IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received December 24, 2019, accepted January 8, 2020, date of publication January 13, 2020, date of current version January 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966008

A Multi-Objective Optimization and Hybrid Heuristic Approach for Urban Bus Route Network Design

CHAO WANG, ZHIRUI YE, AND WEI WANG

Jiangsu Key Laboratory of Urban ITS, School of Transportation, Southeast University, Nanjing 211189, China
Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing 211189, China

Corresponding author: Chao Wang (wangchao1990@seu.edu.cn)

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB1601000.

ABSTRACT This study proposed a hybrid optimization model for urban bus transit route network design problem (TRNDP). Although several mathematical methods had been developed to make the problem tractable, the methods relied on excessive assumptions, which resulted in over-simplification or idealization of the problems. In light of these considerations, a multi-level and multi-mode network design method was introduced in this study. A multi-level network consisted of three levels: skeleton network, arterial network, and feeder network. The different levels of network were associated with different modes of public transportation (such as subway, light rail transit, trolleybus, BRT, normal bus, and community branch) based on the features of bus routes, city sizes, etc. In addition, according to the respective features of urban transit route network structures, we developed different optimization models for different levels. Finally, the proposed methodology was applied to case studies of the city of Zhaoyuan in China and the transit network of Mandl benchmark. The results showed that the total travel time for the proposed method was significantly lower than that of the competing method, with a 21.51% reduction. In addition, the proposed method provided 85.23% direct travelers, 14.65% travelers with one transfer, 0.12% travelers with two transfers, and no unsatisfied demand, which were better than the results from the compared method.

INDEX TERMS Network design, public transportation, multi-objective optimization, hybrid heuristic methods.

NOMENCLATURE

t	actual time cost;
t_0	free flow time cost;
α, β	BPR parameters;
$\beta_{i,i+1}$	safety score, $\beta_{i,i+1} \in [0, 1]$;
$t_{i,i+1}$	time cost between node i and $i+1$;
N	number of lanes;
$q_{i,i+1}$	traffic volume between node i and $i+1$;
$c_{i,i+1}$	design traffic capacity between node i and $i+1$;
$a_{i,i+1}, b_{i,i+1}$	safety influence coefficient;
$y_{i,i+1}$	0-1 variable (whether the nodes i, $i+1$ are in the same route);
$x^k_{i,i+1}$	0-1 variable (whether the nodes i, $i+1$ are in the route k);
l	length of line;
l_{max}	upper bound of the line length;
ψ_1	factor of intersection density;
ψ_2	factor of speed limit;
ψ_3	factor of bus stop density;
ψ_4	factor of saturation;
$\mu_{0,n}$	coefficients of the above influencing factors;
$v_{i,i+1}$	free flow speed between node i and $i+1$;
$l_{i,i+1}$	line length between node i and $i+1$;
Q_{ij}	passenger volume between node i and j;
v_{ij}	free flow speed;
A_{ij}, B_{ij}	acceleration and deceleration;
l_{ij}	line length;
q_{ij}	traffic volume;
c_{ij}	design traffic capacity;
P_{up}	number of boarding passengers;
P_{down}	number of alighting passengers;
N_{up}	number of doors for boarding;

The associate editor coordinating the review of this manuscript and approving it for publication was Razi Iqbal.

N_{down} number of doors for alighting;
t_{oc} additional delay at bus stop;
t_{up} service time for each boarding passengers;
t_{down} service time for each alighting passengers;
f_k departure frequency of the kth line;
λ_a conversion coefficient of transfer time;
$Q_{i,j-1}$ passenger volume between node i and j-1;
λ_b conversion coefficient of operation expense;
λ_c conversion coefficient of vehicle emission;
d_a intersection delay;
h_{ij} vehicle emission factor in the road segments;
h_a vehicle emission factor in the intersections;
T_a signal cycle length;
g_a green ratio;
Q_k maximum flow in a cross section;
x_{ij}^k 0-1 variable (whether the nodes i, j are in the route k);
$(\varphi_k)_{max}$ upper bound of full-load ratio;
C_k rated load;
β_{min} lower limit of the safety score;
E_d set of all links entering the intersection;
S upper bound of operation vehicles;
m number of bus lines;
M_{max} maximum number of lines;
β_{ij} safety score between node i and j;
SI_{ij} severity index between nodes i and j;
$(Z_1)_{ij}$ total number of fatal crashes between node i and j;
$(Z_2)_{ij}$ total number of serious injury crashes;
$(Z_3)_{ij}$ total number of other injury crashes;
$(Z_4)_{ij}$ total number of property-damage-only crashes;
SI maximum value of all severity indexes between nodes i and j;
T_1 in-vehicle travel time;
T_2 dwell time at bus stop;
T_3 passengers' waiting time;
T_4 passengers' transfer time;
T_5 bus operating cost;
T_6 vehicle emission cost;
$\omega_1, \omega_2, \ldots, \omega_6$ weight coefficients for the above time (cost);
d passengers' walking distance;
$x_{i,i+1}$ 0-1 variable, and is used to judge if there is overlap of feeder network ($x_{i,i+1} = 0$ represents that there is no overlap of feeder network, and vice versa);
R_s radius of service;
ρ remaining coefficient of pheromone, $\rho \in (0, 1)$;
τ_{ij} pheromone concentration between node i and j;
τ'_{ij} initial pheromone concentration;
D released intensity of pheromone concentration;
p_{ij}^u probability for an ant u crawling from node i to j;
ϑ' relative importance of pheromone concentration;
ϕ' relative importance of visibility;
S_u set of nodes accessible to the ant u;
$\bar{P}_{mj}$ average maximum number of passengers in period j;
d_j capacity of a vehicle (number of seats plus the maximum allowance standees);
ξ_j load factor during period j, $0 < \xi_j \leq 1$;
C_k rated load ($= \xi_j \times d_j$).

I. INTRODUCTION

A public transportation network is one of the basic components of transit system planning. Increasingly, research scholars focus on the study of public transportation networks, which can be referred to as the transit route network design problem (TRNDP) [1]–[3]. As noted by Ceder and Lam [4], the systematic decision sequence for the bus planning process is comprised of five efficient steps: 1) designing the network of routes; 2) setting frequencies; 3) developing timetables; 4) scheduling vehicles; and 5) scheduling crews. This study focused on step 1, i.e. transit route network design (TRND) [5]. It is the basis of the other four steps.

In early research on TRNDP, several categories of methodologies have been identified: manual approach, systems analysis approach, market analysis project (MAP) approach, and systems analysis with interactive graphic approach [6], [7]. Recent research has focused on the development and application of mathematical methodologies. Compared to the above methods, the mathematical optimization approach is capable of producing an optimum system and more reliable results [8]. For instance, Zhao and Zeng presented a meta-heuristic method for optimizing transit networks, including route network design, vehicle headway, and timetable assignment. The results showed that the methodology was capable of producing improved solutions to real-life transit network design problems in reasonable amounts of time and computing resources [9]. In a large city, operating a bus transit service on a route network based on destination-oriented or point-to-point approach, which considers all possible routes with the node set, is cumbersome and impractical. Alternatively, a Hub and Spoke network, which combine the destination-oriented and direction-oriented approaches, could be a more efficient choice [10].

In general terms, the mathematical optimization approach involves three aspects: 1) maximization or minimization objective functions; 2) feasibility constraints; and 3) solution methodologies. Maximization or minimization objective functions of TRNDP involves the following considerations: direct travelers (zero-transfer) [11]; number of transfers [4], [12]; unsatisfied demand [8]; operator cost [12]–[14]; passenger in-vehicle and out-of-vehicle time [15], [16]; transfer time [17], [18]; and fleet size [19]. The multi-objective approach is mostly used as it can generate solutions reflecting tradeoffs among conflicting objectives. Feasibility constraints usually include: limits on route lengths [14], [18]; allowable fleet size [12], [14]; operating frequencies [8], [14], [19]; unsatisfied demand [13]; load factor [13]; number of routes [14]; and percentage of demand [17], [18]. As for the solution methodology, existing research demonstrates a variety of methods to formulate and solve the TRNDP. In general, two types of methods have been used for the problem: 1) conventional methods based on analysis [20] and mathematical programming [21]; and 2) heuristic methods based on traditional heuristics and metaheuristics [22]. As reported [23], the TRNDP is a complex problem and conventional methods cannot be efficiently applied. Therefore, researchers tended to utilize the heuristic methods, such as genetic algorithm [24], ant colony optimization [25], simulated annealing [26], [27], and tabu search [28]. These methods outperformed the conventional methods with respect to computation speed and optimization results [29].

Although numerous mathematical methods have been developed to make the problem tractable, the methods relied on excessive assumptions, which resulted in oversimplification or idealization of the problems. For instance, the transit network design was generally considered from the perspective of a single level or mode while neglecting the level and modal interactions, which was not capable of handling real world design of transit networks [1], [3]. In real-world applications, two-level networks (trunk + feeder) were commonly used for TRND [30]. However, urban areas have rapidly increased in size, and various modes of public transportation are now available on different types of urban roads with different levels of service (LOSs). Currently, most big cities in China have several types of public transportation in service, and they also have different LOSs in terms of speed and capacity. Hence, the traditional two-level TRND does not meet the needs of current situations. In light of this, a multi-level and multi-mode optimization model is proposed, which divides urban transit systems into three levels (skeleton network, arterial network, and feeder network). When properly designed, a multi-level (especially three-level) transit network can have better performance than a two-level network for saving travel time and maximizing direct and transfer demand [5], such as in Rome [15], Winnipeg [12], and Dalian [31]. Compared with those studies, the objectives and contributions of this study focus on the following aspects: 1) In previous studies, transit network design was only considered from the perspective of multi-level and did not consider modes. In this study, a different level of network is associated with different modes of public transportation according to the features of bus routes, city sizes, etc. (Table 1). For these types of cities, the multi-level network includes three levels: skeleton network, arterial network, and feeder network. In this study, a skeleton network consists of major transit corridors, which connect the relatively prosperous districts in the city and meet the demand for direct travelers. Such routes are usually covered by transportation modes with high capacity and speed such as subway, light rail transit, and trolleybus [31]. As the main system to cover the large areas unserved by the skeleton network, an arterial transit route network provides transit service between skeleton and feeder networks. Skeleton and arterial networks together become the backbone of urban public transportation. A feeder network provides transfer service for skeleton and arterial networks, penetrates into dwelling districts, and fills in the gaps of the arterial network to increase the density of the overall network. Feeder routes are often provided by normal bus and community branch transportation modes. They expand the network accessibility and shorten the walking distance for passengers [31]. 2) In previous studies, the design of a multi-level network was generally developed by a single model, while neglecting the differences in skeleton, arterial, and feeder networks, such as transportation demands, route network functions, road conditions, and line length. According to the respective features of urban transit route network structures, we developed different models (using different solutions) for different levels.

TABLE 1. **Multi-level and multi-mode transit route network varying by city size.**

Network Level	Transit Mode		
	Megalopolis	Metropolis	Medium-sized city
Skeleton Network	Subway/ Light Rail Transit (LRT)	LRT/ Trolleybus	Trolleybus/ BRT
Arterial Network	Trolleybus/ BRT	Trolleybus/ BRT	BRT/ Normal Bus
Feeder Network	Normal Bus/ Community Branch	Normal Bus/ Community Branch	Normal Bus/ Community Branch

Based on the three-level network, transit network design process consists of three sequential steps, as shown in Figure 1. A set of transit route networks, categorized by hierarchy, are optimally developed by using different objective functions, feasibility constraints, and solution methodologies.

Recently, several research studies have focused on the time-dependent accessibility for transportation network design. Tong et al. developed a space-time prism analysis framework to address a new urban network design problem to maximize the system-wide transportation accessibility between major activity locations, subject to a given highway construction budget [32]. Their contributions focused on incorporating space-time accessibility into network design models, while neglecting the congestion effect caused by road

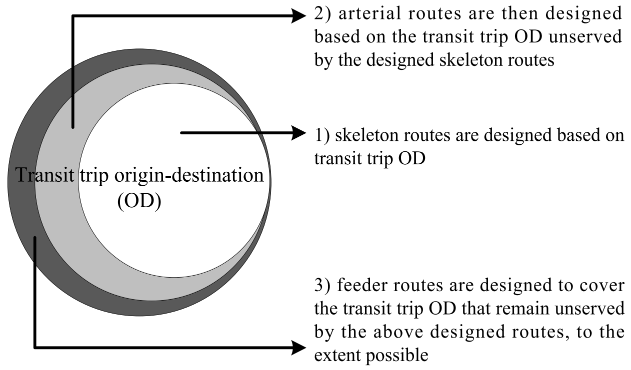


FIGURE 1. Transit network design process using transit trip OD.

capacity constraints. Di et al. studied a new discrete network design problem for metropolitan areas, in which some concepts, such as the accessible flow and travel time budget function, were proposed. It is worth mentioning that their study was actually the first time to investigate the network design problem in the viewpoint of improving the accessibility of the travel flow [33]. Multi-mode and multi-level network design problems should be an interesting topic in their future studies. Chu solved the simultaneous planning problem of network design and timetabling for urban bus systems. An innovative mixed-integer programming (MIP) model was formulated and a parallel branch-and-price-and-cut (BPC) algorithm was proposed to solve the problem [34]. In its computational study, however, a small size of network (only 26 nodes and 84 links) was introduced to test the performance of the proposed method. For other urban transit systems (for instance, urban rail transit network), the study by Yang et al. investigated a collaborative optimization for the last train timetable. By using a space-time network framework, all the involved transportation activities (such as train space-time travel arcs, passenger travel arcs, and transfer arcs) were well characterized in an extended space-time network [35]. In our future research, it would be interesting to study how accessibility-based planning methods could further influence the results of transit route network design.

II. MODEL FORMULATION

With the expansion of urban areas, various modes of public transportation are available on different network levels with different levels of service (LOSs) in terms of speed and volume. If public transit routes on different levels are optimized and designed using the same standards, the resulting transit system may have a low level of efficiency. Therefore, it's critical to adjust urban transit route network structures according to their respective features.

A. SKELETON NETWORK

An urban skeleton network consists of major transit corridors with large passenger demand, meets the demand for direct travelers. In the design of skeleton transit routes, travel time and the number of direct passengers should be adopted as principal elements. Hence, it can be treated as a minimum cost and maximum flow (MCMF) problem. That is, on the premise of the maximization of passenger volume, travel time is minimized. When determining the time cost [36], [37], a common method is Bureau of Public Road (BPR) function [38], which can be expressed as:

$$t = t_0 \left(1 + \alpha \left(\frac{Q}{C}\right)^{\beta}\right) \tag{1}$$

where t is actual time cost; t_0 denotes free flow time cost; Q represents volume; C is capacity; α and β denote BPR parameters. On the basis of BPR function, the following three factors should be included:

1) Road width factor. For skeleton networks, exclusive bus lanes should be considered to assure fast speed and high volume requirements; as a result, there's a constraint to the road width. To simplify the problem, number of lanes N is used to represent the road width factor. Some paths should be excluded: those with road widths narrower than one-way, two-lane or two-way, four-lane.

2) Safety factor. Many potential hazards exist because of the complexity of public transportation, such as road condition, passenger (and driver) condition, and vehicle condition. Safety plays a particularly important role in TRND. However, this factor hasn't received enough attention in existing studies of transit route network optimization. In order to quantify safety effects, a safety score is introduced in this research, which has a negative relationship with traffic crashes. Normalized transformation is used so that safety score $\beta_{i,i+1}$ falls between 0 and 1, i.e., $\beta_{i,i+1} \in [0, 1]$. By statistically analyzing safety data collected from Beijing, Guangzhou and other cities in China [39], [40], an approximately linear or exponential relationship was identified between traffic volume and safety score depending on the volume/capacity ratio. Thus, based on BPR function, the actual time cost between node i and $i+1$ can be expressed as $t = t_{i,i+1}\left(1 + a_{i,i+1}\left(\frac{q_{i,i+1}}{c_{i,i+1} \cdot N}\right)^{b_{i,i+1}}\right)$. N is number of lanes; $q_{i,i+1}$ represents traffic volume between node i and $i+1$; $c_{i,i+1}$ represents design traffic capacity between node i and $i+1$; $a_{i,i+1}$, $b_{i,i+1}$ denote safety influence coefficient. For low volume/capacity ratio, an approximately linear relationship was identified between traffic volume and safety score; and for high volume/capacity ratio, an approximately exponential relationship was identified [39], [40]:

$$a_{i,i+1} = \begin{cases} 1/\beta_{i,i+1} & \forall \dfrac{q_{i,i+1}}{c_{i,i+1} \cdot N} \in [0,\ 0.5) \\ 1 & \forall \dfrac{q_{i,i+1}}{c_{i,i+1} \cdot N} \in [0.5,\ +\infty) \end{cases} \tag{2}$$

$$b_{i,i+1} = \begin{cases} 1 & \forall \dfrac{q_{i,i+1}}{c_{i,i+1} \cdot N} \in [0,\ 0.5) \\ \beta_{i,i+1} & \forall \dfrac{q_{i,i+1}}{c_{i,i+1} \cdot N} \in [0.5,\ 1) \\ 1/\beta_{i,i+1} & \forall \dfrac{q_{i,i+1}}{c_{i,i+1} \cdot N} \in [1,\ +\infty) \end{cases} \tag{3}$$

3) Line length factor. Long lines offer direct connections, thereby increasing the proportion of direct travelers and reducing transfers. However, the operational reliability is often negatively related to the line length, leading to additional waiting time for passengers [41]. Considering all three factors, the time cost t is:

$$t = \sum_{k=1}^{m} \sum_{i=1}^{n-1} \frac{1}{1-l/l_{max}} \cdot t_{i,i+1} \cdot y_{i,i+1} \cdot x_{i,i+1}^{k} \cdot \left(1 + a_{i,i+1} \cdot \left(\frac{q_{i,i+1}}{c_{i,i+1} \cdot N}\right)^{b_{i,i+1}}\right) \quad (4)$$

where $y_{i,i+1}$ is a 0-1 variable (whether the nodes i, $i+1$ are in the same route); $x^{k}_{i,i+1}$ is also a 0-1 variable (whether the nodes i, $i+1$ are in the route k); l is length of line; l_{max} denotes upper bound of the line length. In addition, the estimation of travel time based on BPR function may be influenced by many factors. Thus, in this study, we calibrate the BPR parameters ($\mu_0 = 0.68, \mu_1 = 2.48$) based on intersection density, speed limit, bus stop density, and saturation, instead of using the given value ($\mu_0 = 0.15, \mu_1 = 4$). By utilizing the factors of intersection density (ψ_1), speed limit (ψ_2), bus stop density (ψ_3), and saturation (ψ_4), travel time based on BPR function is [42]:

$$t_{i,i+1} = \mu_{0,0}+\mu_{0,1}\psi_1+\mu_{0,2}\psi_2+\cdots+\mu_{0,n}\left(\frac{q_{i,i+1}}{c_{i,i+1} \cdot N}\right)^{\mu_1} = \frac{l_{i,i+1}}{v_{i,i+1}}\left[1 + \mu_0\left(\frac{q_{i,i+1}}{c_{i,i+1} \cdot N}\right)^{\mu_1}\right] \quad (5)$$

where $v_{i,i+1}$ represents free flow speed between node i and $i+1$; $l_{i,i+1}$ denotes line length between node i and $i+1$; $\mu_{0,0}, \mu_{0,1}, \mu_{0,2}, \cdots, \mu_{0,n}$ represent the coefficients of influencing factors.

Common line problem (i.e., multiple transit lines pass through the same link or node pair) is an important issue in transit network design. Assume that there are two routes through stops i and j, and the time cost between the two stops along route 1 will be shorter than that along route 2. In this situation, passengers from stops i to j will all be assigned to the route section of route 1. However, each route would have a maximum capacity limitation of the passenger demands assigned on route 1. Thus, the remaining passengers have to select route 2. As mentioned above, if the passenger demands on a certain section have reached or exceeded the maximum capacity of the route section, the time cost of the section between stops i and j would be set as very large impedance to ensure that the subsequent assignments do not select this section [31].

B. ARTERIAL NETWORK

As the main system to cover the large areas unserved by the skeleton network, the function of an arterial transit route network is to provide transit service between the skeleton and feeder network. Most existing methods for arterial TRND and optimization are based on an average or fixed travel time. However, in practice, travel time may be influenced by various factors, such as traffic volume, vehicle characteristics, road capacity, and so on. Therefore, all above mentioned factors should be reflected in the representation of travel time.

In this study, the arterial network is optimized from the perspective of passengers, operators, and traffic environment. For passengers, travel time can also be divided into four parts, which include in-vehicle travel time, dwell time, waiting time, and transfer time [43]–[45]. Operators and traffic environment are represented by operation expense and vehicle emission expense, respectively. The following equations show the calculation of all factors (time or equivalent time).

1) Passengers' in-vehicle travel time. Considering bus acceleration process in the exit area of stops and deceleration process in the entry area of stops, bus acceleration and deceleration time should be added into the in-vehicle travel time, which can be expressed as:

$$T_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ij} \times \left(\frac{l_{ij}}{v_{ij}}\left(1+\mu_0\left(\frac{q_{ij}}{c_{ij} \cdot N}\right)^{\mu_1}\right) + \frac{v_{ij}}{1+\mu_0\left(\frac{q_{ij}}{c_{ij} \cdot N}\right)^{\mu_1}} \cdot \left(\frac{1}{2A_{ij}}+\frac{1}{2B_{ij}}\right)\right) \quad (6)$$

where Q_{ij} represents passenger volume between node i and j; v_{ij} is free flow speed; A_{ij}, B_{ij} denote acceleration and deceleration, respectively; l_{ij} denotes line length; q_{ij} is traffic volume. It should be noted that traffic volume q_{ij} can be set to 0 for the grade-separated bus lanes and at-grade bus lanes, because the buses in an exclusive lane will not be disturbed by other vehicles. c_{ij} represents design traffic capacity. In this study, the capacity is simplified as a constant for the ease of analysis and presentation. It is shown in the following table by referring to [46].

TABLE 2. Relationship between design speed and capacity.

Design speed (km/h)	60	50	40	30	20
Capacity (pcu/km/ln)	1800	1700	1650	1600	1400

2) Passengers' dwell time at bus stop:

$$T_2=\sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ij}\left(max\left\{\frac{P_{up}t_{up}}{N_{up}}, \frac{P_{down}t_{down}}{N_{down}}\right\}+t_{oc}\right) \quad (7)$$

where P_{up} is number of boarding passengers; P_{down} denotes number of alighting passengers; N_{up} is number of doors for boarding; N_{down} denotes number of doors for alighting. According to the field surveys, the additional delay at bus stop t_{oc} can be simplified as a constant, with an average value of 6.5s if the bus is delayed by other bus vehicles or if it is difficult to pull out and return to traffic stream from bus stop. The service time for each passenger t_{up} or t_{down} is defined in the following table. It is important to note that when there are passengers standing in the vehicle, the boarding time will increase by 20% [47].

TABLE 3. Passenger boarding and alighting time.

Number of doors	Service time (seconds per passenger)		
	Boarding	Front door alighting	Rear door alighting
1	2.5	3.3	2.1
2	1.5	1.8	1.2
3	1.1	1.5	0.9
4	0.9	1.1	0.7
6	0.6	0.7	0.5

3) Passengers' waiting time:

$$T_3 = \sum_{k=1}^{m} \frac{1}{2f_k} \sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ij} \cdot x_{ij}^k \tag{8}$$

where f_k is departure frequency of the kth line.

4) Passengers' transfer time:

$$T_4 = \lambda_a \sum_{k=1}^{m} \sum_{i=1}^{n} \sum_{j=2}^{n} \left(Q_{i,j-1} \cdot x_{ij}^k - Q_{ij}\right) \tag{9}$$

where λ_a denotes conversion coefficient of transfer time; $Q_{i,j-1}$ denotes passenger volume between node i and j-1.

5) Bus operating cost:

Bus operating cost:

$$T_5 = \lambda_b \sum_{k=1}^{m} f_k \sum_{i=1}^{n} \sum_{j=1}^{n} l_{ij} \cdot x_{ij}^k \tag{10}$$

where λ_b represents conversion coefficient of operation expense.

6) Vehicle emission cost:

$$T_6 = \lambda_c \sum_{k=1}^{m} f_k \sum_{a \in E} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(h_{ij} \cdot l_{ij} + h_a \cdot d_a\right) \tag{11}$$

where λ_c is conversion coefficient of vehicle emission; d_a is intersection delay; h_{ij} is vehicle emission factor in the road segments; and h_a denotes vehicle emission factor in the intersections. The operating speed could affect the vehicle emission factor, and further affect the vehicle emission [48], [49]. In this study, we collected HC, CO, and NOx at different speeds and for vehicle types. The vehicle emission factors in the road segments and intersections are shown as follows:

The queuing delay at an intersection can be calculated with the formula proposed by Doherty [50].

$$d_a = \begin{cases} \frac{T_a}{2}(1-g_a)^2 + \frac{1980}{c_{ij} \cdot g_a} - \frac{q_{ij}}{c_{ij} \cdot g_a - q_{ij}}, & \frac{q_{ij}}{c_{ij} \cdot g_a} \le 0.95 \\ \frac{T_a}{2}(1-g_a)^2 - \frac{198.55 \cdot 3600}{c_{ij} \cdot g_a} + \frac{220 \cdot 3600 q_{ij}}{(c_{ij} \cdot g_a)^2}, & \frac{q_{ij}}{c_{ij} \cdot g_a} > 0.95 \end{cases} \tag{12}$$

where T_a is signal cycle length; g_a denotes green ratio.

TABLE 4. (a) Vehicle emission factor for intersections. (b) Vehicle emission factor for road segments.

(a)

Emission	Average of passenger car and bus vehicle $(g/(veh \cdot h))$
HC	99.53
CO	629.12
NO_X	12.05
Total	740.70

(b)

Average speed (km/h)	Average of passenger car and bus vehicle $(g/(veh \cdot km))$			
	HC	CO	NO_X	Total
10	19.29	120.79	3.72	143.80
20	11.69	76.59	3.34	91.62
30	9.11	61.85	3.21	74.17
40	7.14	45.47	3.23	55.84
50	5.91	34.64	3.26	43.81
60	5.09	27.40	3.28	35.77

Therefore, the arterial TRNDP in this study is analogous to an optimization problem, i.e., minimizing the time cost under the constraint of various factors:

$$min\ T = \sum_{i=1}^{6} \omega_i \cdot T_i \tag{13}$$

$$s.t.\ f_k = \frac{Q_k}{(\varphi_k)_{max} \cdot C_k}, \quad \forall k \in \{1, 2, \cdots, m\} \tag{14}$$

$$l_{ij} = \begin{cases} l_{ij}, & \forall \beta_{ij} \ge \beta_{min} \\ \infty, & \forall \beta_{ij} < \beta_{min} \end{cases} \ (safety\ score\ constraint) \tag{15}$$

$$\sum_{a \in E_d} g_a = 1,\ \forall E_d \subset E (green\ ratio\ constraint) \tag{16}$$

$$f_{min} \le f_k \le f_{max}\ (departure\ frequency\ constraint) \tag{17}$$

$$l_{min} \le l \le l_{max}\ (line\ length\ constraint) \tag{18}$$

$$x_{ij}^k = \begin{cases} 0, & node\ i, j\ in\ the\ route\ k \\ 1, & other \end{cases} \tag{19}$$

$$\sum_{k=1}^{m} 2f_k \cdot T_1 \le S\ (fleet\ size\ constraint) \tag{20}$$

$$m \le M_{max}\ (maximumnumberofroutesconstraint) \tag{21}$$

where Q_k denotes the maximum flow in a cross section; x_{ij}^k is a 0-1 variable (whether the nodes i, j are in the route k); $(\varphi_k)_{max}$ is upper bound of full-load ratio; C_k denotes rated load; β_{min} is lower limit of the safety score; E_d denotes set of all links entering the intersection; S is upper bound of operation vehicles; m represents the number of bus lines; M_{max} denotes the maximum number of lines; β_{ij} is safety score between node i and j. In this study, we employed a crash severity weighting system to quantify the safety score. There is general agreement that more severe crashes should have greater weights in identifying unsafe locations [51]. According to the reference [52], the severity index between nodes i and j SI_{ij} could be computed by Equation (22). Based on the severity index, the safety score between nodes i and j

is specified as follows:

$$SI_{ij} = 3.0\,(Z_1)_{ij} + 1.8\,(Z_2)_{ij} + 1.3\,(Z_3)_{ij} + (Z_4)_{ij} \tag{22}$$

$$\beta_{ij} = 1 - \frac{SI_{ij}}{SI} \tag{23}$$

where $(Z_1)_{ij}$ is total number of fatal crashes between node i and j; $(Z_2)_{ij}$ represents total number of serious injury crashes; $(Z_3)_{ij}$ is total number of other injury crashes; $(Z_4)_{ij}$ represents total number of property-damage-only crashes; SI represents the maximum value of all severity indexes between nodes i and j; β_{ij} is the safety score.

The objective function is to minimize the sum of in-vehicle travel time, bus dwell time, passengers' waiting time, passengers' transfer time, vehicle emission cost, and bus operating cost. The first and second terms of the objective function are the total passenger in-vehicle travel time. The third and fourth parts are the total passenger out-of-vehicle travel time. Note that the walking time factor is considered only in the feeder network model, when passengers access skeleton/arterial lines using feeder lines. When passengers directly use skeleton/arterial lines without transferring from feeder lines, an average of 5 minutes of transfer time per passenger is used based on field surveys. Consequently, walking time was not used as one of the principal elements in the skeleton/arterial network design. The last two components are the costs associated with operations and environment. $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5$, and ω_6 are introduced to reflect tradeoffs between passengers, operators, and traffic environment.

The first constraint is the safety score, which reflects the necessity of safe driving. The green ratio constraint guarantees that the sum of green ratios is 1. The third one shows that the given frequency of service on the proposed bus lines should not exceed a maximum value. The fourth constraint is the line length constraint. Long lines can increase the proportion of direct travelers and reduce transfers. However, the operational reliability is often negatively related to line length, which may lead to additional passenger waiting time. The fifth (fleet size) constraint sets a limit of available public transportation resources in operation vehicles. The maximum number of routes reflects the fact that in solving the TRNDP, transit planners usually set a maximum number of routes based on fleet size. In practice, the local authority (or government) is constrained by a limited budget which does not always allow the implementation of the identified optimal transit network and service design. Similar to previous studies [12], it is assumed that the constraint is set by the number of seats (i.e. required capacity satisfying the passenger demand $\sum_i S_i (C_k)_i$) for simplicity, instead of monetary budget.

C. FEEDER NETWORK

A feeder network is intended to meet the demand for internal trips and provide transfer service for skeleton and arterial networks. The feeder transit routes penetrate into residential districts and fill in the gaps of arterial networks to increase network density. As a result, it expands the accessibility of a transit network and shortens walking distance for passengers. Similar to skeleton networks, the model (optimization of feeder networks) is also a MCMF problem, and is resolved with the labeling method. In addition to safety factor, walking time and route overlap factors are considered.

1) WALKING TIME FACTOR

A transit trip may require walking paths from the origin to a bus stop and from a bus stop to the destination. One task for TRND is to reduce the walking time (or walking distance) for passengers and improve the level of accessibility. In this study, the passengers' walking distance is no more than twice of the serving radius.

2) ROUTE OVERLAP FACTOR

The function of a feeder network is to fill the gaps of a transit route network. To increase the network density, the feeder network should avoid route overlap. The route overlap is verified by checking on the similarity of the itinerary of two different bus lines [27]. Considering safety, walking time, and route overlap factors, time cost for the feeder network is:

$$t = \sum_{k=1}^{m} \sum_{i=1}^{n-1} \frac{1}{1 - min\left\{1, \left[\frac{d}{2R_s}\right]\right\}} \cdot t_{i,i+1} \cdot y_{i,i+1} \cdot x_{i,i+1}^{k} \cdot \left(1 + a_{i,i+1}\left(\frac{q_{i,i+1}}{c_{i,i+1} \cdot N \cdot (1 - x_{i,i+1})}\right)^{b_{i,i+1}}\right) \tag{24}$$

where d is passengers' walking distance; $x_{i,i+1}$ is a 0-1 variable, and is used to judge if there is overlap of feeder network ($x_{i,i+1} = 0$ represents that there is no overlap of feeder network, and vice versa); R_s denotes the radius of service; $a_{i,i+1}, b_{i,i+1}$ denote safety influence coefficient. Similarly, for low volume/capacity ratio, an approximately linear relationship was identified between traffic volume and safety score; and for high volume/capacity ratio, an approximately exponential relationship was identified:

$$a_{i,i+1} = \begin{cases} 1/\beta_{i,i+1} & \forall \frac{q_{i,i+1}}{c_{i,i+1} \cdot N \cdot (1 - x_{i,i+1})} \in [0, 0.5) \\ 1 & \forall \frac{q_{i,i+1}}{c_{i,i+1} \cdot N \cdot (1 - x_{i,i+1})} \in [0.5, +\infty) \end{cases} \tag{25}$$

$$b_{i,i+1} = \begin{cases} 1 & \forall \frac{q_{i,i+1}}{c_{i,i+1} \cdot N \cdot (1 - x_{i,i+1})} \in [0, 0.5) \\ \beta_{i,i+1} & \forall \frac{q_{i,i+1}}{c_{i,i+1} \cdot N \cdot (1 - x_{i,i+1})} \in [0.5, 1) \\ 1/\beta_{i,i+1} & \forall \frac{q_{i,i+1}}{c_{i,i+1} \cdot N \cdot (1 - x_{i,i+1})} \in [1, +\infty) \end{cases} \tag{26}$$

III. SOLUTION METHODOLOGIES

A. LABELING METHOD FOR SKELETON AND FEEDER NETWORKS

Let network $G = (V, E, C, t)$, where V is a set of nodes, E denotes a set of arcs, and C represents capacity of the arcs. $f = \{f(e) : e \in E\}$ represents a feasible flow in the network.

If the feasible flow isn't the maximum flow, there must be an augmenting chain for f in the network, and there must be a minimum cost augmenting chain for f. The cost of the augmenting chain μ can be represented as:

$$t\left(f'\right)-t\left(f\right)=\sum\nolimits_{\mu^{+}} t_{ij}\left(f'_{ij}-f_{ij}\right)-\sum\nolimits_{\mu^{-}} t_{ij}\left(f'_{ij}-f_{ij}\right) \quad (27)$$

where f_{ij} is feasible flow from node i to j; f'_{ij} represents adjustment feasible flow; t_{ij} denotes time cost. The following summary details the process of this method:

First, mark starting point of route V_s as $(0, V_s, 0)$ and mark other nodes as $(\infty, ?, \infty)$. At this time, V_s is an unexamined labeled node, and other nodes are unlabeled. The process by which all unexamined labeled nodes for $\gamma = r\,(r = 0, 1, 2, \cdots)$ are examined is called an iteration. The unexamined labeled nodes V_i are expressed as $(\alpha_i, \beta_i, \gamma_i)$, which are called α label, β label and γ label, respectively. The α label of V_j represents the upper bound of the shortest path from V_s to V_j; β label represents the origin of α label; γ label represents path length from V_s to V_j. If $\left(V_i, V_j\right) \in E, f_{ij} < C_{ij}$ and $\alpha_i + t_{ij} < \alpha_j$, examine the label of V_j. Let $\alpha_j = \alpha_i + t_{ij}, \beta_j = V_i$ and $\gamma_j = \gamma_i + 1$. If $\left(V_j, V_i\right) \in E, f_{ji} > 0$ and $\alpha_i - t_{ji} < \alpha_j$, examine the label of V_j. Let $\alpha_j = \alpha_i - t_{ji}, \beta_j = -V_i$ and $\gamma_j = \gamma_i + 1$. After that, node V_i has been labeled and examined, and V_j is the unexamined one. Then go to the next iteration. The process will be repeated until all nodes are labeled and examined. When all iterations are finished, α label of destination of route V_t denotes the cost of the minimum cost augmenting chain from V_s to V_t. With the reverse tracer method [53], the minimum cost augmenting chain from V_s to V_t can be found based on β label of V_t.

The following two steps are used to solve the MCMF problem. For a minimum cost flow f:

Step 1: Find the augmenting chain of minimum cost by using the labeling method. If all labeled nodes have been examined, but the node V_t hasn't been labeled, then the labeling process is finished. In this case, the minimum cost flow is the MCMF. If an augmenting chain μ for f from V_s to V_t can be found, then go to Step 2.

Step 2: Adjustment process. With the reverse tracer method, the minimum cost augmenting chain can be found based on β label of V_t. The adjustment amount is:

$$\Delta = min\left\{\underset{\mu^{+}}{min}\left(C_{ij} - f_{ij}\right), \underset{\mu^{-}}{min}\left(f_{ij}\right)\right\} \quad (28)$$

$$f'_{ij} = \begin{cases} f_{ij} + \Delta, & \left(V_i, V_j\right) \in \mu^{+} \\ f_{ij} - \Delta, & \left(V_i, V_j\right) \in \mu^{-} \\ f_{ij}, & other \end{cases} \quad (29)$$

The new minimum cost feasible flow f' can be represented as $V\left(f'\right) = V\left(f\right) + \Delta$. After obtaining the new feasible flow, remove all labels and re-label the minimum cost flow.

B. HYBRID HEURISTIC APPROACH FOR ARTERIAL NETWORK

There remains a challenge to be solved. The TRNDP is nondeterministic polynomial (NP-hard) due to its high complexity [17], [54]. An efficient transit route network consists of a set of routes that is "optimal" or "near-optimal" under certain criteria. For NP-hard problems, it is difficult to calculate the optimal solutions [23], [55]. In this situation, we seek a feasible solution that fulfills all the constraints, rather than an accurate optimal solution that is not possible to achieve. Consequently, the hybrid heuristic approach is introduced to address the multi-objective optimization problem.

1) ESTABLISH THE SET OF CANDIDATE ROUTES

The set of candidate routes is the solution space for arterial network design. The process aims to provide a searchable set of feasible routes for network design. In many situations, the optimization of a network is influenced by several factors such as passenger flow demand and actual traffic conditions. In order to extend the solution space, it's necessary to consider both the shortest and sub-shortest routes. The following three steps are used to establish the set of candidate routes:

Step 1: The safety factor is a critical factor in the optimization design of a transit route network. So a path with a high number of traffic crashes and the intersections with severe potential traffic conflicts should be ruled out. It is noted that, if an excluded path is essential for resident trips (the results can be obtained from the resident trip survey), it should be preserved.

Step 2: As the vital corridors for middle-sized zones, arterial networks should meet the demand for setting exclusive bus lanes. Those paths with road widths narrower than one-way, two-lane or two-way, four-lane should be excluded.

Step 3: Find the shortest path based on the shortest path algorithm and add it into the set of candidate routes. Corresponding to the shortest paths, all nodes except for the destination nodes should be chosen once. Two constraints for seeking shortest paths should be considered: 1) It should neither be a loop circuit nor coincide with all nodes contained by the shortest paths; 2) It shouldn't totally coincide with the k-1 shortest paths.

2) TRND BASED ON SIMULATED ANNEALING

Following the establishment of candidate routes, a hybrid method of simulated annealing and artificial ant colony optimization is proposed to handle the TRNDP. Detailed procedures of this method can be summarized as:

Step 1: For each i, j pair (node), the set of passengers' initial paths IP is generated with the k-shortest path algorithm. For each path $ip \in IP$, set $h = 1$ and the times of selection $n_{ip} = 0$.

Step 2: The distribution of travel time can be achieved by analyzing historical data or simulation. It is generally considered that travel times follow a normal distribution [27], and generate random number series of travel time $T_e^{(n)}, e \in E$.

Step 3: For each i, j pair, calculate the travel time of each path in IP and record the path with a shortest travel time. If ip is verified as the shortest path, set $n_{ip} = n_{ip} + 1$.

Step 4: If $h > \bar{h}$, then go to Step 5. Otherwise, set $h = h+1$ and go back to Step 2. $\bar{h}$ is the upper bound of the simulation times.

Step 5: For each i, j pair, sort the list by the value n_{ip} (high to low) in the set IP, and record the top 10 routes with highest value of n_{ip} into the set of candidate routes CR. The number of candidate routes will be quite large if excessive routes are selected for each nodes pair [27], so it is set to 10. Let $H = 1$ and define initial temperature as $\bar{T}^{(1)}$. Choose three routes from the set CR randomly and record them into the optional passenger route set OR.

Step 6: Based on the CR, apply the artificial ant colony optimization mechanism to assign passengers' flow. In the initial state, the pheromone intensity of each link is uniform. The single ant colony is divided into several sequential and mutually independent sub-colonies so as to assign them into the multiprocessors. In the first processor, ants can only use local network information and search the route depending mostly on the distances or randomly from their current spots to the next which has not yet been visited. When solutions are constructed, the global optimum is computed and broadcasted to all the processors in information exchange. Then, every processor updates the pheromone information based on the global optimum [56]. The function of updating pheromone can be expressed as:

$$\tau_{ij} = \rho\tau'_{ij} + \sum\nolimits_{u=1}^{U} \left(D/l\right), \quad \forall u \in (i, j) \tag{30}$$

where ρ is remaining coefficient of pheromone, $\rho \in (0, 1)$; τ_{ij} denotes pheromone concentration between node i and j; τ'_{ij} represents initial pheromone concentration; D represents the released intensity of pheromone concentration.

At the beginning, the ants can only follow some local information. Once some ants have constructed solutions, pheromone information is built. The paths with higher pheromone concentration will attract more ants. In addition, the pheromone concentration decreases as time progress. The probability for an ant u crawling from node i to j is:

$$p^u_{ij} = \frac{\tau^{\vartheta'}_{ij}\left(1/t_{ij}\right)^{\phi'}}{\sum_{S\in S_u} \tau^{\vartheta'}_{is}\left(1/t_{is}\right)^{\phi'}} \tag{31}$$

where ϑ' denotes relative importance of pheromone concentration; ϕ' is relative importance of visibility; and S_u denotes set of nodes accessible to the ant u. Calculate the passenger volume of route k and its departure frequency $f^{(1)}_k$. $f^{(1)}_k$ can be applied to calculate the corresponding objective function, denoted by $Z^{(1)}$.

Step 7: If $\bar{T}^{(H)} < \bar{T}^{min}$, then stop and output the results. $\bar{T}^{min}$ is the lower limit of the given temperature. Otherwise, set $\bar{T}^{(H+1)} = \eta\bar{T}^{(H)}$, $0 < \eta < 1$. This formula is applied to control the iteration times. Let $H = H + 1$.

Step 8: Based on CR, randomly generate $OR^{(H)}$ and assign the passengers' flow. Then $f^{(H)}_k$ and $Z^{(H)}$ can be calculated.

Step 9: If $Z^{(H)} < Z^{(H-1)}$, set $Z^{(H)} = Z^{(H)}$, $f^{(H)}_k = f^{(H)}_k$. Otherwise, set $Z^{(H)} = Z^{(H-1)}$ and $f^{(H)}_k = f^{(H-1)}_k$, then go back to Step 7.

Step 10: The present solution can be output as the best solution if it meets end conditions.

IV. CASE STUDIES

To analyze the efficacy of the proposed algorithm, it is compared to the previous studies using the transit network of Mandl [57] as a benchmark. In addition, to evaluate whether the proposed method can solve real world design of transit networks, an actual size road networks for the city of Zhaoyuan in China is conducted in this study.

A. MANDL BENCHMARK

The Mandl benchmark, which is also called the Swiss network, is a network of 15 nodes and 21 links that has been widely cited by researchers [12], [17], [23]. As shown in Table 5, this network has a total demand of 15,570 total trips per day. The length of each link in kilometers is shown in Figure 2 [57].

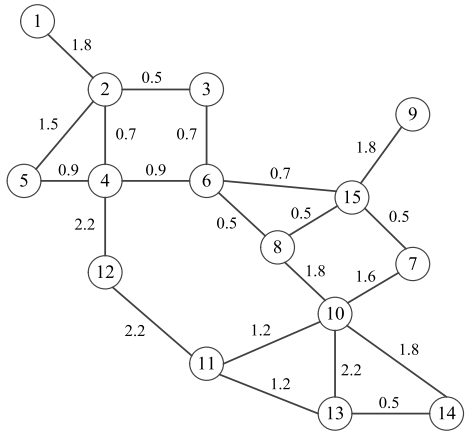


FIGURE 2. Swiss road network.

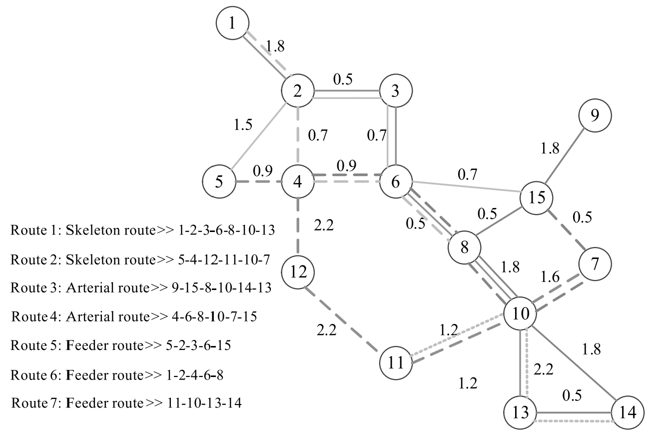


FIGURE 3. Optimal transit route network with seven lines using Mandl benchmark.

Figure 3 shows the results in which the developed procedures result in a transit network of 2 skeleton routes, 2 arterial routes, and 3 feeder routes. It should be noted that the demand matrix should be revised after designing the skeleton routes and arterial routes, respectively.

TABLE 5. Symmetrical travel demand of the Swiss network.

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

The proposed method is compared to the studies of Bagloee and Ceder [12], and Baaj and Mahmassani [17]. In these two competing methods, the former is multi-level TRND method, and the latter presents an artificial intelligence (AI)-based approach. The results show that the number of lines by Baaj and Mahmassani, Bagloee and Ceder, and this study are 7, 12, and 7, respectively. In addition, the length of lines by these three methods is 37.0, 59.9, and 38.5 kms, respectively. By contrast, there are two findings. 1) Compared with the method by Baaj and Mahmassani: the results indicate that the proposed methodology, with almost the same number of lines and length of lines as the method by Baaj and Mahmassani, has better performance for direct travelers, proportion of transfers, and total travel time. 2) Compared with the method by Bagloee and Ceder: although the total travel time of the proposed method is slightly more than that of the solution by Bagloee and Ceder (only 0.99%, less than 1%), the proposed method results in a significant reduction in number and length of lines compared with the solution by Bagloee and Ceder. More number of lines and longer length of lines mean more budget and input. In addition, the proposed methodology satisfies a much larger percentage of demand without any transfers (87.22%), and has no "two transfers." Consequently, the proposed methodology provides the best solution.

B. ZHAOYUAN CASE STUDY

Zhaoyuan is a medium-sized city of Shandong province, China, with approximately 300,000 residents. Based on the guidelines in Table 1, we select BRT, normal bus, and community branch modes for skeleton, arterial, and feeder network level, respectively. The city's current road network consists of 22 traffic zones, 278 nodes, and 482 links. In addition, the collected data set also includes 29,500 questionnaires from a resident trip survey (about 10 percent of the total population). The resident trip survey is an effective measure of identifying indispensable bus lines, which are filtered/deleted from candidate bus lines based on unsatisfying constraints or conditions. Based on the survey, 209,563 transit demands in this city were documented in Table 7. We set different minimum number of travelers of the selected route during each planning stage, $Q_{min} = 3000$ during skeleton route searching, $Q_{min} = 2000$ during arterial route searching, and $Q_{min} = 1000$ during feeder route searching.

TABLE 6. Comparison with other competing methods using Mandl benchmark.

Solution	No. of lines	Length of lines (km)	Percentage of demand (%)			Total travel time (min)
			Zero transfer	One transfer	Two transfers	
Baaj and Mahmassani	7	37.0	80.99	19.01	0	217,954
Bagloee and Ceder	12	59.9	83.66	15.21	0.95	202,255
This study	7	38.5	87.22	12.78	0	204,250

The proposed methodologies for the Zhaoyuan case study are evaluated in two parts. The first one attempts to illustrate the performance of the proposed method as well as the current situation with a simplified road network, in which each traffic zone is represented by a node. The second one aims to test the generality of the proposed method, which is applied to design real transit routes on the actual-size road network in Zhaoyuan.

In order to conduct this case study, several parameters need to be pre-set. In this study, the parameter settings are mainly based on realistic data, such as the upper bound of operation vehicles S: It is important to ensure adequate space to accommodate the maximum number of passengers along the entire route over a given time period (usually an hour). According to the peak-load factor concept, the number of

TABLE 7. Symmetrical travel demand of the network in the Zhaoyuan case study.

OD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1		249	187	656	2625	62	217	2531	2218	1661	150	718	187	406	616	75	94	65	150	208	68	2062
2			2125	125	312	187	601	750	712	4322	52	338	562	500	676	150	145	29	89	301	57	75
3				156	1529	111	372	496	797	1729	225	526	312	62	73	147	218	468	156	68	111	188
4					125	27	145	500	906	718	62	312	625	344	437	74	64	187	203	32	45	37
5						437	1468	2156	6437	3281	125	1250	1656	1062	1594	441	136	131	676	69	83	75
6							156	375	562	375	143	125	62	53	250	51	70	62	80	72	103	49
7								500	1968	7374	4781	1563	219	1001	1218	625	225	610	526	36	186	75
8									5562	3406	531	1781	437	625	500	141	144	160	125	249	634	150
9										10186	1312	4437	1312	1594	2250	1750	343	1375	1062	989	840	1466
10											4062	5031	2218	1749	3343	1968	313	1687	5222	156	824	2631
11												2687	125	281	3000	437	145	128	290	196	528	75
12													687	687	1375	2281	343	250	750	437	853	543
13														1750	437	125	687	625	125	226	67	79
14															2125	813	187	562	500	57	79	250
15																7468	250	1062	981	1125	62	406
16																	281	187	125	1531	62	593
17																		313	258	105	106	62
18																			1343	1938	247	930
19																				3937	138	1625
20																					100	2437
21																						62
22																						

vehicles required for period j is:

$$S_j = \frac{\bar{P}_{mj}}{\xi_j \cdot d_j} \tag{32}$$

where $\bar{P}_{mj}$ is the average maximum number of passengers in period j; d_j represents the capacity of a vehicle (number of seats plus the maximum allowance standees); and ξ_j is the load factor during period j, $0 < \xi_j \leq 1$. The standard ξ_j can be set so that $\xi_j \times d_j$ is equal to a desired fraction of the capacity (e.g. $\xi_j \times d_j$ = rated load C_k) [11]. From the resident trip survey, 14058 transit demands per hour (max load) in this study were documented. To ensure that the total number of passengers doesn't exceed the total capacity of lines, the number of vehicles should be set no less than 14058/50≈281. Thus, the upper bound of operation vehicles is set to 300. Note that long routes increase route cycle time and the probability of breakdown, while short routes result in more transfers. In light of this, the minimum and maximum thresholds of the line length are 6 and 12 km, respectively. Similarly, the minimum and maximum values of the departure frequency are set to 4 and 15 buses per hour, respectively. Besides, it is assumed that weighting values for in-vehicle travel time, dwell time, waiting time and transfer time is 2, and for operation expense and vehicle emission expense is 1 based on the principle of emphasizing passengers' interests. If an input parameter cannot be calibrated by the realistic data, it can be determined according to the relevant published literature. For instance, the conversion coefficient of transfer time and operation expense λ_a and λ_b are equal to 5 min and 1 [27], and relative importance of pheromone concentration ϑ' and visibility ϕ' are both set to 0.5 [56]. In addition to the above parameters, other parameters are set below: R_s = 500m; M_{max} = 15; $N_{down}(N_{up})$ = 1; λ_c = 0.1; $\rho = 0.9$; τ_{ij} (initial) = 10; $D = 2$.

The proposed solution methodologies are applied to address the network design problem with the initial parameter settings. In the skeleton network, the MCMF approach is used

TABLE 8. Optimal bus lines obtained by the proposed solution methodologies for Zhaoyuan.

Route number	Length (km)	Frequency	Node sequence
Skeleton: BRT			
1	11.4	10	11-10-12-15-20-22
2	7.8	9	1-8-9-10
Arterial: Normal bus			
1	10.8	7	17-3-4-13-8-9-10-11-7
2	11.3	8	2-4-13-14-15-16-21
3	9.0	9	17-18-14-9-10-7
4	11.3	7	1-8-12-15-20-21
Feeder: Community branch			
1	7.8	5	19-18-14-12-10-7-11
2	11.2	6	22-20-15-12-9-5-1
3	7.7	5	2-4-13-9-10-12-11
4	7.9	7	7-11-12-15-20-19
5	9.2	4	16-21-20-18-14-13-4
6	9.9	6	3-2-1-5-6

TABLE 9. Examples of route optimization using different methods for Zhaoyuan.

Alternatives (List two selection routes)		Passenger volume	Current situation		Proposed method	
			Travel time	Node sequence	Travel time	Node sequence
From 10 to 19	Route 1	5222	28.340	10-12-14-T-14-18-19	**18.340**	**10-12-14-18-19**
	Route 2		**28.328**	**10-12-15-T-15-20-19**	28.328	10-12-15-T-15-20-19
From 2 to 10	Route 1	4322	**25.748**	**2-4-13-14-12-9-10**	**25.124**	**2-4-13-12-9-10**
	Route 2		38.884	2-1-5-T-5-8-9-10	35.492	2-1-T-1-8-9-10
From 5 to 10	Route 1	3281	**14.432**	**5-8-9-10**	**14.432**	**5-8-9-10**
	Route 2		33.624	5-6-T-6-9-T-9-10	20.496	5-9-T-9-10
From 2 to 3	Route 1	2125	**25.468**	**2-4-T-4-3**	**10.924**	**2-3**
	Route 2		34.692	2-4-13-T-13-17-3	25.468	2-4-T-4-3
From 3 to 10	Route 1	1729	**27.148**	**3-4-17-18-14-12-10**	**27.080**	**3-4-13-8-9-10**
	Route 2		37.080	3-4-13-8-T-8-9-10	34.356	3-4-13-T-13-12-9-10
From 5 to 15	Route 1	1594	**20.504**	**5-8-9-10-12-15**	**17.808**	**5-8-9-12-15**
	Route 2		32.656	5-6-7-T-7-11-12-15	26.816	5-8-13-T-13-14-15
From 3 to 5	Route 1	1529	**31.268**	**3-4-13-8-T-8-5**	**22.596**	**3-2-1-5**
	Route 2		46.696	3-4-T-4-2-T-2-1-5	31.268	3-4-13-8-T-8-5
From 9 to 22	Route 1	1466	**26.060**	**9-10-T-10-12-15-20-22**	**19.284**	**9-12-15-20-22**
	Route 2		—	—	26.060	9-10-T-10-12-15-20-22

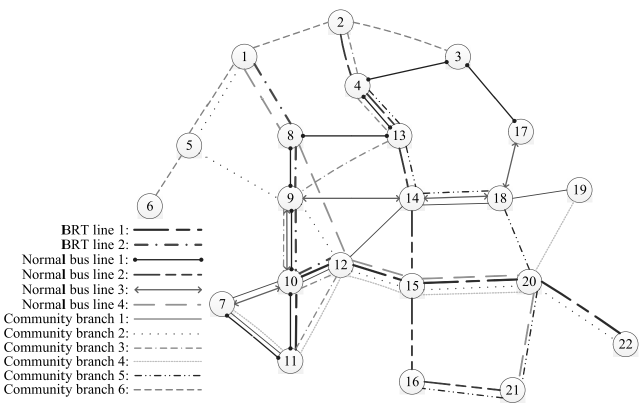


FIGURE 4. Optimal transit network (in accordance with Table 8) for Zhaoyuan case study.

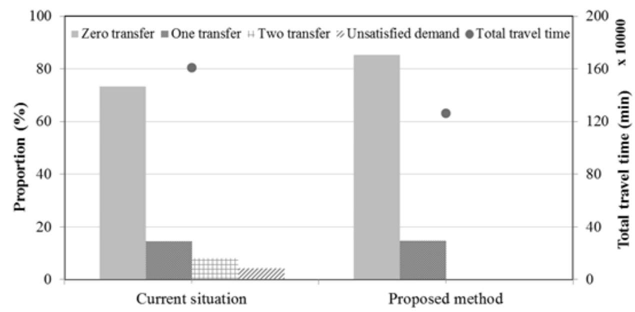


FIGURE 5. Comparison of demand proportion and total travel time.

to find the optimal transit corridors based on transit OD. The time cost will tend to ∞ if the line length is longer than 12 km. Next, the arterial transit routes are designed based on the transit trip ODs excluding those assigned on the skeleton transit routes. A grand total of 325 paths are generated in the k-shortest path algorithm ($k = 10$). Then, the optimal bus lines are chosen during the simulated annealing process. Lastly, feeder transit routes are designed to serve the rest of the transit trip ODs with the MCMF method. As a result, 12 bus lines are designed (see Table 8 and Figure 4). It is noteworthy that there are 4 bus lines on links (9, 10), (10, 12), (4, 13), (12, 15), and (15, 20). They correspond to passenger volume, especially the proportion of direct travelers.

The results from the proposed methodology are compared to those from the current situation (existing network). As shown in Figure 5, the proposed method not only provides a high percentage of direct travelers (85.23%), but also has lower percentages of "two transfers" and unsatisfied demand. In addition, the transit route network using the proposed method offers a substantially shorter "total travel time," with a 21.51% reductions as compared to the current situation.

To better illustrate and compare the performance of the two methods, the eight heavy OD pairs are selected to design bus routes. Table 9 presents the enumeration of the two best selected routes for each OD pair by each method. In the table, the route in boldface represents the optimal one of the two selected routes and "T" means transfer. The optimal routes are chosen based on the travel time and whether or not transfers are needed. As shown in the table, the optimal routes by the proposed method do not require transfers for any of the eight OD pairs, while four of the routes from the current situation require transfers. In addition, the total travel time for the eight OD pairs from the proposed method, with a 21.80% reduction, is significantly lower than the competing case.

The generality of the proposed method is explored to identify an efficient transit route network on the actual-size road networks of Zhaoyuan. The design results are shown in Figure 6, as denoted by different colors. Two skeleton

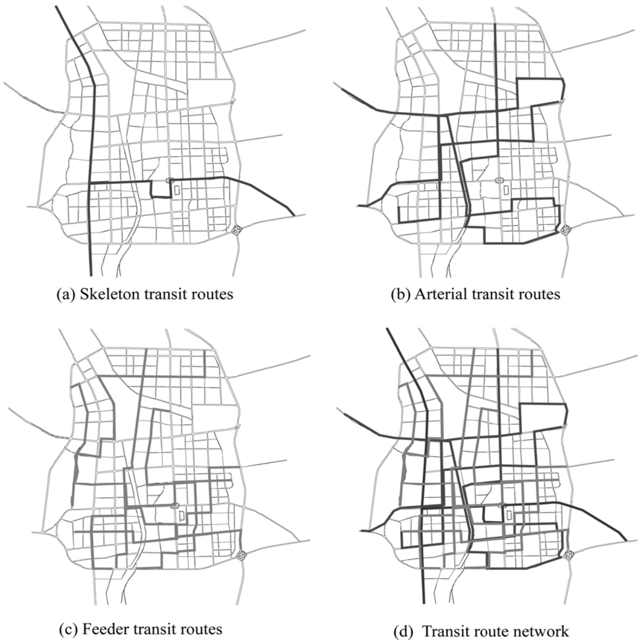


FIGURE 6. Designed transit network and the road network in the city of Zhaoyuan.

transit routes go through the center of the city, and link up the passenger transportation center, the railway station and the city center. Compared with the skeleton transit routes, four arterial and six feeder transit routes have broader coverage. In particular, the feeder routes reach the communities and provide better access for residents.

V. CONCLUSION

This study proposed a multi-level and multi-mode optimization model to design a transit route network. In this method, urban transit routes were divided into three levels (skeleton network, arterial network, and feeder network) based on the features of bus routes and city size. Different levels of the network were correlated with different modes of public transportation. For the skeleton and feeder networks, we utilized a labeling method to solve the MCMF problem. For the arterial network, a multi-objective optimization method and a hybrid heuristic approach were developed.

To highlight the efficacy of the proposed algorithm, it was compared to the studies of Bagloee and Ceder, and Baaj and Mahmassani using the transit network of Mandl as a benchmark. The results showed that the proposed methodology, with almost the same number of lines as the compared method by Baaj and Mahmassani, had the better performance for direct travelers, proportion of transfers, and total travel time. Although the total travel time for the proposed approach was slightly more than that of the solution by Bagloee and Ceder, the proposed method satisfied a much larger percentage of demand without any transfers, and had no two transfers. In addition, the solution by Bagloee and Ceder proposed more routes and longer lines than our method. Considering efficiency and economy, the proposed methodology provided a moderately better solution than other competing methods. An actual size road networks for the city of Zhaoyuan in China was also used for comparison. The total travel time for the proposed routes was significantly lower than that of the current situation, with a 21.51% reduction as compared to the current situation case. In addition, the proposed method provided 85.23% direct travelers, 14.65% travelers with one transfer, 0.12% travelers with two transfers and no unsatisfied demand, which were better than the results from the current situation case. The results indicated the proposed method was capable of handling real world design of transit networks.

REFERENCES

[1] L. Zhang, J. Lu, X. Yue, J. Zhou, Y. Li, and Q. Wan, "An auxiliary optimization method for complex public transit route network based on link prediction," *Mod. Phys. Lett. B*, vol. 32, no. 5, Feb. 2018, Art. no. 1850066.

[2] J. Xiong, B. Chen, Y. Chen, Y. Jiang, and Y. Lu, "Route network design of community shuttle for metro stations through genetic algorithm optimization," *IEEE Access*, vol. 7, pp. 53812–53822, 2019.

[3] X. Fu and W. H. K. Lam, "Modelling joint activity-travel pattern scheduling problem in multi-modal transit networks," *Transportation*, vol. 45, no. 1, pp. 23–49, Jan. 2018.

[4] A. Ceder and N. H. M. Wilson, "Bus network design," *Transp. Res. B, Methodol.*, vol. 20, no. 4, pp. 331–344, 1986.

[5] V. Guihaire and J.-K. Hao, "Transit network design and scheduling: A global review," *Transp. Res. A, Policy Pract.*, vol. 42, no. 10, pp. 1251–1273, Dec. 2008.

[6] T. A. Chua, "The planning of urban bus routes and frequencies: A survey," *Transportation*, vol. 12, no. 2, pp. 147–172, 1984.

[7] Z.-R. Peng and R. Huang, "Design and development of interactive trip planning for Web-based transit information systems," *Transp. Res. C, Emerg. Technol.*, vol. 8, nos. 1–6, pp. 409–425, Feb. 2000.

[8] E. Cipriani, S. Gori, and M. Petrelli, "Transit network design: A procedure and an application to a large urban area," *Transp. Res. C, Emerg. Technol.*, vol. 20, no. 1, pp. 3–14, Feb. 2012.

[9] F. Zhao and X. Zeng, "Optimization of transit route network, vehicle headways and timetables for large-scale transit networks," *Eur. J. Oper. Res.*, vol. 186, no. 2, pp. 841–855, Apr. 2008.

[10] A. Verma, A. Kumari, D. Tahlyan, and A. B. Hosapujari, "Development of hub and spoke model for improving operational efficiency of bus transit network of Bangalore city," *Case Stud. Transp. Policy*, vol. 5, no. 1, pp. 71–79, Mar. 2017.

[11] A. Ceder, *Public Transit Planning and Operation: Theory, Modeling and Practice*, 1st ed. Oxford, U.K.: Elsevier, 2007.

[12] S. A. Bagloee and A. A. Ceder, "Transit-network design methodology for actual-size road networks," *Transp. Res. B, Methodol.*, vol. 45, no. 10, pp. 1787–1804, Dec. 2011.

[13] J. Agrawal and T. V. Mathew, "Transit route network design using parallel genetic algorithm," *J. Comput. Civil Eng.*, vol. 18, no. 3, pp. 248–256, Jul. 2004.

[14] W. Fan and R. B. Machemehl, "Using a simulated annealing algorithm to solve the transit route network design problem," *J. Transp. Eng.*, vol. 132, no. 2, pp. 122–132, Feb. 2006.

[15] S. Carrese and S. Gori, "An urban bus network design procedure," in *Transportation Planning* (Applied Optimization), vol. 64. Boston, MA, USA: Springer, 2002, pp. 177–196.

[16] E. Martins De Sá, I. Contreras, J.-F. Cordeau, R. Saraiva De Camargo, and G. De Miranda, "The hub line location problem," *Transp. Sci.*, vol. 49, no. 3, pp. 500–518, Aug. 2015.

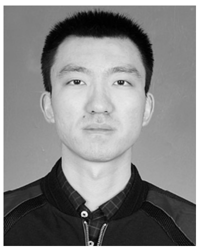
[17] M. H. Baaj and H. S. Mahmassani, "An AI-based approach for transit route system planning and design," *J. Adv. Transp.*, vol. 25, no. 2, pp. 187–209, Mar. 1991.

[18] M. Baaj and H. S. Mahmassani, "Hybrid route generation heuristic algorithm for the design of transit networks," *Transp. Res. C, Emerg. Technol.*, vol. 3, no. 1, pp. 31–50, Feb. 1995.

[19] S. B. Pattnaik, S. Mohan, and V. M. Tom, "Urban bus transit route network design using genetic algorithm," *J. Transp. Eng.*, vol. 124, no. 4, pp. 368–375, Jul. 1998.

[20] I. Constantin and M. Florian, "Optimizing frequencies in a transit network: A nonlinear bi-level programming approach," *Int. Trans. Oper. Res.*, vol. 2, no. 2, pp. 149–164, Apr. 1995.

[21] P. D. Site and F. Filippi, "Bus service optimisation with fuel saving objective and various financial constraints," *Transp. Res. A, Policy Pract.*, vol. 35, no. 2, pp. 157–176, Feb. 2001.
[22] R. Z. Farahani, E. Miandoabchi, W. Szeto, and H. Rashidi, "A review of urban transportation network design problems," *Eur. J. Oper. Res.*, vol. 229, no. 2, pp. 281–302, Sep. 2013.
[23] P. Chakroborty and T. Wivedi, "Optimal route network design for transit systems using genetic algorithms," *Eng. Optim.*, vol. 34, no. 1, pp. 83–100, Jan. 2002.
[24] C. Gerez, L. I. Silva, E. A. Belati, A. J. S. Filho, and E. C. M. Costa, "Distribution network reconfiguration using selective firefly algorithm and a load flow analysis criterion for reducing the search space," *IEEE Access*, vol. 7, pp. 67874–67888, 2019.
[25] A. S. Mohaymany and A. Gholami, "Multimodal feeder network design problem: Ant colony optimization approach," *J. Transp. Eng.*, vol. 136, no. 4, pp. 323–331, Apr. 2010.
[26] F. Zhao and X. Zeng, "Simulated annealing-genetic algorithm for transit network optimization," *J. Comput. Civil Eng.*, vol. 20, no. 1, pp. 57–68, 2006.
[27] Y. Yan, Z. Liu, Q. Meng, and Y. Jiang, "Robust optimization model of bus transit network design with stochastic travel time," *J. Transp. Eng.*, vol. 139, no. 6, pp. 625–634, Jun. 2013.
[28] J. R. Gómez, J. Pacheco, and H. Gonzalo-Orden, "A tabu search method for a bi-objective urban waste collection problem," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 30, no. 1, pp. 36–53, Jan. 2015.
[29] H. Martínez, A. Mauttone, and M. E. Urquhart, "Frequency optimization in public transportation systems: Formulation and metaheuristic approach," *Eur. J. Oper. Res.*, vol. 236, no. 1, pp. 27–36, Jul. 2014.
[30] C. Robert, *The Transit Metropolis: A Global Inquiry*. Washington, DC, USA: Island Press, 1998.
[31] B. Yu, Z.-Z. Yang, P.-H. Jin, S.-H. Wu, and B.-Z. Yao, "Transit route network design-maximizing direct and transfer demand density," *Transp. Res. C, Emerg. Technol.*, vol. 22, pp. 58–75, Jun. 2012.
[32] L. Tong, X. Zhou, and H. J. Miller, "Transportation network design for maximizing space-time accessibility," *Transp. Res. B, Methodol.*, vol. 81, pp. 555–576, Nov. 2015.
[33] Z. Di, L. Yang, J. Qi, and Z. Gao, "Transportation network design for maximizing flow-based accessibility," *Transp. Res. B, Methodol.*, vol. 110, pp. 209–238, Apr. 2018.
[34] J. C. Chu, "Mixed-integer programming model and branch-and-price-and-cut algorithm for urban bus network design and timetabling," *Transp. Res. B, Methodol.*, vol. 108, pp. 188–216, Feb. 2018.
[35] L. Yanga, Z. Dia, M. M. Dessoukyc, Z. Gaoa, and J. Shib. (2019). *Collaborative Optimization of Last-Train Timetable With Passenger Accessibility: A Space-Time Network Design Based Approach*. https://cpb-us-e1.wpmucdn.com/sites.usc.edu/dist/0/249/files/2019/04/Last-train-20190427.pdf.
[36] C. Sun, L. Cheng, S. Zhu, F. Han, and Z. Chu, "Multi-criteria user equilibrium model considering travel time, travel time reliability and distance," *Transp. Res. D, Transp. Environ.*, vol. 66, pp. 3–12, Jan. 2019.
[37] X. Yan, P. Dong, X. Du, T. Zheng, H. Zhang, and M. Guizani, "Congestion game with link failures for network selection in high-speed vehicular networks," *IEEE Access*, vol. 6, pp. 76165–76175, 2018.
[38] J. Ma, D. Li, L. Cheng, X. Lou, C. Sun, and W. Tang, "Link restriction: Methods of testing and avoiding braess paradox in networks considering traffic demands," *J. Transp. Eng., A, Syst.*, vol. 144, no. 2, Feb. 2018, Art. no. 04017076.
[39] W. Du, L. Jiang, and Y. Liufu, "Accident model prediction of urban intersection based on traffic flow," *J. East China Jiaotong Univ.*, vol. 26, no. 5, pp. 19–25, 2009.
[40] J. Chen, Z. Li, W. Wang, and H. Jiang, "Evaluating bicycle–vehicle conflicts and delays on urban streets with bike lane and on-street parking," *Transp. Lett.*, vol. 10, no. 1, pp. 1–11, Jan. 2018.
[41] N. Van Oort and R. van Nes, "Line length versus operational reliability: Network design dilemma in urban public transportation," *Transp. Res. Rec.*, vol. 2112, no. 1, pp. 104–110, 2009.
[42] N. Liu, S. Zhao, and N. He, "Further study of impedance function based on BPR function," *J. Wuhan Univ. Technol. (Transp. Sci. Eng.)*, vol. 37, no. 3, pp. 545–548, 2013.
[43] C. Wang, Z. Ye, E. Chen, M. Xu, and W. Wang, "Diffusion approximation for exploring the correlation between failure rate and bus-stop operation," *Transportmetrica A, Transp. Sci.*, vol. 15, no. 2, pp. 1306–1320, Nov. 2019.
[44] E. Chen, Z. Ye, C. Wang, and M. Xu, "Subway passenger flow prediction for special events using smart card data," *IEEE Trans. Intell. Transp. Syst.*, to be published, doi: 10.1109/tits.2019.2902405.
[45] Q. Cheng, Z. Liu, and W. Szeto, "A cell-based dynamic congestion pricing scheme considering travel distance and time delay," *Transportmetrica B, Transp. Dyn.*, vol. 7, no. 1, pp. 1286–1304, Dec. 2019.
[46] *Ministry of Housing and Urban-Rural Development of the People's Republic of China*, China Archit. Building Press, Beijing, China, 2012.
[47] KFH Group, *Transit Capacity and Quality of Service Manual*, 3rd ed. Washington, DC, USA: Transportation Research Board, 2013.
[48] Y. Pan, S. Chen, F. Qiao, S. V. Ukkusuri, and K. Tang, "Estimation of real-driving emissions for buses fueled with liquefied natural gas based on gradient boosted regression trees," *Sci. Total Environ.*, vol. 660, pp. 741–750, Apr. 2019.
[49] Y. Zhang, Y. Jiang, W. Rui, and R. G. Thompson, "Analyzing truck fleets' acceptance of alternative fuel freight vehicles in China," *Renew. Energy*, vol. 134, pp. 1148–1155, Apr. 2019.
[50] A. R. Doherty, "A comprehensive junction delay formula," Dept. Transp., Chennai, Inida, LTRI Working Paper, 1977.
[51] K. Geurts, G. Wets, T. Brijs, and K. Vanhoof, "Identification and ranking of black spots: Sensitivity analysis," *Transp. Res. Rec.*, vol. 1897, no. 1, pp. 34–42, Jan. 2004.
[52] L. Truong and S. Somenahalli, "Using GIS to identify pedestrian-vehicle crash hot spots and unsafe bus stops," *J. Public Transp.*, vol. 14, no. 1, pp. 99–114, Mar. 2015.
[53] W. Wang and J. Lu, *System Analysis Methods for Highway and Traffic Engineering*. Beijing, China: China Communications Press, 2011.
[54] Y. Lai, F. Yang, L. Zhang, and Z. Lin, "Distributed public vehicle system based on fog nodes and vehicular sensing," *IEEE Access*, vol. 6, pp. 22011–22024, 2018.
[55] Z. Liu, S. Wang, B. Zhou, and Q. Cheng, "Robust optimization of distance-based tolls in a network considering stochastic day to day dynamics," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 58–72, Jun. 2017.
[56] H. Sun, X. Li, Y. Liu, and S. Yu, "Ant colony optimization arithmetic of capacity restraint traffic assignment," *J. Southeast Univ.*, vol. 39, no. 1, pp. 177–180, 2009.
[57] C. E. Mandl, "Evaluation and optimization of urban public transportation networks," *Eur. J. Oper. Res.*, vol. 5, no. 6, pp. 396–404, Dec. 1980.



CHAO WANG received the Ph.D. degree in transportation engineering from Southeast University, in 2018. From 2017 to 2018, he went to Purdue University as a Visiting Ph.D. He is currently a Research Assistant Professor with the School of Transportation, Southeast University. His research interests include public transport, intelligent transportation systems, and transport modeling and optimization.

ZHIRUI YE received the Ph.D. degree in transportation engineering from Texas A&M University, College Station, TX, in 2007. He was a Research Scientist with the Western Transportation Institute, Montana State University. He is currently a Professor with the School of Transportation and the Director of the Transportation Research Institute, Southeast University (SEU). His research interests include, but are not limited to, advanced public transportation, traffic safety, and connected vehicles.

WEI WANG received the M.Eng. and Ph.D. degrees in civil engineering from Southeast University, China, in 1985 and 1989, respectively. He is currently a Professor with the Jiangsu Key Laboratory of Urban ITS, Southeast University. His research interests include urban transportation and intelligent transportation systems. He is a member of the model traffic technology panel of the National High-Tech Research and Development Program of China (863 Program) and also a member of the panel of the National Natural Science Foundation of China.