

Research Article

Method for Determining the Location of Highway Passenger Transportation Hubs Using POI Data and the Dijkstra Algorithm in Large City

Hui Li ¹, Pei Tong ², and Xu Zhang ¹

¹School of Civil Engineering, Henan University of Technology, Zhengzhou 450001, China

²School of Civil Engineering, Zhengzhou University, Zhengzhou 450001, China

Correspondence should be addressed to Pei Tong; tongpei_zzu@163.com

Received 8 July 2022; Revised 11 November 2022; Accepted 30 November 2022; Published 10 December 2022

Academic Editor: Hongtai Yang

Copyright © 2022 Hui Li 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.

In order to reduce the impact of highway passenger transportation hubs on urban internal traffic operations and improve the efficiency of external travel, a method to determine the location of such hubs based on POI data and the Dijkstra algorithm is proposed. Firstly, the process of connecting the starting point of travel in the city to the urban road node and the highway exit is analyzed. Taking the minimum total time cost as the objective function, the mathematical model of the location method of a highway passenger transportation hub is established. Then, an algorithm for solving the mathematical model is proposed. In the solution method, the concept of “equivalent distance” is introduced into the Dijkstra algorithm to calculate the shortest path for passengers traveling on the urban road network. Weighted random numbers are generated to simulate the choice of highway exits when intercity coaches carrying passengers leave the city. Finally, Zhengzhou City is taken as an example for application analysis. The calculation is carried out using a script written in Python. The calculation results of the proposed method are compared with the results of the P -median model and the current locations of hubs. The results show that the proposed method consumes less time than the P -median method. The total time cost is also lower, which can reduce the time for passengers to leave the city. The location of the highway passenger transportation hub is closer to the periphery of the city and highway nodes, which conforms to the engineering practice and development trend of hub location so as to reduce its impact on urban internal traffic.

1. Introduction

Intercity passenger transportation planning is an important part of comprehensive transport system planning. The location planning of highway passenger transportation hubs is one of the main tasks of intercity passenger transportation [1]. According to statistical data from the Ministry of Transport, in the first quarter of 2019, passenger volume on highways was 3.36 billion in China, and intercity passenger transportation accounted for 75%. Passenger transportation hubs, coaches, and transportation lines constitute the intercity passenger transportation system [2]. Passenger transportation hubs are important nodes that connect intercity passenger transportation systems together and transfer intercity passengers to the inner-city passenger

transportation system. The hub provides services for medium- and long-distance passenger transportation between cities, and it organizes and transfers passengers. The rationality of the location of the highway passenger transportation hub is crucial to the realization of the above functions and also determines the overall operational efficiency of the intercity passenger transportation system.

With the expansion of urban scale and the development of the social economy, the location of urban highway passenger transport hubs is facing many problems. Many highway passenger transportation hubs could have been located at the edges of cities. However, with the expansion of the city, the location has changed to the interior of the city. With the rapid development of motorization, the contradiction between highway passenger transport hubs and

urban transportation has become increasingly prominent, which affects urban traffic flow, the time cost of passengers, and the service level of intercity passenger transportation. Therefore, highway passenger transportation hubs located in the center of the city need to be relocated, or new highway passenger transportation hubs need to be built in the peripheral areas of the city. According to the “Code for Intercity Transportation Planning” (GB 50925-2013) in China, highway passenger transportation hubs should be reasonably located based on travel demands, urban layout, and intercity travel patterns. The “Standard for Comprehensive Urban Transport System Planning” (GB/T 51328-2018) in China also provides the layout requirements for highway passenger transportation hubs, which should be consistent with the organization of the overall transportation system in urban areas. However, the method for determining the location of highway passenger transportation hubs is not specified.

1.1. Literature Review. Location problems are an important field of operations research. Because of its extensive application in practice, it has attracted the attention of many scholars [3–9]. The methods for determining the location of passenger transportation hubs are mostly divided into two categories: (a) continuous location models, which mainly include the center-of-gravity approach, and (b) discrete location models, which require that the candidate locations of hubs are limited and countable and the optimal location is determined through evaluation. The center-of-gravity approach, as a classic algorithm for determining the location of hubs, obtains the optimal location of hubs by calculating the center of gravity of planar objects [10]. The disadvantage of this method is that it inevitably selects the hub location in the urban center. Therefore, it is difficult to adapt to the rapid development of urbanization and motorization. Due to limitations in urban space and road resources, undeveloped land on the edges of cities is used to build new highway passenger transportation hubs [11]. Fan et al. [12] proposed a bilevel programming model (BPM) to solve the integrated optimization of multihub locations and network layouts. The objective of the upper-level model is to minimize network operating costs, network transformation costs, and hub construction costs. The lower-level model depicts the distribution of user equilibria for fixed levels of demand. Alumur et al. [3] proposed a linear mixed integer programming model for multimodal transport hub location and hub network design to minimize transport costs and travel time. They proposed a linear mixed integer programming formulation for the most general case of the multimodal hub location and hub network design problem. Zhu et al. [13] analyzed the principles and influencing factors of the location selection of passenger transportation hubs and optimized the mixed integer programming model. The OD matrix and generalized cost matrix are used to solve the model. Travel is divided into two processes. One is from the traffic starting point to the hub, and the other is from the hub to the attraction. Considering the existing infrastructure and environment, the opportunity to develop, the transport

access, and other factors, Martinov [14] uses the weighted sum method to determine the location of the intermodal terminal. Da Graça Costa et al. [15] established a bicriteria model of hub location based on the objective functions of minimizing cost and service time. Taghipourian et al. [16] proposed a dynamic virtual hub location problem using the fuzzy integer linear programming approach to minimize the transportation cost in the network. In recent years, many scholars have also studied the problem of facility locations in logistics networks and intermodal transportation networks, including Mohammadi et al. [17], Correia et al. [18], Lin and Lin [19], and Teye et al. [20].

The study on the location planning of highway passenger transportation hubs is mainly based on the gravity method. However, due to the complexity of the location problem, some studies have not fully considered such factors as population distribution and road network structure. Due to the development of GIS technology in the field of transportation engineering [21, 22] and the availability of POI data [23–25], it is of practical significance to study the new method of intercity highway passenger transport hub location for improving the rationality of hub location decisions and the calculation efficiency of hub location models. The acquisition of basic information is the basis for determining the hub location, and the comprehensiveness of the basic information determines whether the hub location is reasonable and effective. The travel information of travelers can be accurately obtained through the traditional manual questionnaire survey method. However, due to the low sampling rate, this method makes it difficult to truly reflect the population distribution of the whole city, especially for large cities. With the development and popularization of the MAP APP, various types of POI information can be used to reflect the distribution of urban travel demand, greatly reducing the difficulty of information collection and processing.

The objective of this study is to propose a more applicable and effective hub location method for highway passenger transport hubs in large cities. This method is based on POI data and the Dijkstra algorithm, which greatly reduces the complexity of traffic demand analysis in the survey. It also considers the process of passenger travel and the location of exits connecting urban roads and surrounding highways. The proposed method has a good application prospect for large cities with difficulty conducting traffic surveys and measuring traffic pressure. The proposed method can greatly reduce the complexity of traffic demand investigation. It has a good application prospect for large cities with large POI data volumes and heavy traffic pressure inside the city.

1.2. Contributions. (1) The POI data is used to approximate the external traffic travel demand. The sample size is large and distributed throughout the city, greatly reducing the difficulty and cost of a traffic survey. (2) Based on the principle that “passenger transportation hubs should be far away from the city center,” the total time cost from the starting point to the highway passenger transportation hub

and from the highway passenger transportation hub to the exit connecting urban roads and surrounding highways is considered in the proposed mathematical model so as to make the hub location more suitable for the development trend of large cities.

2. Methodology for Determining Highway Passenger Transportation Hub Locations

2.1. Data. It is assumed that the distribution of highways around the city is uniform and reasonable. The departure frequency of intercity highway transportation lines is relatively dense and continuous, and the departure interval is short. From the origin to the exit, passengers mainly go through three important nodes: the trip generation node in the city (residential area, O), then the highway passenger transportation hub, and finally the highway exit in the periphery of the city (D) [12] (Figure 1). That is, the external travel of passengers can be expressed as: origin i —road network node i' near i —road network node k' near k —hub k —road network node k' near k —road network node j' near j —highway exit j .

The POI data are obtained by Python from the API of MAP. The data are classified into three categories: (a) the coordinates of residential areas in the city are used as the origin set (O set) for trip generation and the basis for traffic demand forecasting; (b) the latitude and longitude coordinates of the road network nodes are the foundation for establishing the road network model and calculating path choices; and (c) the information on the locations of exits connecting urban roads to peripheral highways (exits for short) composes the destination set (D set). In the analysis process, the POI data is pretreated by filtering, classifying, and sorting. The coordinates are converted into plane coordinates, and all coordinate operations are performed in the WGS-84 plane coordinate system.

2.2. Mathematical Model. The process by which highway passengers leave the city includes two stages. The first stage is from the origin to the highway passenger transportation hub. The second stage is from the highway passenger transportation hub to the exit, connecting urban roads and peripheral highways. The corresponding travel distance L can be expressed as

$$L = L_{i \rightarrow k, \min} + L_{k \rightarrow j, \min}, \quad (1)$$

where L is the travel distance, $L_{i \rightarrow k, \min}$ is the shortest travel distance from residential area i ($i \in O$) in the origin set O to the highway passenger transportation hub k ($k \in T$, where T is the set of highway passenger transportation hubs), and $L_{k \rightarrow j, \min}$ is the shortest travel distance from highway passenger transportation hub k to exit j ($j \in D$, where D is the set of exits) connecting urban roads to peripheral highways.

$$L_{i \rightarrow k, \min} = L_{i \rightarrow i', \text{near}} + L_{i' \rightarrow k', \min} + L_{k' \rightarrow k, \text{near}}, \quad (2)$$

where $L_{i \rightarrow i', \text{near}}$ is the distance from residential area O in the origin set O to road network node i' near i , $L_{i' \rightarrow k', \min}$ is the shortest distance from i' to highway passenger transportation hub k , and $L_{k' \rightarrow k, \text{near}}$ is the distance from k to road network node k' near k .

$$L_{k \rightarrow j, \min} = L_{k \rightarrow k', \text{near}} + L_{k' \rightarrow j', \min} + L_{j' \rightarrow j, \text{near}}, \quad (3)$$

where $L_{k' \rightarrow j', \min}$ is the shortest distance from k' to road network node j' near exit j , $L_{j' \rightarrow j, \text{near}}$ is the distance from road network node j' to exit j .

In order to minimize the urban traffic pressure caused by highway passenger transportation hubs and reflect urban traffic conditions, the time cost from the origin to the highway passenger transportation hub and the time cost of passengers leaving the city are considered in this paper. Therefore, the path selection of passengers in the urban road network and the route selection of highway passenger transportation leading to the exit are important components of the solution algorithm. The objective function can be expressed as minimizing the total time cost in the WGS-84 plane coordinate system, i.e.,

$$\min f = \sum_{k=1}^g \left(\sum_{i=1}^n Q_{ik} \cdot L_{i \rightarrow k, \min} + \sum_{j=1}^m Q_{kj} \cdot L_{k \rightarrow j, \min} \right). \quad (4)$$

Let

$$f_k = \sum_{i=1}^n Q_{ik} \cdot L_{i \rightarrow k, \min} + \sum_{j=1}^m Q_{kj} \cdot L_{k \rightarrow j, \min}. \quad (5)$$

Then,

$$\min f = \sum_{k=1}^g f_k, \quad (6)$$

where n is the number of observations in the set O , that is, the number of origins, Q_{ik} is the number of passengers from residential area i traveling to passenger transportation hub k , m is the number of exits corresponding to passenger transportation hub k , Q_{kj} is the total number of passengers on highway passenger transportation lines that choose exit j from passenger transportation hub k , and g is the number of highway passenger transportation hubs.

Different from the traditional gravity model and the P -median location model, the goal of this study is to use a completely quantitative method to make the location of highway passenger transportation hub not in the urban center. Compared with hubs located in urban centers, passenger transportation hubs located at the edge of large cities can reduce the travel time cost of passengers and improve the operation efficiency of hubs. The specific reasons can be analyzed as follows:

- (a) As shown in equation (5), the cost of travelers' trips can be divided into two parts: the cost of travel from the origin point to the hub, $\sum_{i=1}^n Q_{ik} \cdot L_{i \rightarrow k, \min}$; and the cost of travel from the hub to the highway exit, $\sum_{j=1}^m Q_{kj} \cdot L_{k \rightarrow j, \min}$. Because the origin points (residential areas) are relatively uniformly distributed

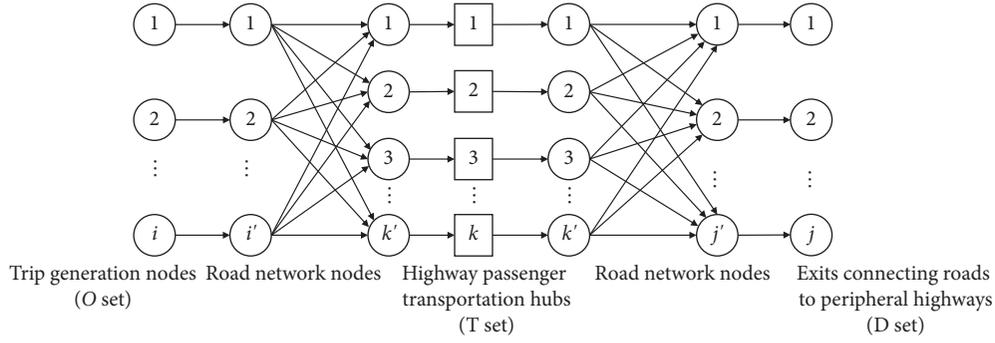


FIGURE 1: Trips from trip generation nodes (O) to highway exits (D).

across the urban area, the hubs are located in the central area, which reduces $\sum_{i=1}^n Q_{ik} \cdot L_{i \rightarrow k, \min}$. However, at the same time, since the hubs are moved from the central areas to the edges of the city, $\sum_{j=1}^m Q_{kj} \cdot L_{k \rightarrow j, \min}$ greatly decreases. Therefore, the total time cost for passengers does not increase.

- (b) The proposed method takes into account the grade of the road, and the resulting locations are closer to the highway exits. This reduces the time costs for travelers to an extent.
- (c) The passenger transportation hub in the central area of the city increases the risk of congestion in the nearby urban road network, which greatly increases the delay of intercity coaches leaving the city, thus increasing the time cost of passengers leaving the city.

3. Solution Algorithm

3.1. *Basic Flow.* The specific steps of the solution algorithm are as follows (Figure 2):

Step 1: the number of highway passenger transportation hubs is determined according to the size of the city, external traffic demand, and highway exits.

Step 2: the highway exits are divided into g groups according to the exit direction and the coordinate positions of exits. Each group of exits corresponds to a passenger transportation hub. The set D is divided into g subsets D_k . D_k is used to select the set of exits serviced by passenger transportation hub k .

Step 3: the initial value of k is set to 1.

Step 4: the value of k is assessed. If $k \leq g$, the calculation proceeds to step 5. If $k > g$, the calculation ends.

Step 5: travel volume Q_{ik} is determined from origin i to highway passenger transportation hub k . The initial iteration point (x_0, y_0) and iteration step s are set.

Step 6: the value of x_k is determined. If x_k is within the calculation area, the calculation goes to step 7. If x_k exceeds the calculation area, the calculation goes to step 10.

Step 7: the value of y_k is assessed. If y_k is within the calculation area, the calculation goes to step 8. If y_k exceeds the calculation area, the calculation goes to step 9.

Step 8: the shortest distances from origin $i (i \in O)$ to (x_k, y_k) and from (x_k, y_k) to exit $j (j \in D_k)$ are calculated. The weight α_{kj} of exit j for (x_k, y_k) is calculated. Based on the weight, the distribution of passenger flow to exit j and the shortest distance from the passenger transportation hub to the exit are determined. The objective function f_k is calculated, and the results are recorded. The formula $y_k = y_k + s$ is set, and the calculation returns to step 7.

Step 9: $x_k = x_k + s$ and $y_k = y_0$. Then, the calculation returns to step 6.

Step 10: all the objective function values f_k are traversed, and the minimum f_k is obtained. Then, the (x_k, y_k) corresponding to the minimum f_k is returned. The formula $k = k + 1$ is set, and the calculation returns to step 4.

Due to the large amount of POI data, a large number of OD pairs are used in the calculation process, which leads to more iterations. By debugging the actual script, the results show that the time complexity of the process of searching for paths is high for a single iteration (e.g., the time complexity of the Floyd algorithm and Dijkstra algorithm are $O(n^3)$ and $O(n^2)$, respectively). The time cost of calculation is mainly due to the search for paths from the origins to the hubs and from the hubs to the exits connecting the urban roads to the peripheral highways. Hence, improving the efficiency of the path search process is the key to improving the overall efficiency of the proposed method. Thus, the Dijkstra algorithm is used to reduce the time cost and intermediate storage utilization of the path search process. For the whole method, the iteration step should not be too small, and an appropriately sized step can effectively reduce the number of iterations and storage utilization.

In engineering practice, compared with urban internal traffic, the external traffic volume of the city is small. Therefore, path capacity is not considered in the Dijkstra algorithm in this study. It adopts an assignment method similar to "all or nothing" and assumes that travellers choose the shortest path. This assumption may differ from the actual situation to some extent, but it is acceptable for the macro problems of urban hub planning. In existing studies, Alumur et al. have adopted the same assumption, and path capacity is not considered in the multimodal hub location problem [3].

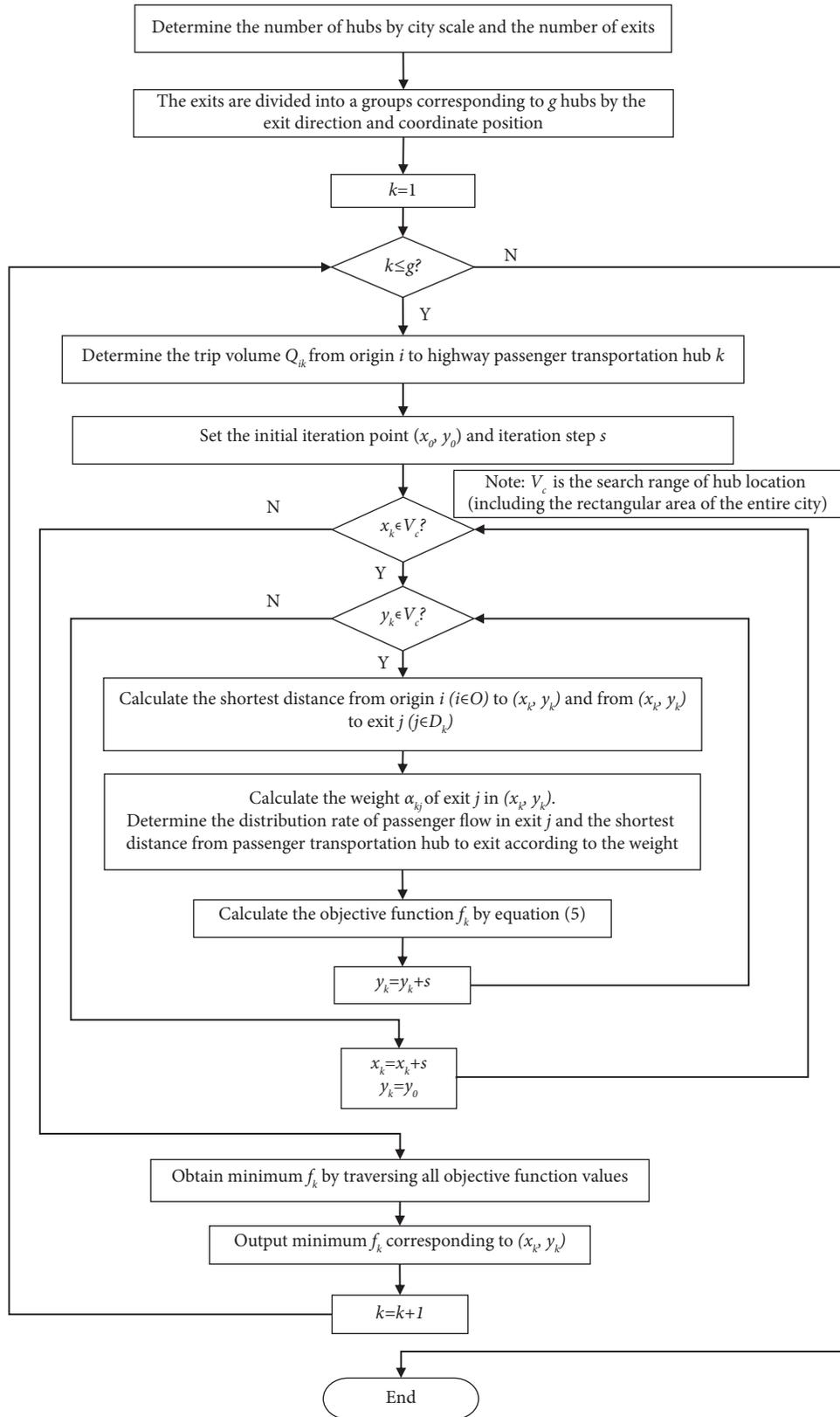


FIGURE 2: Flowchart of the solution algorithm.

3.2. *Path Choice of Passengers.* Passengers' choice of path to highway passenger transportation hubs and their selection of highway passenger transportation lines to exits are important parts of the solution algorithm.

3.2.1. *Shortest Path for Passengers to the Highway Passenger Transportation Hub.* It is assumed that vehicles are used for passenger travel. Finding the shortest path is a classic problem in operational research, and a variety of mature algorithms have been proposed in the existing literature. The Dijkstra algorithm [25] is the basis of the travel path solution because of its high computational efficiency.

The construction of the road network structure matrix M is the core part of the algorithm for identifying the shortest path. The model in this study is based on a path finding problem that considers a traveler's travel distance (or travel time). When the speed is assumed to be consistent, the essence of time and distance in the problem is the same. A more realistic calculation of travel time (subway travel time, etc.) between the departure node and the hub will certainly improve the effectiveness of the model. However, the amount of travel demand data used in this study is huge, and the location of the hub is constantly updated during the calculation process. Due to the computing power of the workstation, it is difficult to accurately fit the travel time between any OD pairs. Therefore, it is assumed that the average travel speed of roads of the same grade is consistent in long-distance travel, and uniformly use the travel distance (travel time) of automobile travel between OD pairs as a reference in the model. The travel time of the road section is calculated based on the speed, and the waiting time at the intersection is a simple approximation according to the different road grades.

Based on a traditional linear distance matrix, the travel time delay of a given intersection is converted into an equivalent distance to calculate the effective distance between nodes. The calculation method is as follows:

$$l_{pq} = l_{pq,pra} + l_{pq,del}, \quad (7)$$

$$l_{pq,pra} = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}, \quad (8)$$

$$l_{pq,del} = n \cdot v \cdot t,$$

where l_{pq} is the effective distance between road network nodes p and q , $l_{pq,pra}$ is the actual road length between road

network nodes p and q , $l_{pq,del}$ is the equivalent distance caused by intersection delays, x_p, y_p, x_q, y_q are the plane coordinates of road network nodes p and q , n is the number of intersections, v is the average travel speed between nodes, and t is the average delay due to the intersection.

Then, the road network structure matrix M can be constructed based on the effective distance l_{pq} .

$$M = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1h} \\ l_{21} & l_{22} & \cdots & l_{2h} \\ \vdots & \vdots & \ddots & \vdots \\ l_{h1} & l_{h2} & \cdots & l_{hh} \end{bmatrix}_{h \times h}, \quad (9)$$

where h is the number of road network nodes.

If there is no direct link between road network nodes p and q ,

$$l_{pq} = \infty. \quad (10)$$

If there is a direct link between road network nodes p and q , the urban road grades are taken into account and equation (7) can be expressed as

$$l_{pq} = \begin{cases} \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}; & \text{if } p \text{ in } P_E \text{ and } q \text{ in } P_E, \\ \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} + n \cdot v \cdot t; & \text{if } p \text{ in } \overline{P}_E \text{ or } q \text{ in } \overline{P}_E, \end{cases} \quad (11)$$

where P_E is the set of nodes along the urban expressways, \overline{P}_E is the set of nodes along other grades of urban roads.

The Dijkstra algorithm and the effective distance are used to calculate the link set $P_{i' \rightarrow k'}$ and $P_{k' \rightarrow j'}$ of the optimal travel path. The travel path is from the road network node i' near origin i to the road network node k' near passenger transportation hub k , and from k' to the road network node j' near exit j . The effective distance of the optimal travel path can be expressed as

$$L_{i' \rightarrow k', \min} = \sum_{pq \in P_{i' \rightarrow k'}} l_{pq}, \quad (12)$$

$$L_{k' \rightarrow j', \min} = \sum_{pq \in P_{k' \rightarrow j'}} l_{pq},$$

where l_{pq} is the effective distance of the shortest path for that link.

Then, equation (4) can be expressed as

$$\min f = \sum_{k=1}^g \left[\sum_{i=1}^n Q_{ik} \cdot \left(\sqrt{(x_i - x_{i'})^2 + (y_i - y_{i'})^2} + \sum_{pq \in P_{i' \rightarrow k'}} l_{pq} + \sqrt{(x_k - x_{k'})^2 + (y_k - y_{k'})^2} \right) \right. \\ \left. + \sum_{j=1}^m Q_{kj} \cdot \left(\sum_{pq \in P_{k' \rightarrow j'}} l_{pq} + \sqrt{(x_k - x_{k'})^2 + (y_k - y_{k'})^2} + \sqrt{(x_j - x_{j'})^2 + (y_j - y_{j'})^2} \right) \right], \quad (13)$$

where x_i, y_i are the plane coordinates of origin i , $x_{i'}, y_{i'}$ are the plane coordinates of road network node i' near origin i , x_k, y_k are the plane coordinates of passenger transportation hub k , $x_{k'}, y_{k'}$ are the plane coordinates of road network node k' near passenger transportation hub k , x_j, y_j are the plane coordinates of exit j , and $x_{j'}, y_{j'}$ are the plane coordinates of road network node j' near exit j .

3.2.2. Selection of Highway Passenger Transportation Lines to Exits. There are more signal controls on urban roads, lower speed restrictions, and longer travel distances in cities, which inevitably lead to delays. Therefore, intercity coach drivers prefer to enter the expressway quickly. Travel distance is the main factor for drivers to decide where to exit. The distance between hub k and highway exit j and the probability of exit j being selected are negatively correlated. Therefore, weighted random numbers are generated to simulate the selection of highway exits by highway passenger transportation lines when coaches leave the city. The concept of the selection weight α_{kj} is defined as the probability of a coach at hub k selecting exit j to leave the city. For example, there are two exits (denoted as exit 1 and exit 2) for coaches at hub 1 to leave the city. The distances between hub 1 and exit 1 and exit 2 are 10 km and 20 km, respectively. The probabilities of drivers selecting exit 1 and exit 2 will be 0.66 and 0.33, respectively.

$$\alpha_{kj} = \frac{d_{kj}^{-1}}{\sum_{k=1}^m d_{kj}^{-1}}, \quad (14)$$

$$d_{kj} = \sqrt{(x_j - x_k)^2 + (y_j - y_k)^2},$$

where m is the number of exits, d_{kj} is the distance from passenger transportation hub k to exit j .

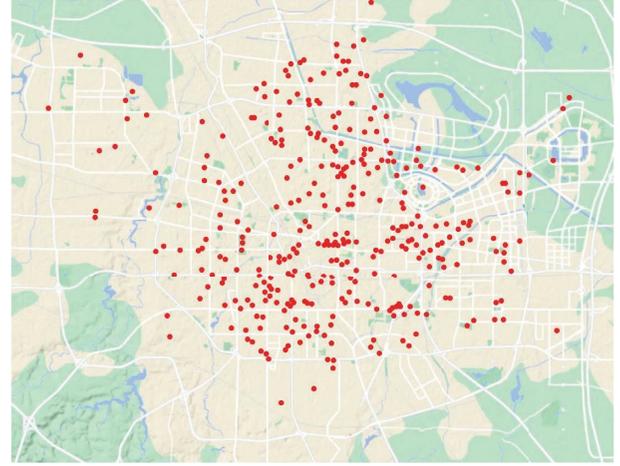
The script is compiled in Python based on POI data. All selection weights $\alpha_{kj} (\forall k \in T, \forall j \in D)$ can be calculated. Using a random number generator, weighted random numbers are generated to simulate the selection of highway passenger transportation lines for freeway exits. Then,

$$Q_{kj} = \alpha_{kj} \cdot \sum_{i=1}^n Q_{ik}. \quad (15)$$

4. Case Study

4.1. Data Acquisition and Urban Road Network Model Construction. Zhengzhou, a large city with more than 10 million people, is taken as an example. The latitude and longitude coordinates of each residential area are obtained based on the open map platform. A total of 732 valid data are screened and retained (Figure 3).

A total of 92 important nodes on urban expressways and major arterial roads within the 4th Ring Expressway in Zhengzhou city are selected to construct the urban arterial road network (Figure 4).



• Origins

FIGURE 3: Distribution of residential areas.



• Nodes on urban expressways and major arterial roads

FIGURE 4: Urban arterial road network.

The road network structure matrix M is constructed using the NumPy package in Python, as shown in equation (16).

$$M = \begin{bmatrix} +\infty & 3853.24 & +\infty & +\infty & +\infty & \cdots & +\infty \\ 3853.24 & +\infty & 4966.84 & +\infty & +\infty & \cdots & +\infty \\ +\infty & 4966.84 & +\infty & 1735.27 & +\infty & \cdots & +\infty \\ +\infty & +\infty & 1735.27 & +\infty & 1994.97 & \cdots & +\infty \\ +\infty & +\infty & +\infty & 1994.97 & +\infty & \cdots & +\infty \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & +\infty \\ +\infty & +\infty & +\infty & +\infty & +\infty & \cdots & +\infty \end{bmatrix}_{91 \times 91}. \quad (16)$$

The set of nodes along urban expressway P_E and the set of nodes along other grades of roads \bar{P}_E are as follows:

$$P_E = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 28, 35, 40, 60, 70, 85, 86, 84, 83, 82, 81, 80, 61, 51, 46, 91, 90, 89, 88, 87, 27, 34, 39, 59, 69, 68, 67, 66, 65, 64, 63, 62, 52, 47, 41, 22, 29, 36, 37, 38, 20, 25, 32, 57, 77\},$$

$$\overline{P}_E = \{17, 18, 19, 21, 23, 24, 26, 30, 31, 33, 42, 43, 44, 45, 48, 49, 50, 53, 54, 55, 56, 58, 71, 72, 73, 74, 75, 76, 78, 79\}.$$

(17)

4.2. Hub Location Calculation and Analysis. For hub location selection, an appropriate set of candidate hub locations is usually determined according to urban land use planning. Then, mathematical modeling methods (such as integer programming) are used to find the optimal result from a limited set of candidate hub locations. The set of candidate hub locations selected based on land-use planning can greatly reduce the computational complexity. The objective of this study is to provide an effective and practical quantitative mathematical method for highway passenger transportation hub location. For large cities, a large number of POIs are used as the travel demand input. Whether the calculation time is too long becomes the main concern in determining whether the proposed method is practical. In order to verify whether the calculation time of the method is acceptable, it is assumed that all land-use properties meet the needs of the hub. It is assumed that external travel demand is equal for each origin and the travel volume in each direction is equal. The average speed and average delay due to intersections are obtained by investigating major arterial road data and road spacing, as shown in Table 1.

The parameters in equation (11) can be set based on the data in Table 1, i.e.,

$$l_{ij} = \begin{cases} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, & \text{if } i \text{ in and } j \text{ in } P_E, \\ 1.34\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, & \text{if } i \text{ in } \overline{P}_E \text{ or } j \text{ in } \overline{P}_E. \end{cases} \quad (18)$$

The locations of all freeway exits are shown in Figure 5. With the WGS-84 plane coordinate system, the origin point for iteration is the southwest boundary point (3835287, 450863). The destination for iteration is the northeast boundary point (3859296, 484729). The calculation area contains the entire urban built-up area. In combination with the principle of reducing the impact of hubs on the urban center and balanced development, the target number of passenger transportation hubs is set to be 4. In practice, there are also four hubs in Zhengzhou city. According to the distribution of freeway exits, the toll stations on the freeway are categorized into four groups: eastern, western, southern, and northern. The group for each direction is taken as the candidate set for coaches going out of the city, which starts from the hub in the same direction as the group. Then, the locations of four hubs are calculated by using 2000 m as the iteration step. The configuration of the running environment and the runtime overhead for the algorithm scripts are recorded in this study. With an ordinary small workstation (Intel (R) Xeon (R) CPU E5-2637v3@3.5 GHz RAM: 64 GB) as the running environment, the run time for the scripts is 5054 s after the optimization of the proposed method [26, 27].

There are relatively few studies on the location of highway passenger transportation hubs. The P -median location model is the most widely used in related studies on

multihub locations. Therefore, the calculation results of the proposed method are compared with the P median location model to verify its rationality. First, the same origin points and urban road network structures are set for the two methods, as shown in Figures 3 and 4. The number of highway passenger transportation hubs for the two methods is both set at 4. The iteration step for the proposed method is set at 2 km. Similarly, the search area is assigned candidate sites for the P -median method at intervals of 2 km. Results of the two methods are shown in Table 2 and Figure 6.

The results obtained by the proposed method are closer to the urban periphery than those obtained by the P -median location method (Table 3 and Figure 6). The results obtained by the P -median location method are more conducive for travelling to the hubs. However, the results of the proposed method reduce the time cost of intercity coaches leaving the city, which has advantages in reducing the total time cost.

By comparing the two methods, the following can be seen:

- (1) The results of the P -median location model are closer to the city interior and more evenly distributed in the middle of the origin points. It is easier for passengers to reach the nearest highway passenger transportation hub.
- (2) The calculation result of the proposed method is closer to the city periphery and the highway exit, which is the result of comprehensive consideration of the process of passenger lines leaving the city and the grade of urban roads. Although the hub's distance from the urban center increases the travel time cost of travelers from the starting point to the highway passenger transportation hub, the final destination of travelers is not the highway passenger hub. Therefore, highway passenger transportation hubs near the periphery of the city greatly reduce the time cost of long-distance intercity coaches leaving the city, and the overall time cost of travelers will not increase.
- (3) The objective of the P -median location method is to obtain the optimal total travel cost for all travelers to the nearest hub. However, it should be noted that this advantage is based on the premise that each hub contains all intercity coach lines, that is, travelers can travel along any line they need from the nearest hub. However, in actual operation, different intercity coach lines are located in different hubs. The proposed method arranges all intercity coach lines in the same direction at the hubs in the corresponding direction, which is more consistent with the layout of intercity coach lines and thus reduces the cost of allocating the same long-distance intercity coach lines to different hubs.

TABLE 1: Mean values of traffic parameters of major arterial roads in the arterial road network.

Parameter	Travel speed (km/h)	Delay of intersection (s)	Road spacing (m)
Mean value	39.1	12.6	400

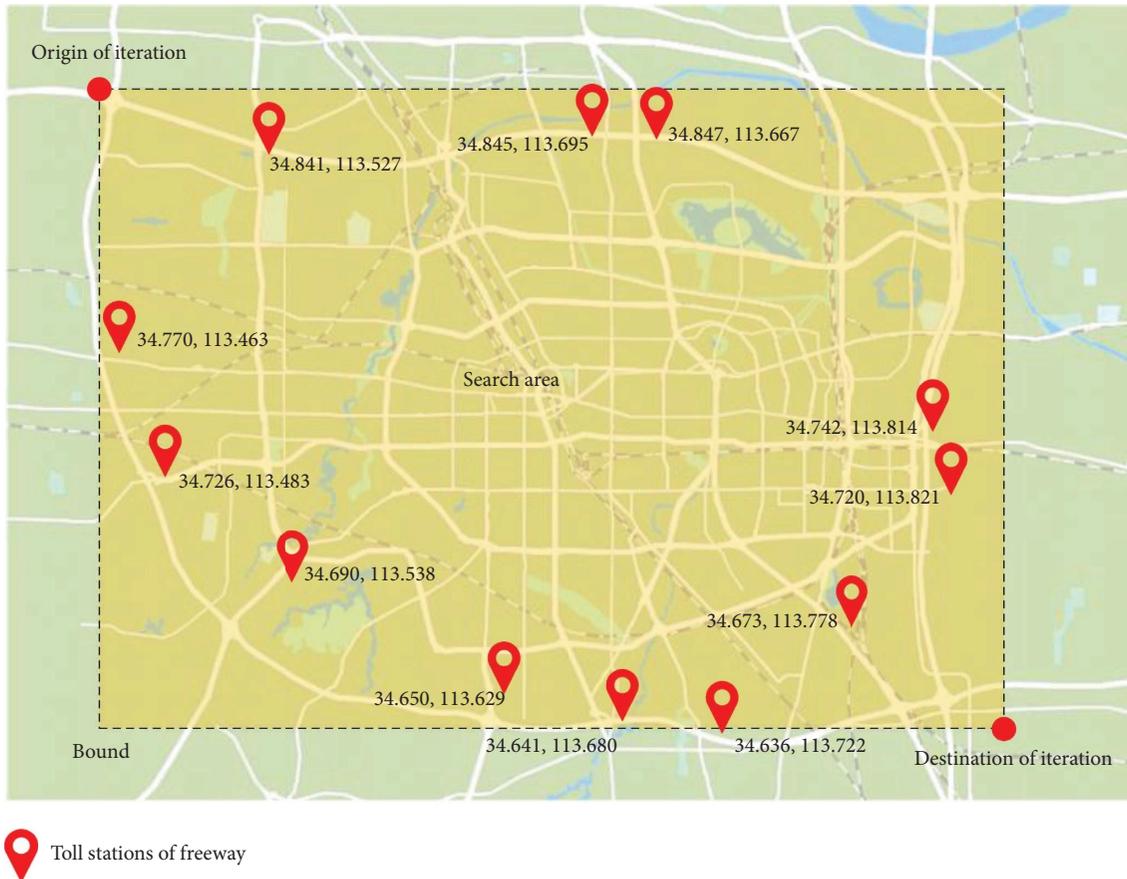


FIGURE 5: Highway exits and the range of calculation area.

TABLE 2: Calculation results of two methods (unit: °).

Method	Eastern hub (longitude, latitude)	Western hub (longitude, latitude)	Southern hub (longitude, latitude)	Northern hub (longitude, latitude)
<i>P</i> -median location method	(113.73, 34.72)	(113.55, 34.79)	(113.64, 34.74)	(113.68, 34.81)
Proposed method	(113.81, 34.74)	(113.53, 34.74)	(113.68, 34.66)	(113.68, 34.84)

The time consumption of the *P*-median method depends on the size of the candidate location set. With the increase in the number of candidate locations, the calculation amount of *P*-median method increases exponentially. In order to verify the effectiveness of the proposed method and its applicability in large cities, the proposed method was compared with the *P*-median method at different scales under the same scenario, assumptions, computing environment, and scripting language. The results are shown in Table 4. The heuristic search algorithm is used for the *P*-median method, and the script structure is a single thread.

As the number of candidate locations increases, the time consumption of *P*-median gradually increases. Compared

with the *P*-median method, the time consumption of the proposed method is lower. Moreover, the proposed method makes the hub location closer to the periphery of large cities, which conforms to the hub location principle. In addition, although the time cost of the proposed method from the starting point to the hub is higher than that of the *P*-median method, it has advantages in the total time cost. Therefore, considering the time consumption and total time cost, the proposed method is feasible in large cities.

By comparing the hub location calculated by the proposed method with the actual hub location in use, some additional conclusions can be drawn. The calculation results of the northern hub are almost consistent with the current hub location, while the calculation results of the eastern,

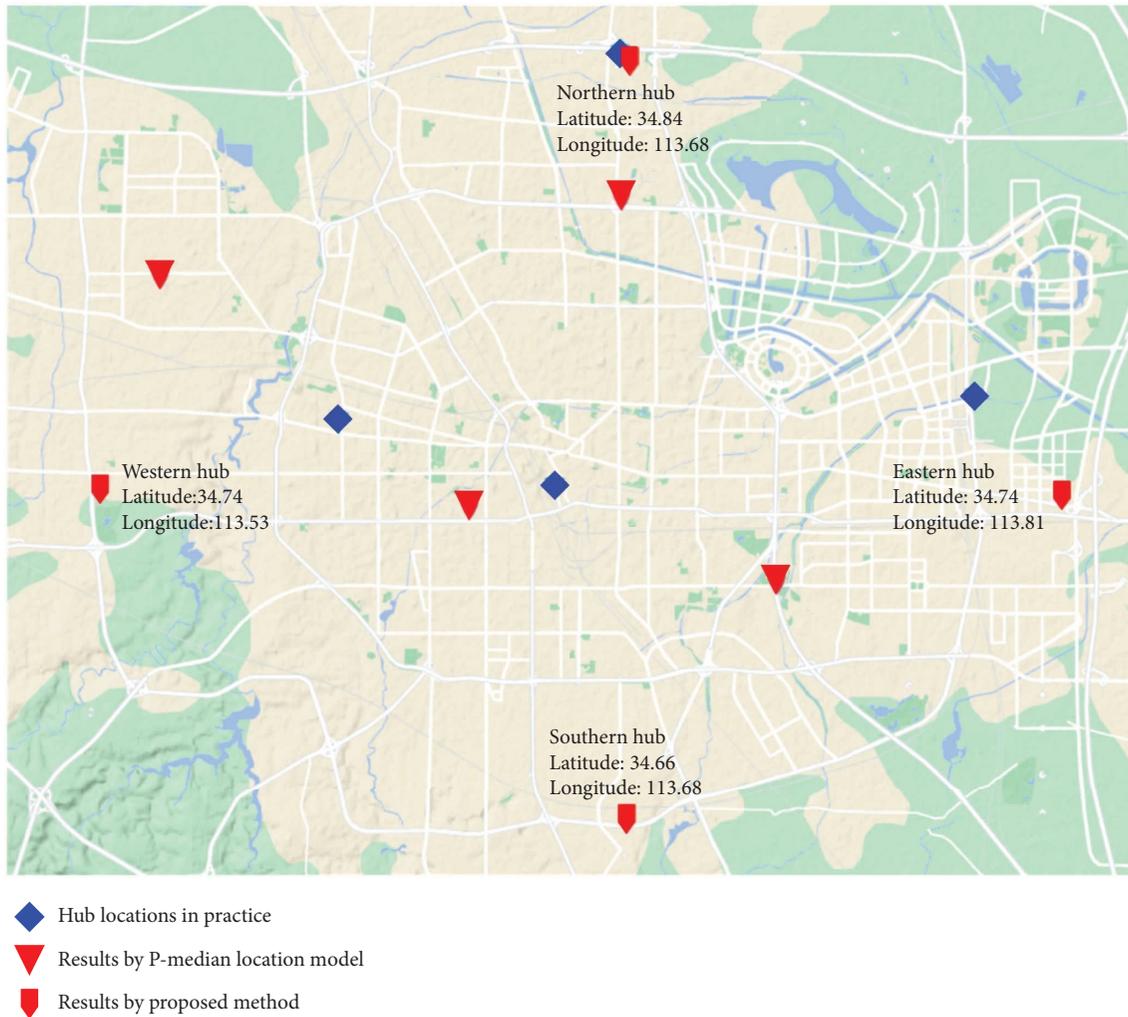


FIGURE 6: Calculated hub locations and the current hub locations in use.

TABLE 3: The time cost of two methods (unit: h).

Method	<i>P</i> -median location method	Proposed method
Time cost to hubs	314.94	1024.06
Time cost from hubs to highway exits	948.76	141.37
Total time cost	1263.70	1165.43

TABLE 4: Comparison of calculation results.

Methods	Proposed method	<i>P</i> -median (10 candidate points)	<i>P</i> -median (20 candidate points)	<i>P</i> -median (30 candidate points)
Calculation time consumption (s)	5054	1287	8941	13514
Distance to the city center	Far	Near	Near	Near
Time cost to hubs (h)	1024.06	314.94	305.64	302.19
Time cost from hubs to freeway exits (h)	141.37	948.76	939.64	944.02
Total time cost (h)	1165.43	1263.70	1245.28	1246.21

western, and southern hubs are closer to the edge of the city than the current hub. This is because the eastern, western, and southern hubs were built relatively early, while the northern hub was newly built. Therefore, it is in line with the trend of urban hub location development to set up highway passenger transportation hubs outside large cities.

5. Discussion and Conclusions

External traffic planning plays an important role in comprehensive urban traffic planning. The location of highway passenger transportation hubs directly affects urban traffic operations. When determining the location of a highway

passenger transportation hub, due to the rapid development of urbanization, the traditional gravity method and P -median location model cannot well adapt to the traffic conditions of large cities. However, the emerging analytic hierarchy process is deeply influenced by the subjective will of decision-makers and planners, and its objective theoretical basis is insufficient. This paper simplifies the use of traffic surveys and demand forecasts based on the distribution of the residential population by using POI data. The mathematical model of highway passenger transport hub location is established by considering the time cost of leaving the city, and the improved Dijkstra algorithm is used to simulate the travel paths of passengers.

A large city, Zhengzhou, is used as a case study. The calculation results are compared with the results of the P -median location model and the current locations of the hubs in practice. It is found that the proposed method is feasible in view of time consumption and total time cost. The hub locations from the proposed method are closer to the periphery of the city and are close to the urban expressway nodes, which accords with the principle of hub location. The proposed method has good applicability and feasibility for large cities with large POI data, heavy traffic pressure, and difficulty in traffic surveying.

However, there are still some limitations to this study. The proposed method uses a large number of residential POI data to replace the traditional traffic generation. Using POI data to simulate traffic demand is approximate and does not consider nonresidential travel, which has some deviation from the actual traffic demand. In the future, different kinds of trips will be considered in traffic demand based on realistic land use. Besides, the average travel speed of roads of the same grade is assumed to be consistent, and there is an error in the realistic calculation of travel time (subway travel time, etc.) between the departure node and the hub. In the follow-up research, more reasonable methods will be introduced to improve the accuracy of travel times. "All or nothing" is adopted in traffic assignments based on the shortest path, and path capacity is not considered. In future extensions, the capacity limit of the path will be an important factor to consider.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Tackle Key Problems in Science and Technology Project of Henan Province (No. 222102240052) and the Foundation for High-Level Talents of Henan University of Technology (No. 2018BS029).

References

- [1] Institute of Transportation Engineers and M. D. Meyer, *Transportation Planning Handbook*, Wiley, New Jersey, NJ, USA, 2016.
- [2] Transportation Research Board, *Transportation System Planning*, Transportation Research Board, Washington, DC, USA, 1984.
- [3] S. A. Alumur, B. Y. Kara, and O. E. Karasan, "Multimodal hub location and hub network design," *Omega*, vol. 40, no. 6, pp. 927–939, 2012.
- [4] H. L. Bi, K. Kang, and X. T. Zhang, "A biobjective and trilevel programming model for hub location problem in design of a resilient power projection network," *Mathematical Problems in Engineering*, vol. 2016, Article ID 5329196, 9 pages, 2016.
- [5] M. Hosseini and S. A. MirHassani, "Refueling-station location problem under uncertainty," *Transportation Research Part E: Logistics and Transportation Review*, vol. 84, pp. 101–116, 2015.
- [6] M. A. Pereira, L. C. Coelho, L. A. N. Lorena, and L. C. de Souza, "A hybrid method for the probabilistic maximal covering location-allocation problem," *Computers & Operations Research*, vol. 57, pp. 51–59, 2015.
- [7] L. Verdonck, P. Beullens, A. Caris, K. Ramaekers, and G. K. Janssens, "Analysis of collaborative savings and cost allocation techniques for the cooperative carrier facility location problem," *Journal of the Operational Research Society*, vol. 67, no. 6, pp. 853–871, 2016.
- [8] T. T. Li, R. Song, S. W. He, M. K. Bi, W. C. Yin, and Y. Q. Zhang, "Multiperiod hierarchical location problem of transit hub in urban agglomeration area," *Mathematical Problems in Engineering*, vol. 2017, Article ID 7189060, 15 pages, 2017.
- [9] T. H. Tran, G. Nagy, T. B. T. Nguyen, and N. A. Wassan, "An efficient heuristic algorithm for the alternative-fuel station location problem," *European Journal of Operational Research*, vol. 269, no. 1, pp. 159–170, 2018.
- [10] Z. Chen and W. He, "Study and application of Center-of-Gravity on the location selection of distribution center," in *Proceedings of the 10th International Conference on Intelligent Systems Design and Applications*, Harbin, China, July 2010.
- [11] M. Diao, Y. Zhu, and J. R. Zhu, "Intra-city access to inter-city transport nodes: the implications of high-speed-rail station locations for the urban development of Chinese cities," *Urban Studies*, vol. 54, no. 10, pp. 2249–2267, 2017.
- [12] B. Fan, Y. Yang, and L. Li, "Integrated optimization of urban agglomeration passenger transport hub location and network design," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, pp. 168–177, 2018.
- [13] Y. Zhu, L. Li, and M. Liu, "Research on location of comprehensive passenger transport hub in small and medium-sized cities based on MIP model," in *Proceedings of the 19th COTA International Conference of Transportation Professionals*, Nanjing, China, June 2019.
- [14] S. Martinov, "Increasing the accuracy of evaluation in selecting a location for establishing an intermodal terminal," in *Proceedings of the MATEC Web of Conferences*, Les Ulis Cedex A, Paris, France, July 2018.
- [15] M. Da Graça Costa, M. E. Captivo, and J. Climaco, "Capacitated single allocation hub location problem-A bi-criteria approach," *Computers & Operations Research*, vol. 35, no. 11, pp. 3671–3695, 2008.

- [16] F. Taghipourian, I. Mahdavi, N. Mahdavi-Amiri, and A. Makui, "A fuzzy programming approach for dynamic virtual hub location problem," *Applied Mathematical Modelling*, vol. 36, no. 7, pp. 3257–3270, 2012.
- [17] M. Mohammadi, R. Tavakkoli-Moghaddam, and A. Siadat, "A game-based meta-heuristic for a fuzzy bi-objective reliable hub location problem," *Engineering Applications of Artificial Intelligence*, vol. 50, pp. 1–19, 2016.
- [18] I. Correia, S. Nickel, and F. Saldanha-da-Gama, "A stochastic multi-period capacitated multiple allocation hub location problem: formulation and inequalities," *Omega*, vol. 74, pp. 122–134, 2018.
- [19] C. C. Lin and S. W. Lin, "Two-stage approach to the intermodal terminal location problem," *Computers & Operations Research*, vol. 67, pp. 113–119, 2016.
- [20] C. Teye, M. G. H. Bell, and M. C. J. Bliemer, "Urban intermodal terminals: the entropy maximising facility location problem," *Transportation Research Part B: Methodological*, vol. 100, pp. 64–81, 2017.
- [21] C. Qi, J. Fang, and L. Sun, "Implementation of emergency logistics distribution decision support system based on GIS," *Cluster Computing*, vol. 22, no. 4, pp. 8859–8867, 2019.
- [22] S. Shahparvari, A. Nasirian, A. Mohammadi, S. Noori, and P. Chhetri, "A GIS-LP integrated approach for the logistics hub location problem," *Computers & Industrial Engineering*, vol. 146, Article ID 106488, 2020.
- [23] T. Zhang, L. Sun, L. Yao, and J. Rong, "Impact analysis of land use on traffic congestion using real-time traffic and POI," *Journal of Advanced Transportation*, vol. 20178 pages, Article ID 7164790, 2017.
- [24] C. Lai and J. Chang, "Image based roadside POI recognition and GPS correction scheme using smart phones," in *Proceedings of the 3rd International Conference on Innovative Computing Technology*, London, UK, June 2013.
- [25] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.
- [26] L. Wang and J. Lu, "A memetic algorithm with competition for the capacitated green vehicle routing problem," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 2, pp. 516–526, 2019.
- [27] S. Gao, M. Zhou, Y. Wang, J. Cheng, H. Yachi, and J. Wang, "Dendritic neuron model with effective learning algorithms for classification, approximation and prediction," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 2, pp. 601–614, 2019.