban heavy truck mobility networks, we construct a transition probability matrix that describes the probability of moving from one node to another. These probabilities are positively correlated with the weights of connections. Next, imagine a random walker moving on the network. At each step, the walker transitions from one node to another based on the transition probabilities in the matrix. The goal is to find a partition that minimizes the expected description length of the walker's movements. The algorithm iteratively optimizes the description length by trying different partitions and evaluating their quality. Once the algorithm converges, the resulting partition provides information about the detected communities.

Community structure is the organization of nodes in communities, with many links joining nodes of the same community and comparatively few links joining nodes of different communities (Fortunato, 2010). In urban heavy truck mobility networks, the result of community detection is the set of partition $\mathscr{C} = (C_1, C_2, \cdots, C_m)$ with m communities, where each community C is the set that contains the nodes in this community. Fig. 3 shows the illustrations of detected community structure of heavy truck mobility networks of four cities.

4.2.2. Community characterization metrics

We characterize the detected community structure of urban heavy truck mobility networks by using complex network metrics (Fortunato and Hric, 2016). We aim to uncover hierarchical organization of heavy truck flows at the levels of community and node, and to understand the spatial distribution of freight location clusters and the roles of individual freight locations.

With regard to community-level metrics, we use metrics include: modularity Q, global Moran's I GMI, community internal strength w_C^{int}, community external strength w_C^{ext}, centrality index Θ_C, cost ϕ_c, weighted conductance $C_{w,C}$. We use modularity Q (Newman, 2004; Newman and Girvan, 2004) as a quality function to estimate the quality of a partition of the network in communities. Next, we elaborate each measure.

The modularity Q is given by

$$Q = \frac{1}{2W}\sum_{ij}\left(w_{ij} - \frac{s_i s_j}{2W}\right)\delta(\mathscr{G}_i, \mathscr{G}_j) \tag{1}$$

where W is the sum of the weights of all links, w_{ij} is the weight of link $E(i,j)$, s_i is the strength of node i. Node strength s_i is defined as the sum of the weights of all the edges connected to a node, i.e., $s_i = \sum_{j\in\Gamma(i)} w_{ij}$, where $\Gamma(i)$ is the neighbor nodes of node i. δ-function yields one if nodes i and j are in the same community ($\mathscr{G}_i = \mathscr{G}_j$), zero otherwise. The z-score of modularity is given by $z = (Q - \langle Q\rangle_{NM})/\sigma_Q^{NM}$, where $\langle Q\rangle_{NM}$ and σ_Q^{NM} are average and standard deviation of modularity of many realizations of the null model, obtained from the original graph by randomly rewiring its links. If $z \gg 1$, network has strong community structure.

We use global Moran's I GMI (Moran, 1950) to measure the spatial concentration of freight locations in the same community. The global Moran's I is a measure of the overall clustering of the spatial data. For a network with N nodes, GMI exceeds $-1/(N-1)$ and z-score of GMI greater than 1 indicate positive spatial autocorrelation. Community internal strength w_C^{int} is the sum of weights of links connecting nodes in community C, given by $w_C^{int} = \sum_{i,j\in C} w_{ij}$. Community external strength w_C^{ext} is the sum of weights of links connecting nodes in community C and the nodes in other communities, given by $w_C^{ext} = \sum_{i\in C, j\notin C} w_{ij}$.

Centrality index Θ_C (Moraes Pereira et al., 2013) is used to measure the concentration of heavy truck flows among the freight locations in a community, given by

$$\Theta_C = H_C \bullet P_C \tag{2}$$

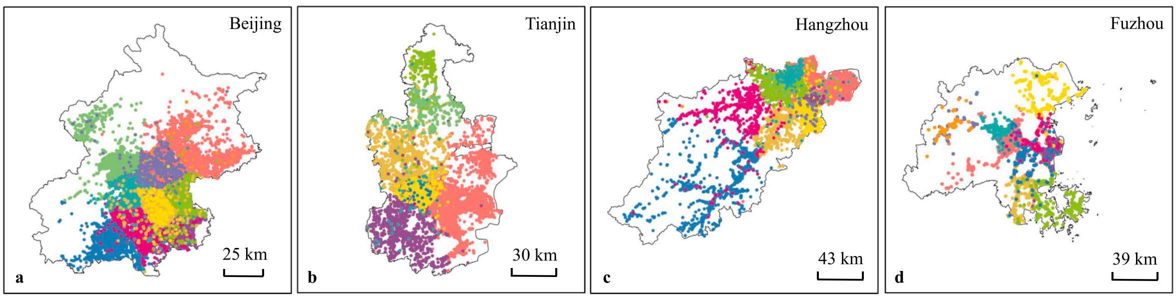

Fig. 3. Results of community detection for heavy truck mobility networks of four cities. The network nodes belong to the same community are marked in the same color.

Where H_C is location coefficient, i.e., $H_C = \sum_1^n |\mu_i - 1/n|/2$, in which n denotes the number of locations in the community C, μ_i denotes the ratio of strength of location i to the total strengths of all locations in community C. P_C is normalized spatial separation index, i.e., $P_C = 1 - V/V_{\max}$, in which $V = \mathbf{S}^{'} \times \mathbf{D} \times \mathbf{S}$ denotes the spatial separation index. $\mathbf{S} = (\mu_1, \mu_2, \mu_3 \cdots \mu_n)^T$ is a column vector composed of μ_i, $\mathbf{D}$ is a distance matrix with elements d_{ij} representing the Euclidean distance between locations i and j. Centrality index Θ_C is within the interval $[0, 1]$, and the closer Θ_C is to 1, the more significantly higher the heavy truck flows generated and attracted by few freight locations in community C. $\Theta_C = 0$ indicates the transportation demands of all freight locations in community C are equal.

Cost ϕ_c is an important metric of spatial networks, given by

$$\phi_c = \ell_T / \ell_T^{MST} \tag{3}$$

where ℓ_T denotes the sum of length of the links between the freight locations in community C,ℓ_T^{MST} denotes the sum of length of the links of minimum spanning tree for the subgraph of community C. Weighted conductance $C_{w,C}$ is defined by the ratio between the community external strength and internal strength: $C_{w,C} = w_C^{ext}/w_C^{int}$.

With regard to node-level metrics, we use metrics include: node degree k_i, node internal strength w_i^{int}, node external strength w_i^{ext}, weighted embeddedness $\xi_{w,i}$, weighted within-module degree $Z_{w,i}$, weighted participation coefficient $P_{w,i}$. Next, we elaborate each node-level metric. Node degree k_i is the sum of edges connected to node, given by $k_i = \sum_j A_{ij}$. When i and j are connected, $A_{ij} = 1$; and vice versa, $A_{ij} = 0$. Node internal strength w_i^{int} is the sum of weights of links connecting node i and the nodes in the same community C that node i is in, given by $w_i^{int} = \sum_{j \in C} w_{ij}$. Node external strength w_i^{ext} is given by $w_i^{ext} = \sum_{j \notin C} w_{ij}$. Weighted embeddedness $\xi_{w,i}$ is the ratio between the internal strength and total strength of a node, given by $\xi_{w,i} = w_i^{int}/s_i$. The larger $\xi_{w,i}$, the stronger the relationship

Table 1
Community characterization metrics and explanation.

Metric	Symbol	Mathematical meanings	Practical meanings	Purposes
Modularity	Q	The difference between observed weights of edges within communities to the expected weights in a random network.	A higher value indicates a stronger community structure, with dense connections within communities and sparse connections between communities.	Evaluating how well the locations in the network are organized into distinct communities.
Global Moran's I	GMI	Normalization of the covariance and variance of the observed edge weights and the spatial weights.	A higher value indicates that spatially adjacent locations more tend to cluster in the same community.	Quantifying the extent of spatial agglomeration of locations.
Community internal strength	w_C^{int}	Summing the weights of the edges that connect nodes within the same community.	A higher value indicates that locations within the community have stronger and more frequent interactions with each other.	Measuring the cohesion of locations within a community.
Community external strength	w_C^{ext}	Summing the weights of the edges that connect nodes belonging to different communities.	A higher value indicates stronger and more frequent interactions between nodes from different communities.	Measuring the integration between different communities.
Centrality index	Θ_C	Multiplying the location coefficient and normalized spatial separation index.	A higher value indicates a more significantly higher the heavy truck flows generated and attracted by few freight locations.	Measuring the concentration of flows among the locations in a community.
Cost	ϕ_c	Ratio of the sum of length of the links between the locations in a community to that of minimum spanning tree.	A higher value indicates a relatively longer transportation routes within the community.	Measuring the spatial organization and accessibility of a community.
Weighted conductance	$C_{w,C}$	Ratio between the community external strength and internal strength.	A higher value indicates that the community has more substantial interactions and exchanges with locations outside the community.	Measuring the internal coherence of a community.
Node degree	k_i	The count of edges connected to a node.	A higher value indicates that a location has more connections with other locations.	Quantifying the connectivity of a node.
Node strength	s_i	Summing up the weights of all the edges connected to a node	A higher value indicates that a location has stronger interactions with other locations.	Measuring the overall importance or influence of a node.
Node internal strength	w_i^{int}	Summing up the weights of all the edges that connect a node to other nodes within the same community.	A higher value indicates that a location has a stronger internal connection within the community.	Identifying the role played by a node
Node external strength	w_i^{ext}	Summing up the weights of all the edges connecting a node to nodes outside its community	A higher value indicates that a location has a stronger external connection with other communities.	Identifying the role played by a node
Weighted embeddedness	$\xi_{w,i}$	The ratio between the internal strength and total strength of a node	A higher value indicates a larger proportion of node's connections within its community	Identifying the role played by a node
Weighted within-module degree	$Z_{w,i}$	Standardized node internal strength using mean and standard deviation	A higher value indicates that a location is more tightly integrated into its community.	Identifying the role played by a node
Weighted participation coefficient	$P_{w,i}$	Summing the squares of the proportions of the weights of node's connections to nodes in other communities, and then subtracting the sum from 1	A higher value indicates that a location has connections with multiple communities and actively participates in their activities.	Identifying the role played by a node

between node i and its community. Weighted within-module degree $Z_{w,i}$ (Guimera and Amaral, 2005) measures how 'well-connected' node i is to other nodes in the community, given by

$$Z_{w,i} = \frac{w_i^{int} - \overline{w}_{int}}{\sigma_{w_{int}}} \tag{4}$$

where w_i^{int} is internal strength of node i, $\overline{w}_{int}$ is the average of internal strengths over all the nodes in node i's community, $\sigma_{w_{int}}$ is the standard deviation of w_{int}. Weighted participation coefficient $P_{w,i}$ (Guimera and Amaral, 2005) measures how 'well-distributed' the links of node i are among different communities, given by

$$P_{w,i} = 1 - \sum_{c=1}^{m} \left(\frac{W_{ic}}{s_i}\right)^2 \tag{5}$$

where s_i is total strength of node i, W_{ic} is the sum of weights of links of node i to nodes in community C. The weighted participation coefficient of a node is therefore close to 1 if its links are uniformly distributed among all the communities and 0 if all its links are within its own community. The larger $P_{w,i}$, the more likely the freight location i is to play the role of community connector. Table. 1 shows the explanation of mathematical meanings, practical meanings and purposes of used metrics.

4.3. Spatial network growth model

In the previous section, we use a variety of community-level and node-level metrics to reveal the structural properties of the real heavy truck mobility networks. Here, we aim to develop a spatial network growth model that can capture the interaction dynamics of freight locations. The model serves as a mathematical tool of simulating the formation process of real-world networks for logistics planning and management. Compared to previous network models (Barrat et al., 2004, 2005), our model, as shown in Fig. 4, considers the spatial agglomeration of industrial clusters and spatial interaction pattern of freight locations, therefore, is more interpretable and realistic.

Our model starts with an initial fully connected seed network containing N_0 nodes, which are randomly selected from among the freight locations in city. Each link in this seed network is given a weight w_0. For simplicity, we set $N_0 = 5$ and $w_0 = 1$ both to constants. At each step, we randomly select a point from the remaining urban freight locations as the new added node i. According to the recent empirical study on heavy truck mobility networks (Yang et al., 2022a), we initialize the degree k_i of new node i from the distribution $p(k) = a(k + \Delta k)^{-\gamma} e^{-k/k_x}$, where a, γ and k_x are the given parameters, and initialize the strength s_i of new node i by $s_i = bk_i^{\alpha}$, where b and α are also the given parameters.

Next, new node i chooses its connecting nodes. In this process, we consider the spatial agglomeration of industrial clusters and spatial interaction pattern of freight locations. In urban freight transport system, a set of companies of specialized industries tend to concentrate within small geographical areas and form industry clusters, i.e., communities, in consideration of raw materials, markets and transport costs (Aljohani and Thompson, 2016; Shakib, 2020; Wu et al., 2022). To explain this phenomenon, the model considers location clustering interactions. The communities of the network are first detected by using Infomap method, and the community with

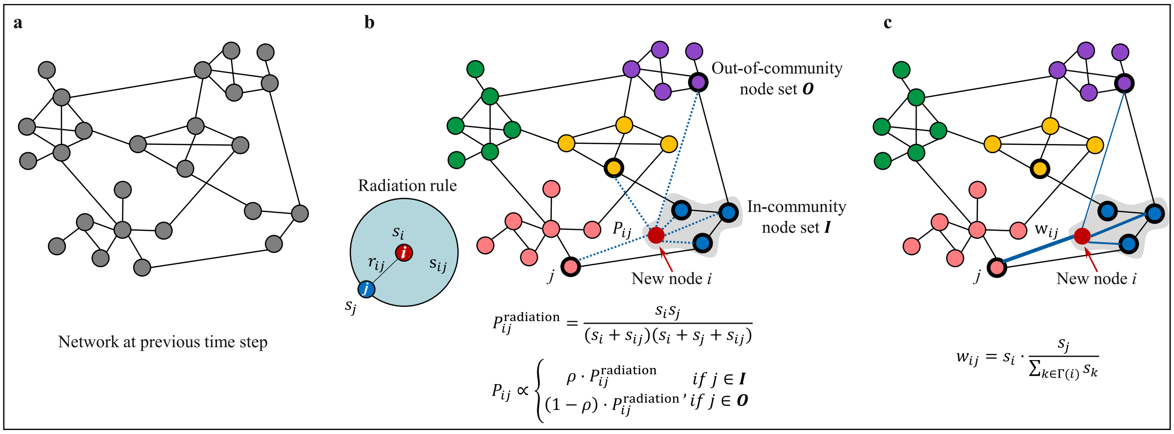

Fig. 4. Model illustration. **a** Network obtained at previous time step. **b** Network growth. In the next time step, the communities of the current network are first detected, and the nodes belong to the same community are marked by the same color. The gray zone denotes the community $\mathscr{G}_i$ that the new node i is in. The nodes in the sets $\boldsymbol{I}$ and $\boldsymbol{O}$ are chosen to be the potential connecting nodes of node i, as linked by dashed lines. The probability of each potential connecting node j being connected by new node i is P_{ij}, in which $P_{ij}^{\text{radiation}}$ is the interaction probability of radiation model, s_{ij} is total strength of nodes located in the circle of radius r_{ij} centred at new node i, as illustrated in the lower left corner. The number of nodes connected by new node i is equal to the initialized degree k_i. **c** Network weight updates. The initialized strength s_i of new node i is proportionally distributed among the links departing from the node i according to the strengths of connected node $\Gamma(i)$.

the shortest average distance from the nodes in it to the new node i is identified to be the community $\mathscr{G}_i$ that the new node i is in, as shown in Fig. 4**a-b**. A distinguishing feature of industry clusters is that the companies in the same industry cluster tend to play similar roles in the urban freight transport system, manifesting themselves as shared business partners. To explain this phenomenon, the model assumes that the nodes (defined as in-community nodes) in community $\mathscr{G}_i$ and the nodes (defined as out-of-community nodes) in other communities but have connections with community $\mathscr{G}_i$ are chosen to be the potential connecting nodes of node i. The probability P_{ij} of each potential connecting node j being connected by new node i is defined as

$$P_{ij} \propto \begin{cases} \rho \cdot P_{ij}^{\text{radiation}} & , if j \in \mathbf{I} \\ (1-\rho) \cdot P_{ij}^{\text{radiation}} & , if j \in \mathbf{O} \end{cases} \tag{6}$$

where $\mathbf{I}$ is in-community node set, $\mathbf{O}$ is out-of-community node set, ρ is embedding coefficient and $P_{ij}^{\text{radiation}}$ is interaction probability of freight locations calculated by radiation model (Simini et al., 2012), given by

$$P_{ij}^{\text{radiation}} = \frac{s_i s_j}{(s_i + s_{ij})(s_i + s_j + s_{ij})} \tag{7}$$

where s_i is the strength of node i, and s_{ij} is total strength of nodes located in the circle of radius r_{ij} centred at new node i (excluding the strength of node i and j) in the network at current step. The radiation model is a universal parameter-free spatial interaction model, widely used to analyze and predict the interactions between locations (Ren et al., 2014). The spatial interaction pattern of radiation model can be summarized as: freight locations tend to establish interactions with the closest and more competitive freight locations, in other words, the probability of interacting with a freight location is positively correlated with its competitiveness and negatively correlated to distance. In our model, the competitiveness of a freight location i is measured by its transportation demand, i.e., generated and attracted heavy truck flows (denoted by node strength s_i). The embedding coefficient ρ in **eq. (6)** controls the relationship between new node i and its community. The larger ρ, the more likely a freight location is to establish interactions within its industry cluster, otherwise the more significant it is to undertake the role of a connector between different industry clusters. When calculating the connection probability P_{ij} between new node i and existing node j, the radius r_{ij} is identical to the spatial distance between them. A larger radius r_{ij} means the probability of establishing a connection between new node i and existing node j is lower. The embedding coefficient ρ is the given model parameter. New node i chooses its connecting nodes $\Gamma(i)$ according to the probability in **eq. (6)**, and the number of connected nodes is equal to the initialized degree k_i of new node i.

In the following, we determine the interaction strength, i.e., link weight, between new node i and each connected node $j \in \Gamma(i)$. Model assumes that the interaction strength between two freight locations is proportional to their competitiveness measured by node strength in the network at current step, as shown in Fig. 4**c**. The weight w_{ij} of each new link $E(i,j)$ is given by

$$w_{ij} = s_i \cdot \frac{s_j}{\sum_{k \in \Gamma(i)} s_k} \tag{8}$$

It is worth noting that the geographic location of this new node n added at the current time step is the same as the corresponding node in the real network, but the strength of this new node is randomly given according to the above pre-defined statistical distribution. The probabilities of potential connections are calculated by considering the strengths of the nodes in the current network. After establishing new edges, the strength of each existing node connected by the new node will be updated too. Therefore, as the network grows, the node strengths are changeable, reflecting the interaction dynamics.

In the next step, another new node is added from the remaining urban freight location POIs, and the current network grows according to the above process. This process will terminate until the number of nodes in the network reaches that of a real network. As the proposed spatial network growth model considers community structure and the spatial interaction pattern of radiation model, we refer to this model as the community-based radiation network model, i.e., CRN model.

4.4. Model validation

In the above, we develop the CRN model to explain the interaction dynamics of freight locations underlying the hierarchical organization of heavy truck flows. In this section, we describe how we validate the model by reproducing the community structure of empirical urban heavy truck mobility networks.

The CRN model contains one key parameter, i.e., embedding coefficient ρ, that needs to be estimated before the model validation. We use a graph similarity-based method (Sala et al., 2010) to estimate model parameter to reproduce the community structure of empirical networks as best as possible. This method estimates the optimal parameter ρ^* by maximising the similarity between model networks and the empirical network. The similarity of two networks is measured by the Canberra distance (Lance and Williams, 1966) between network attribute vectors. We construct the network attribute vector by using two node-level metrics, i.e., weighted embeddedness $\xi_{w,i}$ and weighted participation coefficient $P_{w,i}$. In the model parameter estimation, we obtain multiple optional parameter values of parameter ρ at certain intervals. For each optional parameter value, we generate 100 realizations of CRN model to construct the attribute vector of model networks. The optional parameter value corresponding to the minimum Canberra distance between the attribute vectors of model networks and the empirical network is the estimated optimal parameters ρ^*.

For the real heavy truck mobility network of a city, we can obtain the corresponding model network generated by the CRN model with the estimated parameters ρ^*. The nodes in model network and in real network are identical. We use Infomap method to detect the community structure of model network and that of the real network. To validate the CRN model, we aim to examine whether the model network has similar community structure of the real network, as manifested in three aspects: (1) whether node pairs located in the same community in the real network also clustered in the same community in the model network; (2) whether the community spatial distribution characteristics of the model network match those of the real network; and (3) whether the roles played by nodes with different connectivity in the model network match those of the real network. For the first aspect, we use two similarity metrics, i.e., Wallace index *WI* (Wallace, 1983) and normalized mutual information *NMI* (Danon et al., 2005), to measure whether node pairs that are in the same community in the empirical network are also in the same community in the model network. Wallace index *WI* is a measure based on pair counting depend on the number of pairs of nodes which are classified in the same (different) communities in the two partitions of model network and real network. The normalized mutual information *NMI* is a measure based on information theory (MacKay, 2003). The idea is that, if two partitions are similar, one needs very little information to infer one partition given the other. Both *WI* and *NMI* range from 0 to 1, where 0 indicates complete dissimilarity (no shared node pairs between the communities) and 1 indicates complete similarity (the communities have exactly the same set of node pairs). For the second aspect, we characterize the community spatial distribution of model network by using community-level metrics, and compare them with real results. For the third aspect, we characterize the roles played by nodes with different connectivity in model network by using node-level metrics, and compare them with real results.

Furthermore, we use two benchmark models, i.e., space-constrained growth network model (SGN model) (Barrat et al., 2005) and Barrat-Barthelemy-Vespignani model (BBV model) (Barrat et al., 2004), to validate the advantages of the proposed CRN model. First, SGN model incorporates geographical attributes along with topological and weight (traffic) properties into the network growth process. SGN model is developed for simulate the growth of weighted spatial networks including transportation networks. This model considers the spatial interaction pattern between locations at the micro level, but not location clustering interactions at the meso level, which is incorporated in the CRN model. Second, BBV model is developed for simulating the growth of weighted networks by coupling topology and weight dynamics. This model does not consider either the geographical attributes or the location clustering interactions in the network growth process. We estimate the optimal parameters contained in the SGN model and BBV model respectively by using the data of real network of a city. Next, we use SGN model and BBV model with the estimated optimal parameters to simulate the network of a city and detect the communities of model network respectively. To evaluate the performance of CRN model, we compare the community structure of real network with that generated by CRN model and two benchmark models.

5. Results

5.1. Spatial distribution of freight location clusters

First, the communities of urban heavy truck mobility networks are detected for four cities, i.e., Beijing, Tian, Hangzhou and Fuzhou, in China by using Infomap method (Rosvall and Bergstrom, 2008). The illustration of the detected communities for four cities is shown in Fig. 3. We first uncover the hierarchical organization of heavy truck flows by using community-level metrics (see **Section 4.2**), and then to understand the spatial distribution of freight location clusters.

The global characterization of detected communities of urban heavy truck mobility networks is shown in Table. 2. The network sizes vary across cities, but the networks of these cities have similar structure properties. The z-scores of modularity $Z_Q \gg 1$, indicate urban heavy truck mobility networks have community structure and different sets of freight locations are closely linked to each other to form multiple clusters, e.g., industrial clusters (Shakib, 2020). The global Moran's I *GMI* > 0 and its z-score $Z_{MI} \gg 1$, indicate the spatial autocorrelation of freight locations, i.e., spatially adjacent freight locations tend to be in the same cluster. This phenomenon of spatial agglomeration of industrial clusters is led by factors such as raw materials, markets and transport cost, reflecting urban land use patterns and economic layouts. The results provide supports for formulating policies to better develop the urban agglomeration economies (de Bok and van Oort, 2011). Next, we calculate the internal strength w_C^{int} and external strength w_C^{ext} of each community C, and the distributions of these two metrics with respect to community size n_C are shown in Fig. 5. We can find that both community internal strength and external strength are positively correlated with community size. The larger an industry cluster is, the higher the heavy truck flows not only between the freight locations within this industry cluster, but also between this industry cluster and other industry clusters. The results indicate the hierarchical organization of heavy truck flows across communities with different sizes.

We also calculate the centrality index Θ_C of each community C, and obtain the distributions of this metric with respect to community size n_C (see Fig. 6**a-d**). The positive correlation between Θ_C and n_C suggests that the larger an industry cluster is, the more heterogeneous are the heavy truck flows distributed within it, i.e., heavy truck flows generated and attracted by few freight locations in

Table 2
Global characterization of detected communities of heavy truck mobility networks of four cities in China. Number of network nodes *N*; number of network links *E*; number of communities *NC*; modularity *Q*; z-score of modularity Z_Q; global Moran's I *GMI*; z-score of global Moran's I Z_{GMI}.

City	N	E	NC	Q	Z_Q	GMI	Z_{GMI}
Beijing	15,983	207,264	11	0.51	43	0.43	106
Tianjin	10,697	216,724	12	0.47	30	0.65	142
Hangzhou	10,611	160,718	13	0.42	51	0.51	104
Fuzhou	3391	26,016	10	0.45	38	0.67	82

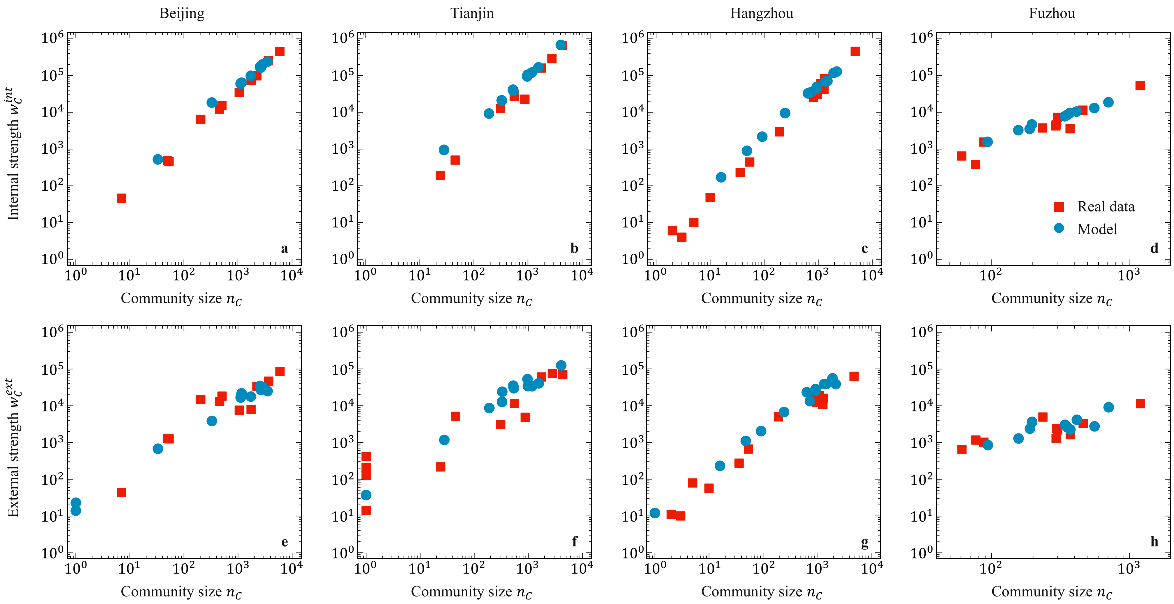

Fig. 5. Community strength characteristics observed from empirical networks of four cities and reproduced by the proposed model. **a-d** Distributions of internal strength w_C^{int} with respect to community size n_C. The legend "real data" represents the distributions of the metrics of the real networks, and each point represents each detected community. The legend "model" represents those of the model networks. **e-h** Distributions of external strength w_C^{ext} with respect to n_C.

larger industry cluster are more significant. Moreover, we calculate the cost ϕ_c of each community C, and obtain the distributions of this metric with respect to community size n_C (see Fig. 6**e-h**). We can find the costs of larger industry clusters are higher, suggesting the subgraphs of small industry clusters tend to be characterized with tree-like features and freight locations in larger industry clusters are more closely linked to each other. This also suggests the scale effects of larger industry clusters are more remarkable. Finally, we obtain the distributions of weighted conductance $C_{w,C}$ with respect to community size n_C (see Fig. 6**i-l**). The results indicate the larger an industry cluster is, the more concentrated heavy truck flows are within it, exhibiting relatively fewer interactions between it and other industry clusters. This can be explained by the fact that companies in larger industry clusters tend to form more established manufacturing supply chains, and therefore their interactions with the companies in other industry clusters are not prominent (Jote et al., 2013).

The above analysis reveals the spatial distribution of freight location clusters. In the following, we explore the cluster-specific industry concentration by analyzing which categories of locations tend to be in the same community and how they interact each other. In urban heavy truck mobility networks, the nodes contain four categories of locations, i.e., enterprise locations, shopping locations, logistics location and terminal locations. We examine which categories of locations tend to be in the same community by applying a chi-squared test (Pearson, 1900), which enables us to analyze the observed frequencies of category co-occurrence within communities and compare them to the expected frequencies under the assumption of independence. If the observed co-occurrence frequency of a category pair significantly deviated from the expected frequency in a community, it indicates these two categories of locations tend to be grouped together in this community. To obtain the expected frequencies, we randomly shuffle the category labels of all nodes while keeping network structure unchanged, and calculate the co-occurrence frequencies of category pairs in each community as the expected frequencies. We apply the chi-squared test for all detected communities, and the test results of four typical communities in the network of Beijing are shown in Fig. 7**a-d**. We also apply the chi-squared test to uncover which categories of locations tend to interact directly, i.e., establishing connections, in a community, and the process is similar to the co-occurrence test. The direct interaction test results of selected four typical communities in the network of Beijing are shown in Fig. 7**e-h**. The results suggest that the spatial organization of different categories of locations varies across different communities.

For the first community, we observe a notable co-location pattern within this community, where the categories of enterprises and shopping (markets and stores) tend to be grouped together (see Fig. 7**a**). This clustering suggests a shared proximity and potential interdependencies between these categories, implying a cohesive concentration of economic activities related to the movement and distribution of goods. The interaction test indicates that the locations categorized as shopping tend to exhibit direct interactions within this community (see Fig. 7**e**). This finding highlights the significance of shopping locations as important nodes for direct exchanges and transactions within this community. For the second community, we observe the tendency for the categories of companies and logistics to appear together in this community (see Fig. 7**b**), while the locations of logistics and terminals exhibit direct interactions (see Fig. 7**f**). Enterprises and logistics facilities have interdependent operations within the supply chains. By co-locating, companies can have direct access to logistics services and resources, leading to streamlined operations in this community. The direct interactions of logistics facilities and freight terminals enable efficiently managing the sorting, consolidation and transfer of goods to save transport costs. For

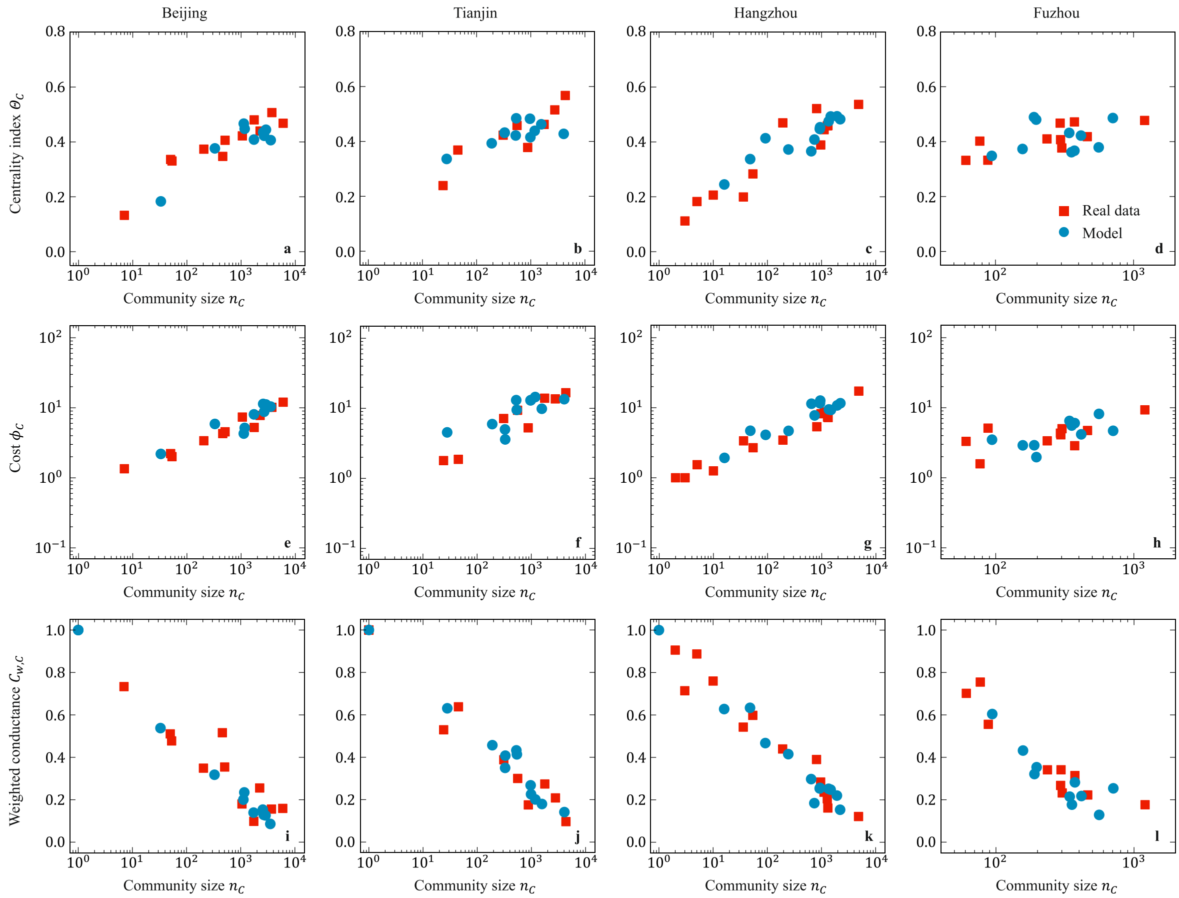

Fig. 6. Community spatial distribution characteristics observed from empirical networks of four cities and reproduced by the proposed model. **a-d** Distributions of centrality index Θ_C with respect to community size n_C. The legend "real data" represents the distributions of the metrics of the real networks, and each point represents each detected community. The legend "model" represents those of the model networks. **e-h** Distributions of cost ϕ_C with respect to n_C. **i-l** Distributions of weighted conductance $C_{w,C}$ with respect to n_C.

the third community, the co-occurrence of shopping and logistics categories indicates a close physical proximity (see Fig. 7**c**), suggesting the importance of logistics support for the efficient functioning of retail activities. Additionally, the direct interactions between logistics and enterprises (see Fig. 7**g**) highlight the collaboration and interdependence between logistics providers and businesses in managing transportation and distribution processes. For the fourth community, logistics holds a significant position within the community, acting as a connector between freight companies, markets and stores (see Fig. 7**dh**).

Overall, these findings provide insights into the spatial organization, interdependencies and functional relationships among industry-related locations within each community. They highlight the importance of co-location, direct interactions and collaborative efforts in facilitating efficient and coordinated operations within the urban heavy truck mobility networks.

5.2. The roles of individual freight locations

In this section, we uncover the hierarchical organization of heavy truck flows by using node-level metrics (see **Section 4.2**), and to understand the roles of individual freight locations. We first calculate the internal strength w_i^{int} and external strength w_i^{ext} of each node i, and the distributions of these two metrics with respect to node degree k_i are shown in Fig. 8. We can find the positive correlations between these two metrics and k_i, suggesting the hierarchical organization of heavy truck flows across freight locations with different connections. A freight location with high internal strength has many connections to other locations within the same community, indicating it is a major center of freight activity within its community. Similarly, a freight location with high external strength has many connections to locations outside of its community, indicating it serves as a major gateway for heavy truck flows to and from other communities. The freight locations with high both internal and external strength play a critical role in the movement of goods and materials, and are essential for the functioning of freight transport system. In addition, we can observe the variations in the slope of the distributions of these two metrics with respect to node degree across different cities, especially for big cities (e.g., Beijing) and small cities (e.g., Fuzhou). The higher slope of the distributions in a big city implies that there is a greater heterogeneity in the roles and influence of individual freight locations. High-connectivity nodes play a relatively more crucial role in terms of their internal and

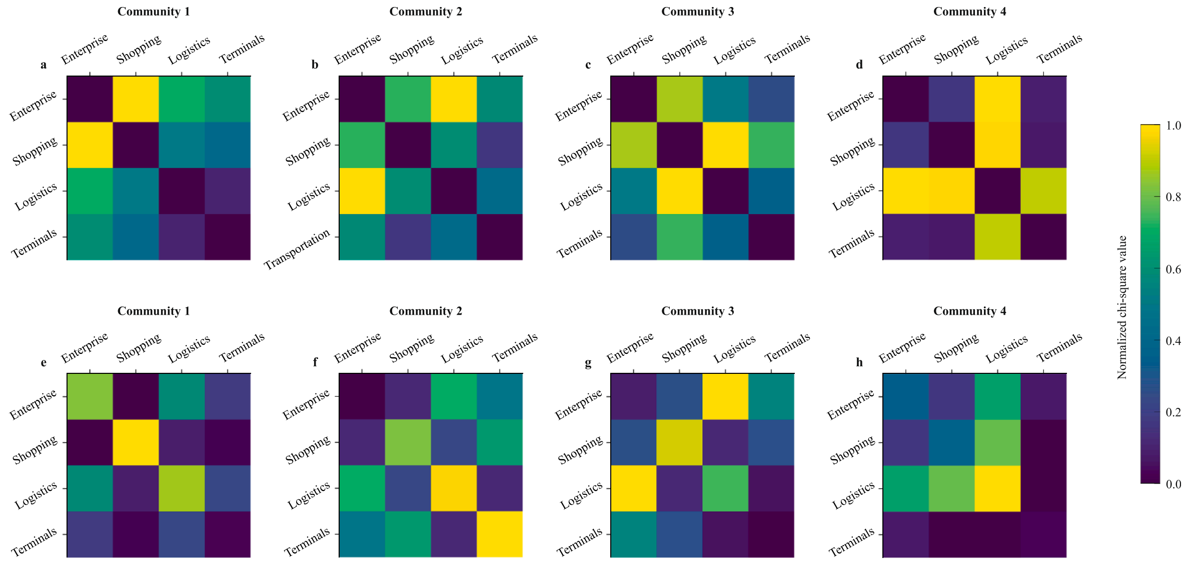

Fig. 7. The analysis of cluster-specific industry concentration for four typical communities in the heavy truck mobility network of Beijing. **a-d** Category co-occurrence test. There is a 4 × 4 matrix for all category pairs for a community. The colorbar represents the normalized chi-square value for each category pair in a community. **e-h** Category direct interaction test.

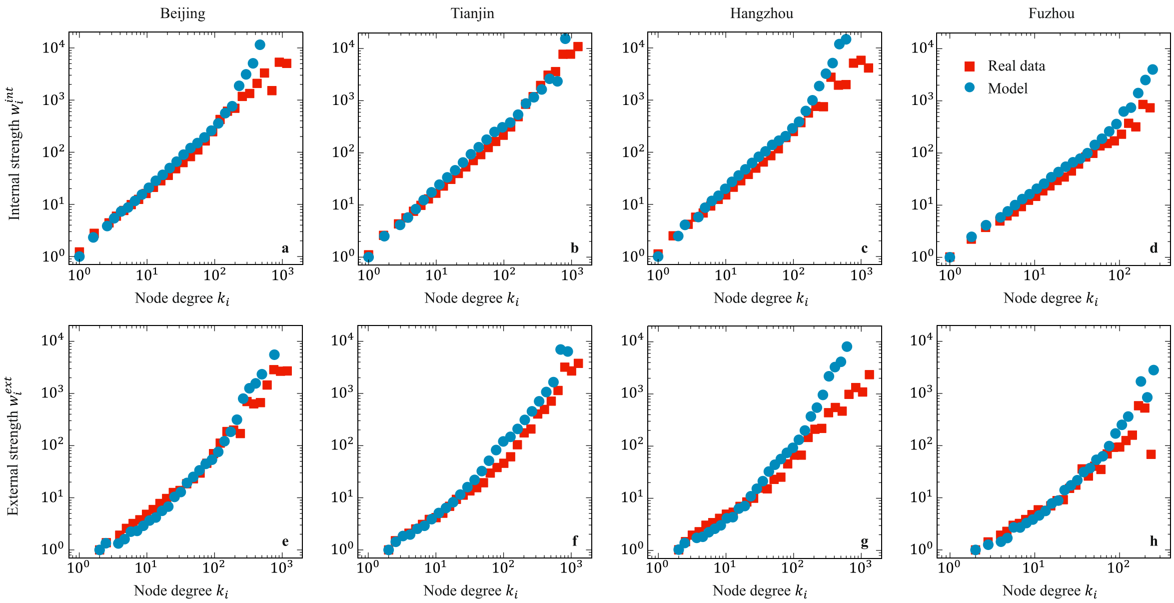

Fig. 8. Node strength characteristics observed from empirical networks of four cities and reproduced by the proposed model. **a-d** Distributions of internal strength w_i^{int} with respect to node degree k_i. The legend "real data" represents the distributions of the metrics of the real networks, and each point represents each node. The legend "model" represents those of the model networks. **e-h** Distributions of external strength w_i^{ext} with respect to k_i.

external strength. The results also suggest that the big city may have a more efficient and structured transportation system to support the key economic hubs that generate and attract a massive volume of truck flows from different regions in the city.

To further explore the roles of freight locations, we calculate the weighted embeddedness $\xi_{w,i}$, weighted within-module degree $Z_{w,i}$ and weighted participation coefficient $P_{w,i}$ of each node i, and obtain the distributions of them with respect to k_i (Fig. 9). We can find that the more connections a freight location has, the higher its weighted embeddedness (see Fig. 9**a-d**), suggesting that high-connectivity freight locations serve as community hubs as they have many strong connections to other freight locations within the same community. In contrast, low-connectivity freight locations tend to be peripheral players in the community, have a small number of weak connections with others. We can also derive similar findings from the distribution of weighted within-module degree $Z_{w,i}$ with respect to k_i in Fig. 9**e-h**. Moreover, Fig. 9**i-l** shows that weighted participation coefficient $P_{w,i}$ is positively correlated to k_i, suggesting

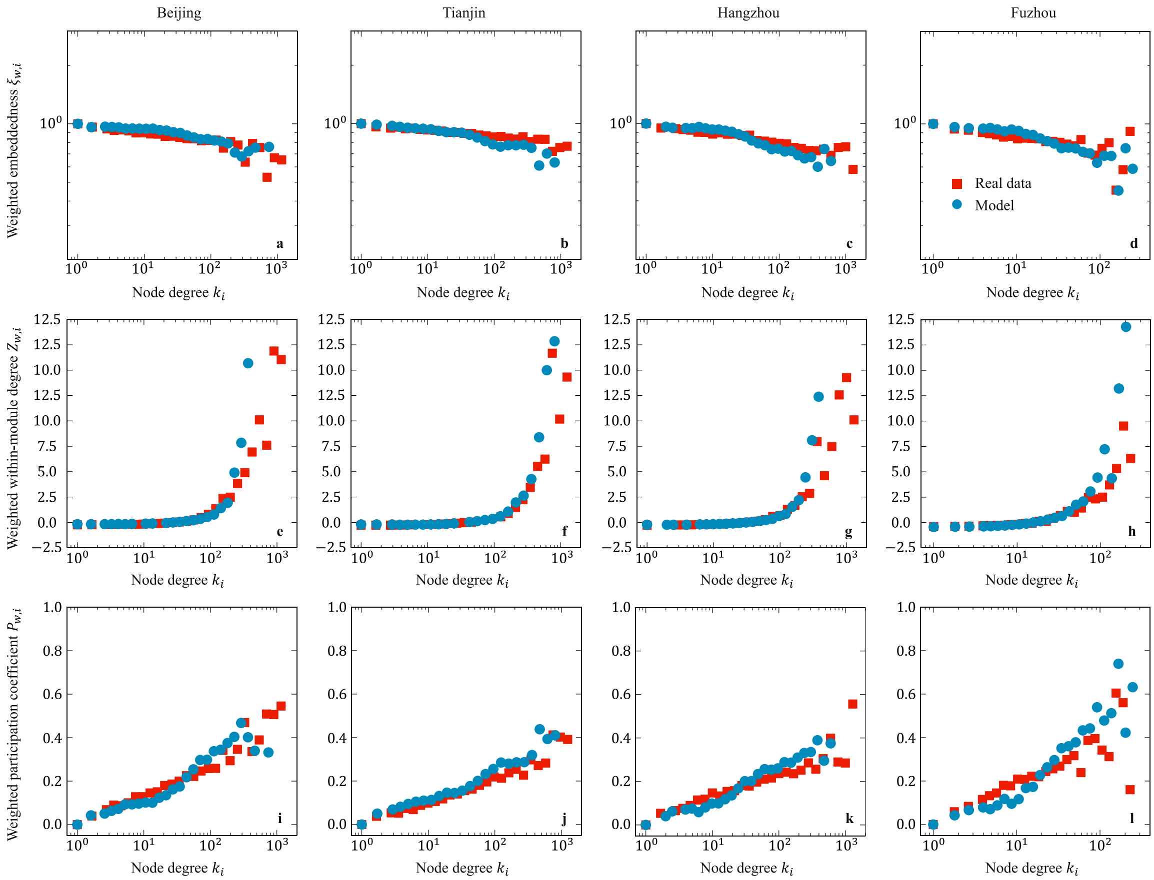

Fig. 9. Node role characteristics observed from empirical networks of four cities and reproduced by the proposed model. **a-d** Distributions of weighted embeddedness $\xi_{w,i}$ with respect to node degree k_i. The legend "real data" represents the distributions of the metrics of the real networks, and each point represents each node. The legend "model" represents those of the model networks. **e-h** Distributions of weighted within-module degree $Z_{w,i}$ with respect to k_i. **i-l** Distributions of weighted participation coefficient $P_{w,i}$ with respect to k_i.

that high-connectivity freight locations also serve as community connectors and they are connected to multiple communities in the network, however, low-connectivity freight locations mainly interact with others within the same community, playing a "peripheral" or "kinless" role. Comparing the roles played by nodes with different connectivity across the four cities, we can find that in larger cities like Beijing and Tianjin, high-connectivity nodes tend to have higher weighted within-module degrees. This suggests that the heavily connected nodes within these cities are more tightly interconnected within their respective clusters. It implies the presence of strong local connections and a higher level of specialization within these bigger cities. In contrast, in small cities like Fuzhou, the weighted within-module degree of high-connectivity nodes is lower, indicating a less dense local structure within the small city. On the other hand, the weighted participation coefficient of low-connectivity nodes is higher in small cities, indicating their importance in facilitating connections between different clusters. These nodes also play a crucial role in integrating different parts of the network and enabling efficient truck flows between various locations.

The above analysis reveals the correlation between the connectivity of locations and the roles they play. Next, we explore industry-specific location roles across different communities. For each detected community, we calculate the weighted within-module degree $Z_{w,i}$ and weighted participation coefficient $P_{w,i}$ for four different categories of nodes. The metric of $Z_{w,i}$ measures how 'well-connected' node i is to other nodes in the community and the metric of $P_{w,i}$ measures how 'well-distributed' the links of node i are among different communities. Fig. 10 shows the distributions of these two metrics for different categories of locations in four typical communities that are identical to those shown in Fig. 7. For ease of expression, we refer to the location with the highest $Z_{w,i}$ as a community hub and the location with the highest $P_{w,i}$ as a community connector. The results suggest that the roles of different categories of locations as community hubs and connectors vary significantly across different communities. For the first community (see Fig. 10**a**), a shopping node, such as market or store, plays a crucial role as community connector, facilitating connections and interactions between different communities. An enterprise node assumes the role of hub within its community, serving as a central point for freight consolidation, storage and distribution. For the second community (see Fig. 10**b**) and fourth community (see Fig. 10**d**), logistics nodes assume both community connector and hub. They ensure the seamless movement of goods between different communities and enable efficient coordination of logistics activities within their own community, facilitating the integration of supply chains and contribute to the

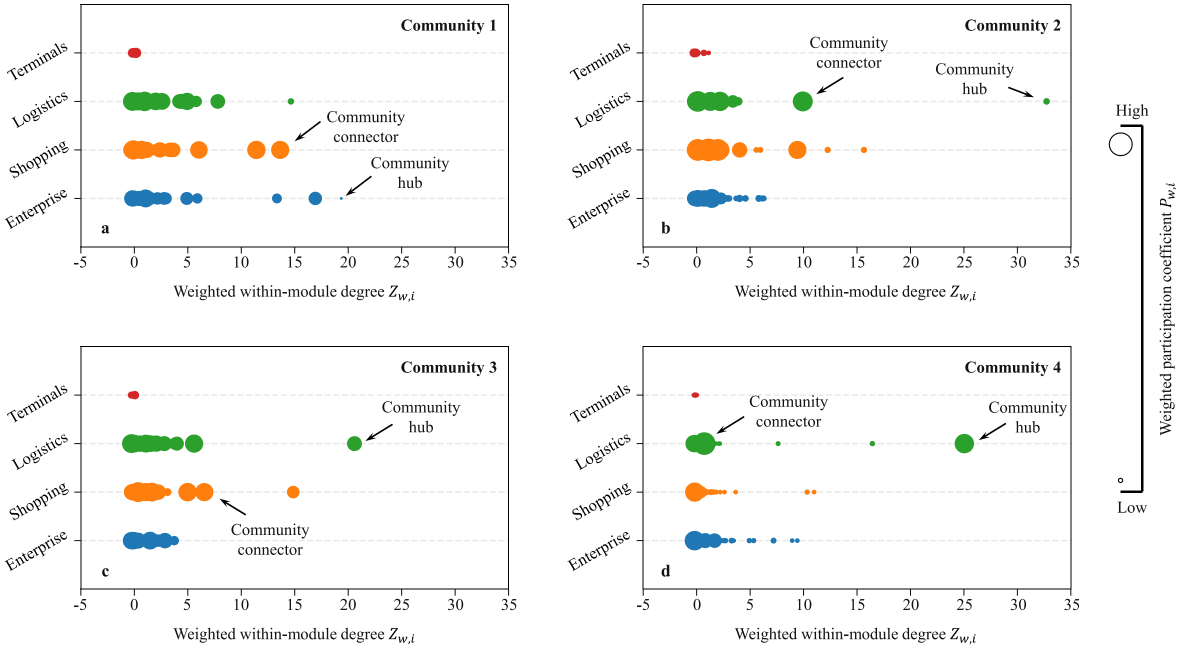

Fig. 10. The analysis of industry-specific location roles across four typical communities in the heavy truck mobility network of Beijing. A point indicates a location in the community. The horizontal axis represents the weighted within-module degree of locations and the vertical axis represents categories of locations. The size of each point represents weighted participation coefficient of each location. We refer to the location with the highest $Z_{w,i}$ as a community hub and the location with the highest $P_{w,i}$ as a community connector, as indicated in the figures.

overall efficiency of freight transportation. For the third community (see Fig. 10**c**), a shopping location acts as community connector, attracting trucks from various sources to supply the demands of locations within the community. A logistics node acts as community hub, receiving shipments from various origins, consolidate them and redistribute them to their final destinations in the community. These results suggest that the roles assumed by different categories of locations can vary greatly within different communities due to the specific characteristics, industry composition and logistical requirements of each community. By strategically locating community connectors and hubs, communities can optimize their logistics operations, reduce congestion, and improve overall transportation efficiency.

Taking together the above and the analysis, we can explore the whole picture of urban freight transport system, as shown in Fig. 11.

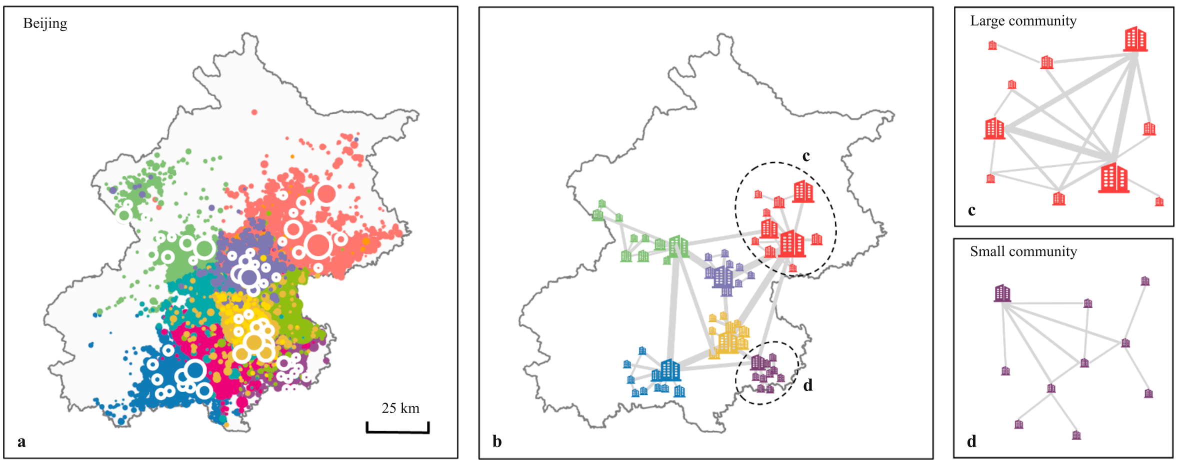

Fig. 11. Illustration of the community structure of empirical network of Beijing. **a** Detected communities and spatial distribution of typical high-connectivity nodes. Different communities are marked with different colors, and the nodes with white boundaries denote the top 10 nodes with the highest degree in each community. **b** Spatial interactions of typical nodes inside and between different communities. The width of solid line denotes interaction strength, i.e., heavy truck flows. The two communities marked with dotted lines correspond to panels **c** and **d**. **c** Spatial interactions between typical nodes in a large community. **d** Spatial interactions between typical nodes in a small community.

Spatially adjacent freight locations tend to cluster in the same community (see Fig. 11**a**). The larger a community, the higher heavy truck flows into and out of that community, and the interactions between two communities are mainly undertaken by high-connectivity freight locations, also known as community connectors or freight hubs (see Fig. 11**b**). In the large community, high-connectivity freight locations are closely linked to each other and interact with most of other locations (see Fig. 11**c**). However, in the small community, the interactions between freight locations tend to be characterized with tree-like structure, and low-connectivity freight locations play a "peripheral" or "kinless" role (see Fig. 11**d**). Heavy truck flows are heterogeneously distributed across communities with different sizes and freight locations with different connections. The roles of enterprise locations, shopping locations, logistics location and terminal locations vary across communities, highlighting their diverse functions and interdependencies. Understanding these industry-specific node roles and their interactions within communities is crucial for effective freight transportation planning and policy-making.

5.3. Model results and analysis

The above analysis uncovers the hierarchical organization of heavy truck flows across communities with different sizes and nodes with different connections, helping us understand the community structure of empirical urban heavy truck mobility networks. In this section, we explore the interaction dynamics of locations and explain how the structural properties of real networks are formed during the network growth by interpreting the model results. We first validate the performance of CRN model by evaluating how well it can reproduce the observed structure properties of real networks, and then provide explanations for their formation process based on the model mechanisms.

For the empirical heavy truck mobility network of a city, we estimate the optimal model parameter ρ^* (see **Section 4.4**) and generate the model network by the CRN model with the estimated ρ^*. Next, we detect the communities of the model network by using Infomap method, and calculate the community-level and node-level metrics of the detected community structure. The global characterization of model networks of four cities in China is shown in Table. 3, which suggests that the CRN model can reproduce not only the community structure of empirical networks (z-score of modularity $Z_Q{}' \gg 1$; Wallace index *WI* and normalized mutual information *NMI* are high), but also the spatial autocorrelation of freight locations (global Moran's I $GMI' > 0$ and its z-score $Z_{GMI}{}' \gg 1$). This is mainly attributed to the spatial agglomeration of industrial clusters and spatial interaction pattern of freight locations considered in the CRN model. In the network growth, new added freight location shares the similar partners of the freight locations in the closest existing industry cluster, i.e., a community, and establish new connections under the interaction pattern driven by spatial distance (associated with transport costs) and freight location competitiveness. Therefore, the CRN model reproduces the observed phenomenon that spatially adjacent freight locations tend to cluster in the same community, and provides an explanation for this.

Besides, the community and node strength characteristics of model networks are in excellent agreement with the empirical results. This is mainly attributed the initialization rule that the strength and degree of a new node obey the power-law relationship, and the weight update rule that the more competitiveness (measure by node strength) a freight location has, the more truck flows it will generate and attract. Therefore, the CRN model reproduces the positive correlations of both community internal strength w_C^{int} and external strength w_C^{ext} with respect to community size (see Fig. 5), and of both node internal strength w_i^{int} and external strength w_i^{ext} with respect to node degree (see Fig. 8). Additionally, we obtain the distributions of centrality index Θ_C (see Fig. 6**a-d**), cost ϕ_c (see Fig. 6**e-h**) and conductance $C_{w,C}$ (see Fig. 6**i-l**) with respect to community size, and they are again in excellent agreement with the indices from the data. This is mainly attributed to the spatial interaction pattern of freight locations considered in the CRN model, which assumes that freight locations tend to establish interactions with the closest and more competitive freight locations. In this way, the CRN model defines an agglomeration economy rule so that a few freight locations are dominant in their communities, and this is more pronounced in larger communities. Therefore, the CRN model can reproduce the community spatial distribution characteristics of empirical networks, as shown in Fig. 6. Similarly, under the agglomeration economy rule, high-connectivity freight locations serve as community hubs and connectors, and low-connectivity freight locations tend to be peripheral players in the community. Therefore, the CRN model can also reproduce the node role characteristics of empirical networks, as shown in Fig. 9.

To evaluate the performance of CRN model, we also compare the community structure of real network with that generated by CRN model and two benchmark models, i.e., SGN model and BBV model. We calculate two similarity metrics, i.e., Wallace index *WI* and normalized mutual information *NMI*, to assess whether node pairs located in the same community in the real network also clustered in the same community in the model network. The calculated *WI* and *NMI* for CRN model and two benchmark models are shown in Fig. 12**a-b**. The results suggest that the BBV model has the lowest performance due to the fact that this model does not consider the

Table 3
Global characterization of model networks for four cities in China. Estimated embedding coefficient ρ; number of network nodes N'; number of network links E'; number of communities NC'; modularity Q'; z-score of modularity $Z_Q{}'$; global Moran's I GMI'; z-score of global Moran's I $Z_{GMI}{}'$; Wallace index *WI*; normalized mutual information *NMI*.

City	ρ	N'	E'	C'	Q'	$Z_Q{}'$	GMI'	$Z_{GMI}{}'$	*WI*	*NMI*
Beijing	0.49	15,983	205,539	11	0.57	56	0.46	113	0.5697	0.5938
Tianjin	0.42	10,697	210,342	12	0.43	35	0.64	130	0.4957	0.4409
Hangzhou	0.39	10,611	158,817	13	0.50	33	0.56	118	0.6189	0.5990
Fuzhou	0.32	3391	26,751	10	0.34	42	0.70	99	0.5143	0.6198

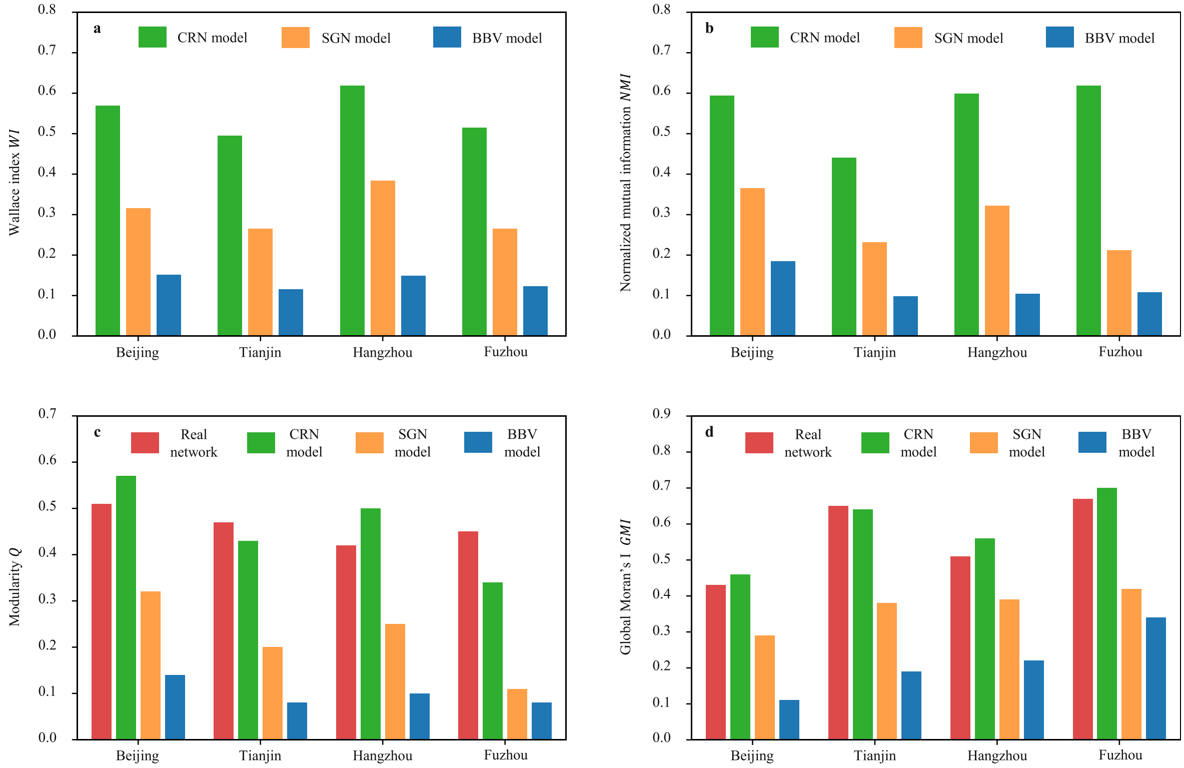

Fig. 12. Comparing the performance of proposed CRN model with two benchmark models.

geographical attributes and location clustering interactions. The SGN model outperforms the BBV model due to its integration of spatial effects on location interactions. However, the SGN model does not consider location clustering interactions, so it cannot explain the spatial agglomeration of industrial clusters. Therefore, the performance of SGN model is lower than that of proposed CRN model. Next, we calculate the metrics of modularity Q and global Moran's I *GMI* for real networks and model networks generated by CRN model and two benchmark models, as shown in Fig. 12**c-d**. The results suggest that the significance of network community structure generated by the CRN model is significantly higher than benchmark models. The above analysis highlights the necessity of integrating space effects and location clustering interactions in modeling urban freight networks, demonstrating the advantages of the proposed model.

To further elucidate the model mechanisms, we analyze the effects of the variation of the key parameter, i.e., embedding coefficient ρ, on the community structure of model networks. We take the empirical network of Fuzhou as the context for our analysis. We give the fixed initialization parameters fitted from real data, and set this key parameter to different values, i.e., $\rho = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. For each parameter value, we generate 100 model networks by the CRN model, and detect the communities of these model networks. Next, we calculate the community-level and node-level metrics of these model networks on average, and find the metrics significantly affected by the parameter ρ, as shown in Fig. 13. The results suggest that the larger the parameter ρ, the more closely the nodes in the community are connected to each other (see Fig. 13**a**) and the fewer the interactions between communities (Fig. 13**b**). Moreover, the

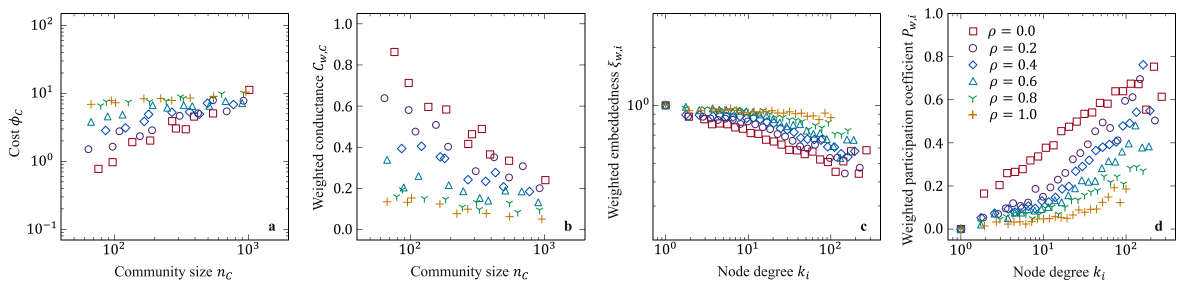

Fig. 13. Results of the model analysis for different values of ρ. **a** Distributions of cost ϕ_C with respect to community size n_C. **b** Distributions of weighted conductance $C_{w,C}$ with respect to community size n_C. **c** Distributions of weighted embeddedness $\xi_{w,i}$ with respect to node degree k_i. **d** Distributions of participation coefficient $P_{w,i}$ with respect to node degree k_i.

parameter ρ in the CRN model controls the roles of individual nodes, i.e., the larger the parameter ρ, the more nodes tend to establish interactions with the nodes in the same community (see Fig. 13**c**), while the less significant the role of community connector undertaken by high-connectivity nodes (see Fig. 13**d**). Taken together, we can adjust the embedding coefficient ρ to control the roles of nodes, and to shape the overall community structure of model networks. The CRN model captures the essential interaction dynamics of freight locations, and can reproduce a wide range of community structure characteristics.

5.4. Practical implications

5.4.1. Policy making

Understanding the community structure of urban heavy truck mobility networks can provide valuable insights for policymaker, and can inform the development of more effective policies. For example, detecting the network communities can help to identify regions that are isolated or disconnected from the rest of the network. These isolated regions may be at a disadvantage in terms of access to markets and may require targeted interventions to improve transportation connections (Ren et al., 2022). Additionally, key nodes, or freight hubs, that play a critical role in connecting different regions and facilitate heavy truck flows can also be identified. These hubs may be important targets for investment in transportation infrastructure, such as ports, airports and logistics warehouses, to improve the efficiency of transportation networks and reduce transportation costs (Wang et al., 2022). Moreover, identifying the communities that have higher heavy truck flows can help to guide investment in infrastructure and safety measures. This can help to reduce the negative impacts of heavy truck traffic on local communities and improve overall safety on the road (Alkhoori and Maghelal, 2021).

The analysis of cluster-specific community structure suggests that the spatial organization and system roles of different categories of locations varies across different communities, highlighting their diverse functions and interdependencies. These findings can provide specific industry-related implications. For example, the finding that different category pairs tend to be grouped together in different communities encourages industry co-location and proximity. Zoning and land use policies (Diriye et al., 2022; Gallagher et al., 2022) can be implemented to encourage mixed-use development, allowing enterprises, shopping locations and logistics facilities to coexist in the local areas. This can be achieved by designating specific zones or areas where these activities are allowed and promoting the development of infrastructure and services that support them. The economic zones or clusters that specialize in specific industries or sectors can also be created. These clusters can attract businesses related to the movement and distribution of goods, fostering collaboration, knowledge sharing and innovation within the industry. In addition, the finding that logistics nodes tend to assume community connectors and hubs inspires the development of policies that strengthen the logistics infrastructure (Cedillo-Campos et al., 2022; Netirith and Ji, 2022). One of the effective measures is to increase the investment in the development and improvement of logistics infrastructure, such as warehouses, distribution centers and intermodal facilities. These facilities should be strategically located to serve multiple communities efficiently. Enhancing infrastructure can help streamline operations, reduce transportation costs and improve supply chain efficiency.

The CRN model explains the interaction dynamics of freight locations, providing valuable policy implications for regulating urban freight transport system. One aspect of the interaction dynamics is embodied that freight locations tend to establish interactions with the closest and more competitive freight locations. This inspires the regulations and incentives to manage freight companies in a strategic way to reduce the needs for long-distance truck trips, so that the transportation costs can be reduced. This can be achieved by identifying and locating freight hubs in strategic locations, planning for co-location of freight facilities, optimizing routes and delivery schedules, and implementing consolidated delivery systems (Aljohani and Thompson, 2020). These strategies can help to improve the efficiency of the freight industry, while also creating cost savings for businesses. Another aspect of the interaction dynamics of freight locations is embodied that the locations in the same industrial cluster tend to share partners. This encourages the development of agglomeration economies by identifying business sectors that have similar production processes, supply chain requirements and market conditions (Abegaz and Nene, 2022). Government can play a key role in the development of industrial clusters by providing financial incentives, tax breaks, and other forms of supports to attract and retain businesses in the target business sectors (Yin et al., 2022). Overall, the interaction dynamics of freight locations explained by the CRN model help regulators develop targeted policies to promote the efficiency and sustainability of urban freight transport systems, and create a more livable urban environment for citizens.

In addition, we conduct a simulation analysis using Fuzhou as the context city to explore the effects of key parameter $\rho(\rho = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0)$ on the structure of model networks in **Section 5.3**. The simulation results suggest that parameter ρ controls the location roles, with larger values of ρ indicating a stronger tendency for locations to establish interactions within their own communities rather than acting as community connectors. This leads to a higher density of links within each community, resulting in an increased total transportation distance or cost within the communities. From another perspective, a higher density of links may also indicate a higher efficiency, due to the reduced possibility of cargo transshipment. Therefore, the simulation analysis highlights the implications for achieving a trade-off between efficiency and cost within industry clusters by leveraging the roles of important locations. Policymakers can explore strategies to influence location roles within communities. This can involve policies and incentives (Ali et al., 2014; van den Heuvel et al., 2013) to encourage industries and supply chains to concentrate within specific communities, promoting intra-community interactions. By strategically clustering industries and supply chains within specific communities, policymakers can promote proximity, collaboration and efficient resource utilization. This localization strategy can minimize the cost of transportation and logistics, enhancing community efficiency while managing costs.

5.4.2. Model application

The primary purpose of the CRN model is to provide a mathematical tool for simulating the formation process of urban freight networks. Such simulations can be used to explore various scenarios, test the impact of different policy interventions, and make informed decisions regarding urban freight transport infrastructure investments.

Firstly, The CRN model offers an effective tool for conducting simulations that allow stakeholders in urban freight transportation to explore a wide range of scenarios and policy interventions. These simulations can include varying parameters such as population growth, changes in industrial clusters, alterations in transportation infrastructure, and shifts in consumer demand. By adjusting these variables within the model, urban planners and policymakers can gain insights into how different factors impact the formation and evolution of freight networks. For instance, they can assess how the expansion of a specific industrial cluster might affect transportation needs or how improvements in transportation infrastructure can enhance network efficiency. This capability empowers decision-makers to make data-driven choices when considering policy interventions and investments.

Secondly, CRN model can be used to test the impact of various policy interventions on urban freight networks. Policymakers can use the model to simulate the outcomes of proposed changes in regulations, congestion pricing strategies, emission reduction initiatives, or shifts to more sustainable transportation modes. By observing the simulated effects on network structure, traffic patterns, and efficiency metrics, decision-makers can better understand the potential consequences of their policy choices before implementing them in the real world. This proactive approach minimizes the risk of unintended negative consequences and enables the development of well-informed policies that align with sustainability, economic, and logistical goals.

Thirdly, decision-makers can use the CRN model to evaluate different scenarios of infrastructure development (Sakai et al., 2020), including the construction of new transportation hubs, expansion of road networks, and deployment of advanced logistics technologies. By simulating the impact of these investments (Lin, 2020) on network performance and efficiency, stakeholders can make informed decisions about where and how to allocate resources. This ensures that infrastructure investments are strategically directed toward areas where they will have the greatest positive impact, enhancing the overall effectiveness and sustainability of urban freight transport systems.

In conclusion, the CRN model's simulation capabilities offer significant benefits for decision-makers in the field of urban freight transportation. It provides a versatile and adaptable framework for exploring scenarios, testing policy interventions, and optimizing infrastructure investments. By harnessing the insights generated through these simulations, stakeholders can make informed decisions that promote the efficiency, sustainability, and resilience of urban freight networks in a rapidly evolving urban landscape (Namatama, 2020).

6. Discussion and conclusion

The advent of the big data era provides us the opportunity to understand and regulate urban freight transport systems by uncovering and modeling the hierarchical organization of heavy truck flows from the perspective of complex networks. In this paper, we construct urban heavy truck mobility networks by using massive heavy truck GPS data and freight location POI data. We detect the communities of empirical networks by using the Infomap method and characterize their community structure by using the community-level and node-level metrics. We uncover the hierarchical organization of heavy truck flows across communities with different sizes, and across nodes with different connections. Additionally, these empirical results reveal the spatial distribution of freight location clusters and the roles of individual freight locations, helping us understand the structure the urban freight transport systems. Moreover, we develop a spatial network growth model, named CRN model, to explain the interaction dynamics of freight locations underlying the hierarchical organization of heavy truck flows.

The development of CRN model takes a bottom-up approach (Barthelemy, 2011; Zhai et al., 2019), starting from individual nodes and simulating their interactions over time. This approach allows for the emergence of network structures and interactions based on local decisions and behaviors. The CRN model is developed based on the assumptions derived from the phenomena of real-world freight systems, including spatial agglomeration of industrial clusters and spatial interaction pattern of locations. The performance of CRN model can be proven by evaluating how well it can reproduce the observed structure characteristics of real networks, and this process also validates the credibility of model assumptions. One of the purposes of CRN model is to explain and understand the underlying unobservable processes and mechanisms that drive the formation and evolution of real networks. For example, the CRN model delves into how freight locations interact and influence each other as the network evolves over time, with a focus on their spatial proximity and competition dynamics. By considering the concept of spatial agglomeration, the CRN model captures the tendency of specialized industries to cluster in close geographic proximity, forming industry-specific communities. The CRN model highlights how freight locations within a community tend to play similar roles, such as shared business partners, contributing to the local dynamics of the network. In addition, the CRN model explores how collective behaviors at the individual node level, combined with interactions within and between communities, give rise to emergent properties at the network level. Specifically, the CRN model accounts for the influence of node strength, which represents transportation demand, on the likelihood of establishing connections. It also incorporates the spatial interaction pattern, wherein freight locations are more likely to interact with those that are closer and more competitive, considering competitiveness in terms of generated and attracted heavy truck flows. This spatial interaction pattern mirrors real-world behavior and explains how network structure emerges from local interactions. The CRN model demonstrates how the spatial organization of freight locations, their competitive dynamics, and the formation of local communities collectively shape the structural properties of the evolving network.

Compared to top-down modeling approaches (Ghaffarinasab and Kara, 2022; Guo et al., 2022), CRN model has its own unique advantages in data-driven transportation planning. For example, in top-down hub location and traffic assignment, it involves analyzing existing traffic patterns, demand data, and infrastructure capacity to optimize the allocation of resources and improve network performance. The top-down models are valuable for optimizing existing networks based on predefined objectives and criteria, but are often less adaptable to changing conditions and future uncertainties. They optimize network performance based on historical data and predetermined structures, which may not account for emerging trends, evolving demand patterns or the impacts of new policies. In contrast, the bottom-up modeling approach (like the CRN model) can provide insights into how spatial interactions, location clustering and emergent behavior shape the network. It enables flexibility to assess the impacts of changes in infrastructure, industry distributions or policy interventions. This capability helps decision-makers explore different planning scenarios, optimize network performance and make informed choices to improve traffic management.

For the work applicability, one limitation lies in the simplifications made within our model. The model assumptions may not always align perfectly with the intricacies of real-world urban freight transport systems, which can vary significantly from one location to another. Therefore, our model may not capture the full complexity of every practical scenario. Moreover, the calculated indicators, while valuable for understanding network growth dynamics, may have limitations when applied directly to practical decision-making or policy implementation. Further validation and calibration of the model and indicators against empirical data from specific regions or scenarios would be essential to enhance its practical relevance and applicability.

Given the work in this paper, more meaningful future studies can be done. One of the promising future research direction is exploring the multi-resolution or overlapping communities (Fortunato, 2010) of urban heavy truck mobility networks. We can understand the structure of the urban freight system in greater depth by analyzing the communities from multi-resolution perspectives, i.e., "smaller communities in a big community", and by analyzing the communities with spatial overlap, i.e., "two communities contain some of the same nodes". Another promising future research direction is exploring the communities of freight multi-modal networks and developing an applicable network model. We can provide more reliable policy guidance by analyzing the more realistic urban freight network considering multi-modal shipping mode, including trains, planes and minivans. In addition, we can obtain the cargo-specific information of truck trips, e.g., whether the truck is loaded or empty, the types and weights of transported goods, by identifying the trip purposes (Gingerich et al., 2016) or by acquiring other secondary data. A directed network that captures the specific directional flows can be constructed to capture the true interaction relationships between locations. It enables a more detailed analysis of the key suppliers, consumers and the specific directionality of goods between locations. This refinement in the network analysis can provide valuable insights into the spatial dynamics and optimize decision-making processes related to logistics and transportation planning.

CRediT authorship contribution statement

Yitao Yang: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft. **Bin Jia:** Conceptualization, Supervision, Funding acquisition, Project administration, Writing – review & editing. **Xiao-Yong Yan:** Conceptualization, Supervision, Funding acquisition, Methodology, Writing – review & editing. **Danyue Zhi:** Methodology, Writing – review & editing. **Dongdong Song:** Methodology, Writing – review & editing. **Yan Chen:** Writing – review & editing. **Michiel de Bok:** Supervision, Writing – review & editing. **Lóránt A. Tavasszy:** Supervision, Writing – review & editing. **Ziyou Gao:** Funding acquisition, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under grant number 72288101, 72242102, 71971015 and 72271019, and the Central Leading Local Science and Technology Development Fund Project (236Z0805G).

References

Abegaz, M., Nene, G., 2022. Export agglomeration economies in Sub-Saharan Africa manufacturing and service sectors. Quarterly Review of Economics and Finance 84, 40–51.
Ali, A., Klasa, S., Yeung, E., 2014. Industry concentration and corporate disclosure policy. J. Account. Econ. 58 (2–3), 240–264.
Aljohani, K., Thompson, R.G., 2016. Impacts of logistics sprawl on the urban environment and logistics: Taxonomy and review of literature. J. Transp. Geogr. 57, 255–263.
Aljohani, K., Thompson, R.G., 2020. A multi-criteria spatial evaluation framework to optimise the siting of freight consolidation facilities in inner-city areas. Transp. Res. A Policy Pract. 138, 51–69.
Alkhoori, F.A., Maghelal, P.K., 2021. Regulating the overloading of heavy commercial Vehicles: Assessment of land transport operators in Abu Dhabi. Transp. Res. A Policy Pract. 154, 287–299.
Amer, A., Chow, J.Y.J., 2017. A downtown on-street parking model with urban truck delivery behavior. Transp. Res. A Policy Pract. 102, 51–67.
Anda, C., Medina, S.A.O., Axhausen, K.W., 2021. Synthesising digital twin travellers: Individual travel demand from aggregated mobile phone data. Transportation Research Part c: Emerging Technologies 128.
Arenas, A., Diaz-Guilera, A., Perez-Vicente, C.J., 2006. Synchronization reveals topological scales in complex networks. Phys. Rev. Lett. 96 (11), 114102.

Bachir, D., Khodabandelou, G., Gauthier, V., El Yacoubi, M., Puchinger, J., 2019. Inferring dynamic origin-destination flows by transport mode using mobile phone data. Transportation Research Part c: Emerging Technologies 101, 254–275.
Bai, X.W., Ma, Z.J., Zhou, Y.M., 2023. Data-driven static and dynamic resilience assessment of the global liner shipping network. Transportation Research Part e: Logistics and Transportation Review 170.
Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512.
Barrat, A., Barthelemy, M., Vespignani, A., 2004. Weighted evolving networks: Coupling topology and weight dynamics. Phys. Rev. Lett. 92 (22), 228701.
Barrat, A., Barthelemy, M., Vespignani, A., 2005. The effects of spatial constraints on the evolution of weighted complex networks. J. Stat. Mech: Theory Exp. P05003.
Barthelemy, M., 2011. Spatial networks. Physics Reports-Review Section of Physics Letters 499 (1–3), 1–101.
Bassolas, A., Barbosa-Filho, H., Dickinson, B., Dotiwalla, X., Eastham, P., Gallotti, R., Ghoshal, G., Gipson, B., Hazarie, S.A., Kautz, H., Kucuktunc, O., Lieber, A., Sadilek, A., Ramasco, J.J., 2019. Hierarchical organization of urban mobility and its connection with city livability. Nat. Commun. 10, 4817.
Bombelli, A., Santos, B.F., Tavasszy, L., 2020. Analysis of the air cargo transport network using a complex network theory perspective. Transportation Research Part e: Logistics and Transportation Review 138.
Brandes, U., Gaertler, M., Wagner, D., 2003. Experiments on graph clustering algorithms, In: DiBattista, G., Zwick, U. (Eds.), *Algorithms - Esa 2003, Proceedings*, pp. 568-579.
Cedillo-Campos, M.G., Pina-Barcenas, J., Mario Perez-Gonzalez, C., Mora-Vargas, J., 2022. How to measure and monitor the transportation infrastructure contribution to logistics value of supply chains? Transp. Policy 120, 120–129.
Chen, Y.M., Chen, X.Y., Liu, Z.H., Li, X., 2020. Understanding the spatial organization of urban functions based on co-location patterns mining: A comparative analysis for 25 Chinese cities. Cities 97.
Chen, H., Chen, B., Ai, C., Zhu, M., Qiu, X., 2022. The evolving network model with community size and distance preferences. Physica A 596.
Cheng, J., Lian, F., Yang, Z., 2022. The impacts of port governance reform on port competition in China. Transportation Research Part e: Logistics and Transportation Review 160, 102660.
Cheung, T.K., Wong, C.W., Zhang, A., 2020. The evolution of aviation network: Global airport connectivity index 2006–2016. Transportation Research Part e: Logistics and Transportation Review 133, 101826.
Chi, G., Thill, J.-C., Tong, D., Shi, L., Liu, Y., 2016. Uncovering regional characteristics from mobile phone data: A network science approach. Pap. Reg. Sci. 95 (3), 613-+.
Cidell, J., 2010. Concentration and decentralization: the new geography of freight distribution in US metropolitan areas. J. Transp. Geogr. 18 (3), 363–371.
Condon, A., Karp, R.M., 2001. Algorithms for graph partitioning on the planted partition model. Random Struct. Algoritm. 18 (2), 116–140.
Cong, H.B., Zou, D.L., 2017. The research on the mechanism and spatial-temporal differentiation of the coupling coordination development based on industrial cluster agglomeration. Cluster Computing-the Journal of Networks Software Tools and Applications 20 (1), 195–213.
Cui, J., Dodson, J., Hall, P.V., 2015. Planning for Urban Freight Transport: An Overview. Transport Reviews 35 (5), 583–598.
Dablanc, L., Giuliano, G., Holliday, K., O`Brien, T., 2013. Best Practices in Urban Freight Management: Lessons from an International Survey. Transp. Res. Record (2379), 29–38.
Danon, L., Diaz-Guilera, A., Duch, J., Arenas, A., 2005. Comparing community structure identification. J. Stat. Mech: Theory Exp. P09008.
de Bok, M., van Oort, F., 2011. Agglomeration economies, accessibility and the spatial choice behavior of relocating firms. J. Transp. Land Use 4 (1), 5–24.
de Oliveira, L.K., Lopes, G.P., de Oliveira, R.L.M., Bracarense, L.d.S.F.P., Pitombo, C.S.,, 2022. An investigation of contributing factors for warehouse location and the relationship between local attributes and explanatory variables of Warehouse Freight Trip Generation Model. Transp. Res. A Policy Pract. 162, 206–219.
Demissie, M.G., Kattan, L., 2022. Estimation of truck origin-destination flows using GPS data. Transportation Research Part e: Logistics and Transportation Review 159.
Dernir, E., Bektas, T., Laporte, G., 2014. A review of recent research on green road freight transportation. Eur. J. Oper. Res. 237 (3), 775–793.
Diriye, A.W., Jama, O.M., Chong, R., Abdi, A.M., 2022. Value of cultural worldviews and message framing for the acceptability of sustainable land use zoning policies in post-conflict Somalia. Journal of Environmental Planning and Management 65 (14), 2587–2608.
Ferrari, P., 2014. The dynamics of modal split for freight transport. Transportation Research Part e: Logistics and Transportation Review 70, 163–176.
Ferrari, P., 2015. Dynamic cost functions and freight transport modal split evolution. Transportation Research Part e: Logistics and Transportation Review 77, 115–134.
Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486 (3–5), 75–174.
Fortunato, S., Hric, D., 2016. Community detection in networks: A user guide. Phys. Rep. 659, 1–44.
Gallagher, R., Sigler, T., Liu, Y., 2022. Urban "Blandscapes": How the Practical Implementation of Planning Policy Reduces Land Use Diversity. Urban Policy Res. 41 (3), 295–313.
Ghaffarinasab, N., Kara, B.Y., 2022. A conditional beta-mean approach to risk-averse stochastic multiple allocation hub location problems. Transportation Research Part e: Logistics and Transportation Review 158.
Ghanei, S., Contreras, I., Cordeau, J.-F., 2023. A two-stage stochastic collaborative intertwined supply network design problem under multiple disruptions. Transportation Research Part e: Logistics and Transportation Review 170, 102944.
Gingerich, K., Maoh, H., Anderson, W., 2016. Classifying the purpose of stopped truck events: An application of entropy to GPS data. Transportation Research Part c: Emerging Technologies 64, 17–27.
Grobar, L.M., 2008. The economic status of areas surrounding major US container ports: evidence and policy issues. Growth Chang. 39 (3), 497–516.
Guimera, R., Amaral, L.A.N., 2005. Functional cartography of complex metabolic networks. Nature 433 (7028), 895–900.
Guo, Z., Zhao, P., Senousi, A.M., Liu, X., Mansourian, A., 2023. Exploring the structural characteristics of intra-urban shared freight network and their associations with socioeconomic status. Travel Behav, Soc, p. 32.
Guo, L., Zheng, J., Du, H., Du, J., Zhu, Z., 2022. The berth assignment and allocation problem considering cooperative liner carriers. Transportation Research Part e: Logistics and Transportation Review 164.
Hanaki, N., Peterhansl, A., Dodds, P.S., Watts, D.J., 2007. Cooperation in evolving social networks. Manag. Sci. 53 (7), 1036–1050.
Henry, E., Furno, A., Faouzi, N.-E.-E., Rey, D., 2022. Locating park-and-ride facilities for resilient on-demand urban mobility. Transportation Research Part e: Logistics and Transportation Review 158, 102557.
Huang, Z., Wang, P., Zhang, F., Gao, J., Schich, M., 2018. A mobility network approach to identify and anticipate large crowd gatherings. Transp. Res. B Methodol. 114, 147–170.
Jacobs, W., Koster, H., Hall, P., 2011. The Location and Global Network Structure of Maritime Advanced Producer Services. Urban Stud. 48 (13), 2749–2769.
Jote, N., Beshah, B., Kitaw, D., Mangano, G., De Marco, A., 2013. A review on the integration of supply chain management and industrial cluster. International Journal of Marketing Studies 5 (6).
Kale, R., Evers, P.T., Dresner, M.E., 2007. Analyzing private communities on Internet-based collaborative transportation networks. Transportation Research Part e: Logistics and Transportation Review 43 (1), 21–38.
Kossinets, G., Watts, D.J., 2006. Empirical analysis of an evolving social network. Science 311 (5757), 88–90.
Krutein, K.F., Goodchild, A., 2022. The isolated community evacuation problem with mixed integer programming. Transportation Research Part e: Logistics and Transportation Review 161.
Lance, G.N., Williams, W.T., 1966. Computer programs for hierarchical polythetic classification ("similarity analyses"). Comput. J. 9 (1), 60–64.
Lancichinetti, A., Fortunato, S., 2011. Limits of modularity maximization in community detection. Physical Review E 84 (6).
Lancichinetti, A., Fortunato, S., Radicchi, F., 2008. Benchmark graphs for testing community detection algorithms. Phys. Rev. E 78 (4), 046110.
Li, C., Chen, G., 2006. Modelling of weighted evolving networks with community structures. Physica A 370 (2), 869–876.
Li, R., Dong, L., Zhang, J., Wang, X., Wang, W.-X., Di, Z., Stanley, H.E., 2017. Simple spatial scaling rules behind complex cities. *Nature*. Communications 8.

Lim, K.G., Nomikos, N.K., Yap, N., 2019. Understanding the fundamentals of freight markets volatility. Transportation Research Part e: Logistics and Transportation Review 130, 1–15.
Lin, X., 2020. Multiple pathways of transportation investment to promote economic growth in China: a structural equation modeling perspective. Transportation Letters-the International Journal of Transportation Research 12 (7), 471–482.
Liu, X., Gong, L., Gong, Y.X., Liu, Y., 2015. Revealing travel patterns and city structure with taxi trip data. J. Transp. Geogr. 43, 78–90.
Louail, T., Lenormand, M., Picornell, M., Cantu, O.G., Herranz, R., Frias-Martinez, E., Ramasco, J.J., Barthelemy, M., 2015. Uncovering the spatial structure of mobility networks. *Nature*. Communications 6.
Louf, R., Jensen, P., Barthelemy, M., 2013. Emergence of hierarchy in cost-driven growth of spatial networks. *Proceedings of the National Academy of Sciences of the United States of America* 110(22), 8824-8829.
MacKay, D.J., 2003. Information theory, inference and learning algorithms. Cambridge University Press.
Mesa-Arango, R., Ukkusuri, S.V., 2015. Demand clustering in freight logistics networks. Transportation Research Part e: Logistics and Transportation Review 81, 36–51.
Mohammed, R.A., Nadi, A., Tavasszy, L., de Bok, M., 2023. Data Fusion Approach to Identify Distribution Chain Segments in Freight Shipment Databases. Transp. Res. Rec.
Moraes Pereira, R.H., Nadalin, V., Monasterio, L., Albuquerque, P.H.M., 2013. Urban Centrality: A Simple Index. Geogr. Anal. 45 (1), 77–89.
Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37 (1–2), 17–23.
Mori, T., Smith, T.E., 2015. On the spatial scale of industrial agglomerations. J. Urban Econ. 89, 1–20.
Mullen, C., Marsden, G., 2015. Transport, economic competitiveness and competition: A city perspective. Journal of Transport Geography 49, 1–8.
Murali, P., Ordonez, F., Dessouky, M.M., 2016. Modeling strategies for effectively routing freight trains through complex networks. Transportation Research Part c: Emerging Technologies 70, 197–213.
Namatama, N., 2020. An assessment of stakeholders' participation in land use planning process of Luapula Province Planning Authority. Land Use Policy 97, 104735.
Nassar, R.F., Ghisolfi, V., Annema, J.A., van Binsbergen, A., Tavasszy, L.A., 2023. A system dynamics model for analyzing modal shift policies towards decarbonization in freight transportation. Res. Transp. Bus. Manag. 100966.
Netirith, N., Ji, M., 2022. Analysis of the Efficiency of Transport Infrastructure Connectivity and Trade. Sustainability 14 (15), 9613.
Newman, M.E.J., 2004. Analysis of weighted networks. Phys. Rev. E 70 (5), 056131.
Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. Phys. Rev. E 69 (2), 026113.
Nguyen, T.V., Zhang, J., Zhou, L., Meng, M., He, Y., 2020. A data-driven optimization of large-scale dry port location using the hybrid approach of data mining and complex network theory. Transportation Research Part e: Logistics and Transportation Review 134.
Ouyang, Z.Y., Leung, E.K.H., Huang, G.Q., 2022. Community logistics for dynamic vehicle dispatching: The effects of community departure "time" and "space". Transportation Research Part e: Logistics and Transportation Review 165.
Pamucar, D., Deveci, M., Gokasar, I., Martinez, L., Koppen, M., 2022. Prioritizing transport planning strategies for freight companies towards zero carbon emission using ordinal priority approach. Comput. Ind. Eng. 169.
Pearson, K., 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Phil. Mag. 50, 157–175.
Ren, Y., Ercsey-Ravasz, M., Wang, P., Gonzalez, M.C., Toroczkai, Z., 2014. Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. *Nature*. Communications 5.
Ren, Y., Tian, Y., Xiao, X., 2022. Spatial effects of transportation infrastructure on the development of urban agglomeration integration: Evidence from the Yangtze River Economic Belt. J. Transp. Geogr. 104, 103431.
Rives, A.W., Galitski, T., 2003. Modular organization of cellular networks. *Proceedings of the National Academy of Sciences of the United States of America* 100(3), 1128-1133.
Rodrigue, J.-P., 2020. The geography of transport systems, 5th ed. Routledge.
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.
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.
Sakai, T., Beziat, A., Heitz, A., 2020. Location factors for logistics facilities: Location choice modeling considering activity categories. J. Transp. Geogr. 85, 102710.
Sala, A., Cao, L., Wilson, C., Zablit, R., Zheng, H., Zhao, B.Y., 2010. Measurement-calibrated graph models for social network experiments, *Proceedings of the 19th international conference on World wide web*. Association for Computing Machinery, Raleigh, North Carolina, USA, pp. 861–870.
Sala, L., Wright, S., Cottrill, C., Flores-Sola, E., 2021. Generating demand responsive bus routes from social network data analysis. Transportation Research Part c: Emerging Technologies 128, 103194.
Shakib, M.D., 2020. Using system dynamics to evaluate policies for industrial clusters development. Computers & Industrial Engineering 147, 106637.
Simini, F., Gonzalez, M.C., Maritan, A., Barabasi, A.-L., 2012. A universal model for mobility and migration patterns. Nature 484 (7392), 96–100.
Siripirote, T., Sumalee, A., Ho, H.W., 2020. Statistical estimation of freight activity analytics from Global Positioning System data of trucks. Transportation Research Part e: Logistics and Transportation Review 140, 101986.
Tavasszy, L., De Jong, G., 2013. Modelling freight transport. Elsevier.
van den Heuvel, F.P., de Langen, P.W., van Donselaar, K.H., Fransoo, J.C., 2013. Regional logistics land allocation policies: Stimulating spatial concentration of logistics firms. Transp. Policy 30, 275–282.
Wallace, D.L., 1983. A method for comparing two hierarchical clusterings: comment. J. Am. Stat. Assoc. 78 (383), 569–576.
Wang, Y., Liu, H., Fan, Y., Ding, J., Long, J., 2022. Large-scale multimodal transportation network models and algorithms-Part II: Network capacity and network design problem. Transportation Research Part e: Logistics and Transportation Review 167, 102918.
Wang, Y., Li, Y., Huang, Y., Gong, D., 2023. Analyzing the impacts of logistics suburbanization on logistics service accessibility: Accessibility modeling approach for urban freight. Transp. Policy.
Wang, S., Yan, R., 2023. Fundamental challenge and solution methods in prescriptive analytics for freight transportation. Transportation Research Part e: Logistics and Transportation Review 169.
Watts, D.J., 2004. Small worlds: the dynamics of networks between order and randomness. Princeton University Press.
Wu, Y., Wei, Y.D., Li, H., Liu, M., 2022. Amenity, firm agglomeration, and local creativity of producer services in Shanghai. Cities 120, 103421.
Xia, F., Wang, J.Z., Kong, X.J., Zhang, D., Wang, Z.B., 2020. Ranking Station Importance With Human Mobility Patterns Using Subway Network Datasets. IEEE Trans. Intell. Transp. Syst. 21 (7), 2840–2852.
Yan, X.-Y., Wang, W.-X., Gao, Z.-Y., Lai, Y.-C., 2017. Universal model of individual and population mobility on diverse spatial scales. *Nature*. Communications 8.
Yang, Y., Jia, B., Liu, E., Yan, X.-Y., de Bok, M., Tavasszy, L.A., Gao, Z., 2022a. Structure and evolution of urban heavy truck mobility networks. *arXiv preprint arXiv: 2212.03672*.
Yang, Y., Jia, B., Yan, X.-Y., Jiang, R., Ji, H., Gao, Z., 2022b. Identifying intracity freight trip ends from heavy truck GPS trajectories. Transportation Research Part c: Emerging Technologies 136.
Yang, Y., Jia, B., Yan, X.-Y., Li, J., Yang, Z., Gao, Z., 2022c. Identifying intercity freight trip ends of heavy trucks from GPS data. Transp. Res. Part e: Logistics and Transportation Review 157, 102590.
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.
Yin, Y., Yan, M., Zhan, Q., 2022. Crossing the valley of death: Network structure, government subsidies and innovation diffusion of industrial clusters. Technol. Soc. 71, 102119.

Zhai, W., Bai, X., Peng, Z.-R., Gu, C., 2019. A bottom-up transportation network efficiency measuring approach: A case study of taxi efficiency in New York City. J. Transp. Geogr. 80.
Zhang, J., Liu, F., Tang, J.F., Li, Y.H., 2019. The online integrated order picking and delivery considering Pickers' learning effects for an O2O community supermarket. Transportation Research Part e: Logistics and Transportation Review 123, 180–199.
Zhang, W., Thill, J.-C., 2017. Detecting and visualizing cohesive activity-travel patterns: A network analysis approach. Comput. Environ. Urban Syst. 66, 117–129.
Zhao, D., Balusu, S.K., Sheela, P.V., Li, X., Pinjari, A.R., Eluru, N., 2020. Weight-categorized truck flow estimation: A data-fusion approach and a Florida case study. Transportation Research Part e: Logistics and Transportation Review 136.
Zheng, J.F., Qi, J.W., Sun, Z., Li, F., 2018. Community structure based global hub location problem in liner shipping. Transportation Research Part e: Logistics and Transportation Review 118, 1–19.