

Research Article

Identifying the Key Nodes and Sections of Urban Roadway Network Based on GPS Trajectory Data

Jing Wang ¹, Chunjiao Dong ¹, Chunfu Shao ¹, Shichen Huang ¹ and Shuang Wang ²

¹Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, China

²School of Traffic and Transportation, Beijing Jiaotong University, 3 Shangyuan Cun, Haidian District, Beijing 100044, China

Correspondence should be addressed to Chunjiao Dong; cjdong@bjtu.edu.cn

Received 18 October 2020; Revised 20 February 2021; Accepted 3 March 2021; Published 16 March 2021

Academic Editor: Bin Yu

Copyright © 2021 Jing Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a novel approach to identify the key nodes and sections of the roadway network. The taxi-GPS trajectory data are regarded as mobile sensor to probe a large scale of urban traffic flows in real time. First, the urban primary roadway network model and dual roadway network model are developed, respectively, based on the weighted complex network. Second, an evaluation system of the key nodes and sections is developed from the aspects of dynamic traffic attributes and static topology. At the end, the taxi-GPS data collected in Xicheng District of Beijing, China, are analyzed. A comprehensive analysis of the spatial-temporal changes of the key nodes and sections is performed. Moreover, the repetition rate is used to evaluate the performance of the identification algorithm of key nodes and sections. The results show that the proposed method realizes the expression of topological structure and dynamic traffic attributes of the roadway network simultaneously, which is more practicable and effective in a large scale.

1. Introduction

An urban roadway network is composed of multiple intersections and roadway sections. The studies show that the importance of each intersection and roadway section in the urban roadway network is different, and large-scale congestion of the roadway network is often caused by the congestion of several key intersections and sections. If the top 5% important intersections and sections in the roadway network are attacked, the roadway network will be paralyzed [1–3]. Therefore, if the supporting and vulnerable nodes and sections can be identified, traffic planners and traffic managers can alleviate the traffic pressure by reasonably planning the topological structure of the urban roadway network. Besides, the effectively protection and scientific management can be carried out to improve the survivability and reliability of the urban roadway network, thereby avoiding the occurrence of large-scale traffic congestion.

In recent years, with the maturity of GPS technology, GPS data provide a new data source for urban traffic research. Based on the deep learning theory, Ma et al. [1]

predicted the evolution of traffic congestion with taxi-GPS data. Based on the taxi-GPS data, Kong et al. [2] constructed a method combining support vector machine and the fuzzy comprehensive evaluation model, which realized the identification and prediction of traffic congestion. Feng et al. [3] proposed an identification method of critical roads based on the combination of GPS trajectory data and the directed weighted complex network. Compared with the traditional data, such as video monitoring data and the data collected by the loop coil vehicle sensor, GPS trajectory data have the advantages of easy process, easy access, high quality, and low cost [4]. The GPS trajectory data can be automatically collected and sent to the data center in real time, and the high mobility of GPS trajectory data makes it possible to monitor the large-area roadway network. Among the vehicles that equipped with GPS devices, taxis, the random sample, are distributed in every corner of the city [5]. In other words, taxi-GPS trajectory data are high-quality data for the study of urban traffic problems. Based on this, the paper proposes a new method using taxi-GPS trajectory data to identify the key nodes and sections of the roadway

network, which takes both topological structure and dynamic traffic attributes into consideration. The results can provide theoretical support for traffic managers and traffic planners to help (1) guide the uniform distribution of traffic flow, (2) formulate traffic management policies such as the planning of tidal lanes and off-peak travel policies, and (3) construct the emergency rescue plan for key nodes and sections to avoid the traffic paralysis of the city.

The rest of this paper is organized as follows. In Section 2, we conduct a comprehensive literature review about identifying the key nodes and sections of the roadway network. Then, the proposed identification method is presented in Section 3. In Section 4, the experiment results are described. Finally, the paper is concluded in Section 5.

2. Literature Review

With the increasingly serious of traffic congestion, the research of key nodes is extended to identify the key intersections in the roadway network. Hawick and James [6] used some quantifiable metrics of the complex network such as Dijkstra distance, mean degree, and Euclidean centroid to identify the key nodes of the roadway network. Zhang et al. [7] proposed a traffic shockwave model using the speed of traffic congestion spreading out to evaluate the importance of nodes in the roadway network. Jayaweera [8] proposed a new centrality measure called DelayFlow incorporating travel time delay and commuter flow volume to identify the key nodes in the urban roadway network, and with the experiment on Singapore's subway network, the method was proved more relevant to the urban roadway network than the network topology-based centrality measures. On the basis of the complex network model, the authors in [8] used centrality and betweenness measures to evaluate the importance of nodes which directly influenced the traffic congestion in roadway networks in Sri Lanka. Taking both node centrality and traffic flow into consideration, Xu et al. [9] introduced a data-driven framework to identify the key nodes with the use of comprehensive vehicle trajectories and geographic information. The authors in [10] proposed a method named the improved topological potential model considering entropy, which took both centrality and invulnerability into consideration to evaluate the key nodes of the metro network.

Another main research of roadway network is to identify the key roadway sections. Scott et al. [11] proposed a method that considered network flows, link capacity, and network topology to identify the critical sections, and with the results measured by travel time savings, the method is proved yield far greater system-wide benefits than the V/C ratio (under ideal conditions, the ratio of the maximum service traffic volume to the basic capacity). Sullivan et al. [12] used different link-based capacity-disruption values to rank the critical sections in the roadway network, and the research showed that the rank-ordering of the most critical sections in a network can vary dramatically based on both the capacity-disruption level and the overall connectivity of the network. Luathep et al. [13] adopted the relative accessibility index (AI) following the Hansen integral index to estimate the

importance of sections in the roadway network, and the key sections were ranked according to the differences in the AIs between normal and degraded networks. Rupi et al. [14] suggested that the importance of the roadway section was linked to two aspects: the level of usage, and the impact that the closure of the link itself can have on the general functionality of the network as a whole. Kumar et al. [15] used three factors to rank the sections in the roadway network: the link flows at equilibrium, the importance of facilities served, and the number of origin-destination pairs served. Zhang and Khani [16] adopted the Lagrangian relaxation method to identify critical sections by maximizing network accessibility with limited travel time budget and roadway construction cost budget. Li et al. [17] proposed an approach using the traffic flow betweenness index (TFBI) to identify critical links, which can significantly reduce the computational burden compared with the traditional full-scan method.

The urban roadway network is composed of static topological structure and dynamic traffic attributes [3]. The abovementioned methods most have only considered one of the two aspects, and there is a lack of a systematically analysis of both static topology and dynamic traffic attributes. Besides, in the existing literature, the undirected complex network model is often used to model the urban roadway network. However, in practice, the vehicles often run in two directions, analyzing the differences between two directions is of great significance to understand the tidal phenomenon of urban traffic. To tackle this issue, this paper proposes a directed weighted complex network model which considers both dynamic traffic attributes and static topology to identify key nodes and sections of the roadway network.

3. Methodologies

3.1. Modelling Urban Roadway Network. There are two approaches to construct the urban roadway network model using the complex network: the primary network and the dual network. In the primary network, the intersection is abstracted as node and the roadway section is abstracted as edge. In the dual network, the intersection corresponds to the edge, and the roadway section corresponds to the node. The primary network has certain limitations in describing the dynamic traffic attributes of the roadway network, such as more complex to reflect the key sections of the urban roadway network [18]. Nevertheless, the primary network is the preferred method for researchers because of the simplicity and directness. The dual network can describe the spatial connection relationship of the roadway sections in detail [19]. Therefore, this paper selects the primary method to identify the key nodes and the dual network to identify the key sections. The way to construct the urban roadway network model can be illustrated as follows.

3.1.1. The Primary Roadway Network Model. Assume that there are n intersections in the urban roadway network. The directed weighted complex network model can be defined as follows: $G = (V, E, R)$, where $V = \{v_1, v_2, \dots, v_i, \dots, v_n\}$ is

the set of nodes. $E = \{e_{11}, e_{12}, \dots, e_{ij}, \dots, e_{ji}, \dots, e_{mm}\}$ is the set of edges that is formed in order by pairs of nodes. $e_{ij} = (v_i, v_j)$ represents the directed edge which starts at v_i and ends at v_j . $R = \{r_1, r_2, \dots, r_i, \dots, r_n\}$ is the weight set of nodes, and r_i is the weight of node v_i . Figures 1(a) and 1(b) illustrate the process to build urban roadway network model with a primary method [20]. Generally, the weight of the directed complex network can reflect the dynamic traffic attributes of the roadway network, such as grade of roadway sections, traffic capacity, and traffic speed.

3.1.2. The Dual Roadway Network Model. The dual roadway network model can be defined as follows: $G^{\text{dual}} = (V^{\text{dual}}, E^{\text{dual}}, W)$. V^{dual} is the set of nodes in the dual network corresponding to the edges in the primary network, E^{dual} is the set of edges in the dual network corresponding to the nodes in the primary network, and $W = \{w_{e_{ij}} | e_{ij} \in E\}$ is the set of weights of the nodes in the dual network; $w_{e_{ij}}$ represents the weight of edge e_{ij} in the primary network. The relationship between R and W is shown as

$$r_i = \sum_{j \in N_i} w_{e_{ji}} + \sum_{j \in N_i} w_{e_{ij}}, \quad (1)$$

where N_i is the set of adjacent nodes of node v_i .

When constructing the dual network, the corresponding link will not be created if there is a turn prohibition between v_i and v_j . The set of possible turns T can be defined as

$$T = \{(s, d, x) | (s, d, x) \in V, (s, d) \in E, (d, x) \in E\}. \quad (2)$$

The turn prohibition $P \subset T$. Therefore, the edge set E^{dual} of the dual network can be defined as

$$E^{\text{dual}} = \{(f, g, Q^{fg}) | (f, g) \in V^{\text{dual}}, f = (s, d), g = (d, x), (s, d, x) \notin P\}. \quad (3)$$

The expression is shown in Figures 1(c) and 1(d). $(f, g) \in V^{\text{dual}}$ means the edges of the primary network corresponding to the nodes of the dual network. The dual link (f, g, Q^{fg}) is established when the following three conditions are satisfied [20, 21]:

- (1) The original node of the dual network is the edge of the primary network (s, d)
- (2) The destination node of the dual network is the edge of the primary network (d, x)
- (3) There is no turn prohibition between edge (s, d) and (d, x)

3.2. Evaluation Index System of Key Nodes and Sections. This paper considers both static topology and dynamic traffic attributes when identifying key nodes and sections of the urban roadway network. For dynamic traffic attributes, we define two indexes, the level of traffic congestion and the grade of node, to build the node weight of the roadway network model, and for static topology, we construct three

matrices based on node efficiency and shortest path. Figure 2 shows the evaluation index system of key nodes and sections.

The definition of each index is introduced as follows.

3.2.1. Dynamic Traffic Attributes. In actual roadway network, the expressway is more important than the arterial roadway, the minor arterial roadway, and the local roadway. Therefore, different grades of roadway sections should have different weights. The grade of node of the dual network can be defined as S , and Table 1 [22] shows the value of S .

The nodes and sections which often occur congestion are more important in the roadway network. According to the Urban Road Traffic Performance Index issued by Beijing, China, in 2011, the traffic states of roadway sections are divided into five degrees: smooth, basically smooth, light congestion, moderate congestion, and severe congestion. This paper takes “moderate congestion” as the criterion to judge the congestion of roadway sections. Table 2 [23] shows the boundary speed of each grade of roadway sections.

GPS trajectory data include car number, longitude, latitude, time stamp, and speed. Therefore, based on the GPS trajectory data, we can calculate the average speed of roadway sections in different directions in a period of time. The average speed of roadway section i during the k th time period can be defined as \bar{v}_i^k :

$$\bar{v}_i^k = \frac{\sum_{j=1}^{m_i^k} v_j^k}{m_i^k}, \quad (4)$$

where, v_j^k is the instantaneous speed of the j th vehicle in the k th time period. m_i^k is the car number during the k th time period on roadway section i .

The level of traffic congestion of roadway section which is also the node congestion level of the dual roadway network can be defined as CE_i :

$$CE_i = \sum_{k=1}^K \beta_o CE_i^k, \quad (5)$$

$$\beta_o = \begin{cases} 1, & CE_i^k < 0, \\ 0, & CE_i^k \geq 0, \end{cases}$$

$$CE_i^k = \left(\frac{\bar{v}_i^k}{V^*} - 1 \right) \times 100\%, \quad (6)$$

where V^* is the congested speed of the roadway section which equals the boundary speed of moderate congestion shown in Table 1. K is the number of time periods.

Therefore, the weight of node i in the dual network can be represented as

$$w_i = \begin{cases} S * CE_i^k, & \text{when the roadway section is congested,} \\ S, & \text{when the roadway section is not congested.} \end{cases} \quad (7)$$

Correspondingly, the weight of nodes in the primary network can be calculated by formula (1).

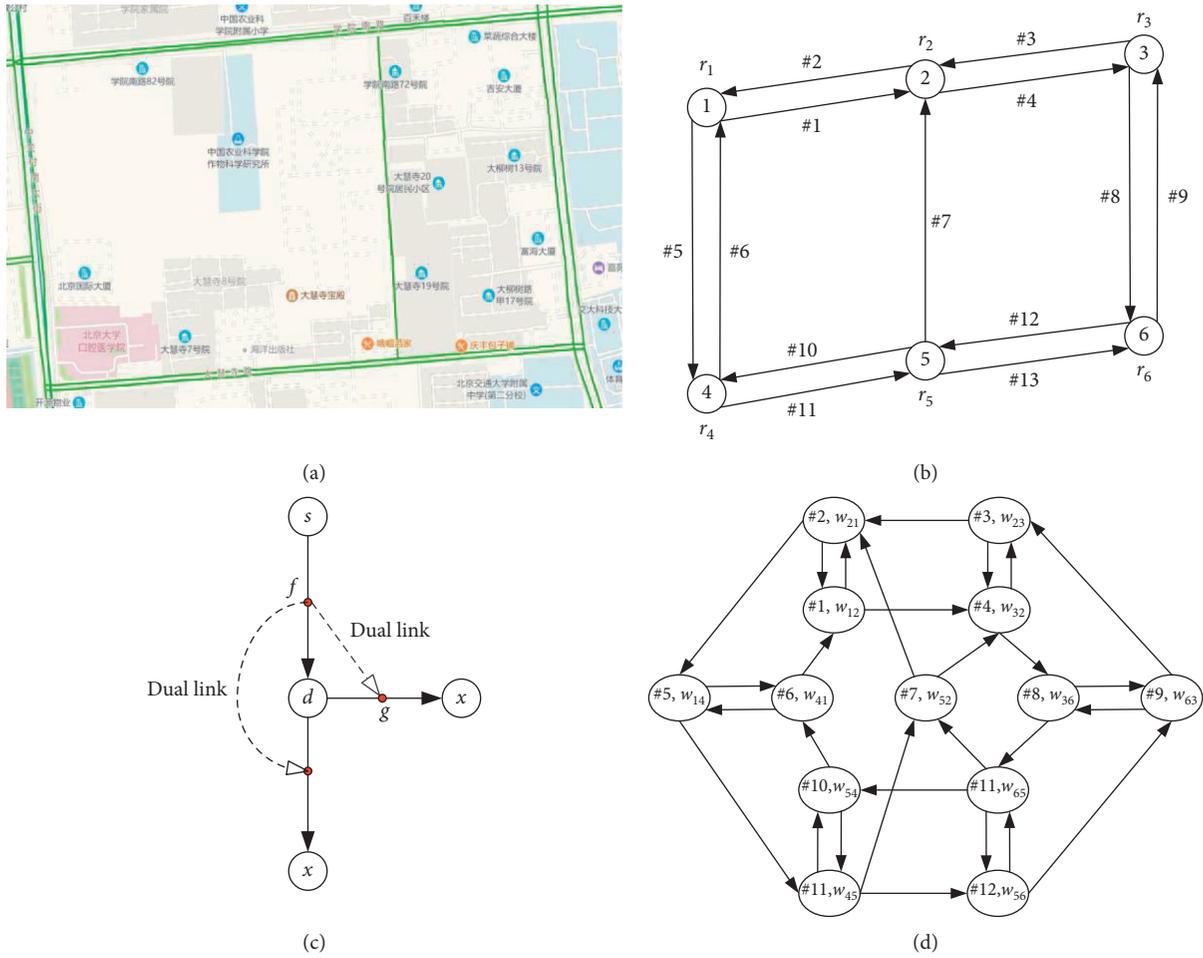


FIGURE 1: Construction process of the roadway network. (a) The actual roadway network. (b) The primary network model. (c) The establishment rule of the dual link. (d) The dual network model.

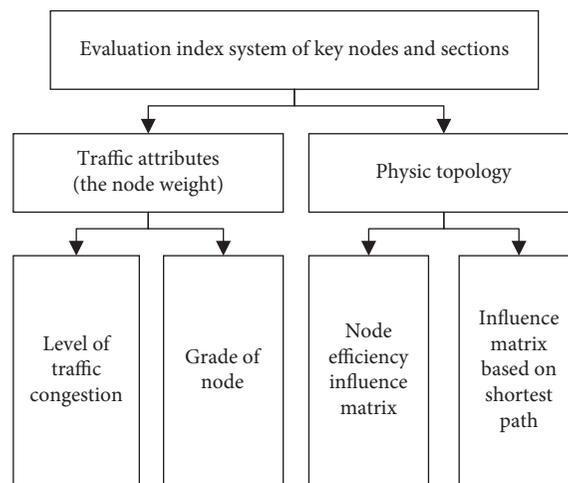


FIGURE 2: Evaluation index system of key nodes and sections.

3.2.2. *Static Topology.* Nodes in the network are influenced and restricted by each other. Therefore, the importance of each node can be obtained according to the influence

ranking. From the perspective of traffic flow, the influence among nodes depends on the shortest path and the number of shortest paths [24]. It should be noted that the

TABLE 1: Weight value corresponding to each grade of roadway section.

Grade of roadway	Expressway	Arterial roadway	Minor arterial roadway and local roadway
Value	10	8	5

TABLE 2: Classification of traffic congestion, unit: km/h.

		Traffic state				
		Smooth	Basically smooth	Light congestion	Moderate congestion	Severe congestion
Grade of roadway	Expressway	$V > 65$	$50 < V \leq 65$	$35 < V \leq 50$	$20 < V \leq 35$	$V \leq 20$
	Arterial roadway	$V > 40$	$30 < V \leq 40$	$20 < V \leq 30$	$15 < V \leq 20$	$V \leq 15$
	Minor arterial roadway and local roadway	$V > 35$	$25 < V \leq 35$	$15 < V \leq 25$	$10 < V \leq 15$	$V \leq 10$

dependency relationship among nodes not only exists in adjacent nodes, but also the nodes which can be reached by effective paths [24–26]. In summary, this paper uses the influence matrix based on node efficiency and the influence matrix based on the shortest path as the evaluation indexes to measure the importance of nodes.

The efficiency between node v_i and v_j is the reciprocal of their distance which can be defined as p_{ij} . If $i = j$ or node i and j are not directly connected, $p_{ij} = 0$. Therefore, the efficiency matrix P can be defined as [27, 28]

$$P = \begin{bmatrix} 0 & p_{12} & \cdots & p_{1n} \\ p_{21} & 0 & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & 0 \end{bmatrix}. \quad (8)$$

According to [28], the influence matrix based on shortest path can be defined as

$$\text{TIP} = \begin{bmatrix} 0 & \frac{A_{12}^{d_{12}}}{A_{12}^{d_{12}} + A_{22}^{d_{12}} + \cdots + A_{n2}^{d_{12}}} & \cdots & \frac{A_{1n}^{d_{1n}}}{A_{1n}^{d_{1n}} + A_{2n}^{d_{1n}} + \cdots + A_{nn}^{d_{1n}}} \\ \frac{A_{21}^{d_{21}}}{A_{11}^{d_{21}} + A_{21}^{d_{21}} + \cdots + A_{n1}^{d_{21}}} & 0 & \cdots & \frac{A_{2n}^{d_{2n}}}{A_{1n}^{d_{2n}} + A_{2n}^{d_{2n}} + \cdots + A_{nn}^{d_{2n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{n1}^{d_{n1}}}{A_{11}^{d_{n1}} + A_{21}^{d_{n1}} + \cdots + A_{n1}^{d_{n1}}} & \frac{A_{n2}^{d_{n2}}}{A_{12}^{d_{n2}} + A_{22}^{d_{n2}} + \cdots + A_{n2}^{d_{n2}}} & \cdots & 0 \end{bmatrix}, \quad (9)$$

$$\text{SIP} = \begin{bmatrix} 0 & \frac{A_{12}^{d_{12}}}{A_{11}^{d_{12}} + A_{12}^{d_{12}} + \cdots + A_{1n}^{d_{12}}} & \cdots & \frac{A_{1n}^{d_{1n}}}{A_{11}^{d_{1n}} + A_{12}^{d_{1n}} + \cdots + A_{1n}^{d_{1n}}} \\ \frac{A_{21}^{d_{21}}}{A_{21}^{d_{21}} + A_{22}^{d_{21}} + \cdots + A_{2n}^{d_{21}}} & 0 & \cdots & \frac{A_{2n}^{d_{2n}}}{A_{21}^{d_{2n}} + A_{22}^{d_{2n}} + \cdots + A_{2n}^{d_{2n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{n1}^{d_{n1}}}{A_{n1}^{d_{n1}} + A_{n2}^{d_{n1}} + \cdots + A_{nm}^{d_{n1}}} & \frac{A_{n2}^{d_{n2}}}{A_{n1}^{d_{n2}} + A_{n2}^{d_{n2}} + \cdots + A_{nm}^{d_{n2}}} & \cdots & 0 \end{bmatrix}$$

where TIP is the target node-centered influence power matrix, and for its element $A_{ij}^{d_{ij}} / \sum_{k=1}^n A_{kj}^{d_{ij}}$ ($i, j =$

$1, 2, \dots, n$), $A_{ij}^{d_{ij}}$ represents the number of shortest paths from node v_i to node v_j ; that is, the number of paths with

length is d_{ij} . The matrix TIP fixes the target node and considers the influence of other nodes on the target node. SIP is source node-centered influence power matrix, and for its element $A_{ij}^{d_{ij}} / \sum_{k=1}^n A_{ik}^{d_{ij}}$ ($i, j = 1, 2, \dots, n$), $\sum_{k=1}^n A_{ik}^{d_{ij}}$ is the number of paths whose length is d_{ij} from node v_j to the other nodes. The matrix SIP fixes the source node and considers the influence of the source node on other nodes.

In summary, the matrix M which reflects the influence of the static topology can be defined as

$$M = (m_{ij})_{n \times n} = \lambda_1 P + \lambda_2 \text{SIP} + \lambda_3 \text{TIP}, \quad (10)$$

where $\lambda_1, \lambda_2, \lambda_3$ represent the weight of efficiency matrix, the weight of target node-centered influence power matrix, and the weight of source node-centered influence power matrix in turn. The relationship among $\lambda_1, \lambda_2, \lambda_3$ is shown as

$$\lambda_1 + \lambda_2 + \lambda_3 = 1. \quad (11)$$

In the paper, $\lambda_1, \lambda_2, \lambda_3$ are calculated by the improved analytic hierarchy process [29]. Finally, the comprehensive importance of nodes Y_i can be expressed as

$$Y_i = w_i \sum_{j=1}^n m_{ij}, \quad i = 1, 2, 3, \dots, n. \quad (12)$$

After normalization, the comprehensive importance value of nodes SY_i can be expressed as

$$SY_i = \frac{Y_i}{\sum_{i=1}^n Y_i}. \quad (13)$$

Generally, the steps of the method to identify the key nodes and sections can be summed up as Figure 3.

4. Experiment and Discussion

The floating car data are the real traffic GPS data collected by about 40,000 taxis in Beijing, China, over a period of a week (January 12–18, 2015). The traffic data are recorded once per minute approximately. The format of GPS data is shown in Figure 4. We first preprocess the data in order to eliminate noisy sample points, of which perceived positions are chaotic. In the second step, the sample points with the same vehicle ID are linked to each other according to their time correlate on. Then, we capture the floating car trajectories on the urban roadway network space [2]. Finally, map matching is carried out according to the floating car trajectories, the latitude and longitude of the vehicles, and the urban geographic information [30].

We select Xicheng District as the research object which is one of the major six urban areas in Beijing with a high population density and a developed economy. The administrative area of Xicheng District is 50.70 square kilometers. The roadway network of Xicheng District is shown in Figure 5(a). As there are many pedestrian streets and small local roadways in the roadway network, filtering of the roadway network is required. The filtering rules are as follows: (1) pedestrian streets are not considered; (2) local roadways with small traffic flow are neglected; (3) small intersections are neglected, and only large intersections are

considered. Finally, the roadway network topology of Xicheng District is simplified as Figure 5(b) which contains 310 roadway sections (two-way) and 91 intersections. There are 46 expressways, 206 arterial roadways, and 58 minor arterial roadways (local roadways).

4.1. Analysis of Congestion in Roadway Network. In this paper, “moderate congestion” is taken as the criterion to judge the traffic congestion of each grade of roadway section. In other words, it is considered as congestion when the average speed of vehicles is under 35 km/h on expressway, under 20 km/h on arterial roadway or under 10 km/h on minor arterial roadway and local roadway. To clearly understand the congestion situation of Xicheng District, the congestion rate of each grade of roadway section is defined as the proportion of congested roadway sections to the total roadway sections per hour. Figures 6 and 7 illustrate the changes of congestion rate of different grade of roadway section in Xicheng District from 6:00 to 24:00 on January 12 (Monday) and January 17 (Saturday). It should be noted that the starting time of the hour interval is used to represent the time period. For example, 6:00 in the paper represents the period from 6:00 to 7:00.

To better understand the distribution law of the congestion rate, we made Figure 8 showing the change curve of the congestion rate of each grade of roadway section in Xicheng District during a week. It can be seen that the working days and weekends present a similar trend.

4.2. Identification of Key Roadway Sections. Python is used to solve the identification algorithm proposed in this paper. The results show that the weight of the efficiency matrix $\lambda_1 = 0.82$, the weight of target node-centered influence power matrix $\lambda_2 = 0.09$, and the weight of source node-centered influence power matrix $\lambda_3 = 0.09$. The key nodes and sections of Xicheng District are identified within five seconds by using the packages imported from Python.

Firstly, the key sections in Xicheng District are analyzed. Figures 9 and 10 highlight the top 50 key sections in the morning and evening peak hours on January 12 (Monday) and January 17 (Saturday). The network with red color shows the first 10 sections with the highest rank and the yellow marks the next 40 sections.

The following rules can be obtained from Figures 9 and 10:

- (a) Whether it is on Monday or Saturday, the key sections are almost expressways and arterial roadways. This is because the grade of roadway sections and the congestion level of expressways and arterial roadways in peaks hours are both higher than those of minor arterial roadways and local roadways.
- (b) The distribution of key sections between the morning and evening peak hours on Monday and Saturday is relatively similar. From the perspective of the whole Xicheng District, the importance value of the roadway sections in the center of the city is generally higher than of the edge of the city because

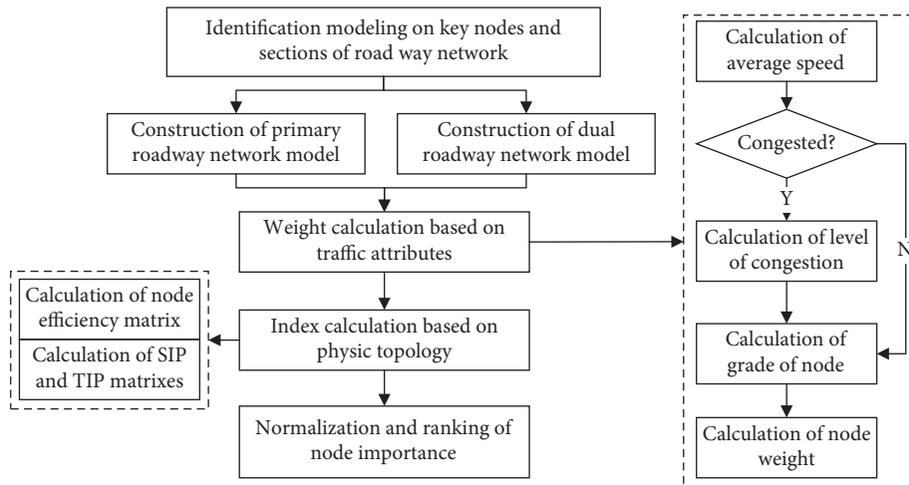


FIGURE 3: Identification process of key nodes and sections of the roadway network.

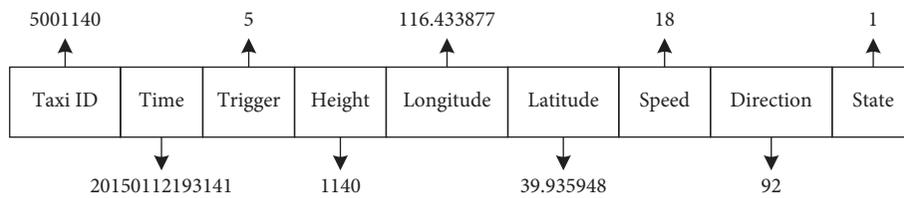


FIGURE 4: Sample of original GPS data.

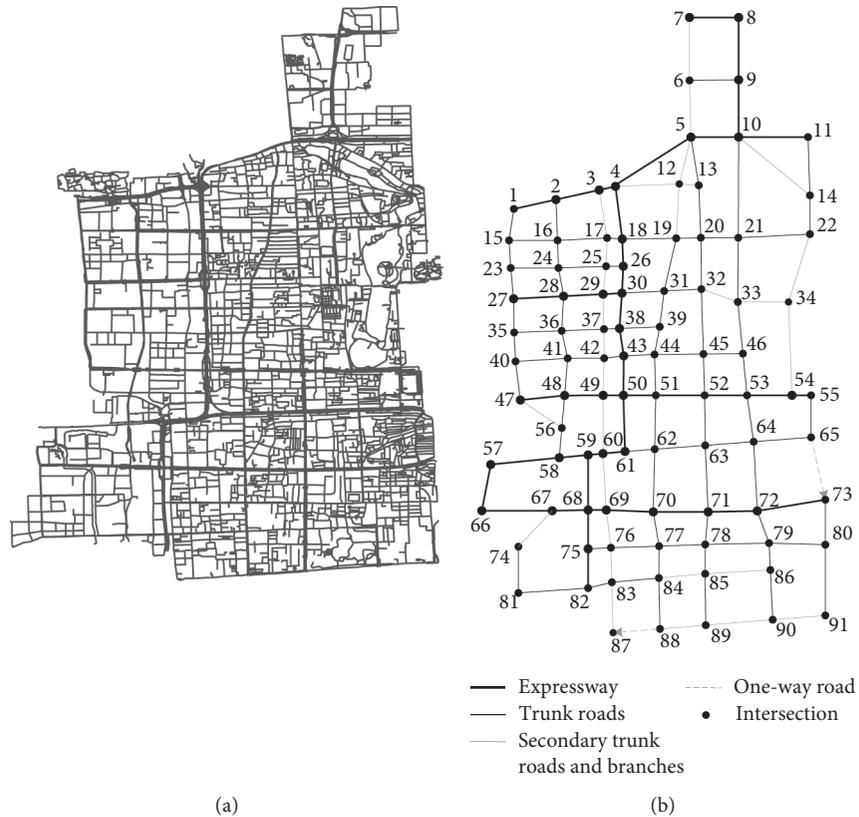


FIGURE 5: Network topology construction of Xicheng District. (a) Roadway map of Xicheng District. (b) Network topology of Xicheng District.

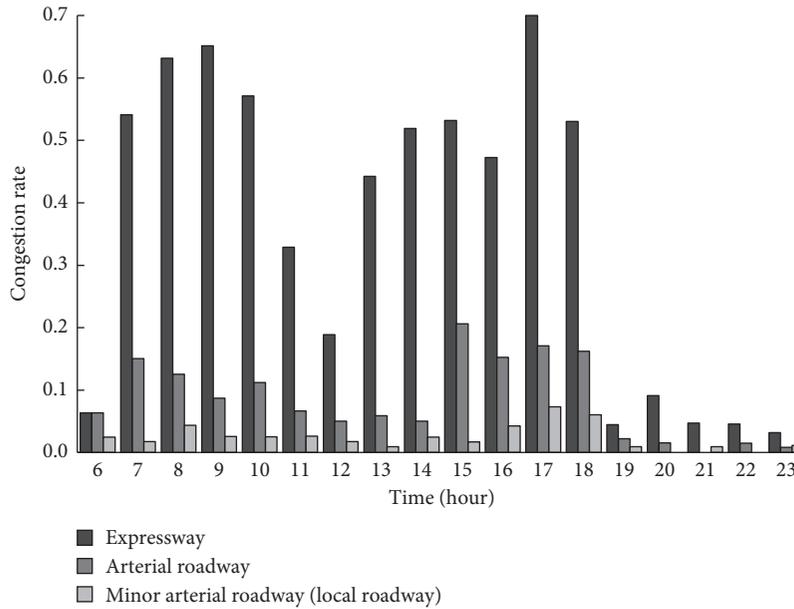


FIGURE 6: The traffic congestion ratio of each grad of roadway sections on Monday.

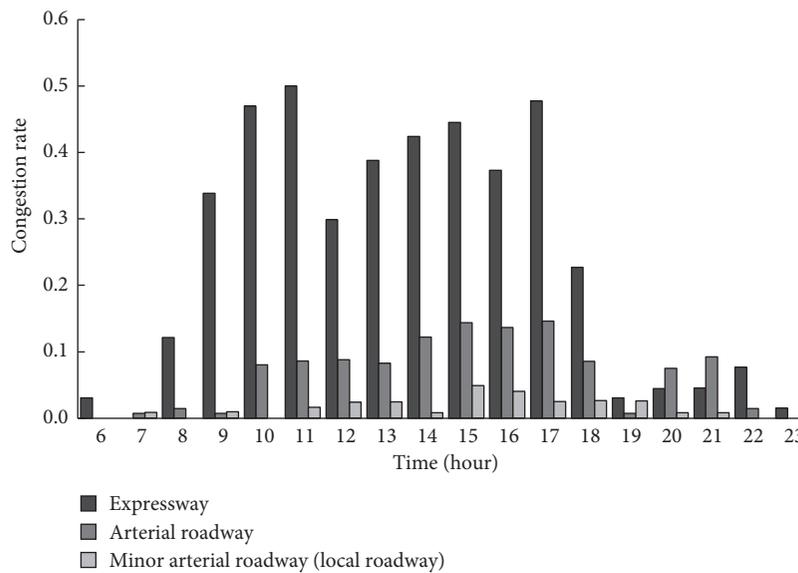


FIGURE 7: The traffic congestion ratio of each grad of roadway sections on Saturday.

the node efficiency of the central section is higher than that of the edge part. Most of the sections centered on the Xidan business district are in the top 50 key sections due to the high congestion level of the roadway sections near the Xidan business district.

(c) Whether it is on Monday or Saturday, the importance values of the key 50 sections change slightly between the morning and evening peak hours. To analyze this change, we made Table 3 showing the importance value of the top 50 key sections in the morning and evening peak hours on Monday. Each section is numbered according to the node serial number in Figure 5(b). For example, 1-2 represents

the section from node 1 to node 2, and 2-1 represents the section from node 2 to node 1.

It can be seen that the 50 key roadway sections in the morning and evening peak hours on Monday are all expressways and arterial roadways. A comparative analysis of the distribution of the 50 key roadway sections between the morning and evening peak hours shows that there are 35 same roadway sections. In addition, there are 16 opposite roadway sections (e.g., section 1-2 and roadway section 2-1) between the morning and evening peak hours, and the rank of the opposite roadway sections most has changed between the morning and evening peak hours. For example, in the morning peak hours, the roadway section 50-43 ranks 2nd,

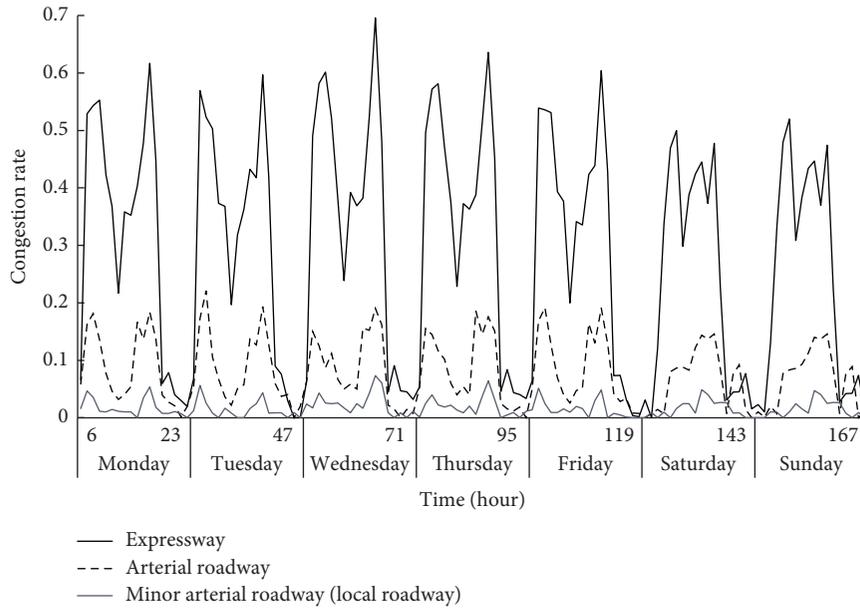


FIGURE 8: Changing curve of traffic congestion ratio of each grad of roadway during a week.

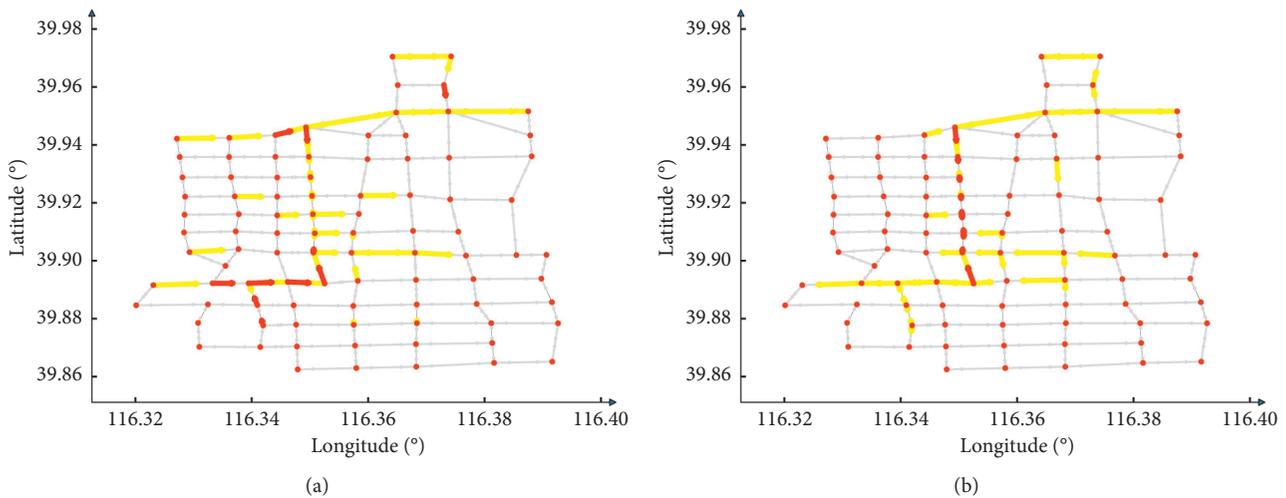


FIGURE 9: Identification results of key roadway sections in peak hours on Monday. (a) 7:00–9:00. (b) 17:00–19:00.

and its opposite roadway section 43–50 ranks 26th. In the evening peak hours, the roadway section 50–43 drops to 4th, and the roadway section 43–50 ranks up to 3rd.

Figure 11 shows the change of the top 50 key sections from 6:00 to 22:00 on Monday. The top 50 key sections are colored according to the normalized importance. The larger the importance value, the darker the color is. The sections ranked after 50 is uniformly colored as light gray. The key sections of the whole day on Monday are still distributed in expressways and arterial roadways, and the importance value of each section is changed with time. It can be seen that, due to the changes of travel behavior, the importance value of roadway sections changes as well; that is, the importance value of roadway sections is directly affected by travel behavior.

4.3. Identification of Key Nodes. The identification of key nodes in Xicheng District is studied. Figures 12 and 13 highlight the top 20 key nodes in morning and evening peak hours on January 12 (Monday) and January 17 (Saturday). The network with red color shows the first 10 nodes with the highest rank and the yellow marks the next 10 nodes.

It can be seen that most of the key nodes in the morning and evening peak hours are connected with expressways and arterial roadways on both Monday and Saturday. It is caused by the higher grade of roadway sections and congestion level of expressways and arterial roadways. The distribution of key nodes is similar to that of key sections, and the importance value of the nodes in the center of the city is generally higher than that of edge areas. Table 4 shows the importance value of the top 20 key nodes in the morning and evening peak hours on Monday.

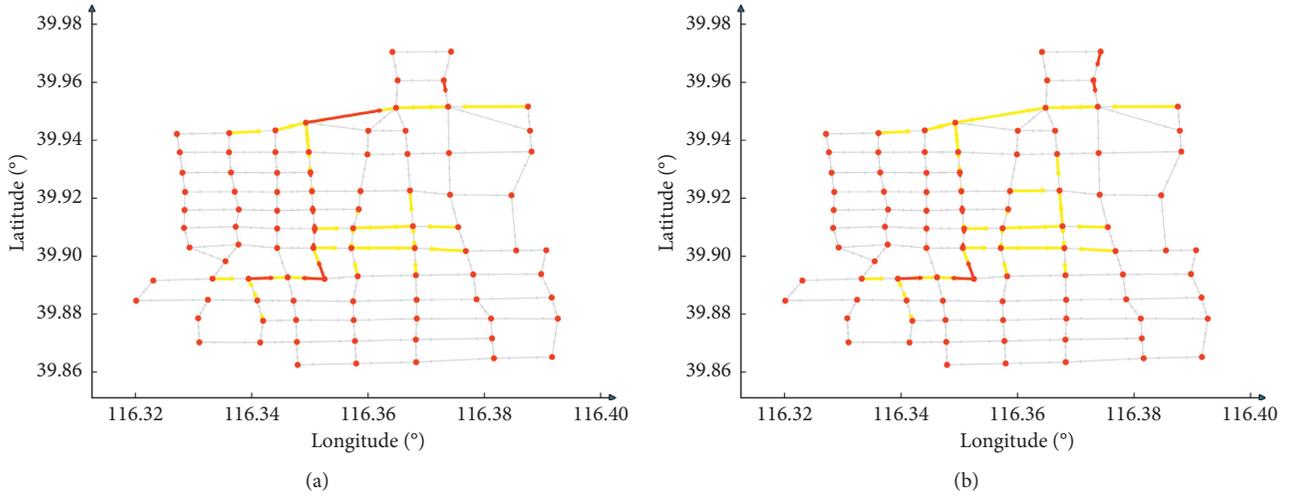


FIGURE 10: Identification results of key roadway sections in peak hours on Saturday. (a) 9:00–11:00. (b) 17:00–19:00.

TABLE 3: Top 50 key roadway sections in peak hours on Monday.

Rank	Morning		Evening		Rank	Morning		Evening	
	Road id	Sr	Road id	Sr		Road id	Sr	Road id	Sr
1	61–50 _E	1.000	38–43 _E	1.000	26	43–50 _E	0.551	59–58 _E	0.605
2	50–43 _E	0.926	26–30 _E	0.996	27	10–11 _E	0.545	75–82 _E	0.543
3	9–10 _E	0.884	43–50 _E	0.992	28	61–60 _E	0.534	58–57 _E	0.541
4	68–59 _E	0.865	50–43 _E	0.985	29	30–38 _E	0.533	10–11 _E	0.511
5	3–4 _E	0.859	61–50 _E	0.977	30	50–61 _E	0.515	8–7 _E	0.498
6	59–60 _E	0.854	30–38 _E	0.964	31	26–18 _E	0.502	9–8 _E	0.486
7	75–68 _E	0.848	43–38 _E	0.95	32	77–70 _A	0.484	37–38 _A	0.458
8	4–18 _E	0.847	4–18 _E	0.932	33	37–38 _A	0.469	53–52 _A	0.442
9	58–59 _E	0.836	18–26 _E	0.909	34	78–71 _A	0.468	51–62 _A	0.442
10	60–61 _E	0.827	38–30 _E	0.889	35	5–10 _E	0.458	52–51 _A	0.433
11	38–30 _E	0.826	50–61 _E	0.849	36	59–68 _E	0.453	18–4 _E	0.432
12	4–5 _E	0.812	59–60 _E	0.839	37	8–7 _E	0.451	5–10 _E	0.429
13	43–38 _E	0.812	9–10 _E	0.837	38	51–52 _A	0.451	61–62 _A	0.429
14	26–30 _E	0.786	61–60 _E	0.801	39	60–59 _A	0.445	50–49 _A	0.429
15	18–26 _E	0.751	60–61 _E	0.783	40	43–44 _A	0.438	20–32 _A	0.427
16	57–58 _E	0.742	3–4 _E	0.772	41	31–32 _A	0.435	44–51 _A	0.426
17	10–5 _E	0.738	5–4 _E	0.772	42	47–48 _A	0.424	75–68 _E	0.425
18	2–3 _E	0.72	26–18 _E	0.759	43	28–29 _A	0.421	63–71 _A	0.421
19	5–4 _E	0.714	4–5 _E	0.748	44	38–39 _A	0.418	58–59 _E	0.418
20	18–4 _E	0.714	30–26 _E	0.74	45	4–3 _E	0.414	68–59 _E	0.418
21	30–26 _E	0.713	10–5 _E	0.727	46	7–8 _E	0.413	63–62 _A	0.417
22	11–10 _E	0.711	59–68 _E	0.695	47	52–53 _A	0.412	7–8 _E	0.412
23	8–9 _E	0.623	60–59 _E	0.685	48	44–51 _A	0.407	51–50 _A	0.407
24	1–2 _E	0.578	11–10 _E	0.667	49	50–51 _A	0.407	52–63 _A	0.407
25	38–43 _E	0.556	68–75 _E	0.619	50	52–51 _A	0.407	44–43 _A	0.399

Note: the subscript represents the grade of roadway section. *E* is expressway, and *A* is arterial roadway.

It can be found that the distribution of the key nodes between the morning and evening peak hours is roughly the same, and most of the top 20 key nodes are the nodes with higher node grades. Moreover, the rank of the top 20 key nodes just changes slightly or remains the same between the morning and evening peak hours. For example, node 10 ranks 1st in the morning and the 4th in the evening. The ranks of node 30, 68, and 62 remain the same between the morning and evening peak hour. Therefore, we can draw a conclusion that the change of congestion rate of the roadway

sections has little to do with the ranking of key nodes. This is because the nodes connect multiple roadway sections at the same time. Even if the tidal phenomenon of traffic congestion happens on opposite roadway sections, the average rate of congestion of the nodes just changes slightly.

4.4. Identification with Traditional Methods. In this part, the identification results of our proposed model are compared with those of the traditional evaluation method based on

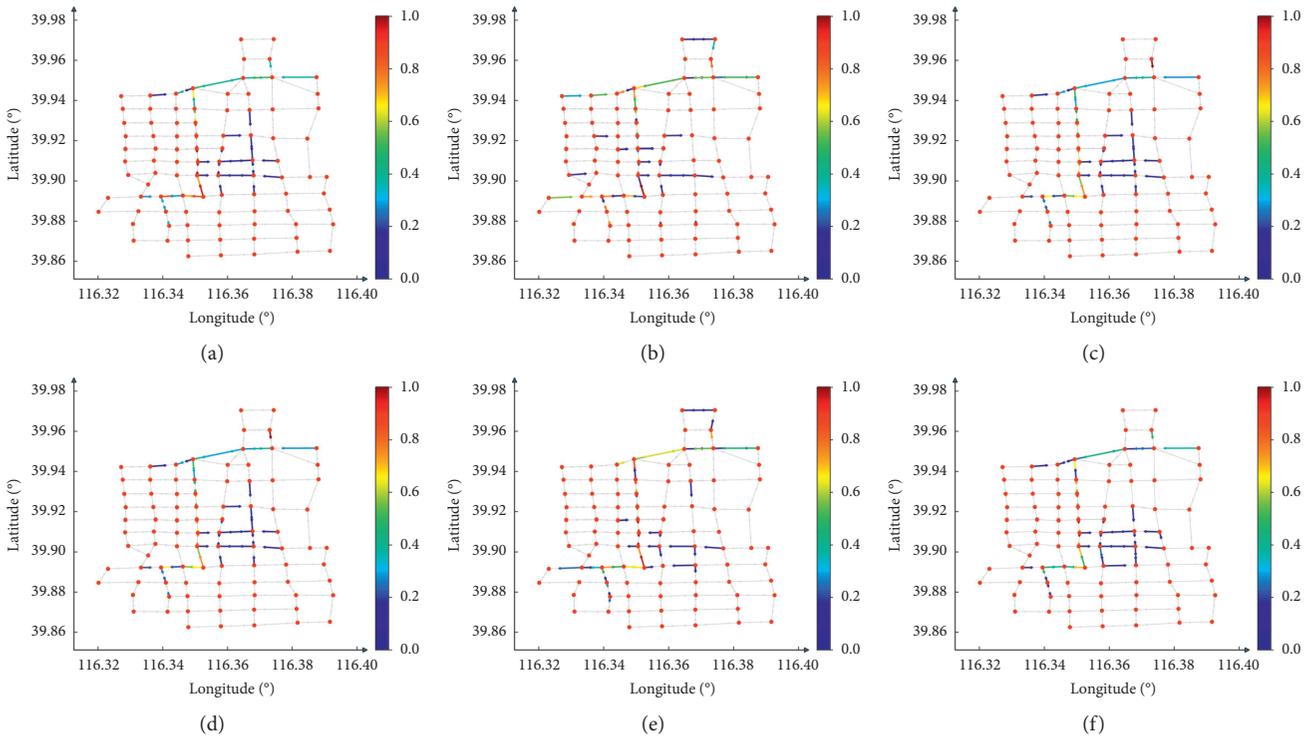


FIGURE 11: Distribution of the top 50 key roadway sections in different periods on Monday: (a) 6:00. (b) 8:00. (c) 12:00. (d) 14:00. (e) 17:00. (f) 22:00.

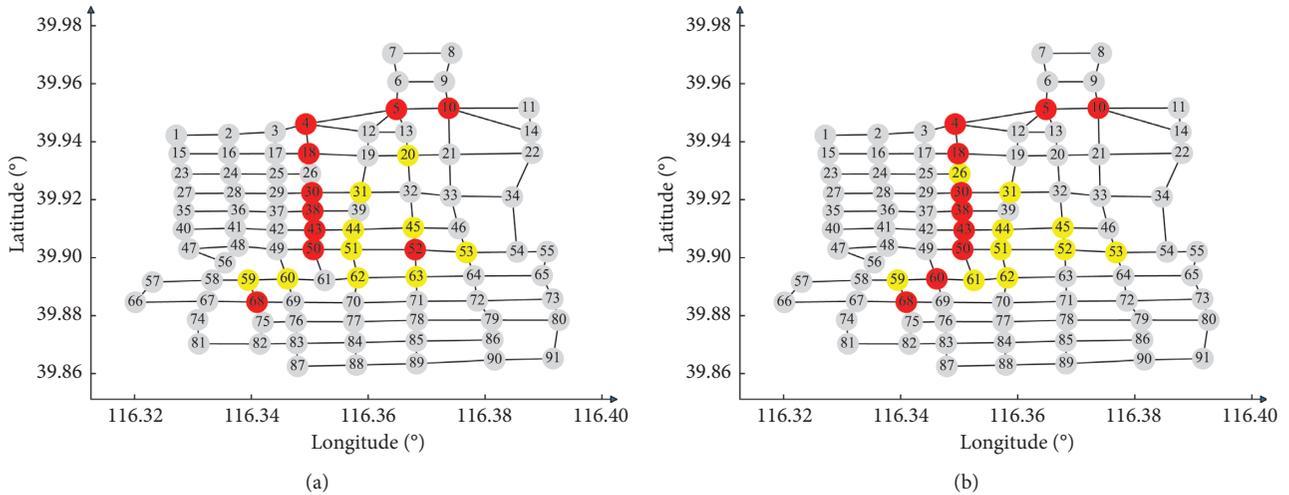


FIGURE 12: Identification results of key nodes in peak hours on Monday. (a) 7:00-9:00. (b) 17:00-19:00.

node betweenness [31] and the evaluation method based on the level of traffic congestion (based on formula (6) and (1)). Figure 14 shows the identification results of the methods based on node betweenness and level of traffic congestion, in which the data at 17:00 on Monday are used.

The results show that the key sections based on node betweenness are mainly located in the center of the city, reflecting the topological status of each section. While the key sections based on the level of traffic congestion are mainly expressways and arterial roadways. Moreover, taking

the data of the whole Monday as an example, the repetition rate of the identification results of the top 50 key roadway sections among the proposed method of this paper and the above two methods is calculated, respectively, at hourly intervals, which is shown in Figure 15.

In the whole day, the repetition rate of the identification results proposed by our model and that of the node betweenness is distributed between 0.2 and 0.5, of which the average value is 0.31. While the repetition rate of the proposed model and the traffic congestion is distributed

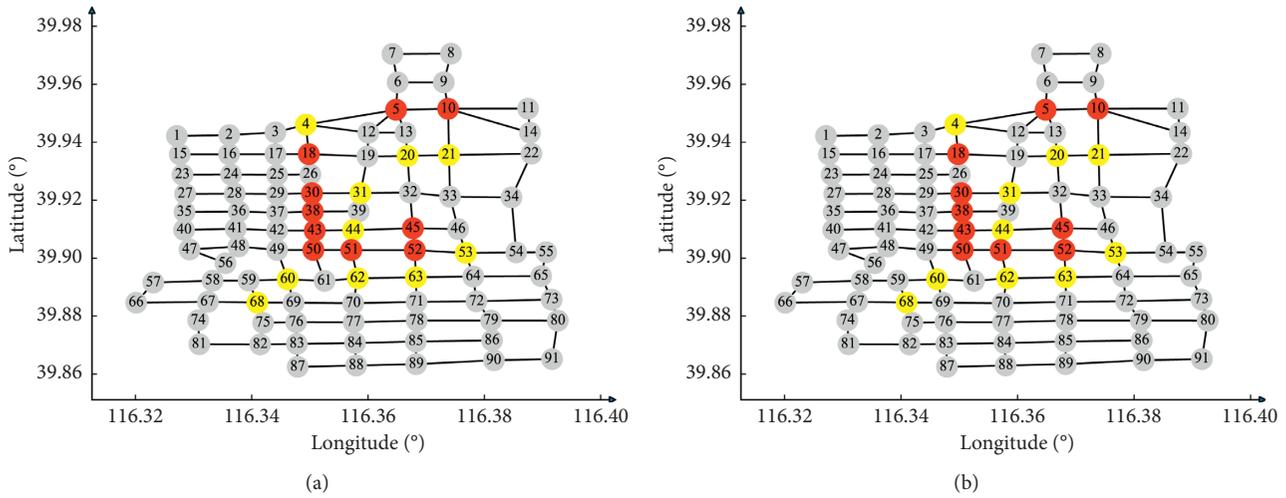


FIGURE 13: Identification results of key nodes in peak hours on Saturday. (a) 9:00–11:00. (b) 17:00–19:00.

TABLE 4: Top 20 key nodes in peak hours on Monday.

Rank	Morning		Evening		Rank	Morning		Evening	
	Node id	Sr	Node id	Sr		Node id	Sr	Node id	Sr
1	10 ₈₆	1.000	43 ₇₂	1.000	11	59 ₆₀	0.703	51 ₅₆	0.697
2	5 ₇₆	0.887	50 ₇₂	0.997	12	51 ₆₄	0.701	26 ₆₄	0.692
3	50 ₇₂	0.877	38 ₇₂	0.967	13	44 ₆₀	0.694	52 ₅₆	0.686
4	43 ₇₂	0.863	10 ₈₆	0.957	14	45 ₆₄	0.692	59 ₆₄	0.684
5	30 ₇₂	0.859	30 ₇₂	0.948	15	31 ₆₄	0.689	44 ₆₀	0.669
6	4 ₇₀	0.853	18 ₇₂	0.849	16	53 ₆₄	0.688	45 ₆₄	0.668
7	38 ₇₂	0.838	5 ₇₆	0.826	17	20 ₆₄	0.662	31 ₆₄	0.654
8	18 ₇₂	0.832	4 ₇₀	0.809	18	62 ₆₄	0.659	62 ₆₄	0.642
9	68 ₇₂	0.800	68 ₆₀	0.733	19	63 ₆₁	0.649	53 ₆₁	0.641
10	52 ₆₄	0.707	60 ₇₂	0.725	20	60 ₇₂	0.643	61 ₅₆	0.636

Note: the subscript represents the grade of node.

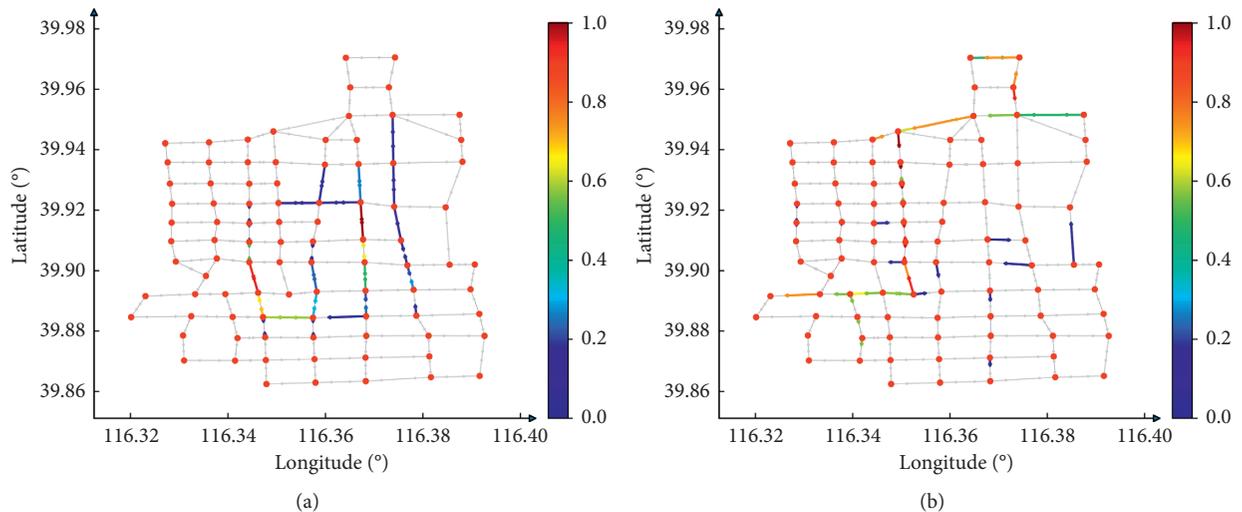


FIGURE 14: Results of key roadway sections of the comparative method. (a) Identification results based on node betweenness. (b) Identification results based on traffic congestion.

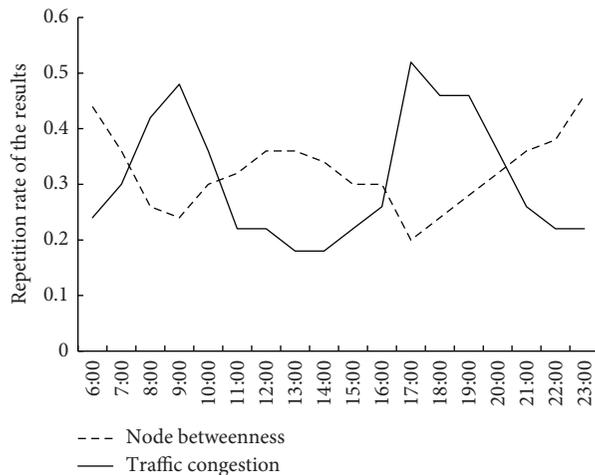


FIGURE 15: Repetition rate of the identification results of key roadway sections.

between 0.2 and 0.6, of which the average value is 0.32. In the traffic peak hours, the repetition rate of the traffic congestion method reaches the peak value, which is 0.52. At 6:00, 12:00, and 22:00, the repetition rate of node betweenness reaches the peak value, which is 0.46. It shows that with the aggravation of traffic congestion, the repetition rate of the traffic congestion increases accordingly. As traffic congestion improves, the repetition rate of the identification results of the model proposed by this paper and those based on the node betweenness increases. In conclusion, the identification model proposed in this paper realizes the expression of both topological structure and dynamic traffic attributes of the roadway network and therefore is more practicable and effective in a large scale.

5. Conclusion

In the paper, a method was proposed to identify the key nodes and sections of the roadway network. In the method, the taxi-GPS data were used to map the floating car sample points with the same vehicle ID to the car trajectories according to their time correlation. First, the weighted complex network is used to construct the roadway network model. Then, the evaluation index system of key nodes and sections is developed from the perspectives of dynamic traffic attributes and roadway network topology. Finally, Xicheng District of Beijing, China, is selected as the experiment area and the method proposed by this paper is compared with the conventional methods. The results show that the results proposed by our model are highly similar to those of the method based on traffic congestion in peak hours. While the results proposed by our model bear similarity to those of the method based on topological structure in other time. In other words, the method proposed in this paper realizes the expression of topological structure and dynamic traffic attributes of the roadway network simultaneously, which is more practicable and effective in a large scale.

There are still some problems need to be further analyzed in this paper. Urban traffic is composed of subway, bus, taxi and private car. Only taxi-GPS data are used in this paper, and integrated data of various traffic modes should be used into the study of key nodes and sections in the future.

Data Availability

Some or all data, models, or code generated or used during the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Key R&D Program of China (grant no. 2019YFF0301400).

References

- [1] X. L. Ma, H. Y. Yu, Y. P. Wang, and Y. H. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," *Plos One*, vol. 10, no. 3, p. 17, 2015.
- [2] X. Kong, Z. Xu, G. Shen, J. Wang, Q. Yang, and B. Zhang, "Urban traffic congestion estimation and prediction based on floating car trajectory data," *Future Generation Computer Systems*, vol. 61, pp. 97–107, 2016.
- [3] H. Feng, F. Bai, and Y. Xu, "Identification of critical roads in urban transportation network based on GPS trajectory data," *Physica A: Statistical Mechanics and Its Applications*, vol. 535, 2019.
- [4] X. Liu and Y. Ban, "Uncovering spatio-temporal cluster patterns using massive floating car data," *ISPRS International Journal of Geo-Information*, vol. 2, no. 2, pp. 371–384, 2013.
- [5] P. Wu, K. Liu, S. Hao, Q. Zhang, and Y. Tan, "Rapid traffic congestion monitoring based on floating car data," *Journal of Computer Research and Development*, vol. 51, no. 01, pp. 189–198, 2014.
- [6] K. A. Hawick and H. A. James, "Node importance ranking and scaling properties of some complex road networks," *Research Letters in the Information and Mathematical Sciences*, vol. 11, pp. 23–32, 2007.
- [7] X. Zhang, J. P. Wu, and Y. M. Du, "The study of critical nodes of road network based on traffic shockwave model," in *Proceedings of the 2012 7th International Conference on Computing and Convergence Technology (ICCCCT)*, Seoul, South Korea, December 2012.
- [8] I. Jayaweera, "Centrality measures to identify traffic congestion on road networks: a case study of Sri Lanka," *IOSR Journal of Mathematics*, vol. 13, no. 02, pp. 13–19, 2017.
- [9] M. Xu, J. Wu, M. Liu, Y. Xiao, H. Wang, and D. Hu, "Discovery of critical nodes in road networks through mining from vehicle trajectories," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 583–593, 2019.
- [10] Z. Du, J. Tang, Y. Qi, Y. Wang, C. Han, and Y. Yang, "Identifying critical nodes in metro network considering topological potential: a case study in Shenzhen city—China," *Physica A: Statistical Mechanics and Its Applications*, vol. 539, Article ID 122926, 2020.

- [11] D. M. Scott, D. C. Novak, L. Aultman-Hall, and F. Guo, "Network robustness index: a new method for identifying critical links and evaluating the performance of transportation networks," *Journal of Transport Geography*, vol. 14, no. 3, pp. 215–227, 2006.
- [12] J. L. Sullivan, D. C. Novak, L. Aultman-Hall, and D. M. Scott, "Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: a link-based capacity-reduction approach," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 5, pp. 323–336, 2010.
- [13] P. Luathep, A. Sumalee, H. W. Ho, and F. Kurauchi, "Large-scale road network vulnerability analysis: a sensitivity analysis based approach," *Transportation*, vol. 38, no. 5, pp. 799–817, 2011.
- [14] F. Rupi, S. Angelini, S. Bernardi, A. Danesi, and G. Rossi, "Ranking links in a road transport network: a practical method for the calculation of link importance," *Transportation Research Procedia*, vol. 5, pp. 221–232, 2015.
- [15] A. Kumar, K. Haque, S. Mishra, and M. M. Golias, "Multi-criteria based approach to identify critical links in a transportation network," *Case Studies on Transport Policy*, vol. 7, no. 3, pp. 519–530, 2019.
- [16] Y. Zhang and A. Khani, "Identifying critical links in transportation network design problems for maximizing network accessibility," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 2, pp. 237–251, 2020.
- [17] F. Li, H. Jia, Q. Luo, Y. Li, and L. Yang, "Identification of critical links in a large-scale road network considering the traffic flow betweenness index," *PLoS One*, vol. 15, no. 4, Article ID e0227474, 2020.
- [18] S. Porta, P. Crucitti, V. Latora, P. B-planning, and Design, "The network analysis of urban streets: a primal approach," *Environment and Planning B: Planning and Design*, vol. 33, no. 5, pp. 705–725, 2006.
- [19] S. Porta, P. Crucitti, and V. Latora, "The network analysis of urban streets: a dual approach," *Physica A: Statistical Mechanics and Its Applications*, vol. 369, no. 2, pp. 853–866, 2006.
- [20] J. Anez, T. De La Barra, and B. Perez, "Dual graph representation of transport networks," *Transportation Research Part B: Methodological*, vol. 30B, no. 3, pp. 209–216, 1996.
- [21] S. Marshall, "Line structure representation for road network analysis," *Journal of Transport and Land Use*, vol. 9, no. 1, pp. 29–64, 2016.
- [22] Y. Wang, G. Zheng, P. Wang, and S. Zhang, "Identification method of critical nodes in road network with grey relational grade and Entropy method," *Basic Science and Technology*, vol. 17, no. 06, pp. 84–87, 2015.
- [23] People's Republic of China Public Safety Industry Standard, "Urban road traffic performance index," 2011.
- [24] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, 1977.
- [25] F. Shao and B. Cheng, "Optimal routing strategy based on specifying shortest path," *International Journal of Computers Communications & Control*, vol. 9, no. 5, pp. 602–609, 2014.
- [26] Y. Xu and F. Shao, "An optimal routing strategy based on specifying shortest path," *Telkomnika - Indonesian Journal of Electrical Engineering*, vol. 11, no. 10, pp. 6224–6231, 2013.
- [27] D. A. Griffith, M. M. Fischer, and J. LeSage, "The spatial autocorrelation problem in spatial interaction modelling: a comparison of two common solutions," *Letters in Spatial and Resource Sciences*, vol. 10, no. 1, pp. 75–86, 2017.
- [28] Y. Wang and J. Guo, "Evaluation method of node importance in directed-weighted complex network based on multiple influence matrix," *Acta Physica Sinica*, vol. 66, no. 05, Article ID 050201, 2017.
- [29] Y. Zhu, Z. Meng, and S. Kan, "Determination of weight value by AHP," *Journal of Northern Jiaotong University*, vol. 23, no. 5, pp. 119–123, 1999.
- [30] M. A. Quddus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: state-of-the art and future research directions," *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 5, pp. 312–328, 2007.
- [31] S. Porta, P. Crucitti, and V. Latora, "The network analysis of urban streets: a primal approach," *Environment and Planning B: Planning and Design*, vol. 33, no. 5, pp. 705–725, 2006.