

## Research Article

# Large-Scale Demand Driven Design of a Customized Bus Network: A Methodological Framework and Beijing Case Study

Jihui Ma,<sup>1</sup> Yang Yang,<sup>1</sup> Wei Guan,<sup>1</sup> Fei Wang,<sup>1</sup> Tao Liu,<sup>2</sup> Wenyuan Tu,<sup>1</sup> and Cuiying Song<sup>1</sup>

<sup>1</sup>MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China

<sup>2</sup>Department of Civil and Environmental Engineering, The University of Auckland, Auckland 1142, New Zealand

Correspondence should be addressed to Yang Yang; 11114218@bjtu.edu.cn

Received 10 January 2017; Revised 10 March 2017; Accepted 16 March 2017; Published 29 March 2017

Academic Editor: Xiaolei Ma

Copyright © 2017 Jihui Ma 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 recent years, an innovative public transportation (PT) mode known as the customized bus (CB) has been proposed and implemented in many cities in China to efficiently and effectively shift private car users to PT to alleviate traffic congestion and traffic-related environmental pollution. The route network design activity plays an important role in the CB operation planning process because it serves as the basis for other operation planning activities, for example, timetable development, vehicle scheduling, and crew scheduling. In this paper, according to the demand characteristics and operational purpose, a methodological framework that includes the elements of large-scale travel demand data processing and analysis, hierarchical clustering-based route origin-destination (OD) region division, route OD region pairing, and a route selection model is proposed for CB network design. Considering the operating cost and social benefits, a route selection model is proposed and a branch-and-bound-based solution method is developed. In addition, a computer-aided program is developed to analyze a real-world Beijing CB route network design problem. The results of the case study demonstrate that the current CB network of Beijing can be significantly improved, thus demonstrating the effectiveness of the proposed methodology.

## 1. Introduction

With the rapid increase in urbanization, many Chinese cities are facing problems associated with urban environments, such as increased traffic congestion, serious environmental pollution, and extreme energy deficiencies. Furthermore, with the gradual increase in the standard of living of urban residents, existing public transportation (PT) has not satisfied the travel demands of passengers, is unable to encourage private car passengers to switch travel modes, and has been unable to improve PT at attractive rates. Therefore, customized buses (CBs) have been established across China as an innovative mode of PT services. Network planning is a major component of CB systems. The rationality and proportionality of network planning play a vital role in the entire CB operational system. Therefore, scientific and systematic research on CB network planning must be conducted. Scientific and reasonable network planning can maximize the use of CB resources, satisfy the travel demands of most

passengers, and improve the service quality of CBs while reducing operating costs for CB operators and improving the attractiveness of CBs. Therefore, the study reported in this paper has great significance for effectively encouraging private car owners to change travel modes, ease traffic congestion, and mitigate the problems of air pollution.

Since the CB concept was first introduced, scholars have conducted a considerable amount of research, the majority of which remains theoretical in nature. In addition, less specific methods have been used for CB research. Kirby and Bhatt [1] analyzed ten specific CB cases and discussed the potential impact of CBs, including easing urban traffic congestion, environmental pollution, and energy consumption. They also developed guidelines for CB passenger recruitment, network planning, operation scheduling, and fares. In addition, the authors in [2] discussed the seven main features of CBs that assured successful operation, such as having more than 50 same points and ends on the long-distance lines, an organization for operations management, constantly adjusting the

lines and scheduling to meet demands, and guaranteeing seats with personalized service. At the time, the customized services were provided by small, private organizations, which faced such problems as increased numbers of commuters, difficulties in the management of operations, and a lack of security and funds for expansion. Therefore, better services could not be provided to passengers. For those reasons, Bautz [3] proposed CBs as a part of urban PT. They compared the cost of CBs to those of different operating modes, and concluded that, compared with private operation organizations, the encouragement and coordination by governmental operations organizations can maximize the benefits of customized PT. McCall [4] analyzed commuter buses that had 47 lines and provided service to more than 2,000 commuters in Ventura, Los Angeles, and Orange Counties. The commuter bus network was a business model whereby the local private sector did not accept subsidies and had only slightly longer travel times than travelling by private car but had considerably lower travel costs than private cars. The successful operation of the commuter buses played a critical role in reducing environmental degradation, but the majority of the PT system required high subsidies for normal operations. McKnight and Paaswell [5] comprehensively analyzed the Chicago CB network. Their results showed that the scope of the Chicago CB service was small, and the main role was to ease peak rail transport travel. A series of steps to expand the PT market were proposed by them. First, the cause of CB travel demand was determined. Then, methods to meet the demands of commuter travel were established using a chart analysis and market research. At the same time, the paper changed rail operations, which were scheduled to support the development of the CBs. Shaheen et al. [6] reviewed the CB development process after the car-sharing concept was proposed and summarized the advantages of CBs: passengers could save costs, and society could reduce the demand for parking spaces and wasted resources. They also suggested that the future development trend of CBs was a constant expansion of the scope of services and the use of more advanced booking technology. Chang and Schonfeld [7] constructed mathematical models for custom transit fixed-route buses and conventional flexible lines. The vehicle size and service space were based on an optimization of the decision variables and total system cost. The objective was to minimize the combination of operational costs and user costs. Martin and Shaheen [8] suggested that passengers encouraged to participate in a car-sharing service could generate significant traffic, land use, and social and environmental benefits, including a reduction in the total number of kilometers of vehicular travel and carbon dioxide emissions. Their study focused on the impact of car sharing on greenhouse gas emissions through theoretical and methodological systems. Potts et al. [9], who studied CB services in the United States and Canada for nearly a decade, proposed different CB modes that were applicable to large, medium, and small cities and rural areas. His research provided guidance on whether and how to open CB lines, combined with the practical situations of CB operators in different areas. Duncan [10] stated that car-sharing services have been vigorously developed in the United States in recent years. Such services enable sustainable

development for an urban transportation system but also could bring greater benefits to cities. The most effective way to increase car-sharing functions was cost savings. Based on the quantification and comparison of the potential cost savings of different travel modes, the car-sharing services are the most cost-effective ones. El Fassi et al. [11] noted that CB operator organizations must continue to improve reticle layouts and increase station capacities to meet growing passenger demands. In addition, policy makers often saw a loss of resources, time, and market share if they made decisions based solely on experience. Thus, providing decision support to policy makers based on a discrete simulation event was proposed. Such decision support could optimize the network, maximize passenger satisfaction, and minimize the number of buses. De Lorimier and El-Geneidy [12] used a multilevel regression analysis method to determine the factors that influenced the effectiveness of decisions to use vehicles based on the CB system in Montreal, Canada. That method provided a reference for building or expanding PT networks for CB operators. Nair and Miller-Hooks [13] constructed a balanced network model to determine the optimal configuration of a CB system. Passengers could take the nearest bus according to their demands. The operator must determine the optimal location of the station, the number of buses, and the station capacity to maximize benefits. Le Vine et al. [14] studied two CB modes: point to point and round trip. This paper suggests that the number of prospective subscribers to a point-to-point CBs in London is between three and four times as large as the comparable number for round-trip CBs. Point-to-point CBs could be used as an alternative to PT, and round-trip CBs could be used as a complement to a point-to-point CB. Liu and Ceder [15] studied the developmental background of Chinese CBs, analyzed China CB network planning and operation processes, and summarized the advantages and disadvantages of CBs and trends in its development in China. In addition, their studies provided a reference for CB operators for policy formulation and academic research.

Compared with the number of studies focusing on the theory of CB, there have been relatively few CB network planning studies. However, there has been increasing research on conventional bus network planning theory and methods. Lampkim and Saalmans [16] studied a specific case. That paper better optimized the entire PT system and achieved more efficient resource utilization by redesigning the network, determined departure frequencies, and developed better timetables and vehicle scheduling than the previous network plan. Ceder and Wilson [17] summarized different bus network planning methods and proposed an easier implementation method in combination with the advantages of a previous method. That paper considered the interests of passengers and operators and presented the design of a new transit network algorithm. Baaj and Mahmassani [18] combined artificial intelligence methods and a genetic algorithm to solve the problem of bus network planning. Tom and Mohan [19] used the operators' cost and total passenger travel time as a total system cost and objective function. First, a series of candidate paths were established. A genetic algorithm was then used to select the line with

minimum cost and determine the departure frequency. Jerby and Ceder [20] proposed that the reasonable planning of a subway PT network could attract passengers to PT. Their method was based on a situation in which a large number of passengers used private cars to travel to a subway transfer station, which caused traffic congestion and a subway station parking space overload. Based on the Rome PT network, Cipriani et al. [21] used a parallel genetic algorithm to solve a complex network topological structure, multimode PT system, large-city PT network planning problem with the characteristics of multi-to-multi bus travel demands. Nikolić and Teodorović [22] used the artificial bee colony algorithm to design a PT network that minimized the total travel time of passengers, maximized passenger satisfaction, and minimized total passenger transfer times as the objective function. Badia et al. [23] studied the PT network planning method of a radial network city. The size of the center area and the departure interval were decision variables. In addition, the operating cost and minimum passenger cost were the objective functions for the preferred line.

In the research of bus data processing methods, there have been relatively few CB data processing methods studies. However, there has been increasing research on conventional bus data processing methods. Trépanier et al. [24] presented a model to estimate the destination location for each individual boarding a bus with a smart card. Experiments carried out with a database programming approach showed that the data must be thoroughly validated and corrected prior to the estimation process. Li [25] investigated statistical inference for a transit route O-D matrix using on-off counts of passengers, created a Markov chain model, and inferred the unknown parameters of the Markov model using Bayesian analysis. After that, Ma et al. [26] developed a Markov chain based Bayesian decision tree algorithm to extract passengers' origin data from recorded SC transaction information. Using the time invariance property of the Markov chain, the algorithm was further optimized and simplified to have a linear computational complexity. Munizaga and Palma [27] presented a methodology for estimating a public transport OD matrix from smartcard and GPS data for Santiago of Chile, obtained detailed information about the time and position of boarding public transportation, and generated an estimation of time and position of alighting for over 80% of the boarding transactions. Ma et al. [28] proposed an efficient and effective data-mining procedure that models the travel patterns of transit riders in Beijing of China and identified the transit riders' trip chains based on the temporal and spatial characteristics of their smart card transaction data. In addition, Ma and Wang [29] attempted to develop a data-driven platform for online transit performance monitoring and Ma et al. [30] developed a series of data-mining methods to identify the spatiotemporal commuting patterns of Beijing public transit riders.

A review of the related literature on CB research demonstrates that CB research has placed more emphasis on the theoretical and practical significance of CBs. Relatively less research has been performed on CB network planning, but there are more references to conventional bus network planning theory. However, as an innovative PT mode, CB has its

own characteristics. This paper proposes an area clustering algorithm based on the travel demands of passengers because CBs currently lack network planning and the CB resource allocation efficiencies are not high. A multiobjective integer programming model is established in combination with the background and significance of CB and comprehensively considering CB operating costs, environmental costs, and traffic congestion costs. Based on the actual CB travel demands in Beijing, this paper reports a case study and demonstrates the rationality and effectiveness of the proposed method. The results of this research provide a new method for CB network planning that will be useful for guiding practical efforts.

The remainder of this paper is organized as follows. In Section 2, the details of the proposed CB network planning method based on area division are proposed. The specific content of the method is described in Section 3. The method is demonstrated through a case study in Section 4. The summary of this paper, limitations, and next steps are discussed in Section 5.

## 2. Details of the Proposed CB Network Planning Method

CB network planning differs from conventional bus network planning in that it is a bottom-up activity based on travel demand. A travel demand survey of passengers is the most important data source for CB network planning, which provides the most important data base for CB network planning. CB travel demand surveys are typically questionnaires provided on the Internet. The contents of the questionnaire are mainly focused on passengers' trip ODs, travel times, and travel purpose. CB provides travel services for passengers based on their travel demands. Therefore, an analysis of passenger travel demand is key to the operation of CBs.

In this paper, considering CB network planning and using concept from a point-to-line layout into a network, the original (O) and destination (D) areas of a line are first divided according to the travel demands of passengers. The operating line scheme is then determined according to the model solution, all operation lines are laid out, and the final CB network is determined.

In this paper, in combination with the conventional bus and existing CB network planning experience, a CB network planning method is proposed based on area division. The specifics of this method are as follows.

(1) *Large-Scale Travel Demand Data Processing.* A travel demand survey was the basis of the transportation planning and provided comprehensive and accurate data for transportation planning. CB travel demand surveys are typically questionnaires provided on the Internet. Because CB operations are based on the actual travel demands of passengers, it is important to determine passengers' trip ODs, travel times, and travel purpose. The passengers' travel demand data were analyzed, and the collected data were quantified to provide data support for network planning.

(2) *CB Line OD Area Division.* CB operators do not use bus stops. The line OD area division is a component of the line and is an important part of network planning. The number of area divisions decreases with increases in the radius of

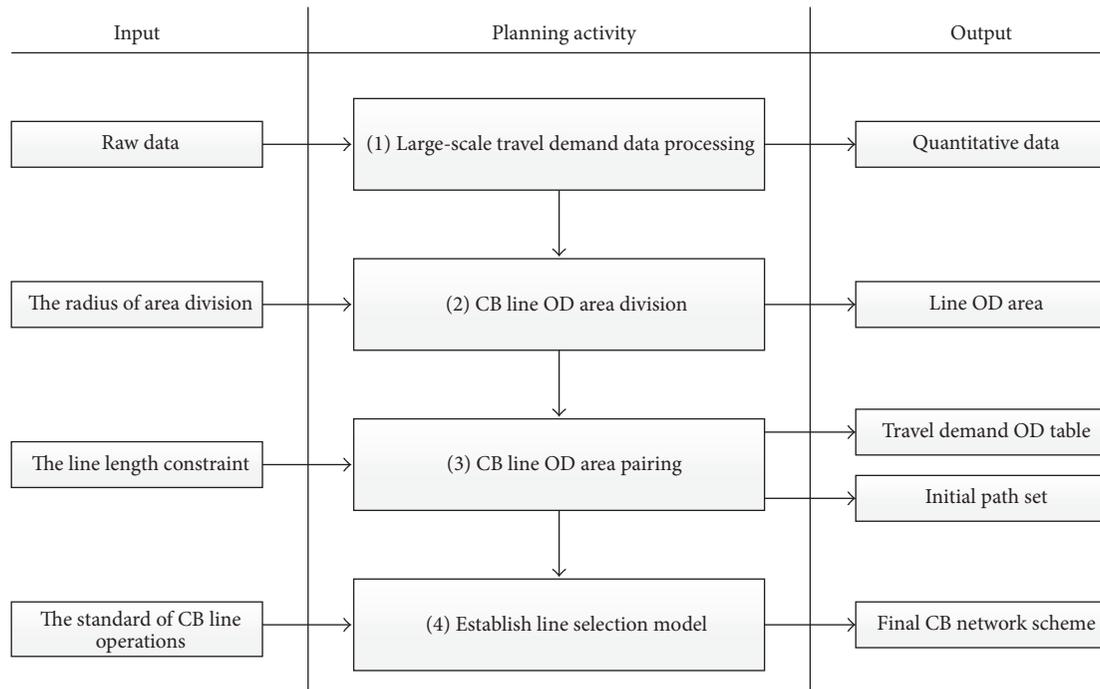


FIGURE 1: CB network planning process.

the line OD area division, which can effectively reduce the operating costs; however, it is difficult to attract passengers because passengers' walking distance is increased. On the contrary, reductions in the radius of the line OD area division can increase operating costs and lead to losses. These two conditions are not conducive to the long-term development of CBs. This paper divides the CB line OD area according to the passengers' position distribution using an area division algorithm, which places similar demands on an area and establishes a reasonable radius that the line OD area must cover. Passengers board the bus at the origin area stops and exit at the destination area stops, which can prevent halfway stops. This scheme will satisfy the demands of large numbers of passengers, make full use of resources, and effectively reduce companies' operating costs.

(3) *CB Line OD Area Pairing.* After the line OD area is determined, different CB lines can be established by pairing the OD areas. Using travel demand data processing, travel demand OD tables can be developed based on the line OD area. By deleting lines that do not conform to the distance constraint, the initial path set can be determined and the OD table can be updated.

(4) *Establish the Line Selection Model.* The objective of CB operations is to provide a comfortable and rapid riding environment for passengers, ease urban traffic congestion, and solve the problem of air pollution. These factors are regarded as the standard of CB line operations. The operating lines and CB network scheme were determined in this study by constructing an objective function for minimizing the sum of the operating, environmental, and traffic congestion costs and using the linear programming method to obtain

a solution. The CB network planning process based on area division is shown in Figure 1.

### 3. CB Network Planning Method

3.1. *Large-Scale Travel Demand Data Processing.* Quantitative data related to CB network planning can be obtained by sorting and calculating a series of raw data collected from the demand survey; these data can provide data support for the planning method. First, according to the passenger OD survey, the latitudes and longitudes of the passenger ODs can be determined. The latitude and longitude coordinates are then translated into planar coordinates using software so that the Euclidean distance between the origins and destinations of each demand can be calculated. Finally, the coordinates are marked on the diagram.

3.2. *CB Line OD Area Division.* Area division algorithms primarily include hierarchical clustering and K-means clustering, and the choice of algorithm is based on the purpose and application of the data analysis [31]. The K-means cluster method needs to set a parameter in advance, indicating the number of classes. Besides, the result is sensitive to the initial core of the data. In the clustering process of CB travel demand, the number of classes is unknown, and the initial core is selected randomly. Agglomerative hierarchical clustering is based on the bottom-up strategy, it aggregates a given data set according to the distance measurement criteria between categories until a certain condition is satisfied. Thus, the agglomerative hierarchical clustering method is more suitable for the CB travel demand clustering.

Agglomerative hierarchical clustering is based on the bottom-up strategy. First, each object is treated as a category, the distance between the objects is calculated, and the initial distance matrix is obtained. Two categories of minimum distances between categories are merged into one category, and the distances between the new category and all other categories are recalculated. The previous step is repeated, and the categories become increasingly larger until all objects are merged into a single category or a certain end condition is satisfied. According to the different distance measurement criteria between categories, hierarchical clustering can be divided into 4 types: (1) the average distance method, (2) the minimum distance method, (3) the maximum distance method, and (4) the barycenter method. Considering the actual definition of the distance between categories, the maximum distance method was selected as the distance measurement criteria for the hierarchical clustering. The distance between the two points that are most distant in the two categories is expressed as the distance between the two areas, which is formulated in (1). When the distance between the two areas is less than the maximum distance, the two categories are combined into a single category.

$$D(A, B) = \max_{i \in A, j \in B} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (1)$$

where  $A, B$  are the sets of points that belong to categories  $A$  and  $B$ , respectively,  $x_i, y_i$  are the horizontal and vertical coordinates of the points that belong to category  $A$ , respectively,  $x_j, y_j$  are the horizontal and vertical coordinates of the points that belong to category  $B$ , respectively, and  $D(A, B)$  is the distance between categories  $A$  and  $B$ .

In this paper, the data set was initially divided using hierarchical clustering. Then, a hierarchical clustering tree was generated, the maximum distances between the categories were used to determine the number of categories  $m$ , the data points were divided into  $m$  categories, and the area division scheme was determined. Additional details are provided below.

(1) By collecting the travel demand data, the origin address set and the destination address set can be obtained. Based on the plane coordinate set of the origins and destinations determined by processing the travel demand data in Section 3.1, the points of the origin and destination sets are marked on the diagram.

(2) The distance between all the points in the origin set and destination set is separately calculated, and the two data sets based on the maximum distance measurement criteria are then separately classified. Two categories that are separated by a distance less than the maximum distance are combined into a large category until all points are merged into a single category, and the hierarchical clustering tree is generated.

(3) According to expert opinion or practical experience, the area division radius can be determined. In addition, the maximum distance between the categories also can be determined. In this way, the number of categories  $m$  is determined according to the maximum distances between the categories.

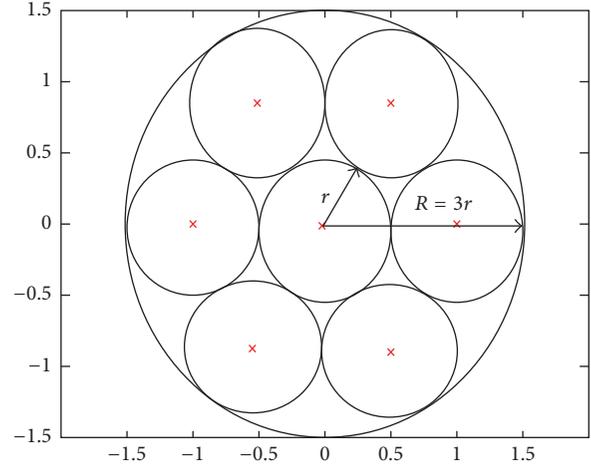


FIGURE 2: Diagram of the method for determining the area division radius.

(4) The two data sets are classified separately according to the number of categories  $m$  in the last step. Each category represents an origin area or destination area. The centroid coordinates of each origin area and each destination area can be calculated using (2). The centroid is then marked on the diagram, and the final area division scheme is determined.

$$(X_m, Y_m) = \left( \frac{\sum_{i=1}^n x_i}{n}, \frac{\sum_{i=1}^n y_i}{n} \right), \quad (2)$$

where  $X_m, Y_m$  are the horizontal and vertical coordinates of the centroid of category  $m$ , respectively,  $x_i, y_i$  are the horizontal and vertical coordinates of a certain point belonging to category  $m$ , respectively, and  $n$  is the total number of points in category  $m$ .

If there are an excessive number of stops in the CB OD areas, the time spent at the stops will be excessively long, which will increase the passenger travel time. Therefore, considering the passengers' walking distances, the radius of the area division cannot be overly large. However, operating costs will increase if the radius of area division is overly small. Therefore, the radius of the CB area division should be reasonably selected. The method for determining the area division radius is shown in Figure 2. The large circle represents a category that is an OD area, whereas the small circle represents the coverage of a stop in the OD area. The radius  $r$  of the small circle indicates the passengers' walking distances, which are between 500 and 1,000 m. No more than 7 stops in an area are appropriate. If the 7 stops are distributed in the area and guarantee the maximum areal coverage, an area coverage radius  $R$  of  $3r$  can be obtained. Therefore, the area coverage radius is 1.5 to 3 km. Then, the range of the maximum category distance is the diameter of the circle, which is 3 to 6 km. The reasonable value of the maximum category distance is typically determined based on the actual situation, which includes the city scale, number of buses, operating scale, and number of commuters.

3.3. *CB Line OD Area Pairing.* After the line OD area is determined, different CB lines can be constituted by pairing

the OD areas. Assuming that there are  $m$  origin areas and  $n$  destination areas,  $m \times n$  bus lines can be connected by pairing. Each path from the origin area to the destination area is represented as a bus line. The travel demand data corresponding to those  $m \times n$  lines are processed. By taking the number of people from the  $i$  origin area to the  $j$  destination area as the OD quantity of this line, the travel demand OD table can be determined by integration. Each cell in the demand OD table represents the travel demand of a line from an origin area to a destination area.

CBs can reduce passengers' commuting transfer and travel times and reduce travel costs to attract private car passengers. If the passenger travel distance is short, traditional buses can satisfy the passenger travel demand at a low price and high non-bus stop rate. If the passenger travel distance is long, the CB can effectively reduce transfer times. In Beijing, the round-trip travel cost of a 20 km indicates that CB costs are 30% of the cost of travelling by private car and 15% of the cost of taking a taxi. Thus, CBs can effectively reduce the travel costs of private car passengers. Therefore, the distance from the origin area to the destination area, which is also the line, should not be overly short. In this paper, the distance between the origin area clustering center and the destination area clustering center is represented as the line length. The length of each line  $l_b$  is given by a logistic function:

$$l_b = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}, \quad (3)$$

where  $l_b$  is the length of each line,  $b = \{1, 2, \dots, m \times n\}$ ,  $X_i, Y_i$  are the horizontal and vertical coordinates of the origin area centroid, respectively,  $i = \{1, 2, \dots, m\}$ ,  $X_j, Y_j$  are the horizontal and vertical coordinates of the destination area centroid, respectively, and  $j = \{1, 2, \dots, n\}$ .

$l_b$  should not be less than the minimum line length constraint:

$$l_b \geq l_{\min}, \quad (4)$$

where  $l_{\min}$  is the minimum line length.

In this paper, the minimum line length was set to  $l_{\min} = 8$  km. An initial path set  $L_0 = \{l_1, l_2, \dots, l_k\}$  is formed by all  $k$  lines that satisfy the minimum line length constraint. After determining the initial path set, the travel demand OD table should be updated. If the line is not included in the initial path set, the value of the line travel demand is then changed to 0.

**3.4. Establishing the Line Selection Model.** The main considerations of CB operations are passenger comfort, economical operation, and social benefits. Because the aforementioned area division algorithm has been designed to ensure that passenger walking distances are not overly long, the line selection model in this section mainly considers the remaining two aspects: economical operation and social benefits. The operation of a CB will increase the operating costs of a company, but with the CB operations, the company can attract more private car passengers who choose CB, which will reduce vehicle pollution emissions and ease traffic congestion. Therefore, there is a certain contradiction between operating costs and

social benefits in CB operations. In this section, considering those two aspects, a CB line selection model that is based on total cost is established. By solving the model, the directions of the CB lines and the number of buses used in each line that minimize the total cost of a single line can be determined.

The model is established based on the following assumptions:

- (1) Passengers travel by either bus or car.
- (2) The linear distance between the origin area clustering center and destination area clustering center is represented by the CB line length.
- (3) Each line uses the same CB vehicle.

The target function presented in this paper consists of three parts. The first is the company's operating cost, including the CB operating cost per kilometer and fixed cost (including drivers' and conductors' salaries, management fees, and maintenance fees). The second is the environmental pollution cost, including the pollution cost of CBs and car emissions. The third is road congestion cost, including the additional time cost to passengers who travel by CB and by car because of traffic congestion.

The company's operating cost is determined using the following logistic function:

$$Z_1 = c_{Gm} \times \left[ \frac{n_{Gi}}{\alpha_G} \right] \times l_i + c_{Go} \times \left[ \frac{n_{Gi}}{\alpha_G} \right], \quad (5)$$

where  $Z_1$  is the total operating cost,  $c_{Gm}$  is the fuel cost per kilometer per vehicle,  $c_{Go}$  is the fixed operating cost per vehicle per day,  $l_i$  is the mileage per line,  $i = \{1, 2, \dots, k\}$ ,  $n_{Gi}$  is the number of people travelling by CBs on each line,  $\alpha_G$  is the maximum load capacity per CB vehicle, and  $\lceil n_{Gi}/\alpha_G \rceil$  is the number of CB vehicles on each line calculated by rounding up.

The environmental cost is given by the following logistic function:

$$Z_2 = c_a \left( W_{Ga} \times \left[ \frac{n_{Gi}}{\alpha_G} \right] \times l_i + W_{Ca} \times \left[ \frac{n_{Ci}}{\alpha_C} \right] \times l_i \right), \quad (6)$$

where  $Z_2$  is the total environmental pollution cost,  $c_a$  is the environmental pollution cost per unit of pollution,  $W_{Ga}$  is the pollutant emissions per bus per kilometer,  $W_{Ca}$  is the pollutant emissions per car per kilometer,  $n_{Ci}$  is the number of people travelling by car on each line,  $\alpha_C$  is the average load capacity per car, and  $\lceil n_{Ci}/\alpha_C \rceil$  is the number of cars on each line calculated by rounding up.

The traffic congestion cost is the additional time cost incurred by all passengers because of the change in the number of buses and cars on the road, which is formulated as

$$Z_3 = c_s \left( \left( \frac{l_i}{v_{Gs}} - \frac{l_i}{v_G} \right) \times n_{Gi} + \left( \frac{l_i}{v_{Cs}} - \frac{l_i}{v_C} \right) \times n_{Ci} \right), \quad (7)$$

where  $Z_3$  is the total congestion cost,  $c_s$  is the value per unit of time,  $v_{Gs}$  is the average running speed of the bus in the case of exclusive bus lanes and traffic congestion,  $v_G$  is the average

running speed of buses during normal running,  $v_{Cs}$  is the average running speed of cars in the case of traffic congestion, and  $v_C$  is the average running speed of cars during normal running.

In summary, the objective function  $Z$  of the line operating standard model can be determined using the linear weighted sum of the three parts as

$$\min Z = \omega_1 Z_1 + \omega_2 Z_2 + \omega_3 Z_3, \quad (8)$$

where  $Z$  is the total cost of a single line and  $\omega_1, \omega_2, \omega_3$  are the operating, environmental, and traffic congestion cost weights, respectively.

The decision variables of the line selection model in this paper are the number of commuters who travel by customized bus and the number of commuters who take a private car, respectively, represented by  $n_{Gi}$  and  $n_{Ci}$ . Actually, in this paper, the total travel demand on each line is known; thus  $n_{Gi}$  and  $n_{Ci}$  should satisfy the following constraints:

$$n_{Gi} + n_{Ci} = n_i, \quad (9)$$

$$n_{Gi}, n_{Ci} \in N, \quad (10)$$

where  $n_i$  is the travel demand on each line and  $N$  is the natural number set.

Equation (9) specifies that the sum of the numbers of people who are travelling by CBs and by car is equal to the travel demand on each line. Equation (10) specifies that the number of people who are travelling by CBs and by car is a natural number.

**3.5. Model Solution.** Solving the CB line selection model is essentially an integer programming problem and a discrete optimization problem. The purpose of the problem is to find a solution that conforms to the objective function from a limited number of possible scenarios. The optimal solution can be obtained by comparing the objective function value and using an enumeration method. However, in practical problems, the solution space is large and consumes large amounts of computational time and memory. Integer programming optimization methods for solving research problems are also gradually maturing with the development of optimization theory. At present, the branch-and-bound method, cutting-plane method, tabu search, and genetic algorithms are commonly used methods. In this paper, the branch-and-bound method [32] was used to solve the CB line selection model, and the final network operation scheme was obtained using a computer.

Integer programming is a branch of linear programming. If the integer constraint is removed, integer programming is transformed into linear programming. This linear programming problem is called the linear programming relaxation problem of integer programming, and all the feasible solutions of the integer programming problem are included in the linear programming relaxation problem. If the solution of the linear programming relaxation problem is expressed as  $Z_0$ , the optimal integer solutions that have been found are expressed by  $Z_i$ , the optimal integer solutions are expressed by  $Z^*$ , the lower bound is expressed by  $Z_l$ ,

and the upper bound is expressed by  $Z_u$ . For the objective function minimization problem, the optimal integer solution must satisfy

$$Z_l = Z_0 \leq Z^* \leq Z_i = Z_u. \quad (11)$$

The branch-and-bound method is based on the above relationship. The method begins by solving the linear programming relaxation problem. The feasible region of the linear relaxation problem is then decomposed into smaller subdomains (branches). The next step is to continuously update the upper and lower bounds by finding better integer solutions, which are obtained by the branches together. This approach can accelerate convergence, simplify the operation, and cause the lower bound to be equal to the upper bound. In this manner, the optimal solution of the integer programming is obtained [33]. The algorithm steps are as follows.

*Step 1.* The linear relaxation problem of integer programming is solved by removing the integer constraints. If there is no feasible solution, then the integer programming problem has no optimal solution, and the procedure is stopped. If an integer solution is obtained, then the solution is the integer programming optimal solution  $Z^*$ , and the procedure is stopped. If a noninteger solution is obtained, then the solution is taken as the lower bound of the integer programming problem, the upper bound is expressed by infinity, and the processing continues to Step 2.

*Step 2.* Choose any variable that does not meet the integer criteria  $n_{Gi}$  or  $n_{Ci}$   $i = 1, 2, \dots, k$  from the solution of the linear relaxation problem. Its value is  $b_i$ . Using  $[b_i]$  represents the largest integer one that does not exceed  $b_i$ . Next, two mutually exclusive inequality constraints,  $n_{Gi} \geq [b_i] + 1$  and  $n_{Gi} \leq [b_i]$  ( $n_{Ci} \geq [b_i] + 1$  and  $n_{Ci} \leq [b_i]$ ), are constructed. The two constraints are added to the linear relaxation problem, two subproblems can be obtained, and the optimal solution can be sought. The smallest result can then be found from the obtained optimal solutions and is used as a new lower bound  $Z_l$ . Finally, the greatest value of the objective function from each subproblem consistent with integer conditions is found and used as a new upper bound  $Z_u$ .

*Step 3.* For the subproblems with nonfeasible solutions and integer solutions, the downward branch is not continued downward and the node is closed. For the subproblems with noninteger optimal solutions, determine whether the optimal value is greater than the upper bound of  $Z_u$ . If the optimal value is greater than or equal to  $Z_u$ , stop the branch; otherwise, branch out and repeat Step 2 until all the nodes are closed. Optimal integer solutions can be obtained at this time with  $n_{Gi}^*$  and  $n_{Ci}^*$  ( $i = 1, 2, \dots, k$ ), and the optimal value is  $Z^* = Z_u$  [34].

The flowchart of branch-and-bound method is shown in Figure 3.

Using the branch-and-bound method to solve the CB line selection model, the numbers of people travelling by CB and car on each line can be obtained. If the number of people travelling by CB is zero, the line is not running; otherwise,

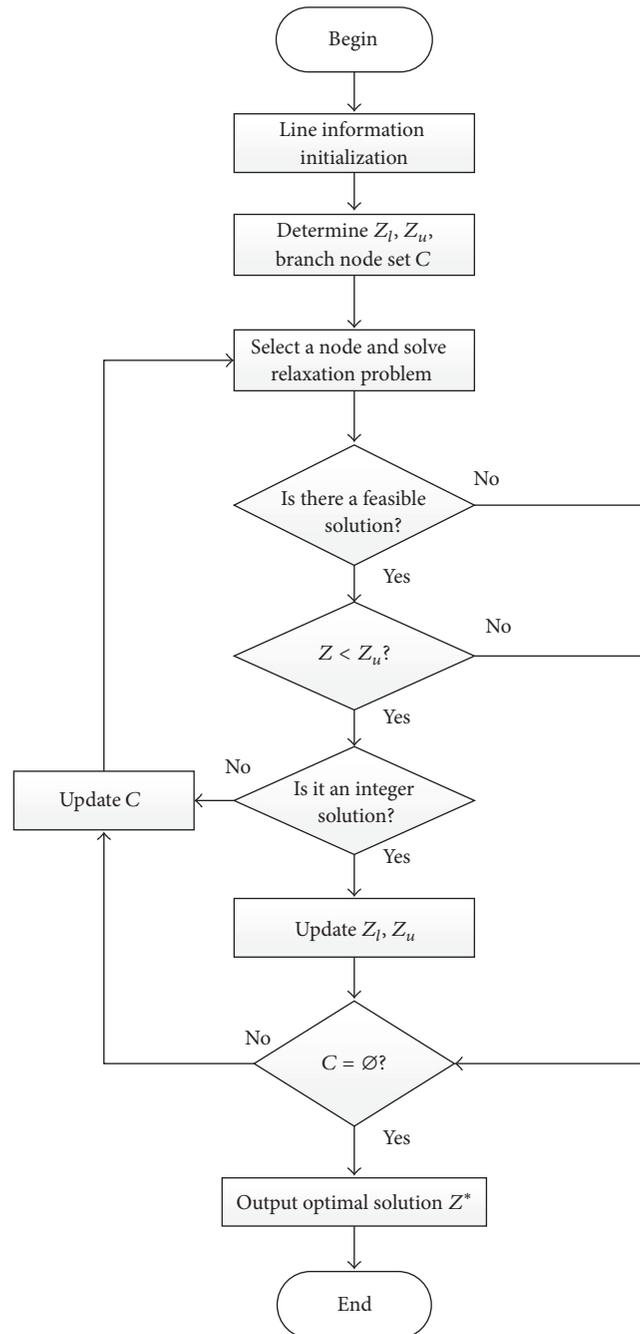


FIGURE 3: The flowchart of branch-and-bound method.

the line is operating, and the required number of vehicles is computed according to the travel demand of passengers.

#### 4. Case Study

In this section, Beijing CB network planning is employed as a case study. The optimization scheme for Beijing CB network planning can be determined using the CB network planning method introduced in this paper. The CB network planning method introduced in this paper was proven to be applicable

and effective compared with the current status of Beijing CB network planning.

*4.1. Large-Scale Travel Demand Data Processing.* The travel demands of passengers were collected using an Internet survey. A total of 15,000 morning peak travel passenger demands for CBs were collected from August 2015 to November 2015. Each travel demand includes a passenger's trip origin and destination. According to the passenger travel demand address data, the origin address and destination

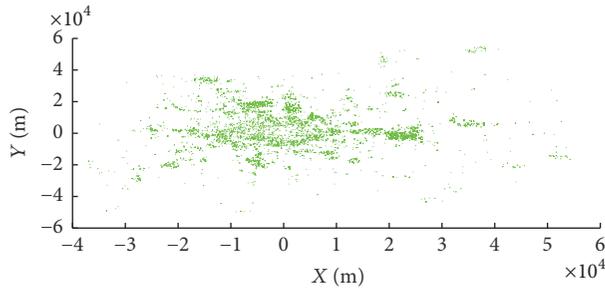


FIGURE 4: Scatterplot of the passenger trip origins.

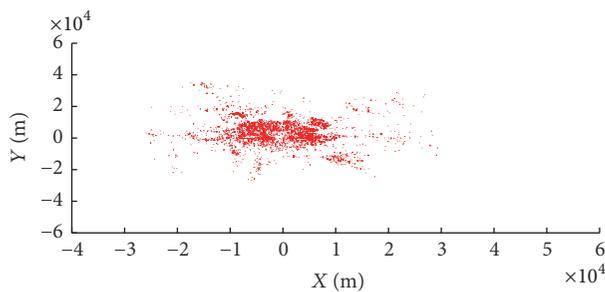


FIGURE 5: Scatterplot of the passenger trip destinations.

address of every passenger were stored as text documents or EXCEL documents. The documents then were imported into the XGeocoding software program and saved, and parsing was initiated. Parsing batch queried the longitude and latitude coordinates of every passenger OD, and the latitude and longitude coordinates of the passengers' locations were stored in text format. Subsequently, the text documents were imported into the COORD coordinate conversion software program, which can transform the latitude and longitude coordinates into plane coordinates that can be easily used to calculate the point-to-point distances. Because the passenger OD coordinate values are large numbers, for convenience, the coordinates of Tiananmen were used as the origin of the coordinates, the Tiananmen coordinates were subtracted from all the coordinates, and the origin and destination demands of the passengers then were marked separately using MATLAB software. The passenger trip origins are shown by green points in Figure 4. The passenger trip destinations are shown by red points in Figure 5.

In Figure 4, the range in the distribution of the CB passenger trip origins was wide and concentrated in a large residential zone. However, the passenger trip destinations were mainly concentrated in a large business zone centered on Tiananmen with a radius of 10 km, as shown in Figure 5. The characteristics of the morning peak travel demand in Beijing City is an agglomeration from rural areas to the urban area, and the travel distances were long. Therefore, the operation of the CB in Beijing City can effectively ease traffic congestion and reduce travel time.

**4.2. Line OD Area Division Based on Hierarchical Clustering.** From the processing results of the passenger travel

demand data in Section 4.1, all ODs were classified using hierarchical clustering according to the maximum distance measurement criteria. In this paper, considering the walking cost of passengers, the passenger walking distance should not exceed 800 m. The area coverage radius was 2.5 km, and the maximum category distance was 5 km. Thus, when the maximum distance between two categories was less than 5 km, the two categories were combined into one category. The final results yielded 84 origin categories and 67 destination categories. The hierarchical clustering tree was then built to represent the clustering results vividly. The origin hierarchical clustering tree is shown in Figure 6. The destination hierarchical clustering tree is shown in Figure 7. In the figure, the distance between categories is represented by the vertical coordinates, the category is represented by the horizontal coordinates, and the number on each horizontal coordinate represents the travel demands that are clustered by this category.

According to the results, the origin and destination were classified using hierarchical clustering. The passenger plane coordinate data for each category were stored as an array. Using (2), the centroid coordinates of each category were calculated, and they are marked on the diagram. The origin clustering results are shown in Figure 8. The destination clustering results are shown in Figure 9. In the diagram, the points in different colors represent the passenger travel demand points in different categories. The crosses represent the centroids of the categories. In the origin hierarchical clustering diagram, each category represents an origin area, and each centroid represents the center of an origin area. In the destination hierarchical clustering diagram, each category represents a destination area, and each centroid represents the center of a destination area.

**4.3. Determination of the Initial Path Set.** Based on the hierarchical clustering results of the origin and destination areas presented in Section 4.2, approximately 5,628 CB lines were obtained by pairing the OD areas. A travel OD demand table was obtained by processing the travel demand data corresponding to the 5,628 lines. In addition, the lengths of the lines were the distances between the centroids of the origin and destination areas. In the travel OD demand table, each cell indicates the travel demand of a line from an origin area to a destination area. In the line length table, each cell indicates the length of a line. The names of the origin area and destination area were derived from the centroid coordinates of each area using COORD and an XGeocoding software transformation.

In this paper, the minimum length of the line operation was determined to be  $l_{\min} = 8$  km. According to the line length constraint, a total of 233 lines that did not satisfy the line length constraint were deleted, and the 5,395 lines that satisfied the line length constraint were used as the initial line set. Using the initial path set, the travel demand OD table was updated, a total of 1,973 travel demands were removed, and 13,321 travel demands remained.

**4.4. Model Solution and Scheme Determination.** Based on the travel OD demand and line length of each line, the

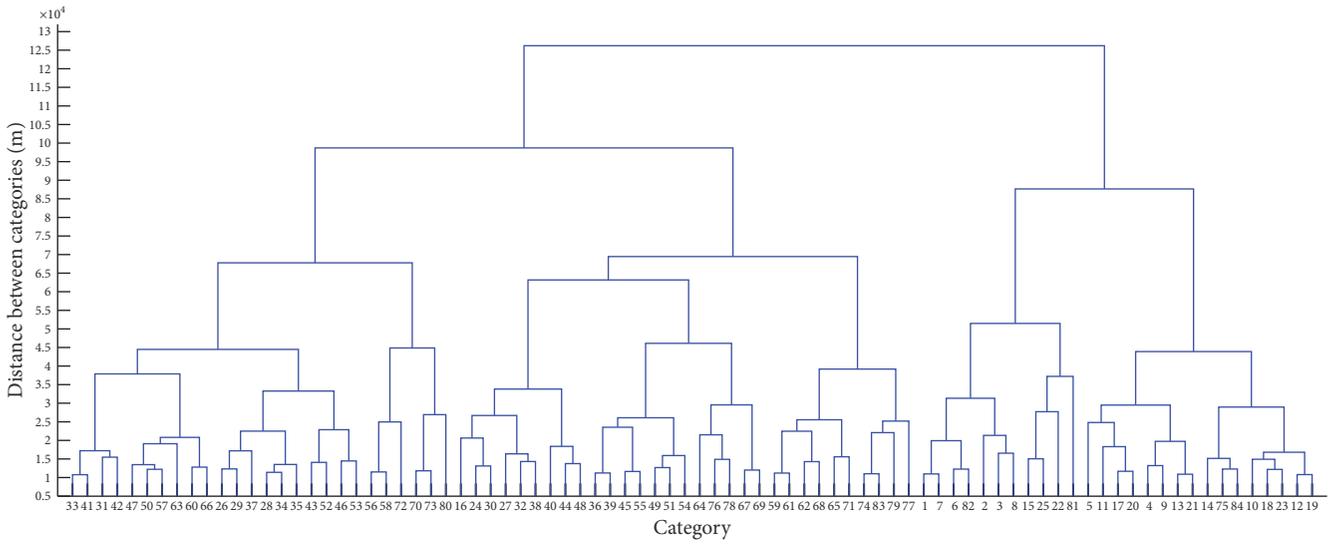


FIGURE 6: Origin hierarchical clustering tree.

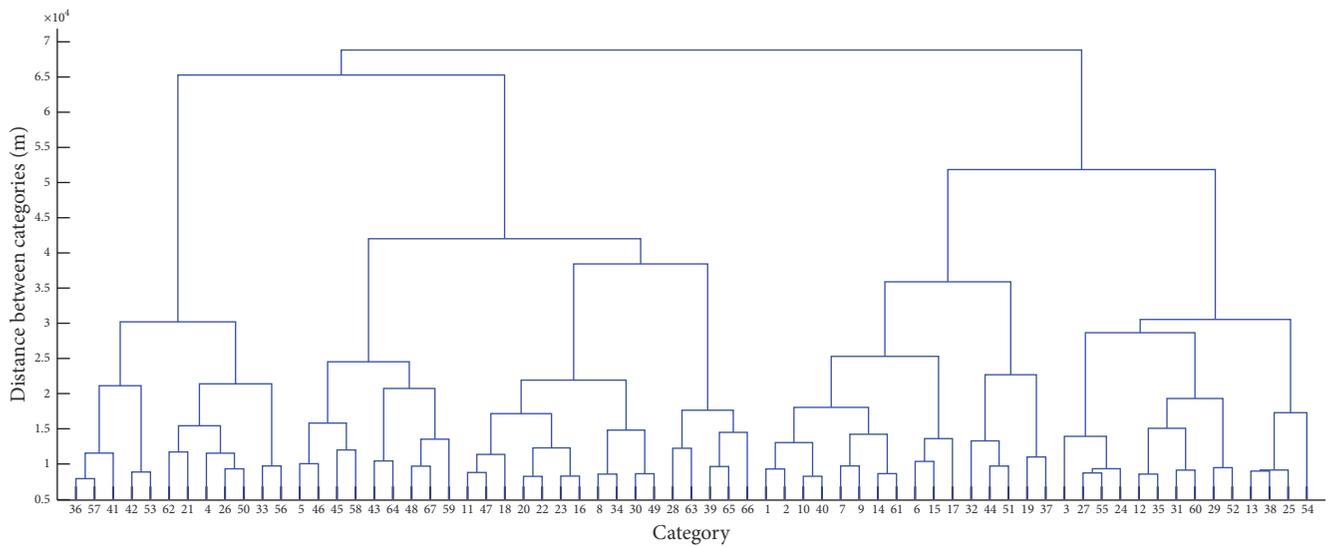


FIGURE 7: Destination hierarchical clustering tree.

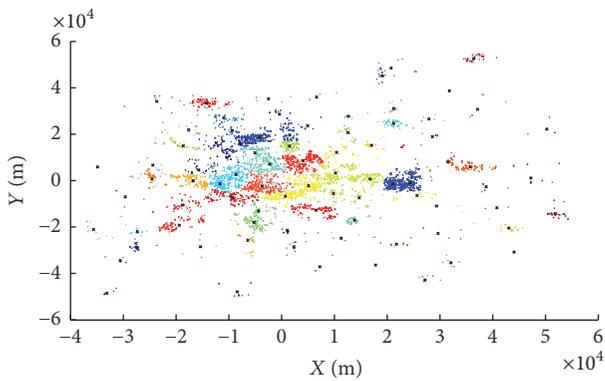


FIGURE 8: Origin hierarchical clustering diagram.

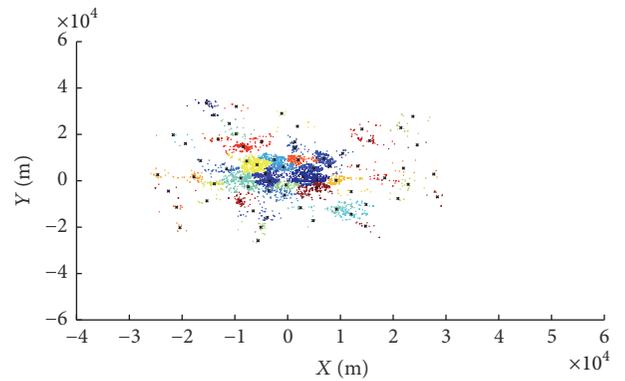


FIGURE 9: Destination hierarchical clustering diagram.

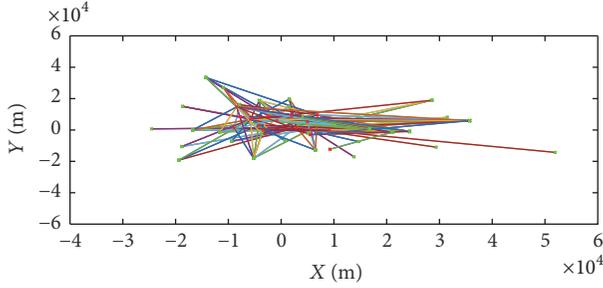


FIGURE 10: Planning line diagram for the Beijing CB network.

final operating scheme for the Beijing CB network was solved using MATLAB according to the line selection model in Section 3.4 and the branch-and-bound method in Section 3.5. The related parameters are as follows: the fuel cost per kilometer per vehicle  $c_{Gm} = 2.1$ , the fixed operating cost per vehicle per day  $c_{Go} = 350$  [35], the environmental pollution cost per unit of pollution  $c_a = 3.3$  [36], the unit time value  $c_s = 36$ , the maximum load capacity per CB vehicle  $\alpha_G = 30$ , the average load capacity per car  $\alpha_C = 2$ , the pollutant emissions per bus per kilometer  $W_{Ga} = 1.2$ , the pollutant emissions per car per kilometer  $W_{Ca} = 0.4$  [37], the average running speed of bus in the case of bus lanes and road congestion  $v_{Gs} = 28$ , the average running speed of buses during normal running  $v_G = 35$ , the average running speed of cars in the case of road congestion  $v_{Cs} = 26$ , the average running speed of cars during normal running  $v_C = 46$  [38], the weight of operating cost  $\omega_1 = 0.3$ , the weight of the environmental cost  $\omega_2 = 0.4$ , and the weight of congestion cost  $\omega_3 = 0.3$ .

According to the 15,000 CB morning peak travel demands of passengers, approximately 123 CB lines were found using the CB network planning method introduced in this paper. In addition, the operating kilometers totaled 2,708.3 km, which required 183 CBs with 30 seats, each of which served 5,009 passengers. The specific line distribution range is shown in Figure 10. In the figure, the green crosses represent the CB origins, the red crosses represent the CB destinations, and the connecting lines between them represent CB operating lines.

**4.5. Comparison and Evaluation.** In reality, using the approximately 100,000 CB morning peak travel demands of passengers from September 2013 to November 2015, the Beijing CB company designed a total of 92 morning peak lines. By comparison, in this paper, a total of 15,000 CB morning peak travel demands of passengers were collected from August 2015 to November 2015, and a total of 123 morning peak lines were designed by using the method proposed in this paper. The OD distributions of the current and planning schemes are compared in Figures 11 and 12. The contrast figure of the CB origin distribution is shown in Figure 11. The contrast figure of the CB destination distribution is shown in Figure 12. In the figure, the black crosses represent the current stations, and the green crosses represent the planning stations. In addition, in combination with the current planning line diagram of the CB and the planning line diagram of the

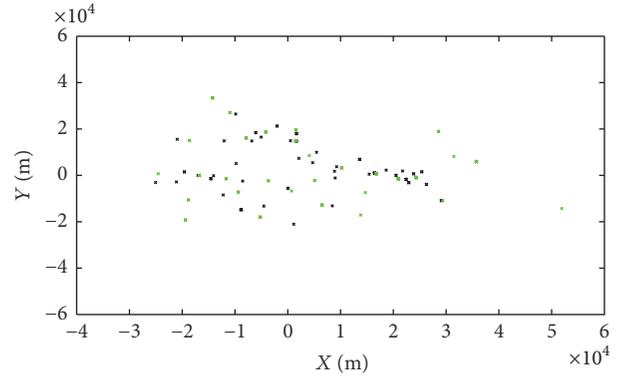


FIGURE 11: Contrast figure of the CB origin distribution.

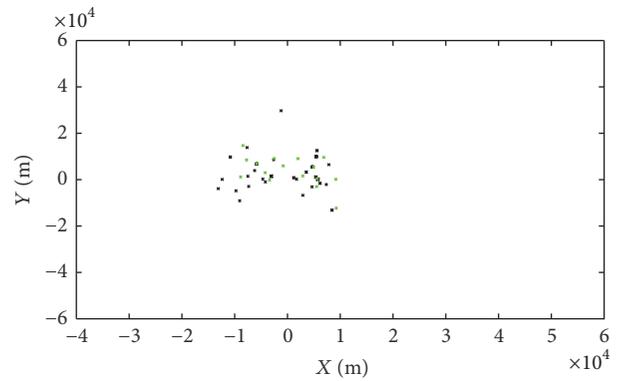


FIGURE 12: Contrast figure of the CB destination distribution.

CB determined using this paper's method, it was found that the planning scheme generated using the method of this paper was basically consistent with the current scheme. This shows that the method proposed in this paper is feasible and effective for solving CB network planning problems.

Figure 11 illustrates that the origins of the planning and current schemes were concentrated in large residential areas, such as Tongzhou, Huilongguan, and Fengtai, whereas the distribution of the planning scheme was more balanced and the stations had a wider coverage. Figure 12 illustrates that the destinations of the planning and current schemes were concentrated over a range that included Tiananmen as its center and had a radius of 10 km, which included the China World Trade Center, Finance Street, and Zhongguancun. In contrast, the station coverage range of the planning scheme was smaller than that of the current scheme.

CB is a new public transportation mode without transfer based on travel demands of passengers. The CB operating line is the shortest line between origin area stops and destination area stops. CB network evaluation differs from conventional bus network evaluation but also has some similar places as conventional bus. In this paper, the following evaluation indexes are established to evaluate the service level of CB network.

(1) The site coverage rate  $\gamma$  is the proportion of CB site areal coverage in the total areal coverage of travel demand,

TABLE 1: Beijing CB network evaluation index contrast.

|                           | Current scheme | Planning scheme |
|---------------------------|----------------|-----------------|
| The number of lines       | 92             | 123             |
| The total length of lines | 1412.44 km     | 2708.30 km      |
| The average line distance | 15.35 km       | 22.02 km        |
| The number of vehicles    | 192            | 241             |
| The site coverage rate    | 23.38%         | 32.71%          |
| The average load factor   | 72.53%         | 77.05%          |
| The service rate          | 30.17%         | 37.60%          |

that is, the proportion of the areal coverage of all original area in the total areal coverage of travel demand, which is formulated in

$$\gamma = \frac{\sum_{i=1}^m s_i}{S_C} \times 100\%, \quad (12)$$

where  $s_i$  is the areal coverage of each original area,  $i = \{1, 2, \dots, m\}$ , and  $S_C$  is the total areal coverage of travel demand.

(2) The average load factor  $\bar{\lambda}$  is the proportion of the number of people travelling by CBs in the maximum number of passengers provided by CBs, which is formulated in

$$\bar{\lambda} = \frac{\sum_{i=1}^k n_{Gi}}{\sum_{i=1}^k b_i \times \alpha_G} \times 100\%, \quad (13)$$

where  $b_i$  is the number of vehicles of each line,  $i = \{1, 2, \dots, k\}$ .

(3) The service rate  $\varphi$  is the proportion of the number of people travelling by CBs in the total travel demand of passengers, which is formulated in

$$\varphi = \frac{\sum_{i=1}^k n_{Gi}}{N} \times 100\%, \quad (14)$$

where  $N$  is the total travel demand of passengers.

The comparative results of the specific network evaluation index are listed in Table 1. Compared with the current network scheme, the travel demand data used by the planning scheme were smaller. However, the total number of lines, total length of the operating lines, and number of people being served were greater. Thus, the planning scheme placed more emphasis on the passenger service rate. In addition, the average line distance of the current scheme was 15.35 km, and the planning scheme was 22.02 km, demonstrating that the planning scheme lines mainly served passengers who travelled medium and long distances.

From the perspective of the company's operations, 192 vehicles are used in the current scheme, with an average load factor of 72.53%; in contrast, 241 vehicles are used in the planning scheme, with an average load factor of 77.05%. Although the average load factor cannot reflect the actual operational situation, as long as the planning is scientific and reasonable regarding the stops and timetable, the actual average load factor should be equal to the numerical value. Therefore, the

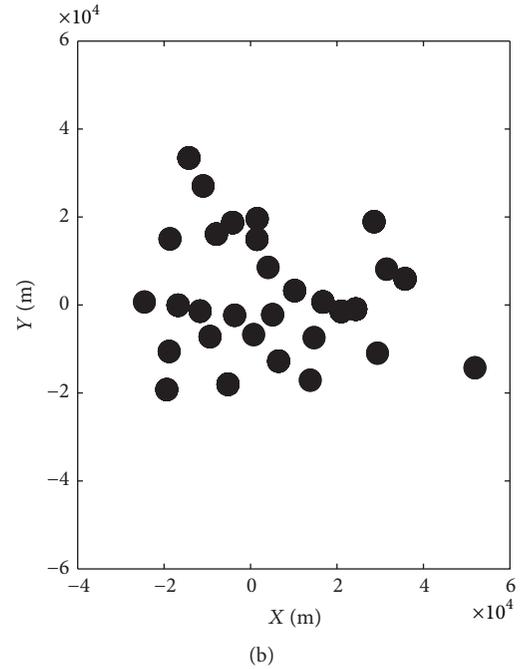
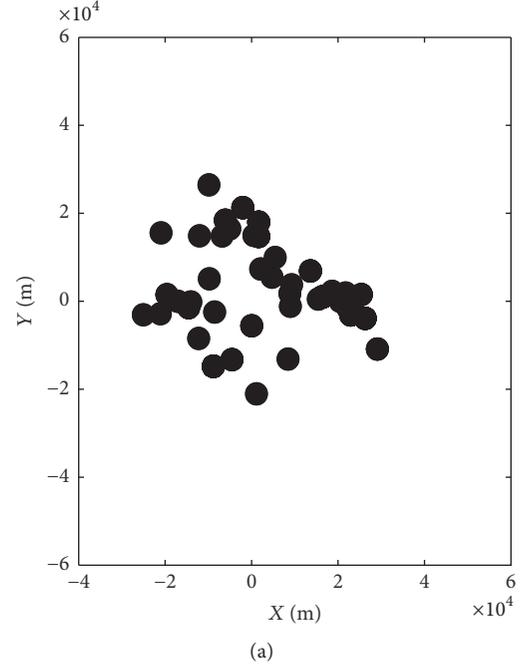


FIGURE 13: Contrast figure of the CB original stations' areal coverage. (a) Current original stations' areal coverage. (b) The planning original stations' areal coverage.

planning scheme can optimize the allocation of resources and effectively reduce the operating company's costs.

The coverage radius for each traffic zone was set to 2.5 km, and the site coverage rate for the original traffic zone was calculated. The contrast in the original stations' areal coverage for the current and planning schemes is shown in Figure 13. Combined with Table 1 and Figure 13, the station distribution

of the planning scheme was more balanced, and the site coverage rate was considerably higher. Therefore, the CV network can more effectively encourage private car owners to change travel modes, ease traffic congestion, and solve the problem of air pollution.

In summary, combined with the actual operational status of the Beijing CB network and the comparison results, a more scientific and reasonable network scheme can be obtained using the method presented in this paper. Compared with the current scheme, the planning scheme can more effectively provide services to passengers, save operating costs, and create positive social benefits.

## 5. Conclusions

This paper presented the status of research on CB network planning methods. On that basis, combined with conventional bus and existing CB network planning experience, a CB network planning method based on area division was proposed. The method uses ideas from a point-to-line layout into a network. First, the line OD area is divided according to the passenger travel demand; O and D of line OD represent the passenger's trip origin and trip destination, respectively. Second, the operating line scheme is determined according to the solution of the model, all operating lines are laid out, and the final CB network is determined. The results of this paper advance the theoretical research on CB network planning and provide a precise and efficient technical method for CB operators who are planning networks. The following aspects should be investigated further to optimize and improve the methods proposed in this paper.

(1) A major feature of CBs is that they have exclusive bus lane rights. When the traffic is highly congested in the morning and evening peak periods, exclusive bus lanes can ensure the effectiveness of the running time and improve the punctuality. In this paper, the line layouts did not consider the specific paths of each line. A straight line between the origin and destination was used as a line, and the important role of exclusive bus lanes in the layout of CB lines was neglected. In future research, the layout of exclusive bus lanes should be considered.

(2) In this paper, a line operation model based on operating cost and social benefits was proposed. The main aspect of the model is the determination of the parameter values of various influencing factors and quantitative methods. By consulting various references to determine the parameter values, the parameters values achieved good results in the study case demonstration. However, whether those parameters values are applicable to all cities and how to assign the weight ratio of each factor should be considered in future research.

(3) In this paper, the vehicle standard was 30 seats. In future research, the vehicle standard should be chosen according to the line travel demand, which can improve the use efficiency of CB vehicles.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The paper is supported by National High-tech R&D Program of China (863 Program): 2015AA124103.

## References

- [1] R. F. Kirby and K. U. Bhatt, *Guidelines on the Operation of Subscription Bus Services*, National Technical Information Service, Alexandria, Va, USA, 1974.
- [2] R. F. Kirby and K. U. Bhatt, "An analysis of subscription bus experience," *Traffic Quarterly*, vol. 29, no. 3, pp. 403–425, 1975.
- [3] J. A. Bautz, "Subscription service in the United States," *Transportation*, vol. 4, no. 4, pp. 387–402, 1975.
- [4] C. H. J. McCall, *COM-BUS: A Southern California Subscription Bus Service*, National Technical Information Service, Alexandria, Va, USA, 1977.
- [5] C. E. McKnight and R. E. Paaswell, *The Potential of Private Subscription Bus to Reduce Public Transit Subsidies*, Urban Mass Transportation Administration, Washington, DC, USA, 1985.
- [6] S. A. Shaheen, D. Sperling, and C. Wagner, "Carsharing in Europe and North America: past, present and future," *Transportation Quarterly*, vol. 52, no. 3, pp. 35–52, 1998.
- [7] S. K. Chang and P. M. Schonfeld, "Optimization models for comparing conventional and subscription bus feeder services," *Transportation Science*, vol. 25, no. 4, pp. 281–298, 1991.
- [8] E. Martin and S. A. Shaheen, "Assessing greenhouse gas emission impacts from carsharing in north america: theoretical and methodological design," in *Proceedings of the 15th World Congress on Intelligent Transport Systems and ITS America Annual Meeting*, pp. 1183–1194, November 2008.
- [9] J. F. Potts, M. A. Marshall, E. C. Crockett, and J. Washington, "A guide for planning and operating flexible public transportation services," TCRP Report 140, 2010.
- [10] M. Duncan, "The cost saving potential of carsharing in a US context," *Transportation*, vol. 38, no. 2, pp. 363–382, 2011.
- [11] A. El Fassi, A. Awasthi, and M. Viviani, "Evaluation of car-sharing network's growth strategies through discrete event simulation," *Expert Systems with Applications*, vol. 39, no. 8, pp. 6692–6705, 2012.
- [12] A. De Lorimier and A. M. El-Geneidy, "Understanding the factors affecting vehicle usage and availability in carsharing networks: a case study of commuauto carsharing system from Montréal, Canada," *International Journal of Sustainable Transportation*, vol. 7, no. 1, pp. 35–51, 2012.
- [13] R. Nair and E. Miller-Hooks, "Equilibrium network design of shared-vehicle systems," *European Journal of Operational Research*, vol. 235, no. 1, pp. 47–61, 2014.
- [14] S. Le Vine, M. Lee-Gosselin, A. Sivakumar, and J. Polak, "A new approach to predict the market and impacts of round-trip and point-to-point carsharing systems: case study of London," *Transportation Research Part D: Transport and Environment*, vol. 32, pp. 218–229, 2014.
- [15] T. Liu and A. Ceder, "Analysis of a new public-transport-service concept: customized bus in China," *Transport Policy*, vol. 39, pp. 63–76, 2015.
- [16] W. Lampkin and P. D. Saalmans, "The design of routes, service frequencies, and schedules for a municipal bus undertaking: a case study," *OR: Operational Research Quarterly*, vol. 18, no. 4, pp. 375–397, 1967.

- [17] A. Ceder and N. H. M. Wilson, "Bus network design," *Transportation Research Part B: Methodological*, vol. 20, no. 4, pp. 331–344, 1986.
- [18] M. H. Baaj and H. S. Mahmassani, "Hybrid route generation heuristic algorithm for the design of transit networks," *Transportation Research, Part C: Emerging Technologies*, vol. 3, no. 1, pp. 31–50, 1995.
- [19] V. M. Tom and S. Mohan, "Transit route network design using frequency coded genetic algorithm," *Journal of Transportation Engineering*, vol. 129, no. 2, pp. 186–195, 2003.
- [20] S. Jerby and A. Ceder, "Optimal routing design for shuttle bus service," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1971, no. 1, pp. 14–22, 2006.
- [21] E. Cipriani, S. Gori, and M. Petrelli, "Transit network design: a procedure and an application to a large urban area," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 3–14, 2012.
- [22] M. Nikolić and D. Teodorović, "Transit network design by bee colony optimization," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5945–5955, 2013.
- [23] H. Badia, M. Estrada, and F. Robusté, "Competitive transit network design in cities with radial street patterns," *Transportation Research Part B: Methodological*, vol. 59, pp. 161–181, 2014.
- [24] M. Trépanier, N. Tranchant, and R. Chapleau, "Individual trip destination estimation in a transit smart card automated fare collection system," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 11, no. 1, pp. 1–14, 2007.
- [25] B. Li, "Markov models for Bayesian analysis about transit route origin-destination matrices," *Transportation Research Part B: Methodological*, vol. 43, no. 3, pp. 301–310, 2009.
- [26] X.-L. Ma, Y.-H. Wang, F. Chen, and J.-F. Liu, "Transit smart card data mining for passenger origin information extraction," *Journal of Zhejiang University: Science C*, vol. 13, no. 10, pp. 750–760, 2012.
- [27] M. A. Munizaga and C. Palma, "Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 9–18, 2012.
- [28] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 1–12, 2013.
- [29] X. Ma and Y. Wang, "Development of a data-driven platform for transit performance measures using smart card and GPS data," *Journal of Transportation Engineering*, vol. 140, no. 12, Article ID 04014063, 2014.
- [30] X. Ma, C. Liu, H. Wen, Y. Wang, and Y. Wu, "Understanding commuting patterns using transit smart card data," *Journal of Transport Geography*, vol. 58, pp. 135–145, 2017.
- [31] M. Duan, *The Research and Application of Hierarchical Clustering Algorithm*, Central South University, Changsha, China, 2009.
- [32] L. Yupo, *Study on Problems to Select Initial Cluster Centers of the K-means Algorithm*, Lanzhou Jiaotong University, Lanzhou, China, 2012.
- [33] P. Zhao, *Course of Management Operations Research*, Tsinghua University Press, Beijing, China, 2008.
- [34] Q. Pingping, *The Application of Branch and Bound Algorithm in the Model of Operational Research*, Yanshan University, Qinhuangdao, China, 2009.
- [35] L. Dongmei, *The Transit Scheme Research on Customized City Bus Service of Xi'an*, Chang'an University, Xian, China, 2014.
- [36] Y. Chen, *A Comparative Analysis on the Transportation Costs between Public Transportation and Cars*, Nanjing Forestry University, Nanjing, China, 2009.
- [37] H. Cai and S. Xie, "Determination of emission factors from motor vehicles under different emission standards in China," *Journal of Peking University: Natural Science Edition*, vol. 03, pp. 319–326, 2010.
- [38] Q.-B. Wu, F. Chen, Y. Huang, and Y.-Y. Hu, "Calculation and analysis of traffic congestion cost in Beijing," *Journal of Transportation Systems Engineering and Information Technology*, vol. 11, no. 1, pp. 168–172, 2011.



**Hindawi**

Submit your manuscripts at  
<https://www.hindawi.com>

