

Research Article

An Improved NSGA-II Algorithm for Transit Network Design and Frequency Setting Problem

Shushan Chai  and Qinghuai Liang 

School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Shushan Chai; sschai@bjtu.edu.cn

Received 15 August 2019; Revised 24 October 2019; Accepted 31 October 2019; Published 16 January 2020

Academic Editor: Alain Lambert

Copyright © 2020 Shushan Chai and Qinghuai Liang. 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.

The transit network design and frequency setting problem is related to the generation of transit routes with corresponding frequency schedule. Considering not only the influence of transfers but also the delay caused by congestion on passengers' travel time, a multi-objective transit network design model is developed. The model aims to minimize the travel time of passengers and minimize the number of vehicles used in the network. To solve the model belongs to a NP-Hard problem and is intractable due to the high complexity and strict constraints. In order to obtain the better network schemes, a multi-population genetic algorithm is proposed based on NSGA-II framework. With the algorithm, network generation, modechoice, demand assignment, and frequency setting are all integrated to be solved. The effectiveness of the algorithm which includes the high global convergence and the applicability for the problem is verified by comparison with previous works and calculation of a real-size case. The model and algorithm can be used to provide candidates for the sustainable policy formulation of urban transit network scheme.

1. Introduction

The urban transit planning process includes the following phases: network design, frequency setting, timetable development, vehicle, and driver scheduling [1, 2]. The network design with corresponding frequency is related to the scientific configuration of urban transit facilities and the improvement of network service level, which will affect the development of the city and the evolution of urban space morphology. This paper focuses on the network design and frequency setting problem.

The cross-sectional passenger volume and the congestion situation of a line, which depends on the frequency setting, should be considered in the network design. Furthermore, both the demand of users and the profit of managers should be considered at the same time, thus it is a complicated multi-objective transit network design problem (MTNDP) [3, 4]. The characteristics of the related studies are listed in the Table 1. As can be seen, generally, the problem is formulated by minimizing the travel cost [5–8], minimizing the construction and operation cost [7, 9], and maximizing the passenger attraction or coverage [5, 6, 9, 10] etc. The studies about the

problem can be classified into the entire network design and the routes design based on existing network [7]. At present, the multi-mode networks integration method [11] and the multi-phase integration method in a network mode [12, 13] are engaged. The effective estimation of passenger flow state [14], the congestion and the capacity limitation of transit network [15] will have a great influence on the network design. The optimal matching between the distribution of passenger flow and the network structure can lead to less travel time and higher transport efficiency.

The solving of MTNDP is usually treated as an intractable problem for its high degree of complexity, the multiple objectives against each other and strict constraints, etc. The constraints and variables will explode with the expansion of network size. With the rapid development of computer technology and operational research, mathematical computing software and heuristic algorithm have been applied in transit network planning [6, 9, 10, 16]. However, for large-scale network, its nonlinear problem cannot be easily solved by conventional optimization software. As the classical approximative method, heuristic methods are often used, which includes constructive algorithm, local improvement algorithm, and the

TABLE 1: The characteristics of the related studies on transit network design problem.

Paper	Objectives	Demand	Methodology
G. Laporte et al. (2004)	Trip covering	Mode choice depended	Conventional (CPLEX)
W. Fan et al. (2008)	The sum of operator cost, user cost, and unsatisfied demand costs	Mode choice depended	Heuristic (systematic tabu search)
Á. Marín et al. (2009)	Travel cost, attractiveness	Mode choice depended	Conventional (piecewise linearization)
B. García-Archilla et al. (2011)	Trip coverage	Mode choice depended	Heuristic (GRASP algorithms)
G. Gutiérrez-Jarpa et al. (2013)	Travel cost, traffic capture	Attraction depended between OD pairs	Conventional (branch-and-cut method with CPLEX)
J. S. C. Chew et al. (2013)	Travel cost, operation cost	Fixed	Heuristic/matheuristic (genetic algorithm)
M. Nikolić et al. (2013)	The number of satisfied passengers, transfers, and travel time	Fixed	Metaheuristic (bee colony optimization)
D. Canca et al. (2014)	Net profit of the railway network	Mode choice depended	Heuristic (combination of branch-and-bound and local nonlinear optimization)
G. Gutiérrez-Jarpa et al. (2017)	Construction cost, time savings, patronage	Mode choice depended	Conventional (ϵ -constraint method with CPLEX)
F. López-Ramos et al. (2017)	Travel cost, costs related to construction and exploitation	Fixed	Heuristic/matheuristic (corridor generation algorithm and line splitting algorithm)
D. Canca et al. (2017)	Discounted profit	Mode choice depended	Matheuristic (adaptive large neighborhood search algorithm)
A. T. Buba et al. (2018)	Passenger cost, unmet demand	Fixed	Matheuristic (differential evolution)
X. Feng et al. (2019)	Travel time	Fixed	Heuristic/matheuristic (genetic algorithm)
S. Bhushan Jha et al. (2019)	Travel time, operation cost	Fixed	Heuristic/matheuristic (genetic algorithm, particle swarm optimization)

combinations of them [17–19]. However, many traditional heuristic algorithms have poor global convergence, and the results sometimes depend largely on the initial solutions. If the lines are selected from a given line pool to form the network, the rationality of these lines will largely depend on the quality of the line pool [20]. At present, the metaheuristic methods, such as genetic algorithm [8, 21, 22], simulated annealing algorithm [13], tabu search algorithm [23], swarm intelligence algorithm [24] and other metaheuristics [25, 26] are usually applied in the transit network design problem.

However, the generating of the network and the setting of the frequency, which are interrelated with each other, are separated in some methods. And it is particularly important to use the algorithm with high global convergence ability to search for the better network scenarios when the exact solutions cannot be obtained. In this paper, an effective algorithm based on improved NSGA-II is put forward, in which the network generation, the mode choice, the passenger demand assignment considering passenger trip rule, and the frequency setting are all integrated. The algorithm has high global convergence and good applicability to the problem, which are verified by numerical experiments and case application.

The structure of this paper is as follows: The mathematical model of MTNDP is presented in Section 2. Section 3 describes the multi-objective algorithm for the MTNDP, followed by

numerical experiments in Section 4 and case application in Section 5. Section 6 presents the conclusions.

2. Mathematical Model

2.1. Notations. Let N be the set of all nodes which represent the stations, E be the set of edges which represent the sections in the transit network, $(i, j) \in E$ represent a section between adjacent stations $i \in N$ and $j \in N$, $i \neq j$. The notations in the model are defined as follows.

First, variables associated with the connection among nodes in the network are defined as follows.

- (i) $x_{ij}^l = 1$ if line l contains the section (i, j) , 0 otherwise.
- (ii) $y_i^l = 1$ if line l contains the station i , 0 otherwise.
- (iii) $x_{ij} = 1$ if the network contains the section (i, j) , 0 otherwise.
- (iv) $h_l = 1$ if line l is included, 0 otherwise.
- (v) r is an effective path between starting and ending points.
- (vi) R is the set of effective paths, $r \in R$.
- (vii) $N(i)$ is the set of neighbour nodes of node i .
- (viii) T^l is the route cycle time in line l , min.

- (ix) N_n^l is the total number of stations in line l .
- (x) P_{\max}^l is the maximum cross-sectional passenger volume of line l at peak hour.
- (xi) t_h^l is the average headway of vehicles in line l .

Then, the variables associated with the changes of travel path r are defined as follows.

- (i) $f_{ij}^{serl} = 1$ if the passenger flow corresponding to the OD pair se uses the section (i, j) of line l along the path r , 0 otherwise.
- (ii) $f_k^{serll'}$ = 1 if the passenger flow corresponding to the OD pair se transfers from line l to line l' at station k along the path r , 0 otherwise.
- (iii) N^r is the total number of stations in path r .
- (iv) M^r is the transfer times needed along the path r .
- (v) s_h^r is the headway of vehicles at the starting point of path r .
- (vi) t_{hk}^r is the headway of vehicles taken by passengers travelling after the k^{th} transfer in path r .
- (vii) g_{se}^r is the transit passenger volume travelling along the path r between the starting point s and the ending point e .

Finally, the known quantities in the model are defined as follows.

- (i) N_L is the number of lines in the transit network.
- (ii) $N_{L \max}$ is the maximum limited number of lines in the network.
- (iii) $N_{L \min}$ is the minimum limited number of lines in the network.
- (iv) $N_{n \max}^l$ is the maximum limited number of stations in line l .
- (v) $N_{n \min}^l$ is the minimum limited number of stations in line l .
- (vi) D_{\max}^l is the maximum limited length of line l .
- (vii) D_{\min}^l is the minimum limited length of line l .
- (viii) M_{\max} is the maximum limited transfer times in a path.
- (ix) V^l is the number of seats of each vehicle in line l .
- (x) μ^l is the maximum limited load factor of the vehicles in line l .
- (xi) g_{se} is the transit passenger demand between the starting point s and the ending point e .
- (xii) C_{ij} is the capacity of section (i, j) .
- (xiii) d_{ij} is the length of section (i, j) .
- (xiv) d_{se} is the shortest distance between the starting point s and the ending point e .
- (xv) $t_{i,i+1}$ is the travel time of section $(i, i+1)$.
- (xvi) t_{mi}^r is the dwell time of vehicles at the i^{th} nontransfer station in path r .
- (xvii) t_{zk}^r is the walking time for the k^{th} transfer in path r .
- (xviii) $t_{h \min}$ is the minimum limited headway of vehicles.

2.2. Model Formulation. The generalised travel cost function on paths in transit network is formulated to calculate the passenger travel time, including the waiting time at starting point, in-vehicle travel time and transfer time. The in-vehicle travel time includes the travel time in sections and the dwell time at intermediate stations. The transfer time includes the walking time and the waiting time at transfer stations. Assume that all the passengers arrive at the stations randomly and are able to board the first vehicle on the route they encounter, the average waiting time of passengers can be considered as half of the headway [27–29]. Let T_1^r be the in-vehicle travel time, T_2^r be the total transfer time, then they can be expressed by Equations (1) and (2), respectively.

$$T_1^r = \sum_{i=1}^{N-M^r} t_{mi}^r + \sum_{i=1}^{N-1} t_{i,i+1}, \quad r \in R, \quad (1)$$

$$T_2^r = \sum_{k=1}^{M^r} (t_{zk}^r + 0.5t_{hk}^r), \quad r \in R. \quad (2)$$

Passengers have different sensitivities to the transfer times in the same path. Under the same travel time, the more the transfer times are, the less likely the path would be selected by passengers. In most previous works, such as in [8, 25, 30–32], the fixed penalty coefficient was used to describe the influence of transfers on the travel time. In this paper, the sensitivity coefficient α of transfer times and the adjustment parameter β are introduced to adjust the transfer penalty, so the total transfer time can be expressed by

$$T_2^{r'} = \sum_{k=1}^{M^r} [(t_{zk}^r + 0.5t_{hk}^r)(1 + \alpha k)^\beta], \quad r \in R, \quad (3)$$

where, α represents passenger's sensitivity to the k^{th} transfer during the travel. β is the adjustment parameter, generally greater than 1. The larger the β , the greater the difference for sensitivity levels between different transfer times in the path.

Excessive passengers waiting at station and excessive load factor of incoming vehicle will have a great impact on the passenger travel delay, which can cause an increase in travel time. In order to describe the impact of congestion on travel time while entering a station or transferring, the travel delay coefficient φ_i of the station is defined, as shown in

$$\varphi_i = 1 + \gamma \left(\frac{v_i^r}{c_i^r} \right)^\lambda, \quad i \in N, r \in R, \quad (4)$$

where, v_i^r and c_i^r are the cross-sectional passenger flow and the capacity of next section of the station i in path r , respectively. γ and λ are the adjusting coefficients.

Let C_{se}^r be the generalized travel cost of the path r between the two points s and e , then the C_{se}^r can be expressed by

$$C_{se}^r = 0.5s_h^r \varphi_s + \sum_{k=1}^{M^r} [(t_{zk}^r + 0.5t_{hk}^r)(1 + \alpha k)^\beta \varphi_k] + \sum_{i=1}^{N-M^r} t_{mi}^r + \sum_{i=1}^{N-1} t_{i,i+1}, \quad se \in OD, r \in R. \quad (5)$$

The mode choice model is used to calculate the transit passenger demand g_{se} between OD pair se . Assume that any segment of the basic road network can be used by other modes, such as private car, bicycle, and walk. Since the travel cost is usually the most important influencing factor of mode choice, it is assumed that the mode choice probability depends exclusively on the generalised travel costs using different modes. Compared to the absolute utility, the passengers are more concerned about their relative utility. So the choice probability of the transit can be simply calculated by

$$p_{se}^m = \frac{\exp\left(-C_{se \min}^m / \bar{C}_{se \min}\right)}{\sum_i \exp\left(-C_{se \min}^i / \bar{C}_{se \min}\right)}, \quad se \in OD, \quad (6)$$

where p_{se}^m is the choice probability of public transport for passengers traveling between the OD pair se , $C_{se \min}^m$ is the shortest time traveling between the OD pair se using public transport, $C_{se \min}^i$ is the shortest travel time using the i^{th} mode, $\bar{C}_{se \min}$ is the average shortest travel time of all the modes used.

From the perspective of passengers, the minimum travel time of passengers is regarded as one of objectives to improve the network service level and travel efficiency, which is shown in

$$\min Z_1 = \sum_{se \in OD} \sum_{r \in R} g_{se}^r C_{se}^r. \quad (7)$$

From the perspective of managers, the minimum number of vehicles used in the network is regarded as another objective to reduce operating costs and improve the economic benefit, which is shown in Equations (8)–(10).

$$\min Z_2 = \sum_{l \in L} \left[\frac{T^l}{t_h^l} \right], \quad (8)$$

where, the headway t_h^l is limited to integer for the convenience of operation and management, as shown in Equation (9), and the maximum cross-sectional passenger volume is calculated by Equation (10).

$$t_h^l = \left\lfloor \frac{60V^l \mu^l}{P_{\max}^l} \right\rfloor, \quad l \in L, \quad (9)$$

$$P_{\max}^l = \max \left\{ \sum_{s,e \in OD} \sum_{r \in R} g_{se}^r f_{ij}^{serl}, \sum_{se \in OD} \sum_{r \in R} g_{se}^r f_{ji}^{serl} \right\}, \quad (10)$$

$$l \in L, (i, j), (j, i) \in E.$$

Multiple constraints are included related to network topology relationship, connectivity between stations, limited transfer times, network size, etc., as shown in

$$x_{ij}^l \leq y_i^l, \quad l \in L, i \in N, (i, j) \in E, \quad (11)$$

$$x_{ij}^l \leq y_j^l, \quad l \in L, j \in N, (i, j) \in E, \quad (12)$$

$$x_{ij}^l = x_{ji}^l, \quad l \in L, (i, j) \in E, \quad (13)$$

$$x_{ij}^l = x_{ji}^l, \quad l \in L, (i, j) \in E, \quad (14)$$

$$x_{ij}^l \leq x_{ij}, \quad l \in L, (i, j) \in E, \quad (15)$$

$$f_{ij}^{serl} \leq x_{ij}^l, \quad l \in L, (i, j) \in E, se \in OD, \quad (16)$$

$$\sum_{j \in N(i)} x_{ij}^l \leq 2, \quad l \in L, i \in N, (i, j) \in E, \quad (17)$$

$$h_l + \sum_{(i,j) \in E} x_{ij}^l = \sum_{i \in N} y_i^l, \quad l \in L, i \in N, \quad (18)$$

$$\sum_{i \in B} \sum_{j \in B} x_{ij}^l \leq |B| - 1, \quad B \subset N, |B| \geq 2, l \in L, (i, j) \in E, \quad (19)$$

$$\sum_{(i,j) \in E} |x_{ij}^l - x_{ij}^{l'}| \neq 0, \quad l \neq l', l, l' \in L, \quad (20)$$

$$\sum_{l \in L} y_i^l = |N_c|, \quad i \in N, \quad (21)$$

$$g_{se} = \sum_{r \in R} g_{se}^r, \quad se \in OD, \quad (22)$$

$$f_k^{serll'} = \left\lfloor \frac{1}{2} [f_{ik}^{serl} + f_{kj}^{serl'}] \right\rfloor, \quad k \in N, r \in R, l, l' \in L, \quad (23)$$

$$se \in OD, (i, k), (k, j) \in E,$$

$$\sum_{k \in N} f_k^{serll'} \leq M_{\max}, \quad r \in R, l, l' \in L, se \in OD, (i, k), (k, j) \in E, \quad (24)$$

$$\sum_{k \in N(i)} f_{ik}^{serl} - \sum_{k \in N(i)} f_{ki}^{serl} = \begin{cases} 1, & i = s, \\ -1, & i = e, \\ 0, & \text{else,} \end{cases}$$

$$r \in R, l \in L, i \in N, se \in OD, \quad (25)$$

$$\sum_{l' \in L, l' \neq l} \sum_{k \in N} \sum_{r \in R} f_k^{serll'} + \sum_{(s,i) \in E} \sum_{r \in R} f_{si}^{serl}$$

$$= \sum_{l \in L, l' \neq l} \sum_{k \in N} \sum_{r \in R} f_k^{serll'} + \sum_{(j,e) \in E} \sum_{r \in R} f_{je}^{serl}, \quad se \in OD, \quad (26)$$

$$f_{ij}^{serl} + f_{ji}^{serl} \leq 1, \quad r \in R, se \in OD, l \in L, (i, j) \in E, \quad (27)$$

$$\sum_{se \in OD} \sum_{r \in R} g_{se}^r f_{ij}^{serl} \leq C_{ij}, \quad (i, j) \in E, l \in L, \quad (28)$$

$$N_{n \min}^l \leq N_n^l \leq N_{n \max}^l, \quad l \in L, \quad (29)$$

$$D_{\min}^l \leq \sum_{\{i,j\} \in E} d_{ij} x_{ij}^l \leq D_{\max}^l, \quad l \in L, \quad (30)$$

$$N_{L \min} \leq N_L \leq N_{L \max}, \quad (31)$$

$$\text{Integer } s_h^r, t_{hk}^r, t_h^l \geq t_{h \min}. \quad (32)$$

The terms (11)–(15) are the network topology constraints, indicating the relationship between nodes and edges. The term (16) indicates that the flow corresponding to the OD pair se can use the section (i, j) of a line l along the path r only if the section (i, j) belongs to the line l . The term (17) indicates that line l has no branch at point i . The terms (18) and (19) indicate that line l is continuous and loop-free, where $B \in N$ is an arbitrary node set. Constraint (20) makes different lines have different nodes or connections. The term (21) indicates that the isolated nodes are not allowed in the network, where $N_c \subset N$ is a candidate node set. The term (22) is the network connectivity constraint, which indicates that if there is passenger demand between two stations, the two stations must be connected in the network. Constraints (23) and (24) limit the transfer times for a path. The term (25) indicates that the passenger flow corresponding to the OD pair se only has the inflow at the starting point s and the outflow at the ending point e , and the inflow is equal to the outflow at the intermediate stations of the path r . The term (26) indicates that the passenger flow entering the line l is equal to that exiting the line l . The term (27) indicates that reflow is not allowed in the path. The term (28) limits the capacity of sections. Constraint (29) limits the number of stations in line l . Constraint (30) limits the length of line l . Constraint (31) limits the number of lines in the network. Constraint (32) limits the minimum value of the integral headway.

3. Improved NSGA-II Algorithm for the MTNDP

The multi-objective network design model developed in this paper belongs to the NP-hard problem mathematically. NSGA-II [33], which is the nondominated sorting genetic algorithm with elite strategy, is one of very effective algorithms to solve the multi-objective problem. Compared with NSGA [34], the algorithm reduces the complexity and has the advantage of better robust performance. In NSGA-II, there are three important concepts that need to be explained as follows. For detailed description of the algorithm, please refer to the literature [33] by Deb et al.

(1) Fast nondominated sort

If the objective functions of an individual is no worse than the others, the individual is the nondominated solution. In the iterative process, all the individuals are assigned to several fronts according to the noninferior relationships. The individuals in the next front is dominated by that in the previous front, and each individual i is assigned a rank value P_{rank}^i .

(2) Crowding distance

In order to evaluate the distribution density of other individuals around a particular individual, the average distance between the two sides of the individual needs to be calculated, which is called the crowding distance of the individual. The distance for the two boundary individuals is assigned to infinity, and for the other individuals, the distances is calculated by the Equation (33).

$$cd^i = \sum_k cd_k^i = \sum_k \frac{f_k^{i+1} - f_k^{i-1}}{f_{k, \max}^i - f_{k, \min}^i}, \quad (33)$$

where, cd_k^i is the crowding distance of the k^{th} objective of the i^{th} individual. f_k^{i+1} and f_k^{i-1} are the k^{th} objective function values of the $(i+1)^{th}$ and $(i-1)^{th}$ individuals, respectively. $f_{k, \max}^i$ and $f_{k, \min}^i$ are the maximum and minimum values of the k^{th} objective function of the i^{th} individual.

(3) Crowded comparison

For two individuals i and j , if the ranks P_{rank}^i and P_{rank}^j are same, the crowding distances cd^i and cd^j are compared. In the comparison, if $P_{rank}^i < P_{rank}^j$ or $P_{rank}^i = P_{rank}^j$ but $cd^i > cd^j$, the i^{th} individual is superior to the j^{th} individual.

In this paper, a multi-population genetic algorithm is proposed based on the framework of NSGA-II. Figure 1 presents the flowchart of the algorithm proposed in this paper for the MTNDP.

3.1. Multi-Population Dynamic Coding. In the process of chromosome coding, not only the connections among stations, but also the lines to which the stations belong should be considered. In order to effectively store the topological information and improve the convenience of coding for transit networks, a multi-population dynamic coding method is proposed.

In this paper, the integer coding method is adopted, and the concepts of “main population” and “affiliated population” are proposed. The information about connections among stations is stored in the chromosomes of “main population”, and the information about lines to which the stations belong is stored in the corresponding chromosomes of “affiliated population”, as shown in Figure 2.

In Figure 2, alleles on chromosome in the ‘main population’ represent the stations while the alleles in its “affiliated population” represent the line labels, and the stations that belong to the same line are connected sequentially. An individual which is a characteristic entity of a chromosome represents a network scheme. If a station appears multiple times in a chromosome of “main population”, it represents a transfer station. The chromosomes in the same population have different lengths because of different number of transfer stations in networks. In the iteration process, the chromosomes in the “main population” and its “affiliated population” whose lengths are dynamically changed have one-to-one correspondence.

3.2. Generation of Initial Populations. For large network, it is difficult to generate a large number of initial solutions with high diversity under strict constraints. At present, the main initial solution generation procedures can be classified as follows.

Random connection method of adjacent nodes. In this method, the routes are generated by the connection between adjacent nodes, such as the probability-based IRSG procedure proposed by Jha et al. [26]. The method usually has high calculation efficiency, but cannot guarantee that the travel time between OD pairs will not be too long along the routes.

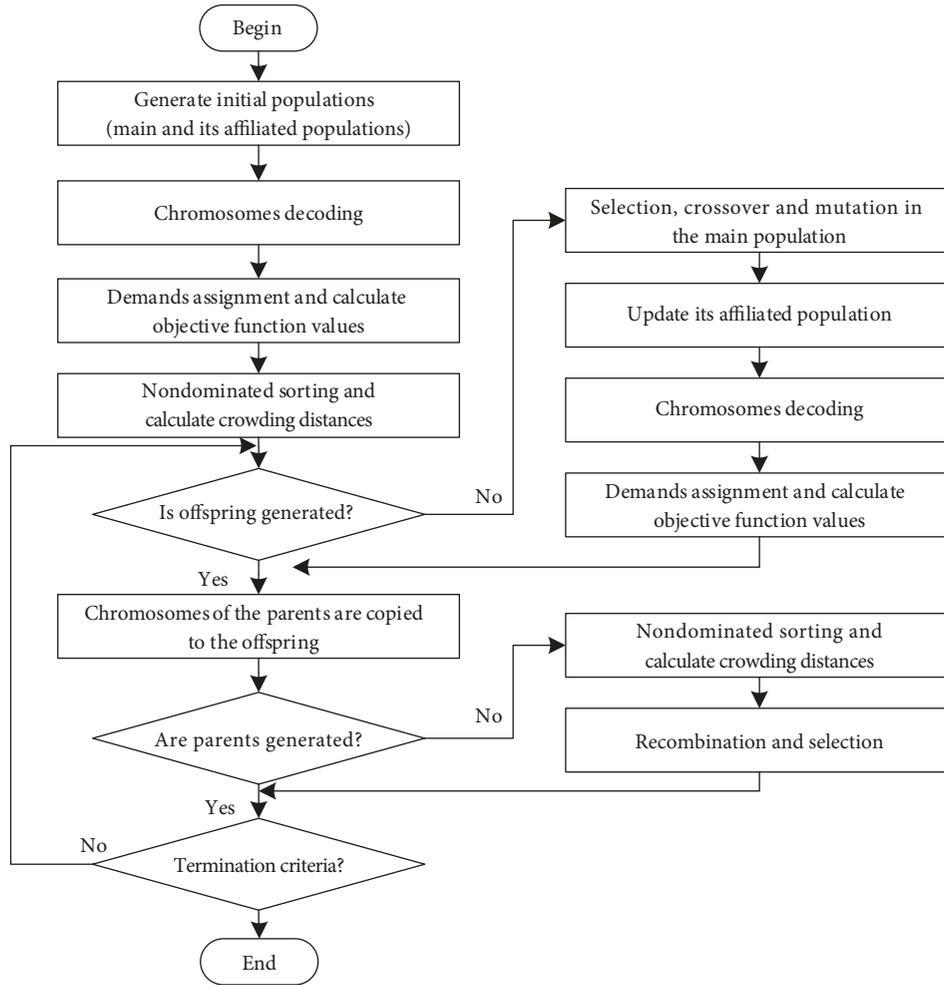


FIGURE 1: Flowchart of the proposed algorithm for the MTNDP.

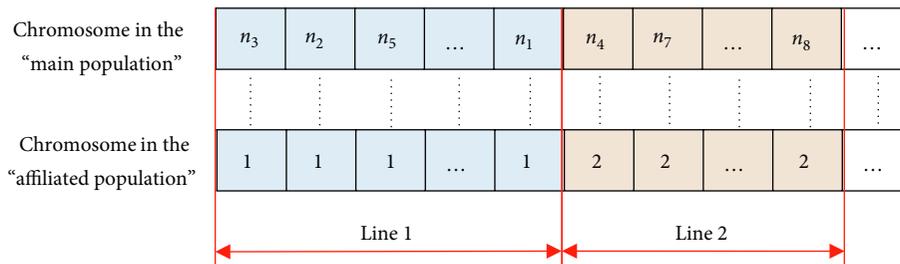


FIGURE 2: Multi-population chromosomes coding.

Shortest-path-based method. In this method, the routes are generated based on the shortest paths searching between OD pairs, and another procedure dealing with constraints is usually needed. For example, the construction and repair procedures used by Ahmed et al. [35] and Szeto and Jiang [36], the initial solution generation procedure based on greedy algorithm used by Nikolić and Teodorović [24] and Nayeem et al. [37], and the Floyd’s algorithm and feasibility check used by Chew et al. [22]. With the method, the direct demand between departure and destination stations can be met, but more vehicles on the network will be needed [38].

Multipath-based method. In this method, the routes are generated based on effective paths, such as the RGA and RFA procedures for initial solution skeleton proposed by Mahdavi Moghaddam et al. [38]. With the method, the initial solutions can be generated from both the perspectives of passengers and managers, but the number of effective paths between OD pairs to be searched will have a significant impact on the calculation time of the procedure.

In the existing procedures, the isolated nodes (i.e., nodes are in the connected basic road network but not served by the routes) cannot be effectively circumvented, and the

characteristics of transit network structure are not well reflected in the initial solutions, such as the phenomenon of “rich club”, that is, nodes with larger degree values (“rich nodes”) tend to be connected to each other [39]. For transit network, the endpoints of lines have the smallest average degree value while the transfer stations have the largest. Therefore, in order to generate initial schemes with actual network characteristics, an initial population generation method is proposed in this paper.

(1) Generation of the endpoint set of lines

The node with smaller degree is more likely to be selected as the endpoint which represents the departure or the terminal station of a line. In this paper, the line endpoints are randomly selected based on the roulette method, and the selection probability can be calculated by the Equation (34).

$$P_{SEi} = \frac{2}{|N_c|} - \frac{D_i}{\sum_{i \in N_c} D_i}, \quad (34)$$

where, P_{SEi} is the probability that node i is selected as the endpoint of a line, satisfying $\sum_{i \in N_c} P_{SEi} = 1$, D_i is the degree value of candidate node i .

(2) Generation of lines

The remaining candidate nodes are randomly selected as intermediate stations of a line after the endpoints are selected. Firstly, the K -shortest path search method in nonweighted road network is used to get the top K_l paths between two endpoints of a line. All the top K_l paths should meet the constraint about nodes number which can be expressed by the distance between the two endpoints of each path, as shown in Equation (35).

$$N_{n,min}^l \leq C_{se}^d + 1 \leq N_{n,max}^l \quad r \in K_l, \quad (35)$$

where, C_{se}^d is the distance between endpoints s and e in the nonweighted network.

Then, a path is selected from the K_l paths as the first line according to a certain probability. The larger the average degree value of the intermediate nodes in the path, the greater the probability that the path is selected. The roulette method is also used and the probability can be expressed by Equation (36).

$$P_{Er} = \frac{\overline{D_{Mr}}}{\sum_{r \in K_l} \overline{D_{Mr}}}, \quad (36)$$

where, P_{Er} is the probability that the candidate path r is selected, satisfying $0 < P_{Er} < 1$ and $\sum_{r \in K_l} P_{Er} = 1$, $\overline{D_{Mr}}$ is the average degree value of intermediate nodes in the path r .

After the first line is generated, the paths between the two endpoints of next line are searched with the same method to constitute the set K'_l . The paths in K'_l that contains the isolated nodes in the road network are stored in the new set as the candidate paths of the next line. So, with the increase of generated lines, the number of candidate paths and isolated nodes decrease gradually.

Finally, if isolated nodes are not allowed and still exist in the generated network, they will be randomly inserted into

Input: Integer $Population_size$, $N_{n,min}^l$, $N_{n,max}^l$ and $Times_tried_{max}$, Set of degree values D , Set All_nodes .

```

1: while PPop < Population_size do
2:   for each l ∈ L do
3:     Select endpoints s, e of each line with probability
       of PSEi by roulette
4:   for each s, e ∈ SEnodes set do
5:     Search for K-shortest path that satisfies
       Nn,minl ≤ Csed + 1 ≤ Nn,maxl
6:   Remaining_nodes ← All_nodes
7:   Times_tried = 0
8:   while Remaining_nodes ≠ ∅ and
       Times_tried < Times_triedmax
9:     do Select path k1 ∈ K1 as line l1 with probability
       of PEr by roulette
10:    Remaining_nodes ←
       {node|node ∈ Remaining_nodes, node ∉ k1}
11:    for each l ∈ L - l1 do
12:      if Remaining_nodes ∩ {node|node ∈ Kl} ≠ ∅ then
13:        Paths_chosen ←
       {Path|Path ∈ Kl, Path ∩ Remaining_nodes ≠ ∅}
14:        else Paths_chosen ← Kl
15:        Select path kl ∈ Paths_chosen as line l with
       probability of PEr by roulette
16:        Remaining_nodes ←
       {node|node ∈ Remaining_nodes, node ∉ kl}
17:    if Remaining_nodes ≠ ∅ then
18:      for each node ∈ Remaining_node
19:        Randomly inserted into line l as an inter-
       mediate node
20:        Times_tried = Times_tried + 1
21:    if not meeting connectivity constraint then
22:      Restart endpoints and lines generation procedure
23:    Chromosomes coding and stored in the set
       Initial_population
24:    Pop = Pop + 1
25:  return Initial_population

```

ALGORITHM 1

the generated lines as intermediate nodes. In order to maintain the diversity of species, the repetitive chromosomes in the population are removed to ensure their uniqueness. The pseudo code for the initial population generation algorithm is shown in Algorithm 1.

3.3. Demand Assignment. For comparison, the assumption about paths selection is the same as in the literature [40] by Arbex and Cunha, that is, assuming that passengers only select the paths with no more than two transfers. The passenger initially selects the direct path to avoid transfer. If the direct path is not available, only a single-transfer path will be selected. If there is no single-transfer path, then the two-transfer path is selected. Ultimately, if there is no two-transfer path, it is considered that the passengers' travel demand cannot be covered.

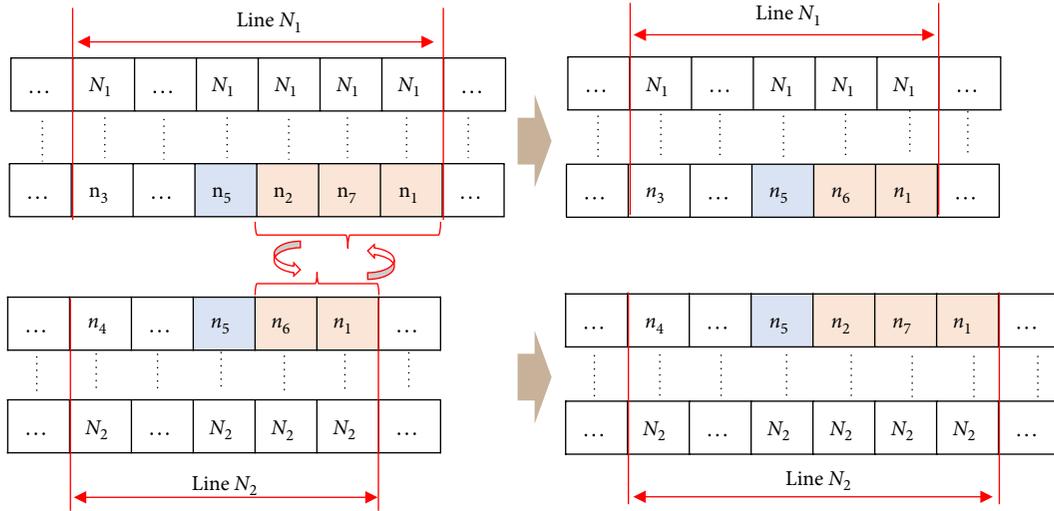


FIGURE 3: Single-line crossover.

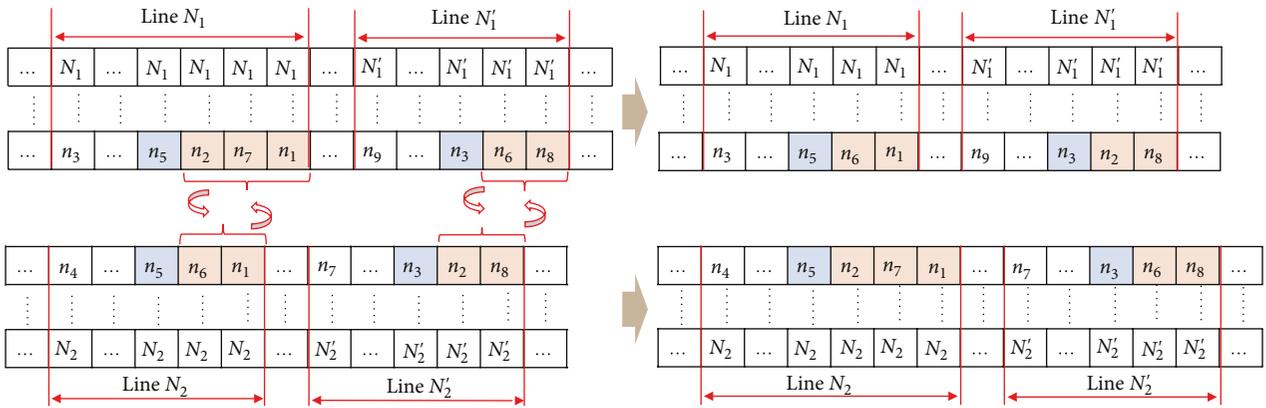


FIGURE 4: Multi-line crossover.

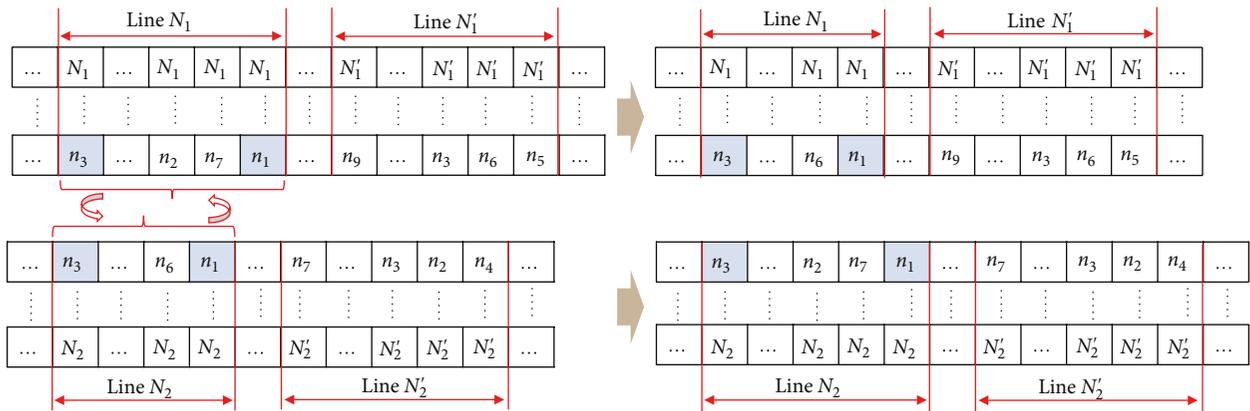


FIGURE 5: Exchange of lines with the common endpoints.

The multipath incremental assignment method is used in this paper. The OD passenger flow is divided into multiple parts equally that is presented in multiple OD matrices and distributed to the network successively. Before an OD matrix is assigned, the selection of paths and lines for passengers are calculated according to the assignment result of the previous OD matrix. In the initial state, there is no passengers on each line and the waiting

time is set to an initial nonzero value. If there are multiple paths, the relative utility of the path r is used to calculate the selection probability p_{se}^r , as shown in Equation (37).

$$p_{se}^r = \frac{\exp(-\theta C_{se}^r / \bar{C}_{se})}{\sum_{q \in R} \exp(-\theta C_{se}^q / \bar{C}_{se})}, \quad r \in R, \quad (37)$$

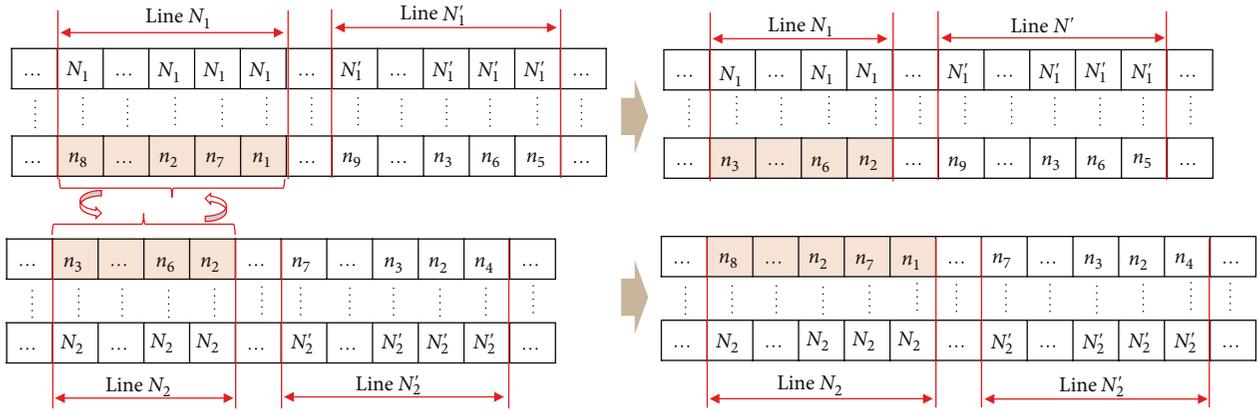


FIGURE 6: Exchange of lines without common endpoints.

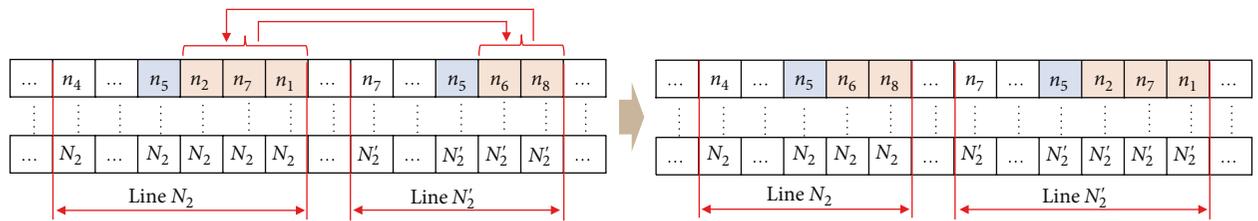


FIGURE 7: Intra-chromosome crossover.

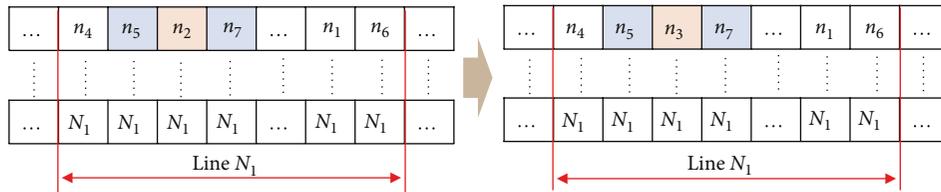


FIGURE 8: Mutation.

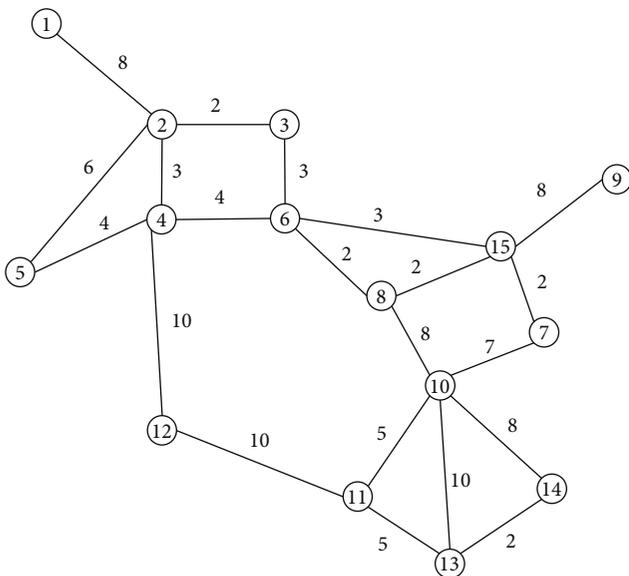


FIGURE 9: Mandl's network.

where, θ is the calibration parameter, $\overline{C_{se}}$ is the average travel cost for the effective paths between the OD pair se .

3.4. Two Types of Crossover. According to the characteristics of transit network, the crossover operators of chromosomes can be classified into two types: inter-chromosome crossover and intra-chromosome crossover.

(1) Inter-chromosome crossover

Inter-chromosome crossover is the replacement and recombination of nodes and connections between different parent individuals. According to the number of changed lines in a chromosome, it can be divided into single-line crossover and multi-line crossover, as shown in Figures 3 and 4. Taking the single-line crossover between two parents as an example, the steps are as follows.

Step 1. If there are common endpoints between lines belonging to the two parents, the common endpoints are selected to constitute the set EN_{com} and the two lines are selected to constitute the set PL_1 and PL_2 respectively, then turn to Step 2. Otherwise, turn to Step 5.

Step 2. If the lines $pl_1 \in PL_1$ and $pl_2 \in PL_2$ have common intermediate nodes, the intermediate nodes are selected to constitute the set IN_{com} and turn to Step 3. Otherwise, turn to Step 4.

Step 3. Randomly select two lines $pl_1 \in PL_1$ and $pl_2 \in PL_2$ which have the endpoint $en_{com} \in EN_{com}$ and intermediate node $in_{com} \in IN_{com}$. All the nodes from in_{com} to en_{com} of the two lines exchange with each other, and update the corresponding chromosomal alleles in the “affiliated population”. If the offspring individuals meet the other constraints, the crossover ends. Otherwise, turn to Step 4.

Step 4. The two lines which have common endpoints are exchanged as a whole with each other between the two parents, and update the corresponding chromosomal alleles in the “affiliated population” (as shown in Figure 5). If the offspring individuals meet the other constraints, the crossover ends. Otherwise, turn to Step 5.

Step 5. Randomly select a line from each parent and exchange them as a whole, then update the corresponding chromosomal alleles in the “affiliated population” (as shown in Figure 6).

By executing the inter-chromosome crossover, not only the network size which includes the total number of stations and the total mileage of the network, but also the line characteristics which include the starting and ending points, the alignment, the mileage, the number of stations of each line and the transfer relationship between lines can be adjusted to get different network schemes.

(2) Intra-chromosome crossover

Intra-chromosome crossover is the replacement and recombination of nodes and connections between lines in the same parent individual, as shown in Figure 7. It can't change the network size, but the line characteristics. The steps of the intra-chromosome crossover in a parent are as follows:

Step 1. Search for the transfer nodes in the parent to constitute the set TR_{com} .

Step 2. Randomly select a node $tr_{com} \in TR_{com}$, and search for the lines with node tr_{com} to constitute the set PL .

Step 3. Randomly select two lines $pl_1, pl_2 \in PL_2$. All the nodes from tr_{com} to an endpoint of the two lines exchange with each other, and update the corresponding chromosomal alleles in the “affiliated population”.

3.5. Adaptive Mutation. Each iteration can result in Pareto solutions and may contain only partial solutions of the original problem. The nondominated solutions may be dominated by new solutions generated after successive iterations. In this paper, if new nondominated solutions cannot be generated by multiple iterations, the mutation operator will be triggered to jump out of the local optimal solutions, as shown in Figure 8. The steps for mutation of a parent are as follows.

Step 1. Randomly select a line from the parent and three adjacent nodes $Node_1, Node_2,$ and $Node_3$ in order from the selected line.

Step 2. Search for the nodes that can be connected to the nodes $Node_1$ and $Node_3$ to constitute the set NEN_{com} .

Input: Integer $Population_size$, Genetic probabilities $CrossoverI_probability$ $CrossoverII_probability$ and $Mutation_probability$.

```

1: Selection based on tournament selection process
2: while  $Offspring\_size < Population\_size$ 
3:   do Randomly select two parent chromosomes
4:    $P1 \leftarrow Random\_probability \in (0, 1)$ 
5:   if  $P1 < CrossoverI\_probability$  then
6:     Execute inter-chromosome crossover
7:     if Offspring chromosomes meet the constraints then
8:        $Is\_successful\_crossoverI \leftarrow 1$ 
9:     else  $Is\_successful\_crossoverI \leftarrow 0$ 
10:     $P2 \leftarrow Random\_probability \in (0, 1)$ 
11:    if  $P2 < CrossoverII\_probability$  or  $Is\_successful\_crossoverI = 0$  then
12:      Execute intra-chromosome crossover
13:      if Offspring chromosomes meet the constraints then
14:         $Is\_successful\_crossoverII \leftarrow 1$ 
15:      else  $Is\_successful\_crossoverII \leftarrow 0$ 
16:       $P3 \leftarrow Random\_probability \in (0, 1)$ 
17:      if  $P3 < Mutation\_probability$  or  $Is\_successful\_crossoverII = 0$  then
18:        Execute mutation
19:        if Offspring chromosomes meet the constraints then
20:           $Is\_successful\_mutation \leftarrow 1$ 
21:        else  $Is\_successful\_mutation \leftarrow 0$ 
22:         $Offspring\_set \leftarrow Offspring\_set \cup Offspring\_chromosomes$ 
23:      return  $Offspring\_set$ 

```

ALGORITHM 2

Step 3. Randomly select a node $nen_{com} \in NEN_{com}$ to replace the node $Node_2$. The corresponding chromosome in the “affiliated population” remains unchanged.

3.6. Dealing with Strict Constraints. Offspring population is generated by selection, crossover and mutation of parents. It is necessary to ensure the diversity of the offspring. However, the generated offspring may not satisfy the strict constraints after genetic operation. Therefore, an auxiliary algorithm is designed to deal with the strict constraints.

The genetic operators are executed according to certain probabilities. In this paper, intra-chromosome crossover will be executed when the inter-chromosome crossover is finished with ineligible offspring, and the mutation operator will also be executed when the two types of crossover operators are finished but none of eligible offspring is generated. If the size of the offspring does not meet the requirement, the genetic operators will be executed cyclically, and the new eligible offspring will be merged until the size meets the requirement. The pseudo code for the auxiliary algorithm is shown in Algorithm 2.

TABLE 2: OD matrix for Mandl's network.

O/D	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	400	200	60	80	150	75	75	30	160	30	25	35	0	0
2	400	0	50	120	20	180	90	90	15	130	20	10	10	5	0
3	200	50	0	40	60	180	90	90	15	45	20	10	10	5	0
4	60	120	40	0	50	100	50	50	15	240	40	25	10	5	0
5	80	20	60	50	0	50	25	25	10	120	20	15	5	0	0
6	150	180	180	100	50	0	100	100	30	880	60	15	15	10	0
7	75	90	90	50	25	100	0	50	15	440	35	10	10	5	0
8	75	90	90	50	25	100	50	0	15	440	35	10	10	5	0
9	30	15	15	15	10	30	15	15	0	140	20	5	0	0	0
10	160	130	45	240	120	880	440	440	140	0	600	250	500	200	0
11	30	20	20	40	20	60	35	35	20	600	0	75	95	15	0
12	25	10	10	25	15	15	10	10	5	250	75	0	70	0	0
13	35	10	10	10	5	15	10	10	0	500	95	70	0	45	0
14	0	5	5	5	0	10	5	5	0	200	15	0	45	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

TABLE 3: Parameter values of the model used in Mandl's network.

Parameters	Values
N_L	4, 6, 7, 8
$N_{\min}^l, l \in L$	3
$N_{\max}^l, l \in L$	15
$V^l, l \in L$	40
$\mu^l, l \in L$	1.25
$c_i^r, i \in N, r \in R$	$60V^l\mu^l$
$t_{mi}^r, i \in N, r \in R$	30 s
$t_{zk}^r, k \in N, r \in R$	2 min
H	1.5
α	1, 2, 3, 4
β	2
γ	0.15
λ	4
θ	1.866

4. Numerical Experiments

4.1. Parameter Values. The road network of Switzerland given by Mandl in [32] is used as the basic road network, as shown in Figure 9. In the road network, each node represents a city and each edge represents a road segment between two cities. The number on each edge represents the travel time between two cities. The OD matrix of the network are shown in Table 2.

For comparison, the number of seats V^l in each bus, the maximum limited load factor μ^l and the minimum number of stations N_{\min}^l allowed in a line are the same as those used in [40] by Arbex and Cunha. In this calculation, because the passenger demand is static, the mode choice model calculated by Equation (6) is not included. Since the headway of vehicles is not constrained to an integer in the literatures for comparison, the Equation (9) is calculated without rounding down, and the constraint (32) is not considered in this calculation. The parameter values in the model are shown in Table 3. The probabilities of inter-chromosome crossover, intra-chromosome crossover

TABLE 4: Transfer penalty coefficients under different sensitivities.

α	Corresponding transfer penalty coefficient	
	1 transfer	2 transfer
1	4.00	9.00
2	9.00	25.00
3	16.00	49.00
4	25.00	81.00

and mutation is set to 0.9, 0.9, and 0.1, respectively. The maximum number of iterations is set to 1000. Table 4 presents the transfer penalty coefficients under different sensitivities.

4.2. Results and Discussion. Figure 10 presents the comparison among buses, user cost, and direct trips of the networks for the scenarios with 4 lines under different transfer sensitivities α . The results show that the larger the α , the smaller the passengers' willingness to transfer in general, and the higher the user cost under the same direct rate. The solutions with better objective values can be obtained when $\alpha = 1$ under the given transfer sensitivities, but the direct rates of the solutions are obviously lower. Taking into account the objective values and the direct rates at the same time, the smaller number of buses or the lower user cost can be obtained when $\alpha = 3$ under a certain direct rate. Taking the four solutions including the solution P_1 with $\alpha = 1$, the solution P_2 with $\alpha = 2$, the solution P_3 with $\alpha = 3$ and the solution P_4 with $\alpha = 3$ (as shown in Figure 10) as examples, the network scheme corresponding to the solution P_1 requires the least number of buses and the lowest user cost, but the direct rate which is 90.95% has the least advantage. There is little difference for the direct rate among the solutions P_2 , P_3 , and P_4 , but the solution P_2 requires 3 more buses than the solution P_4 , and 0.1×10^5 min more cost than the solution P_3 .

Therefore, the scenarios with different line numbers are calculated when $\alpha = 3$. Figures 11–14 present the comparison between the optimal Pareto solutions obtained by the algorithm and the results in previous literatures. As can be

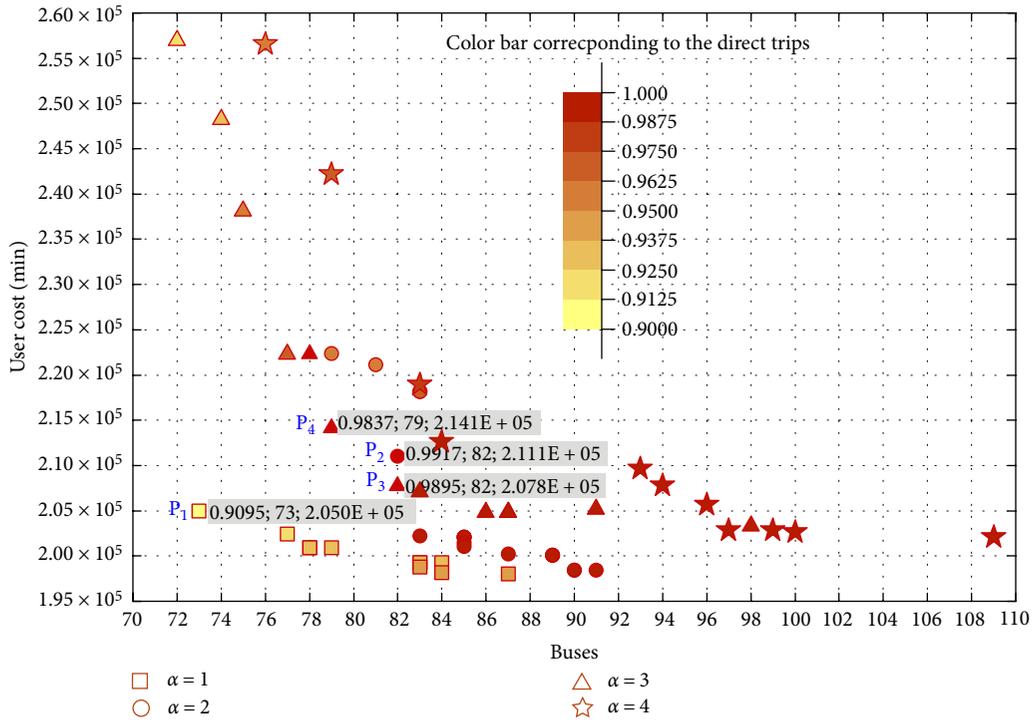


FIGURE 10: Pareto solutions for the scenarios with 4 lines under different values of α .

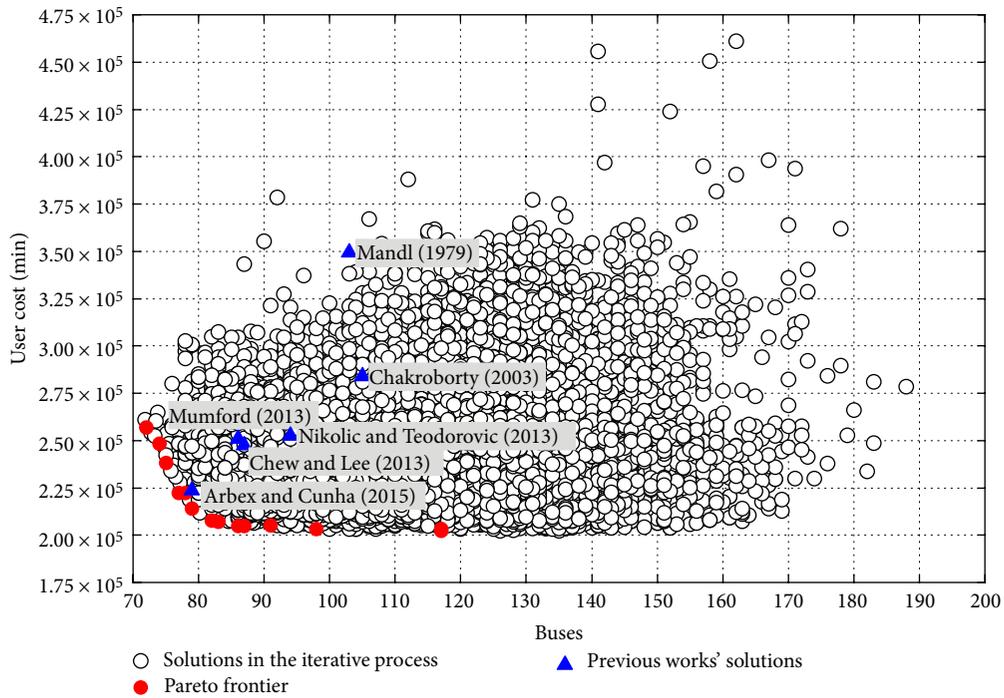


FIGURE 11: Pareto frontier for the scenario with 4 lines when $\alpha = 3$.

seen, the solutions obtained by the algorithm can dominate the results calculated in the previous literatures, which indicates the better solutions can be obtained by the algorithm. The objectives obtained in [40] by Arbex and Cunha is the closest to the solution obtained by the algorithm, but with the increase of transfer sensitivity, the advantage of the

algorithm in the user cost and direct rate is gradually emerging.

The solutions which can dominate the solutions in the previous literatures are selected, as shown in Table 5. The following indicators are usually used in many literatures to test the quality of results:

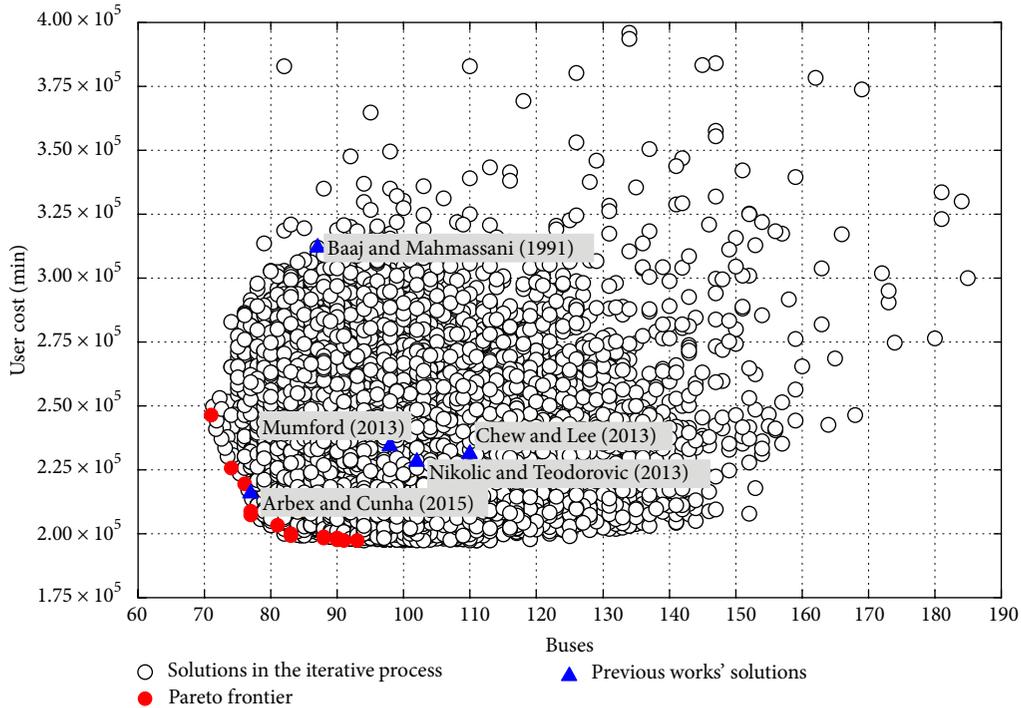


FIGURE 12: Pareto frontier for the scenario with 6 lines when $\alpha = 3$.

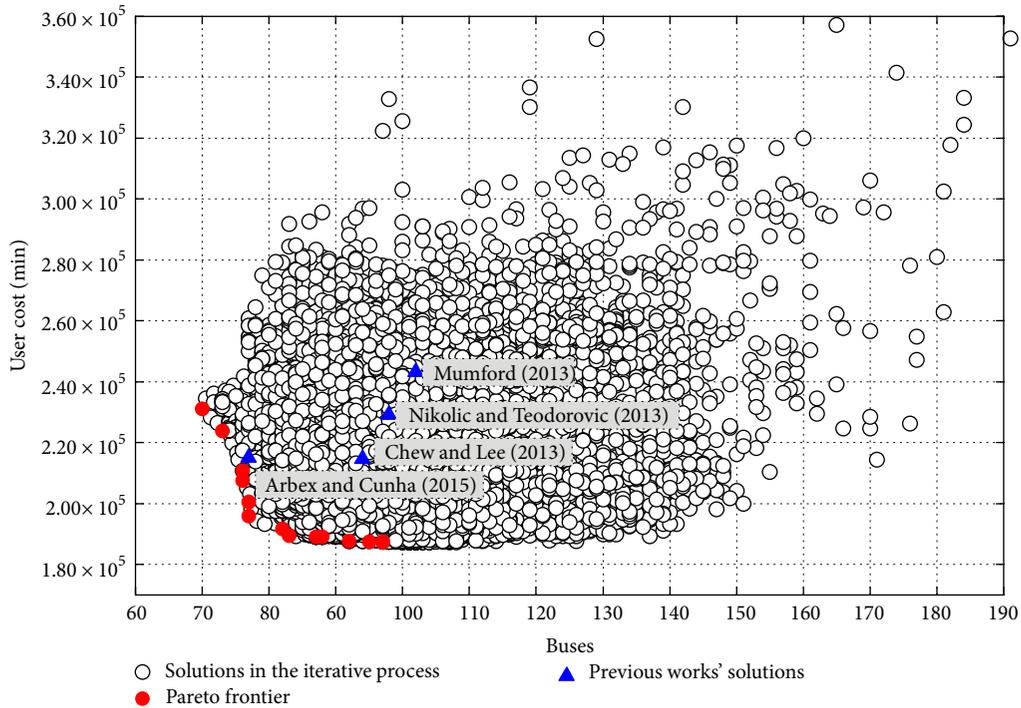


FIGURE 13: Pareto frontier for the scenario with 7 lines when $\alpha = 3$.

d_0 : Percentage of demand satisfied without any transfer.
 d_1 : Percentage of demand satisfied with one transfer.
 d_2 : Percentage of demand satisfied with two transfers.
 d_{un} : Percentage of demand unsatisfied.
 AIVTT: Average in-vehicle travel time in minutes per transit user, min (mpu).

AUC: Average user cost in minutes per transit user min (mpu), comprising travel time and transfer penalties.

As can be observed from Table 5, the selected nondominated solutions obtained by the algorithm requires fewer buses, lower user cost, and lower transfer rate. Therefore, the algorithm has high convergence in solving the MTNDP.

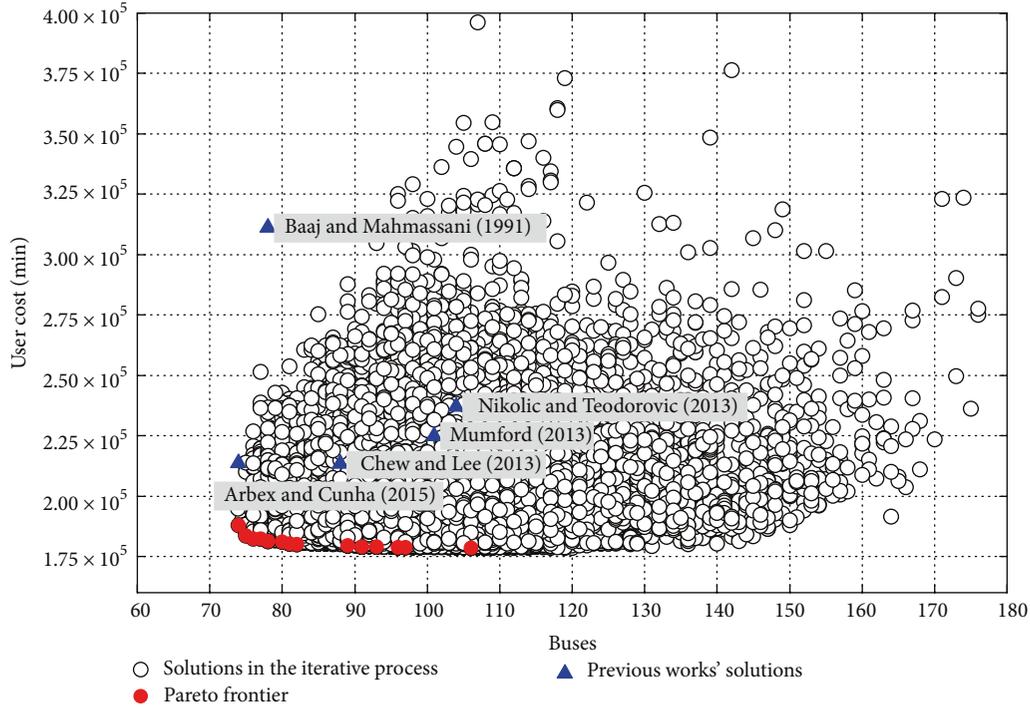


FIGURE 14: Pareto frontier for the scenario with 8 lines when $\alpha = 3$.

TABLE 5: Comparison between selected nondominated solutions and the results in previous literatures.

Line number	Sources in literatures [3, 21, 22, 24, 32, 35, 40, 41]	Buses	User cost (min)	d_0 (%)	d_1 (%)	d_2 (%)	d_{un} (%)	Maximum headway (min)	Average headway (min)	AIVTT (min)	AUC (min)	
4	Mandl (1980)	103	349230.26	69.94	29.93	0.13	0.00	6.67	3.48	11.40	22.43	
	Chakroborty (2003)	105	284009.56	89.98	10.02	0.00	0.00	9.00	4.06	13.10	18.24	
	Mumford (2013)	86	251015.35	91.14	8.86	0.00	0.00	9.33	4.43	10.79	16.12	
	Chew and Lee (2013)	87	247498.31	92.74	7.26	0.00	0.00	5.11	3.64	11.47	15.90	
	Nikolic and Teodorovic (2013)	94	252596.87	91.91	8.09	0.00	—	4.32	3.41	11.71	16.22	
	Arbex and Cunha (2015)	79	223506.75	98.27	1.73	0.00	0.00	8.60	4.60	11.13	14.35	
	Ahmed et al. (2019)	—	—	—	91.84	8.15	0.00	0.00	—	—	—	—
<i>Solution with the algorithm (Scenario with $\alpha = 3$)</i>	79	214118.67	98.37	1.63	0.00	0.00	9.24	4.98	12.97	13.75		
6	Baaj and Mahmassani (1991)	87	311983.89	78.61	21.39	0.00	0.00	11.33	4.11	11.14	20.04	
	Mumford (2013)	98	234358.26	96.08	3.92	0.00	0.00	8.00	5.06	11.77	15.05	
	Chew and Lee (2013)	110	231258.66	98.14	1.86	0.00	0.00	8.80	4.86	12.18	14.85	
	Nikolic and Teodorovic (2013)	102	228122.02	97.24	2.76	0.00	0.00	10.29	5.25	11.64	14.65	
	Arbex and Cunha (2015)	77	215781.79	98.20	1.80	0.00	0.00	9.56	6.42	11.55	13.86	
	Ahmed et al. (2019)	—	—	—	97.17	2.82	0.00	0.00	—	—	—	—
	<i>Solution 1 with the algorithm (Scenario with $\alpha = 3$)</i>	77	207437.46	98.72	1.28	0.00	0.00	14.46	6.87	12.71	13.32	
<i>Solution 2 with the algorithm (Scenario with $\alpha = 3$)</i>	77	208802.92	98.55	1.45	0.00	0.00	12.94	6.65	12.72	13.41		

TABLE 5: Continued.

Line number	Sources in literatures [3, 21, 22, 24, 32, 35, 40, 41]	Buses	User cost (min)	d_0 (%)	d_1 (%)	d_2 (%)	d_{um} (%)	Maximum headway (min)	Average headway (min)	AIVTT (min)	AUC (min)
7	Mumford (2013)	102	243087.48	98.01	1.99	0.00	—	6.80	5.32	12.91	15.61
	Chew and Lee (2013)	94	214543.20	99.10	0.90	0.00	—	10.00	7.89	11.04	13.78
	Nikolic and Teodorovic (2013)	98	229039.11	98.84	1.16	0.00	—	17.50	7.00	11.99	14.71
	Arbex and Cunha (2015)	77	214989.45	98.52	1.48	0.00	—	12.80	7.58	11.91	13.81
	Ahmed et al. (2019)	—	—	98.84	1.15	0.00	0.00	—	—	—	—
	<i>Solution 1 with the algorithm (Scenario with $\alpha = 3$)</i>	77	195888.10	99.55	0.45	0.00	0.00	18.87	9.30	12.37	12.58
	<i>Solution 2 with the algorithm (Scenario with $\alpha = 3$)</i>	77	200534.10	99.14	0.86	0.00	0.00	15.53	9.02	12.46	12.88
8	Baaj and Mahmassani (1991)	78	311059.22	79.96	20.04	0.00	0.00	10.67	4.65	11.00	19.98
	Mumford (2013)	101	224805.54	99.10	0.90	0.00	0.00	15.00	6.91	11.95	14.44
	Chew and Lee (2013)	88	213344.37	99.04	0.96	0.00	—	31.00	9.67	10.72	13.70
	Nikolic and Teodorovic (2013)	104	236835.24	98.97	1.03	0.00	—	29.00	9.66	12.65	15.21
	Arbex and Cunha (2015)	74	213682.25	98.65	1.35	0.00	—	16.00	7.48	11.24	13.72
	Ahmed et al. (2019)	—	—	99.16	0.83	0.00	0.00	—	—	—	—
	<i>Solution with the algorithm (Scenario with $\alpha = 3$)</i>	74	187988.18	99.81	0.19	0.00	0.00	34.67	14.24	11.98	12.07

TABLE 6: Parameter values of the model used in the case.

Parameters	Values
N_L	10
$N_{\min}^l, l \in L$	5
$N_{\max}^l, l \in L$	16
$V^l, l \in L$	40
$\mu^l, l \in L$	1.35
$t_{mi}^r, i \in N, r \in R$	30 s
$t_{zk}^r, k \in N, r \in R$	2 min
H	1.5
α	3
β	2
γ	0.15
λ	4
θ	1.866
t_{hmin}	3 min

5. Case Application with Real-Size Network

The model and the algorithm are used for a real-size bus network calculation in Baotou city, Inner Mongolia. The predicted population density in long term and the trunk roads in the city centre area are shown in Figure 15. In the centre area, 10 bus lines will be planned on the trunk roads and the locations of 50 stops are determined in the trunk roads. In this case, the mode choice among bus, private car, bicycle and walk

is included. It should be mentioned that as the travel distance increases, the probability of walking decreases the fastest until it drops to zero. The parameter values used in the case are shown in Table 6.

Figure 16 presents the Pareto frontier for the case obtained by the algorithm, and the percentages of direct and transfer passengers in each solution are shown in Figure 17. We can see that with the increase of buses on the network, the proportion of transfers decreases in general, and most of passengers can reach their destinations within one transfer. The two single objectives can achieve optimal in the boundary solutions respectively, as shown in Figure 16, and the bus networks for the two boundary solutions are shown in Figures 18 and 19, respectively. The boundary solution with fewer buses has more stations and smaller average cross-sectional passenger volume, and the maximum load rate of the lines for both solutions are less than 1, as shown in Tables 7 and 8. All the optimal Pareto solutions can be used as candidates for the city's bus network schemes. As there is no "one best solution" in the Pareto solutions and the determination of the network scheme is also influenced by the policies such as urban politics, economy and environment, the choice of the best compromise scheme is not within the scope of the paper.

6. Conclusions

Considering both the travel efficiency of passengers and the benefit of managers, the multi-objective network design model

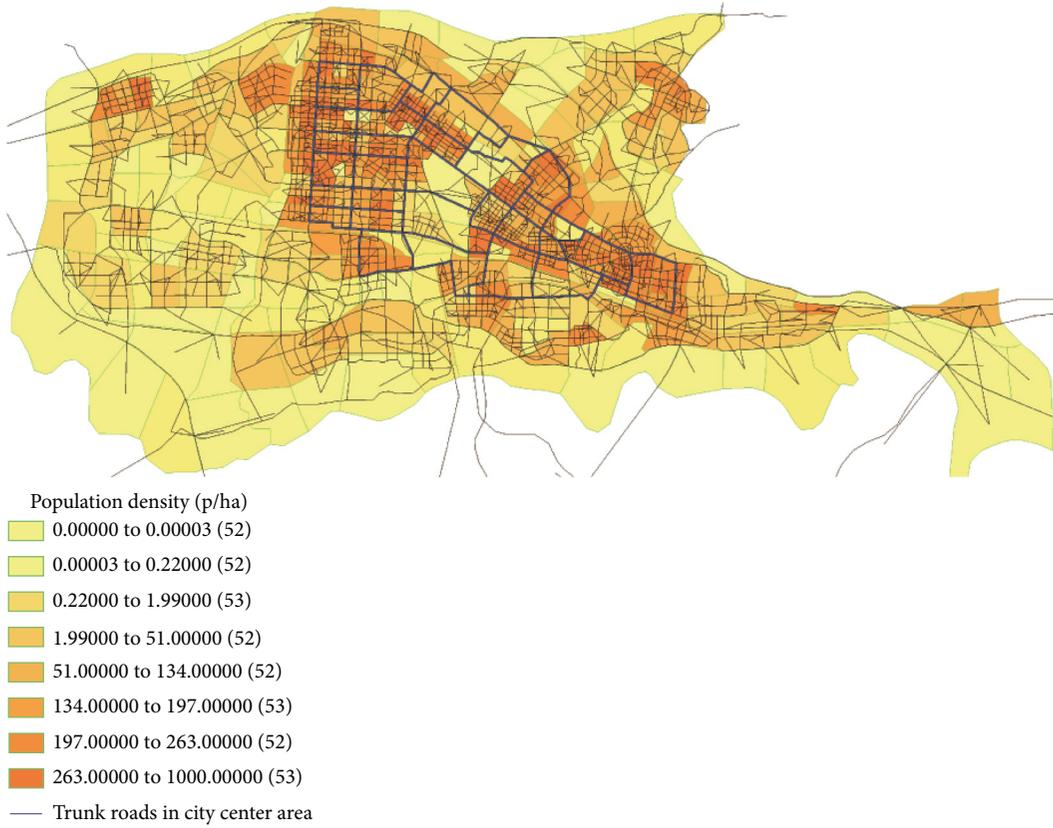


FIGURE 15: Population density and the trunk roads network in Baotou city center area.

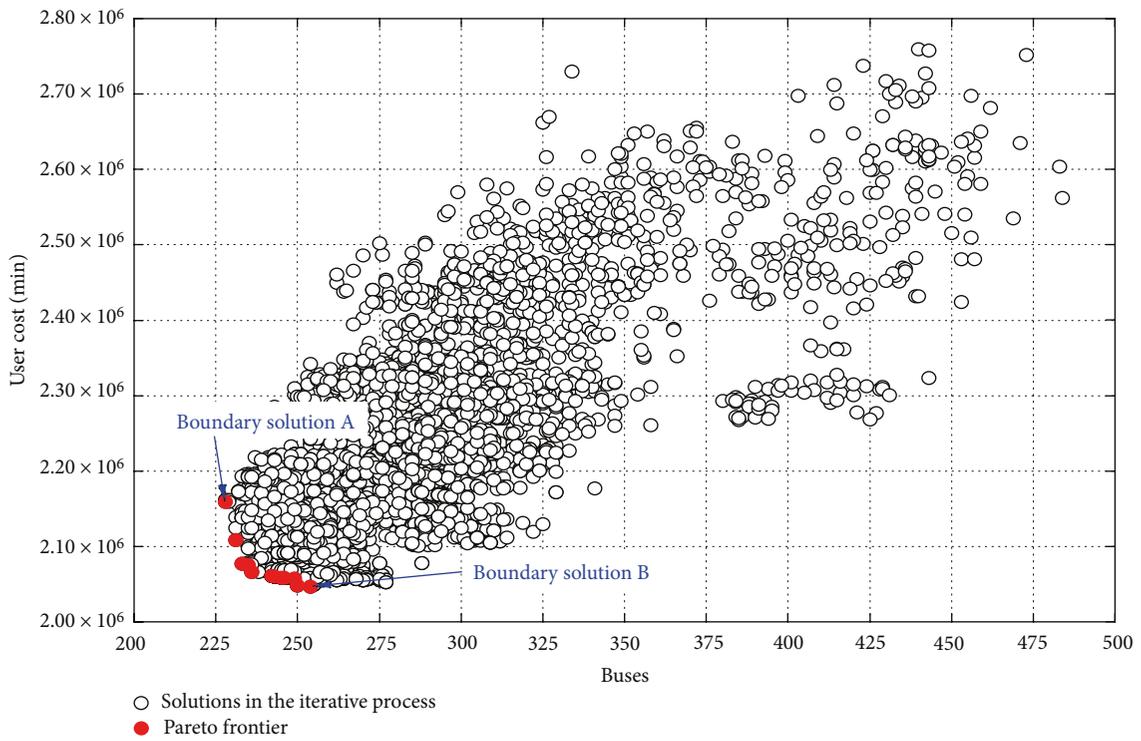


FIGURE 16: Pareto frontier for the case obtained by the algorithm.

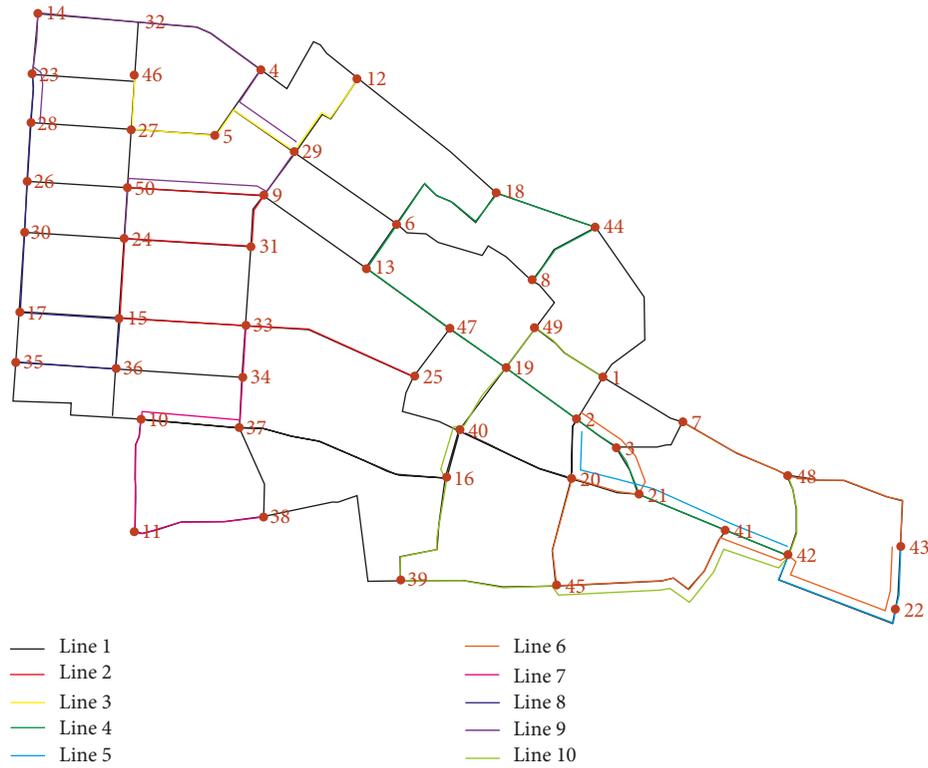


FIGURE 19: The bus network for boundary solution B.

TABLE 7: Detailed results of the lines for the boundary solution A.

Line label	Node sequence	Fleet	Headway (min)	Maximum load factor	Minimum load factor
1	2-20-40-16-37-10	29	3	0.83	0.21
2	25-33-15-24-31-9-50	38	3	0.93	0.06
3	46-27-5-29-12	8	7	1.00	0.21
4	48-42-41-21-3-2-19-47-13-6-18-44-8	27	6	0.90	0.15
5	7-48-43-22-42-41-21-3-2-19-47-13-6-18-44-8	46	5	0.95	0.03
6	42-41-45-20-21-3-2-19-47-13-6-18-44-8	18	11	1.00	0.02
7	33-34-37-10-11-38	18	4	0.95	0.53
8	35-36-15-17-30-26-28-23	20	4	0.89	0.14
9	28-23-14-32-4-29-9-31-24	23	5	0.93	0.09
10	42-41-45-39-16-40-19-49-1	27	5	0.99	0.09

TABLE 8: Detailed results of the lines for the boundary solution B.

Line label	Node sequence	Fleet	Headway (min)	Maximum load factor	Minimum load factor
1	2-20-40-16-37-10	29	3	0.95	0.22
2	25-33-15-24-31-9-50	38	3	0.93	0.01
3	46-27-5-29-12	8	7	0.88	0.21
4	42-41-21-3-2-19-47-13-6-18-44-8	49	3	0.88	0.10
5	43-22-42-41-21-20-2	14	6	0.96	0.02
6	7-48-43-22-42-41-45-20-21-3-2	8	22	0.98	0.22
7	33-34-37-10-11-38	18	4	0.98	0.54
8	35-36-15-17-30-26-28-23	20	4	0.92	0.14
9	28-23-14-32-4-29-9-50-24	19	6	0.92	0.01
10	48-42-41-45-39-16-40-19-49-1	25	6	0.98	0.34

is developed with strict constraints. For passenger travel behaviour, the transfer times and passenger congestion can affect the selection of travel paths in the model. As the MTNDP is highly complex and difficult in solving, an effective algorithm is proposed based on improved NSGA-II. The high global convergence of the algorithm is tested by comparing with previous works while its applicability for the MTNDP is verified by the calculation of a real-size case.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

- [1] A. Ceder and N. H. M. Wilson, "Bus network design," *Transportation Research Part B: Methodological*, vol. 20, no. 4, pp. 331–344, 1986.
- [2] V. Guihaire and J.-K. Hao, "Transit network design and scheduling: a global review," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 10, pp. 1251–1273, 2008.
- [3] M. H. Baaj and H. S. Mahmassani, "An AI-based approach for transit route system planning and design," *Journal of Advanced Transportation*, vol. 25, no. 2, pp. 187–210, 1991.
- [4] A. Mauttone and M. E. Urquhart, "A route set construction algorithm for the transit network design problem," *Computers & Operations Research*, vol. 36, no. 8, pp. 2440–2449, 2009.
- [5] Á. Marín and R. García-Ródenas, "Location of infrastructure in urban railway networks," *Computers & Operations Research*, vol. 36, no. 5, pp. 1461–1477, 2009.
- [6] G. Gutiérrez-Jarpa, C. Obreque, G. Laporte, and V. Marianov, "Rapid transit network design for optimal cost and origin-destination demand capture," *Computers & Operations Research*, vol. 40, no. 12, pp. 3000–3009, 2013.
- [7] F. López-Ramos, E. Codina, Á. Marín, and A. Guarnaschelli, "Integrated approach to network design and frequency setting problem in railway rapid transit systems," *Computers & Operations Research*, vol. 80, pp. 128–146, 2017.
- [8] X. Feng, X. Zhu, X. Qian, Y. Jie, F. Ma, and X. Niu, "A new transit network design study in consideration of transfer time composition," *Transportation Research Part D-Transport and Environment*, vol. 66, pp. 85–94, 2019.
- [9] G. Gutiérrez-Jarpa, G. Laporte, V. Marianov, and L. Moccia, "Multi-objective rapid transit network design with modal competition: the case of concepción, chile," *Computers & Operations Research*, vol. 78, pp. 27–43, 2017.
- [10] G. Laporte, Á. Marín, J. A. Mesa, and F. A. Ortega, "An integrated methodology for the rapid transit network design problem," in *Algorithmic Methods for Railway Optimization*, pp. 187–199, Springer, Berlin, Heidelberg, 2007.
- [11] D. Huang, Z. Liu, X. Fu, and P. T. Blythe, "Multimodal transit network design in a hub-and-spoke network framework," *Transportmetrica A: Transport Science*, vol. 14, no. 8, pp. 706–735, 2018.
- [12] D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa, "A general rapid network design, line planning and fleet investment integrated model," *Annals of Operations Research*, vol. 246, no. 1–2, pp. 127–144, 2014.
- [13] D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa, "An adaptive neighborhood search metaheuristic for the integrated railway rapid transit network design and line planning problem," *Computers & Operations Research*, vol. 78, pp. 1–14, 2017.
- [14] P. Shang, R. Li, J. Guo, K. Xian, and X. Zhou, "Integrating lagrangian and eulerian observations for passenger flow state estimation in an urban rail transit network: a space-time-state hyper network-based assignment approach," *Transportation Research Part B: Methodological*, vol. 121, pp. 135–167, 2019.
- [15] J. Liu and X. Zhou, "Capacitated transit service network design with boundedly rational agents," *Transportation Research Part B Methodological*, vol. 93, pp. 225–250, 2016.
- [16] A. Schiewe, "Integrated algorithms for cost-optimal public transport planning," Georg-August University, 2019.
- [17] G. E. Cantarella, G. Pavone, and A. Vitetta, "Heuristics for urban road network design: lane layout and signal settings," *European Journal of Operational Research*, vol. 175, no. 3, pp. 1682–1695, 2006.
- [18] B. García-Archilla, A. J. Lozano, J. A. Mesa, and F. Perea, "GRASP algorithms for the robust railway network design problem," *Journal of Heuristics*, vol. 19, no. 2, pp. 399–422, 2011.
- [19] C. Liebchen, *On the Benefit of Preprocessing and Heuristics for Periodic Timetabling*, Springer International Publishing, Cham, pp. 709–714, 2018.
- [20] P. Gattermann, J. Harbering, and A. Schöbel, "Line pool generation," *Public Transport*, vol. 9, no. 1–2, pp. 7–32, 2017.
- [21] P. Chakroborty, "Genetic algorithms for optimal urban transit network design," *Computer-Aided Civil and Infrastructure Engineering*, vol. 18, no. 3, pp. 184–200, 2003.
- [22] J. S. C. Chew, L. S. Lee, and H. V. Seow, "Genetic algorithm for biobjective urban transit routing problem," *Journal of Applied Mathematics*, vol. 2013, Article ID 698645, pp. 1–15, 2013.
- [23] W. Fan and R. B. Machemehl, "Tabu search strategies for the public transportation network optimizations with variable transit demand," *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, no. 7, pp. 502–520, 2008.
- [24] M. Nikolić and D. Teodorović, "Transit network design by bee colony optimization," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5945–5955, 2013.
- [25] A. T. Buba and L. S. Lee, "A differential evolution for simultaneous transit network design and frequency setting problem," *Expert Systems with Applications*, vol. 106, pp. 277–289, 2018.
- [26] S. B. Jha, J. K. Jha, and M. K. Tiwari, "A multi-objective metaheuristic approach for transit network design and frequency setting problem in a bus transit system," *Computers & Industrial Engineering*, vol. 130, pp. 166–186, 2019.
- [27] H. Mohring, "Optimization and scale economies in urban bus transportation," *American Economic Review*, vol. 62, no. 4, pp. 591–604, 1972.
- [28] Q. Esteve, S. Lif, M. Cossic, F. Leurent, and X. Xie, "Understanding the transfer function of mass transit hub by automated fare collection data," in *IEEE International Conference on Models & Technologies for Intelligent Transportation Systems*, IEEE, Naples, Italy, 2017.
- [29] M. A. Ramli, V. Jayaraman, K. H. Chee, K. H. Tan, G. L. K. Khoon, and C. Monterola, "Improved estimation of commuter waiting times using headway and commuter boarding information," *Physica A: Statistical Mechanics & Its Applications*, vol. 501, pp. 217–226, 2018.

- [30] A. Fielbaum, S. Jara-Diaz, and A. Gschwender, "Optimal public transport networks for a general urban structure," *Transportation Research Part B: Methodological*, vol. 94, pp. 298–313, 2016.
- [31] K. An and K. L. Hong, "Two-phase stochastic program for transit network design under demand uncertainty," *Transportation Research Part B*, vol. 84, pp. 157–181, 2016.
- [32] C. E. Mandl, "Evaluation and optimization of urban public transportation networks," *European Journal of Operational Research*, vol. 5, no. 6, pp. 396–404, 1980.
- [33] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [34] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary Computation*, vol. 2, no. 3, pp. 221–248, 1994.
- [35] L. Ahmed, C. Mumford, and A. Kheiri, "Solving urban transit route design problem using selection hyper-heuristics," *European Journal of Operational Research*, vol. 274, no. 2, pp. 545–559, 2019.
- [36] W. Y. Szeto and Y. Jiang, "Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach," *Transportation Research Part B*, vol. 67, no. 9, pp. 235–263, 2014.
- [37] M. A. Nayeem, M. K. Rahman, and M. S. Rahman, "Transit network design by genetic algorithm with elitism," *Transportation Research Part C*, vol. 46, pp. 30–45, 2014.
- [38] S. M. H. Mahdavi Moghaddam, K. R. Rao, G. Tiwari, and P. Biyani, "Simultaneous bus transit route network and frequency setting search algorithm," *Transportation Engineering, Part A: Systems*, vol. 145, no. 4, p. 04019011, 2019.
- [39] S. Zhou and R. J. Mondragón, "The rich-club phenomenon in the internet topology," *IEEE Communications Letters*, vol. 8, no. 3, pp. 180–182, 2003.
- [40] R. O. Arbex and C. B. D. Cunha, "Efficient transit network design and frequencies setting multi-objective optimization by alternating objective genetic algorithm," *Transportation Research Part B*, vol. 81, pp. 355–376, 2015.
- [41] C. L. Mumford, "New heuristic and evolutionary operators for the multi-objective urban transit routing problem," in *2013 IEEE Congress on Evolutionary Computation*, pp. 939–946, Cancun, Mexico, June 2013.



Hindawi

Submit your manuscripts at
www.hindawi.com

