IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 4, 2019, accepted August 19, 2019, date of publication August 22, 2019, date of current version September 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2936922

A Two-Stage Algorithm for School Bus Stop Location and Routing Problem With Walking Accessibility and Mixed Load

JINGXUAN REN, WENZHOU JIN, AND WEITIAO WU
School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China
Corresponding authors: Wenzhou Jin (ctwzhjin@scut.edu.cn) and Weitiao Wu (ctwtwu@scut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61473122 and Grant 61703165, and in part by the Science and Technology Program of Guangzhou, China, under Grant 201904010202.

ABSTRACT This paper proposes a School Bus Stop location and Routing Problem with Walking Accessibility and Mixed Load (SBSLRP-WA-ML), where the individual difference of walking accessibilities among students and the possibility of serving students attending different schools with the same bus simultaneously are considered. We first develop a mixed integer programming model for SBSLRP-WA-ML with the objective of minimizing the total commuting time, including walking time from the residence to school, in-vehicle travel time, and service time at stops. A two-stage solution method is then developed. In stage 1, an iterative clustering method based on k-means and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is used to locate bus stops aiming at minimizing the number of stops subject to various walking accessibilities. In stage 2, an improved ant colony optimization algorithm (IACO) integrating two local search operators is devised, which is used to generate bus routes with minimal total commuting time. A number of instances of different sizes are generated to verify the solution approach, and the influential factors with respect to total commuting time are analyzed. The model is also compared to the door-to-door school bus services. Comparison to similar methods and sensitivity analysis of parameters are also conducted to analyze the performance and robustness of the proposed approach.

INDEX TERMS Bus routing, public transportation, school bus, stop location, two-stage algorithm.

I. INTRODUCTION

School bus is a dedicated transportation mode for students' commute. Such services have become the focus of the public attention in China since the issue of Regulation on School Bus Safety Management in 2012, which defines the rules over the school bus, drivers, priority, etc. Due to high operation costs and budget, door-to-door services for students may be difficult in practice due to the practical road network and other realistic reasons. With the human mobility, it is possible to locate a set of stops beforehand to gather and pick up students.

A good design of school bus includes the following steps: collecting the information of students; selecting bus stop locations; assigning students to stops; planning buses routing to pick up and deliver students. Unlike conventional public transit, since the passenger flow is relatively stationary during a school year, the planning of school bus can be undertaken once a year with given demand, which opens up opportunities of setting dedicated stop locations.

To develop an effective combined stop location and routing plan, it is imperative to evaluate the system with a student-oriented systematic performance measurement. Previous research usually investigates the stop location problem by introducing an upper bound on walking distance for all students. This neglects individual difference of walking accessibilities of students and possible unfairness. To address this issue, in this study we introduce a student-oriented systematic measurement and individual walking accessibility in the school bus planning.

Distinctly from the literature where the stop locations are given prior to routing optimization, this paper studies the locating of school bus stops, students' allocation to stops and the bus route optimization simultaneously.

The associate editor coordinating the review of this article and approving it for publication was Rashid Mehmood.

We present a School Bus Stop location and Routing Problem with Walking Accessibility and Mixed Load (SBSLRP-WA-ML), where the individual difference of walking accessibility among students and the possibility of serving students attending different schools with the same bus are considered simultaneously. The main challenge is the formulation of the systematic performance measurement of the system and individual walking accessibility, and the respective solution approach.

The remainder of this paper is organized as follows: In Section 2, we review the related literatures from existing studies and summarize the main contributions of this research. In Section 3, we define the SBSLRP-WA-ML and develop a mathematical model for the proposed problem. The solution algorithm and the computational experiments are presented in Section 4 and Section 5, respectively. Finally, we summarize the study in Section 6.

II. LITERATURE REVIEW AND MAIN CONTRIBUTIONS

Studies focused on school bus have been developed since it was introduced by Newton and Thomas in 1969 [1]. School bus problem is essentially a kind of pickup and delivery problem, while presenting its unique characteristic relative to traditional bus schedule [2]–[5], pickup and delivery problem [6], [7], and traditional vehicle routing problem [8]–[10]. Braca *et al.* [11], Desrosiers *et al.* [12], Park and Kim [13], and Ellegood *et al.* [14] have reviewed the development of school bus problem. The school bus problem includes several sub-problems: data preparation, bus stop selection (student assignment to stops), bus route generation, school bell time adjustment, and route scheduling [12]. Researches abound in the single sub-problems, especially bus route generation [11], [13]–[17] while there are limited studies focused on standalone school bus stop selection problem [18]–[20] in which references [18], [19] considered stop location problem. Since we primarily focus on the school bus location and routing problem, the stop location problem is considered as a sub-problem and separated from stop selection.

Previous studies on school bus stop selection problem generally set the same walking distance for all students as the constraint and the number of bus stops as the objective. Sarubbi *et al.* [18] studied the school bus stop location and allocation problem with the objective of minimizing the number of bus stops using a real and georeferenced data respecting constraint on walking distance, which was solved by an approach based on the equidistant vertices. Galdi and Thebpanya [19] considered school bus stop placement based on GIS. They considered minimum distance between any two stops and constraints related to actual road network other than the walking distance. There also exist literature on the stop selection problem without the consideration of stop placement. For example, Worwa [20] solved the bus stops generation and students' assignment problem aiming at minimizing the number of active stops by a greedy algorithm with a set of candidate stops and the constant maximum walking distance as the main constraint.

There are extensive researches on school bus routing problem with stop selection. Such problems resemble location-routing problem which integrates the facility location problem and vehicle routing problem for goods transportation [21]. Bodin and Berman [22] first proposed routing problem considering bus stop selection in 1979. They adopted the idea of ''ministop'' meaning a location determined by school authorities that can be used as an actual stop and allocated to students as unchanged before the subjective actual stop allocation at stop selection stage. They studied the single-load school bus routing problem using a route-first-cluster-second approach, followed by a scheduling approach for different schools with the objective of minimizing total travel time of all buses. After that, more and more researchers studied the integrated problem due to its complexity and actuality, among which we review the most related ones as follows.

Schittekat *et al.* [23] developed a Greedy Randomized Adaptive Search Procedure (GRASP) + Variable Neighborhood Descent (VND) algorithm for the stop selection and routing problem for a single school scenario with given potential stops to minimize the total travel distance of all active buses. The primary difference of this study from previous works lies in the sequence of the problems being solved. The stop selection and routes generation were solved at the same level followed by the optimal solution students' assignment.

Faraj *et al.* [24] studied the single-load school bus stop selection and routing problem, which was extended by Silva *et al.* [25] to a mixed load problem. Faraj *et al.* [24] solved the sub-problems separately, in which the stop selection sub-problem including bus stops generation and students' assignment was solved through a heuristic for dominant set problem based on a priority related to each vertex, while the routes were generated by a GRASP-like heuristic.

Parvasi *et al.* [26] considered the possibility of demand outsourcing and formulated the school bus routing problem with bus stops selection as a bi-level programming model in which the designer of school bus system as the upper-level decision maker determines the school bus stops and navigate routes to maximize the profit, while at the lower level, student make the mode choice to minimize the traveling costs. The traveling costs include costs of using the school bus system and costs of outsourcing.

When it comes to the solution approaches for school bus stop location and (or) routing problems, according to the latest surveys on school bus routing problem by Park and Kim [13] and Ellegood *et al.* [14], the solution approaches used in school bus routing problem account for considerable part of those of school bus planning problem, while the stop selection is always solved associated with bus routing problem and holds an appreciable part. The algorithm for solving the school bus stop selection and routing problem can be divided into two main strategies in general, that is LAR (Location-Allocation-Routing) and ARL

TABLE 1. Literatur review and contributions of this article.

Ref.	Year	Sub-problem types SL	SS	RG	SC	Constraints	Objectives	Solution approaches
Bodin and Berman [22]	1979		√	√	√	WD, BC	FS, TT_b, RT_s	LAR-based heuristics
Arias-Rojas et al. [27]	2012			√		BC, TW, FS, TT_S	TD_b	ACO
Schittekat et al. [23]	2013		√	√		CS, WD, BC	TD_b	GRASP+VND
Faraj et al. [24]	2014	√	√	√		WD, BC, TD_S	TD_b	SL: heuristic R: GRASP-based algorithm
Huo et al. [29]	2014		√	√		BC, FS,	TD_b	ACO
Silva et al. [25]	2015	√	√	√		WD, BC, LR	TD_b	SL: heuristic R: GRASP-based algorithm
Galdi and Thebpanya [19]	2016	√	√			WD, LD	NS	GIS-based
Sarubbi et al. [18]	2016	√				WD	NS	Pseudo-random constructive heuristics
Yao et al. [30]	2016			√		BC, LR	TD_b	ACO
Worwa [20]	2017		√			CS, WD	NS	Greedy algorithm
Parvasi et al. [26]	2017		√	√		CS, BC	Pro, CO	GA-EX-TS and SA-EX-TS
Mokhtari and Ghezavati [31]	2018			√		BC, TW, TT_S	FS, RT_s	ACO
This paper		√	√	√		WA, BC, LR	TCT	SL: clustering R: IACO

(Allocation-Routing) based on the sequence of stop location, students' allocation and routes generation. The approaches for standalone bus stop selection with or without stop location are described as before.

With respect to the school bus routing problem, the approaches can be divided into heuristics including construction heuristics and improvement heuristics, meta-heuristics (e.g., ant colony optimization, genetic algorithm, tabu search), and exact methods (e.g., column generation, cutting planes). For more details, readers could refer to the references [13] and [14]. Considering we adopt the IACO algorithm in the second stage to generate school bus routes aiming to minimizing the total commute time of all students, we will briefly review the Ant Colony Optimization (ACO) and its application in school bus routing problem. The adoption of ant colony algorithm for solving school bus planning problem shows a growing trend since 2012 when Arias-Rojas *et al.* [27] and Euchi and Mraihi [28] used ACO to solve school bus routing problem.

Arias-Rojas *et al.* [27] clustered stops to several clusters and generated routes for each cluster considered as several traveler salesman problems solved by the application of ACO to minimize total travel distance of all buses. Euchi and Mraihi [28] developed a hybrid metaheuristic integrating artificial ant colony to solve urban bus routing problem and variable neighborhood local search to improve the solutions. The hybrid algorithm was tested on a real school bus routing problem.

In 2014, Huo *et al.* [29] proposed school bus stop locating, students' assignment and routing problem for single-center and single-vehicle scenario and solved it by applying the location-allocation-routing strategy. The stop locating and student assignment are decided by dividing the school district with circles of radius of walking distance, 500 meters, and the center is used as bus stop while students within the circle will be assigned to the stop. The routes are generated by the application of ACO with the aim of minimizing total travel distance of all buses.

Then in 2016, Yao *et al.* [30] developed a two-stage heuristic to aggregate bus stops to several clusters nearest to one school and generate routes through ACO and post improvement realized by crossover and 2-opt exchange to solve the school bus routing problem with mixed load plan. After each routing process, the clustering factor will be adapted according to the performance of routing.

Recently, in 2018, Mokhtari and Ghezavati [31] studied the mixed-load school bus routing problem with the objectives of minimizing the number of active buses and the average riding time of students by a hybrid multi-objective ACO and a novel routing heuristic algorithm.

Table 1 shows the review of some related literature in chronological order and contributions of this article, the abbreviations used in which are explained in Table 2.

In summary, researches on the school bus routing and stop selection problem were dedicated to select a set of stops from given potential stops for students to walk to and wait for buses and generate routes for a set of vehicles to visit the selected stops and pick up the students with a set of students and their geographic information as input. Only a handful of works considered the locating of the bus stops, while the majority considered it as given demand to be input of routing procedure. When it comes to the objectives, to our best knowledge, there is no literature focusing on the total commute time of students from residence to school from the perspective of the whole school bus system. Therefore, in this paper we aim to study the stop location and routing problem form a systematic perspective with the objective to minimize commute time of students, including not only the in-vehicle riding time, but also the walking time from places of residence to stops and the service time at stops.

TABLE 2. Abbreviations used in the literature review.

Categories	Consideration	Abbreviations
Sub-problem types	Bus stop location	SL
	Bus stop selection	SS
	Bus route generation	RG
	Bus route scheduling	SC
Constraints	Upper bound on the walking distance from place of residence to assigned stop for any student	WD
	Bus capacity	BC
	A set of candidate stops to select for students to walk to and buses to visit	CS
	Upper bound on the travel distance for any student	TD_s
	Upper bound on the length for any bus tour	LR
	Lower bound on the distance between any two bus stops	LD
	Time window for students to arrive at school of any school	TW
	Upper bound on the size of the active buses	FS
	Upper bound on the travel time of any student	TT_s
Objectives	Minimizing the number of active buses	FS
	Minimizing total travel time of all buses	TT_b
	Minimizing total in-vehicle riding time for all students	RT_s
	Minimizing total travel distance of all buses	TD_b
	Minimizing the number of active bus stops	NS
	Maximizing the operation profit of the school bus operator	Pro
	Minimizing the travel cost of potential passengers of the school bus system	CO
	Minimizing total commute time of students including walking and ride time	TCT

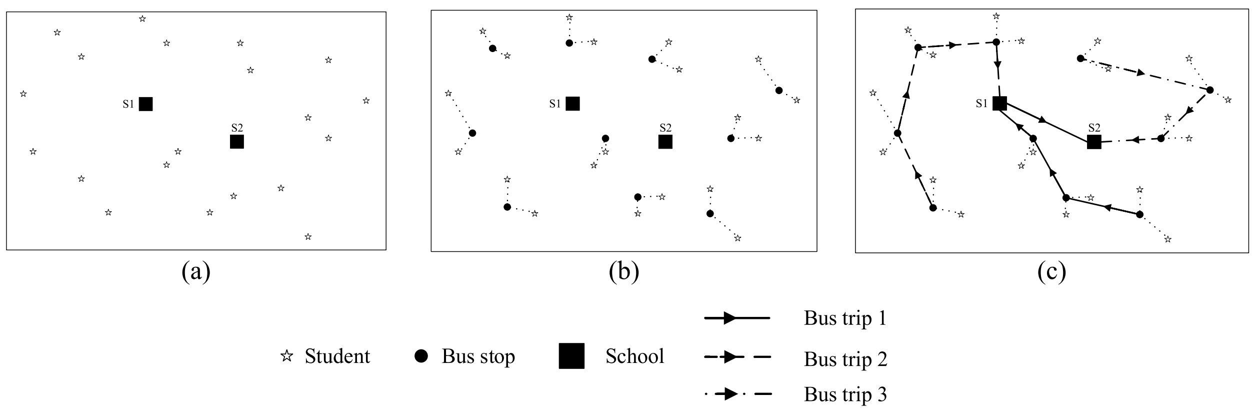


FIGURE 1. An example of SBSLRP-WA-ML (a) school district, (b) school bus stop generation and students' assignment, (c) school bus routes.

More importantly, this paper investigated the integrated model of school bus stop generation, students' assignment, and routes generation considering different walking accessibilities among students. The problem is solved by a two-stage approach including an iterative clustering method based on k-means and DBSCAN to generate stops and assign students, and a customized ant colony algorithm with two local search operators mechanism to generate bus routes to visit stops and schools.

The main contributions of this study are summarized as follows:

- We propose a more realistic and comprehensive objective function in the context of school bus planning-total student commute time from residence to school, which integrates the stop location, students' assignment, and route generation.
- We propose a two-stage solution approach including stops generation and students' assignment through an iterative clustering method based on k-means and DBSCAN at stage 1 and an IACO with two local search operators for routing at stage 2, while considering the interrelationship between two stages at each stage.
- We analyze the distinct walking accessibilities of students attending different schools at distinct educational stage to determine the stop coverage at stop location stage.
- We define and utilize the stop assignment information and school attraction information to consider mixed load and the interrelationship between two stages while developing IACO at stage 2.

III. PROBLEM STATEMENT

In SBSLRP-WA-ML, the input includes a set of students and schools, a fleet of available buses, and the geographic information. The problem is to determine a set of bus stops, assign students to stops, and find out a set of bus tours visiting stops to deliver them to corresponding schools.

Fig.1 illustrates a graphic example of the process of SBSLRP-WA-ML including stop generation, students' assignment, and route generation.

Fig.1 illustrates a graphic example of SBSLRP-WA-ML. Fig.1 (a) shows that there are two schools represented by squares and a set of students denoted by pentagrams who attend School 1 or School 2 in the school district represented by the rectangle. Fig.1 (b) shows the result of stop generation and students' assignment, in which circles represent the bus stops, and the dotted lines represent the students assigned to corresponding stops. Notice that students from different schools can be assigned to a same stop, which leads to the mixed load problem. In Fig.1 (c), three bus trips visiting stops are generated, including bus trip 1, a mixed-load trip which visits school 1 and school 2 in sequence, and bus trip 2 and 3 serve students from school 1 and 2, respectively.

Given a fleet of accessible buses and a set of locations representing students and schools, this paper determines the locations of stops, assignment of students to stops and generation of bus routes for school bus system serving students at multiple grades of multiple schools, which leads to different walking accessibilities and mixed loads. The purpose of this section is to formulate a mixed integer programming model for SBSLRP-WA-ML, which can be easily extended to incorporating special-education students. The main constraints include the walking accessibility for each student, vehicle capacity and fleet size. In particular, the bus navigation along schools must be after visiting all stops. The objective is to minimize the total commute time from home to school for all students.

A. ASSUMPTIONS

The proposed formulation is based on the following assumptions:

- This paper considers providing school bus services at morning period, which means that school buses visit bus stops, pick up students and deliver them to corresponding schools.
- Students assigned to a stop must be served by the same bus.
- A homogeneous fleet of vehicles with the same capacity is available.
- There is no transfer during the services.
- The travel time of a bus from the depot to its first stop to visit and from the last school to visit to the depot is of no importance to students' commute time, so it will not be considered.
- Each student departs from place of residence according to the predefined schedule and arrives at the assigned stop almost at the same time with the bus's arrival, therefore the waiting time at stops can be ignored.
- The buses are running at a constant speed between any two nodes.

B. NOTATIONS

Given the locations of students and schools, the proposed problem is to locate a set of school bus stops, assign students to certain stop to walk to, and generate a set of routes for a set of available buses to visit stops and serve students.

TABLE 3. Parameters and decision variables for SBSLRP-WA-ML.

Parameters	
$S(CX_s, CY_s)$	Coordinates of location of student s or stop s
d_i	The number of students assigned to stop i
D_i^k	The number of students assigned to stops visited by bus k and attending school i
D_i	The number of students to be served who attending school i
l_{ij}	The distance between node i and node j
t_i^s	The service duration at each node i
v_s	The walking speed of student s
a_{ik}	The arrival time of bus k at stop i
L_{ik}	The load of bus k immediately after it leaves stop i
M	A largely positive number
WT_s	Maximum walking time from home to stop for student s
Decision variables	
$I(CX_i, CY_i)$	Coordinates of location of stop i
X_{ijk}	If vehicle k goes from node i to node j, $X_{ijk}=1$; otherwise $X_{ijk}=0$.
Y_{si}	If student s is assigned to stop i, $Y_{si}=1$; otherwise $Y_{si}=0$.

Let $K = \{1, 2, \ldots, k, \ldots, n_k\}$ be the set of buses with the same capacity C and running at equal speed v. From a graph theoretical point of view, the SBSLRP-WA-ML may be stated as follows: Let $G = V \cup A$ be a complete graph representing the transportation network with the node set $V = 0 \cup P^0 \cup P$ and the arc set A, where 0 denotes the virtual depot where each bus starts and finishes its tour. $P^0 = \{P_1^0, \ldots P_i^0, \ldots, P_{n_s}^0\}$ denotes the set of students to be served, in which P_i^0 denotes the set of students' demand to be delivered to school i. $P = P^+ \cup P^-$ denotes the nodes after school bus stop location generation, P^+ denotes the set of bus stops, $P^- = \{1, .., i, \ldots, n_s\}$ denotes the set of schools.

Table 3 shows the parameters and decision variables for the school bus stop location and routing problem with walking accessibility and mixed load.

C. BASIC CONCEPT

1) WALKING ACCESSIBILITY

Since the individual differences are considered in walking accessibility among students while locating school bus stops, we define the walking accessibility as the individual ability to access school bus service conceived as the distance between place of residence and assigned school bus stop and measured in time it takes on foot by each student. Note that the same distance may lead to different walking time durations due to the different walking velocities of students.

In this paper, in order to locate several school bus stops to cluster students to be served, standards should be developed to balance walking accessibilities of students who are assigned to the same bus stop. The standards setting should take both the fairness and efficiency into account. To this end, we set a maximum walking time for each student, based on which the maximum distance between residence to bus stop can be calculated to be the constraint considered in the stop location process. The constraint can be stated as follows: $Y_{si} \cdot l_{si}/v_s \leq WT_s, \forall s \in P^0, i \in P^+$, which represents the time needed to walk from place of residence to assigned stop must be no more than the maximum walking time for each student s.

2) STUDENT'S COMMUTE TIME

As mentioned previously, the objective of SBSLRP-WA-ML is to minimize the total commute time of all students, where a student's commute time represents his (her) total time consumption from place of residence to attending school. According to the service design of the school bus system, the student's commute time generally includes the time spent on walking from home to assigned stop and in-vehicle ride time. The total commute time of all students can be calculated based on the following formula.

$$TCT = \sum_{s\in P^0}\sum_{i\in P^+} Y_{si}\cdot l_{si}/v_s + \sum_{i,j\in P}\sum_{k\in K} X_{ijk}\cdot L_{ik}\cdot l_{ij}/v + t_i^s\cdot\sum_{i\in P}\sum_{k\in K} L_i^k + t_i^s \tag{1}$$

The left-hand side of the formula represents total commute time of all students, which equals the sum of four parts of the right-hand side of the formula from the perspective of the bus operation, including walking time from home to assigned stop (i.e., the first part), and ride time which can be divided into running time between any two nodes in vehicle (i.e., the second part), waiting time at stops due to the pick-up and drop-off of other students (i.e., the third part), and the pick-up or drop-off time by students themselves (i.e., the last part).

D. MATHEMATICAL FORMULATION

The SBSLRP-WA-ML can be formulated as a mixed integer nonlinear programming model as follows:

$$\min\ TCT = \sum_{s\in P^0}\sum_{i\in P^+} Y_{si}\cdot l_{si}/v_s + \sum_{i,j\in P}\sum_{k\in K} X_{ijk}\cdot L_{ik}\cdot l_{ij}/v + t_i^s\cdot\sum_{i\in P}\sum_{k\in K} L_i^k + t_i^s \tag{2}$$

Subject to:

$$l_{ij} = \begin{cases} 0, & i = 0 \text{ or } j = 0 \\ \left|CX_i - CX_j\right| + \left|CY_i - CY_j\right|, & \text{otherwise} \end{cases} \tag{3}$$

$$Y_{si}\cdot l_{si}/v_s \leq WT_s, \quad \forall s\in P^0,\ i\in P^+ \tag{4}$$

$$\sum_{i\in P^+} Y_{si} = 1, \quad \forall s\in P^0 \tag{5}$$

$$\sum_{s\in P^0} Y_{si} = d_i \leq C, \quad \forall i\in P^+ \tag{6}$$

$$\sum_{i\in 0\cup P^+}\sum_{k\in K} X_{ijk} = 1, \quad \forall j\in P^+ \tag{7}$$

$$\sum_{j\in P^+}\sum_{k\in K} X_{ijk} = 1, \quad \forall i\in P^+ \tag{8}$$

$$\sum_{i\in P} X_{ijk} = \sum_{h\in P} X_{jhk}, \quad \forall j\in P,\ k\in K \tag{9}$$

$$Y_{si}\cdot\sum_{h\in P^+} X_{hik} = Y_{si}\cdot\sum_{g\in P} X_{gjk}, \quad \forall s\in P_j^0,\ i\in P^+,\ j\in P^-,\ k\in K \tag{10}$$

$$X_{ijk}\cdot L_{ik} + X_{ijk}\cdot d_j \leq L_{jk} \leq C, \quad \forall i,\ j\in P^+,\ k\in K \tag{11}$$

$$\sum_{i\in P^+} X_{0ik} = \sum_{j\in P^-} X_{j0k} = 1, \quad \forall k\in K \tag{12}$$

$$D_j^k = \sum_{i\in P^+}((\sum_{s\in P_j^0} Y_{si})\cdot\sum_{h\in P} X_{ihk}), \quad \forall k\in K,\ j\in P^- \tag{13}$$

$$D_j = \sum_{s\in P_j^0}\sum_{i\in P^+} Y_{si} = \sum_{k\in K} D_j^k, \quad \forall j\in P^- \tag{14}$$

$$X_{ijk}\cdot L_{jk} = X_{ijk}\cdot L_{ik} - X_{ijk}\cdot D_j^k, \quad \forall i\in P,\ j\in P^-,\ k\in K \tag{15}$$

$$X_{ijk}\cdot a_{ik} + X_{ijk}\cdot t_i^s + X_{ijk}\cdot l_{ij}/v \leq X_{ijk}\cdot a_{jk}, \quad \forall i,\ j\in P,\ k\in K \tag{16}$$

$$a_{ik} + t_i^s + l_{ij}/v \leq a_{jk} + M(1 - X_{ijk}), \quad \forall i,\ j\in P,\ k\in K \tag{17}$$

$$\sum_{i\in P}\sum_{k\in K} X_{0ik} = \sum_{j\in P}\sum_{k\in K} X_{j0k} \leq n_k \tag{18}$$

$$X_{ijk} = 0 \ or\ 1, \quad \forall i,\ j\in P,\ k\in K \tag{19}$$

$$Y_{si} = 0 \ or\ 1, \quad \forall s\in P^0,\ i\in P^+ \tag{20}$$

The objective function (2) minimizes the total commute time from home to school of all students. The constraints considered in the model are interpreted as follows. Constraint (3) represents the Manhattan distance is used to measure the travel distance between any two nodes to better accord with practical road network in school district, which is a symmetric network. Constraint (4) indicates the walking accessibilities of student with respective to the maximum walking time. Constraint (5) ensures that each student is assigned to exactly one stop. Constraint (6) indicates the demand of each stop consists of all students assigned to the stop, and guarantees the amount of demand is no more than bus capacity. Constraint (7) and (8) ensure that each stop is met with exactly once by one bus. Constraint (9) ensures the flow conservation. Constraint (10) guarantees that each student is transported to his (her) school. Constraint (11) states the vehicle capacity. Constraint (12) indicates that all buses start from virtual depot and visiting stops, and end at the virtual depot after servicing schools. Constraint (13) calculates the number of students assigned to stops visited by vehicle k and attending school j. Constraint (14) ensures that all students from each school are assigned to related stops and served by active buses. Constraint (15) indicates that students attending school j get off when bus arrives at the school. Constraints (13), (14), and (15) together ensure that each bus starts from picking up students, ends after serving all schools, and respects the precedence relationship of all stops and all schools. Constraints (16) and (17) compute the tour time of each bus and ensure subtour elimination. Constraint (18) ensures the buses used are no more than the given fleet size. Finally, constraints (19) and (20) ensure that the decision variables related to students' assignment and bus tours are binary.

Notice that different maximum walking time for students leads to the easy extension to school bus planning considering special need. Other than walking accessibility, much of what distinguishes special-education students from general students mainly lies in the need for equipment or accompany in buses, which only needs slight modification of constraints related to vehicle load and capacity to consider various types of capacity or loads associated with equipment or accompany.

IV. SOLUTION APPROACH

The proposed problem integrates the standalone school bus routing subproblem which is similar to vehicle routing problem, a well-known NP-hard problem, and the school bus stop location subproblem, making the problem more sophisticated in terms of space and time. In order to solve the problem in an effective way, the paper presents a two-stage solution approach.

The overall structure of the solution approach is stop location at first stage and routing at second stage. The results of stage 1 affect that of stage 2 in the way that the number and locations of stops and the number of students assigned to each stop directly influence the routing process. Notice that the objective of the problem, namely total commute time of all students, includes walking time from residence to stop which is determined at stage 1 and in-vehicle ride time at stops directly determined by bus routes generated at stage 2. Thus, the interrelation and interact between two stages must be considered at each stage to jointly optimize the walking time and ride time, and finally commute time of all students.

In view of this, we develop an iterative clustering method based on k-means and DBSCAN at stage 1 to determine the number and locations of stops and students' allocation subject to constraint of walking accessibilities of students. The consideration of interrelationship between stages is represented as the objective to minimize the number of generated stops at first stage described in Section 4A.

Due to some practical limitations in stopping restrictions, inaccessible locations may exist where school bus stop is located. The selected stop locations can be determined through negotiation among related authorities due to the particularity and importance of the students being considered and the resulting priority of school bus planning. In case that such negotiation fails, the inaccessible stop locations may be adjusted according to the practical restrictions with minor change in the assigned students due to the dense road network.

After determining the locations of school bus stops, the bus routes are generated and optimized based on the stop locations at stage 2. As described in Section 4B, an improved ant colony optimization routing algorithm is developed taking the interact with stage 1 into consideration which is represented by the stop assignment information, school attraction information and heuristic information of stop and school location during route construction.

A. SCHOOL BUS STOP LOCATION WITH WALKING ACCESSIBILITY

As mentioned before, the objective of the proposed model is to minimize the total commute time of all students without violating various constraints, notably the maximum walking time. In the proposed model, each student has a different walking velocity, which brings great challenge to stop locating considering a great amount of different maximum walking distances. Therefore, to simplify the problem, a school-related average walking velocity is developed, which indicates that students from the same school are considered to walk at the same speed. Through the analysis of walking accessibility, the maximum walking time constraint is replaced with different maximum walking distances for students from distinct schools. Subsequently, the maximum walking distances are used as the neighborhood size of calculate the neighborhood density based on which to determine the adopted to generate bus stops.

1) WALKING ACCESSIBILITY

This paper considers school bus system design in the background of educational stages varying from pre-school to high school in China. Based on the age distribution of students at different grades of different schools, we first review walking velocities of various ages, and calculate the average walking velocity of each common age and each school. Afterwards, the upper bound on walking distance is calculated based on the given upper bound on walking time.

The common age distribution of education stage is described as follows: Let $S = PS \cup ES \cup JS \cup HS$ be a set of schools of different education stages including pre-school (PS), elementary school (ES), junior school (JS) and high school (HS). $A_S = A_{PS} \cup A_{ES} \cup A_{JS} \cup A_{HS}$ denotes a set of age ranges of distinct schools, while $A_{PS} = \{3, 4, 5\}$, $A_{ES} = \{6, 7, 8, 9, 10, 11\}$, $A_{JS} = \{12, 13, 14\}$ and $A_{HS} = \{15, 16, 17\}$ denote the age ranges of students from PS, ES, JS, and HS, respectively, according to the Compulsory Education Law of the People's Republic of China and the actual conditions in most schools. Regardless of change of students such as transfer, wastage, etc., the number of students in each grade in a district is the same denoted by N_g, which means $N_i = N_g, i \in A_s, \forall s \in S$ while N_i denotes the number of students at the age of i years.

Gait development is becoming mature along with the growing age of students, and walking speed changes accordingly [32]. Let v_w^i denote the average velocity of students at the age of i $(i \in A_s, \forall s \in S)$ years. The average walking speed of students from each school can be calculated by Eq. 21.

$$\begin{aligned} v_w^s &= (\sum_{i \in A_s} v_w^i \cdot N_i)/N_s \\ &= (\sum_{i \in A_s} v_w^i \cdot N_g)/(|A_s| \cdot N_g) = \sum_{i \in A_s} v_w^i/|A_s|, \quad \forall s \in S \quad (21) \end{aligned}$$

TABLE 4. Average walking speed of each age involved in this study.

Ages	Average Walking Speed (cm/s)
3	98.5
4	101.54
5	103.38
6	104.49
7	104.99
8	106.53
9	106
10	107.12
11	108.15
12	107.1
13	108.5
14	108.4
15	110
16	130
17	130

Teenagers' walking speed varies from 0.11 m/s to 2.56m/s [33]. The average walking speed of each age involved in this study is calculated through literature study and shown in Table 4 [34]–[37]. Note that there are limited studies about walking velocities of teenagers aged 16 to 17, therefore the average walking speeds of 16-17 age students take the adult walking speed, 1.3 m/s, because the gait development levels off since age 15 [38].

According to the average walking speed at each age, the average walking velocity of students from each school, v_w^s, is calculated by Eq. 21 and shown as follows: $v_w^{PS} = 101.14cm/s$, $v_w^{ES} = 106.21cm/s$, $v_w^{JS} = 108cm/s$, $v_w^{HS} = 123.33cm/s$.

Let $L_w^s = v_w^s \cdot T_w^s, \forall s \in S$ represent the maximum walking distance of students attending each school. Suppose that maximum walking time T_w^s is the identical for all students and equals 10 minutes, and the maximum walking distances for students attending each school are calculated and shown as follows: $L_w^{PS} = 606m$, $L_w^{ES} = 637m$, $L_w^{JS} = 648m$, $L_w^{HS} = 740m$.

2) PROPOSED ITERATIVE CLUSTERING METHOD BASED ON K-MEANS AND DBSCAN

This paper proposes an iterative clustering method based on k-means and DBSCAN to locate a set of stops covering all students in need of school bus within the school district. The basic idea of the proposed method for school bus stops generation lays in the application of iterative k-means clustering with certain initial inputs to determine the location of stops. The inputs including the number of centroids, namely K, and initial locations of cluster centroids are determined by DBSCAN or simple calculation depending on students' distribution. Constraint related to walking accessibility is used to calibrate students' distribution and test the feasibility of solutions for stops location. The objective to minimize the number of bus stops is realized in the way that increasing the number of centroids from initial input as few as possible with the progress of the iteration until constraints for all students are satisfied. The overall structure of the method is motivated by the fact that the density-connected points are clustered into a group by DBSCAN leading to the fewest possible number of clusters considering various walking accessibilities, and k-means clustering locates certain number of cluster centroids.

The decision about method used to obtain the initial inputs for iterative k-means clustering is described as follows in combination with how to run the algorithm. If the densities are too low such that the maximum density is no more than the minimum one which means the overly scattered students' distribution, the initial K is determined by the area of the school district divided by the area of a circle whose radius equals the maximum walking accessibility, and the K centroids are generated uniformly in the school district. With these inputs, k-means clustering is iteratively executed with the increase of the value of K. Otherwise, the integration of DBSCAN and iterative k-means with inputs determined by DBSCAN becomes a process being iteratively executed for several times depending on the neighborhood densities of students details of which can found in the process of method shown later.

The notations used in the algorithm are explained as follows: s represents each school. $Eps(s)$ denotes maximum walking distance used as the radius of a neighborhood for students attending school s. x_i^s denotes each student attending school s. $density(x_i^s)$ denotes the density of student x_i^s of $Eps(s)$-neighborhood. $MinPts$ is a parameter used in DBSCAN as the minimum number of points to form a dense region.

The process of the method is described as follows:

Step 0: Initialize the sets $NumStop$ and $Stop$ to be empty sets, which represent the number and locations of stops generated during iteration.

Step 1: Calculate of $density(x_i^s)$ for all students. And let $MINMinPts$ equal $\max\{\min(density(x_i^s)), s \in S\}$, and $MAXMinPts$ equal $\min\{\max(density(x_i^s)), s \in S\}$.

Step 2: Check if $MAXMinPts > MINMinPts$. If yes, go to Step 3, otherwise, go to Step 4.

Step 3: Calculate the initial K based on the area of the district and the maximum walking accessibility. And randomly generate K uniformly distributed points representing the initial K centroids. Then, go to Step 6.

Step 4: Check if $MAXMinPts$ is larger than the bus capacity, if yes, let $MAXMinPts$ equal the bus capacity. Then, go to Step 5.

Step 5: Select $MinPts$ from the [$MINMinPts$, $MAXMinPts$] and apply the DBSCAN on the students of each school.

Step 5.1: Cluster students of each school by DBSCAN with parameters $Eps(s)$ and $MinPts$.

Step 5.2: Let NC represent the number of clusters obtained by DBSCAN of all schools. Rearrange all students from all schools into a set X in the order border objects, core objects clustered by DBSCSAN, and the noises into a set $NOISE$. Let the first NC objects be the initial K centroids in the next stop. Then go to Step 6.

Step 6: Apply the k-means clustering with the initial K centroids iteratively with the increase of K until all students are assigned to stops within their walking accessibilities.

Step 6.1: Cluster all students to K clusters by k-means clustering with given initial cluster centroids.

Step 6.2: Check whether there is no less than one stop within walking accessibility for students belonging to set $NOISE$. If yes, go to Step 6.3; otherwise, add the location of students to the cluster centroids, and add the number of the students to the value of K.

Step 6.3: Check whether there is at least one stop within walking accessibility for each student. If yes, set number K and the K clusters into the sets *NumStop* and *Stop* respectively, and go to Step 7; otherwise, $K = K + 1$, go back to Step 6.1.

Step 7: Choose the minimum value from the set *NumStop* and the corresponding stop locations in the set *Stop* to be the number of stops and the selected stop locations. Assign each student to its nearest stop.

B. AN IMPROVED ANT COLONY OPTIMIZATION ALGORITHM

Ant Colony Optimization algorithm (ACO) has been widely used in combinatorial optimization problems such as vehicle routing problem, which is similar to the school bus routing problem and proven to be effective since its introduction in 1991 inspired by ants' foraging behavior in real world [39]–[45]. In addition to the population-based characteristic which can expand the search space, the transit rule in ACO approaches to the nature of selecting next stop or school to visit in school bus routing. Based on these considerations, we modify and improve the ACO to adapt to the specific characteristics of the proposed problem.

The application of a general ACO algorithm to the school bus routing problem can be considered as a process of artificial ants (buses) finding a set of routes with minimal cost (the total commute time of all students) traveling through all nodes (stops and schools).

According to the characteristics of the school bus locating and routing problem with walking accessibility and mixed load, there are two main changes while the implementation of the ant colony algorithm on the proposed problem.

Firstly, we improved the transit rule by applying different rules for the selection of bus stops and the schools. The improved pseudorandom proportional rule for selecting school bus stops considering stop assignment information, represented by total walking time of assigned students from their places of residence to the stop, in addition to the pheromone information and heuristic information of stop and school locations. As for the selection of the school to visit while generating the route, the number of in-vehicle students who attending the school being considered as school attraction information, together with pheromone information and heuristic information of stop and school locations are the determinants of the sequence of schools to visit for each bus tour. Secondly, in order to intensify the solution found in the above IACO routing algorithm, we apply two local search operators to the best solution after each iteration to explore the neighborhood of the best solution.

The features of the proposed IACO algorithm are described as follows:

- The number of ants, m, is equal to the number of bus stops, n.
- m ants are randomly placed in bus stops in the initialization phase.
- The transition rule used in the route generation is improved based on the feature of the problem, considering the transit rule to select next bus stop and the transit rule to select next school separately.
- The pheromones are updated globally considering evaporation and using the best-so-far solution after each iteration, while the best-so-far ant means the bus route scheme producing the best objective value since the beginning of the algorithm.
- After each iteration, an inter-route relocation operator and an inter-route exchange operator are applied as local search operators to improve the solution.

1) IMPORTANT CONCEPTS

- *Stop assignment information*

The stop assignment information is used together with pheromone information and heuristic information of stop and school locations to select next stop to visit while generating route.

Denoted by ω_j, the stop assignment information is represented by total walking time of assigned students from their places of residence to the stop j. The relative importance of the stop assignment information is denoted by γ. The contribution of the information to the selection of next stop to visit is represented by total walking time of allocated students for stop j to the power of the relative importance of the information, and denoted by $[\omega_j]^\gamma$.

- *School attraction information*

The school attraction information is used together with pheromone information and heuristic information of school and stop locations to select next school to visit after all stops have been served by a school bus while constructing route.

Denoted by φ_t^k, the school attraction information is represented by the number of students to be delivered to school t by current bus of ant k. The relative importance of the school attraction information is denoted by θ. The contribution of the information to the selection of next school to visit is calculated by the number of in-vehicle k students who attending school t to the power of the relative importance of the information, and denoted by $[\varphi_t^k]^\theta$.

2) ROUTE GENERATION

- *Initialization*

As mentioned above, an ant is placed randomly at a stop in the initial. In order to formulate a set of routes, the ant must travel through all nodes representing stops just once and the schools for several times.

- *Transit rule of selecting next bus stop*

After the initialization, the ant k chooses the next stop to visit by an improved pseudorandom proportional rule shown in Eq. 22 and Eq. 23 until the bus capacity is violated.

$$j = \begin{cases} \arg\max\limits_{j \in N_i^k} \left\{ [\tau_{ij}]^{\alpha}[\eta_{ij}]^{\beta}[\omega_j]^{\gamma} \right\}, & \text{if } q \leq q_0 \\ J, & \text{otherwise} \end{cases} \tag{22}$$

$$p_{ij}^k = \begin{cases} \dfrac{[\tau_{ij}]^{\alpha}[\eta_{ij}]^{\beta}[\omega_j]^{\gamma}}{\sum\limits_{j \in N_i^k} [\tau_{ij}]^{\alpha}[\eta_{ij}]^{\beta}[\omega_j]^{\gamma}}, & \text{if } j \in N_i^k \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

Therein, N_i^k denotes the set of feasible stops of ant k after visiting stop i, which means stops that have not yet been visited by any bus of ant. τ_{ij} and η_{ij} denotes the pheromone information and heuristic information of stop and school locations measured by the reciprocal of the riding time on arc $i \rightarrow j$, respectively. ω_j denotes the stop assignment information as mentioned before. α, β, and γ denote the relative importance of different information.

The balance between the exploration and exploitation is determined by a given constant $q_0 (0 \leq q_0 \leq 1)$. q is randomly generated in the range [0, 1] during the algorithm execution. If $q \leq q_0$, the choice of next stop j to be visited by ant k is determined by Eq. 22, otherwise, Eq. 23 is working.

- *Transit rule of selecting next school*

After all nodes representing bus stops have been visited by a bus represented by an ant, students assigned to these stops need to be delivered to the corresponding attending schools. Since the allowable mixed loading leads to the situation where every bus may serve students attending different schools. Therefore, the sequence of schools to visit should be determined for each school separately as follows.

The bus tour traveling through stops is generated by flipping the ant tour to ensure that the stops with more waiting time are served first. Different from the selection of stops, considering the dense distribution and small number of the schools, the random proportional rule as shown in Eq. 24 is used to decide the next school t to visit. After any stop visited by exactly one bus, a new ant is initialized. The initialization of a new ant and the selection of next stop and school is looped until all stops are visited.

$$t = \arg\max_{t \in N_s^k} \left\{ [\tau_{st}]^{\alpha}[\eta_{st}]^{\beta}[\varphi_t^k]^{\theta} \right\} \tag{24}$$

Therein, N_s^k means schools has not yet been visited by bus tour currently being considered of ant k. The meanings of τ_{ij} and η_{ij} are the same as in the transit rule to select next bus stop. φ_t^k denotes the school attraction information as mentioned before. α, β, and θ denote the relative importance of different information.

```
procedure LocalSearchImprovement
    input: origional solution, i. e., best solution obtained at each IACO iteration
    output: improved solution
    s ⟵ origional solution
    ŝ ⟵ InterrouteRelocation(s)
    ŝ ⟵ InterrouteExchange(ŝ)
    s_best ⟵ ŝ
    return s_best
end – procedure
```

FIGURE 2. Pseudocode for local search improvement.

3) LOCAL SEARCH IMPROVEMENT

The ant colony optimization algorithm described above improves the intra-route solution through the pheromone updating and the rule to choose next stop to complete each tour, which leads to the possibility of cost savings achieved by inter-route improvement. After each iteration, an ant will be chosen according to the roulette choice wheel-based method. The possibility of each ant to be chosen is calculated by the objective value divided by the sum of the objective value of all ants in the iteration. And the local search improvement is only applied for the stops visiting sequence of the solution as shown in Fig.2, and once better solution is generated, the school sequence is determined by applying the proportional rule based on the incumbent solution. In addition, we apply two local search operators to the best solution found at each iteration to explore the neighborhood of the best solution and intensify the solution found in the above IACO routing algorithm.

It's worth noticing that the allowable mixed loading m during the execution of the local search operators, we first rearrange the sequence of the bus stops directly. Afterwards, the visiting sequence of schools in each tour will be adjusted according to new sequence of bus stops to visit.

The neighborhood structures adopted in the local search operators include inter-route relocation and inter-route exchange as shown in Fig.3 and Fig.4, respectively. The circles represent the school bus stops while the squares stand for schools to visit which students assigned to each route attend. Taking two same tours to show how the inter-route relocation and exchange operator work, we suppose that there are 4 stops both in original tour 1 and tour 2 as shown in Fig.3 (a) and Fig.4 (a) where the circle represents the stops and the rectangle represents the schools. Schools related to the stops are described as follows: students assigned to stop 1, stop 7 and stop 9 are all from School 1, students assigned to stop 2 are all from School 2, students assigned to stop 4 are all from School 3, while students at stop 3, stop 6 and stop 8 attend different schools including School 1 and School 2. We suppose that the tours are feasible to perform the local search improvement, which means the demand of each stop along the tours will not violate the capacity constraint. During the execution of the local search, the concept "saving" will be used to measure the quality of the execution of the local search, and it is calculated by the sum of the student

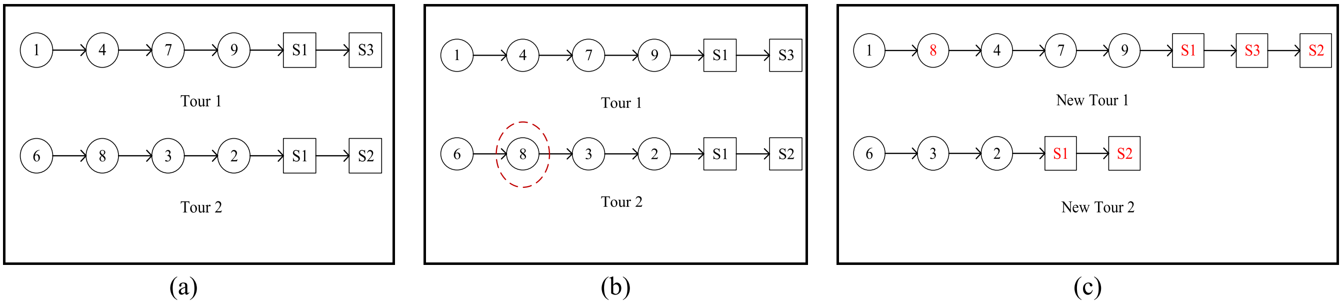


FIGURE 3. Sample of inter-route relocation operator (a) the original two tours, (b) select a stop to apply the inter-route relocation, (c) the new tow tours after relocation.

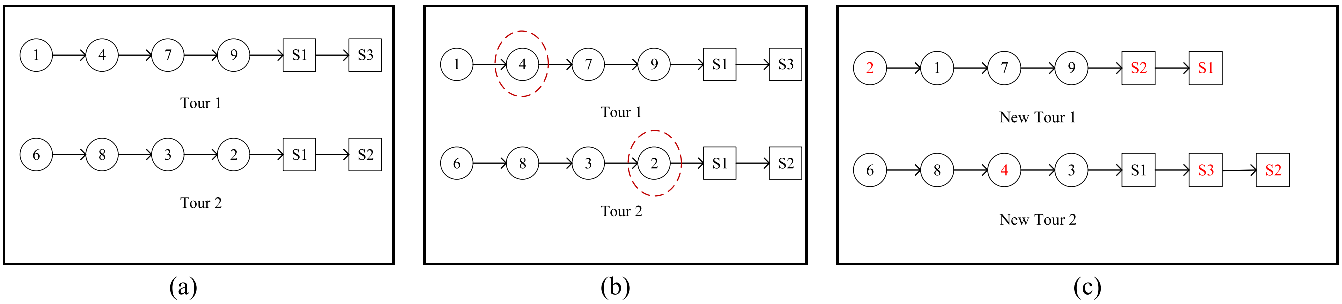


FIGURE 4. Sample of inter-route exchange operator (a) the original two tours, (b) select a stop to apply the inter-route exchange, (c) the new tow tours after exchange.

commute time of the original tours minus the sum of the student commute time of the new ones.

During the execution of the inter-route relocation operator, two original tours shown in Fig.3 (a) are first randomly selected to conduct the local search. Stop 8 in tour 2 is randomly selected to perform the relocation operator as shown in Fig.3 (b). The result of the relocation is shown in Fig.3 (c) where stop 8 is relocated in tour 1 at its best location with the maximum savings and the sequence of visited schools are rearranged according to the relocation.

As for the exchange operator, in Fig.4 (b), stop 4 and stop 2 are randomly selected from tours respectively. Taking the generation of new tour 1 as example, stop 2 from original tour 1 is inserted in the best position with the maximum savings in tour 1 without violating constraints, and the school sequence is rearranged based on new stops sequence. Afterwards, the new tour 1 is generated completely shown in Fig.4 (c).

The new tours replace the original ones only if the relocation or the exchange operator produces better solution, which means less traveling time for all students.

4) PHEROMONE UPDATE

The updating of pheromone trails is conducted after all ants constructing the routes at each iteration, including the decrease due to the evaporation and the increase of the best-so-far ant, which is calculated by Eq. 25 and Eq. 26. Notice that the pheromone trails are only updated on the link among bus stops or between bus stops and schools, which means the pheromone on the links among schools remains unchanged. The reason lies in the fact that the concentrated distribution of the schools leads to the little difference of sequence of schools to visit while it is mainly decided by the number of students in bus who attend each school.

$$\tau_{ij} = (1-\rho)\cdot\tau_{ij} + \rho\cdot\Delta\tau_{ij}^{bs} \tag{25}$$

$$\Delta\tau_{ij}^{bs} = \begin{cases} L_{i,bs}/F^{bs}, & if\ (i,j)\in T^{bs},\ j\in P^{+} \\ 0, & otherwise \end{cases} \tag{26}$$

where ρ determines the speed of evaporation. $\Delta\tau_{ij}^{bs}$ denotes the increase of pheromone produced by the best-so-far solution T^{bs}. $L_{i,bs}$ denotes the bus load on the link (i,j) of T^{bs}. F^{bs} denotes the objective function value of T^{bs}.

V. COMPUTATIONAL EXPERIMENT AND DISCUSSION

This section summarizes the computational experiments on the proposed two-stage algorithm described above. In Section 5A, we outline the generation principle of instances followed by the experimental results of the execution of the proposed algorithm on the instances in Section 5B. In addition, an analysis of influence factors of total commute time of students and comparison to door-to-door school bus service are presented in Section 5C and 5D, respectively. Finally, the proposed solution approach is analyzed through the comparison to similar algorithms and sensitivity analysis to the parameters in Section 5E.

A. INSTANCE GENERATION

Considering that the existing benchmark instances pay more attention to school bus routing, they usually just state demand for each bus stop and neglect students' locations, or generate students' locations based on the stops' locations. This is not in accordance with the proposed problem, thus we create the benchmark representing the problem.

To test the performance of the proposed algorithm, we generate 42 instances of different problem sizes shown in Columns 1 through 6 in Table 5. The numbers in Column 1 indicates the instance ID. The number of students ranges from 75 to 900 as shown in Column 2 and the area of the school district varies from 3 * 3 km^2 to 7 * 7 km^2 as shown in Column 3 denoted by the side length of the district. Column 4 shows the student density of the district which is the number of students divided by the area of the district. We consider a homogenous fleet in each instance with the capacity of the buses 25 or 50 as shown in Column 5 and fleet size shown in Column 6.

Considering the divisions of school districts, students usually attend near-by schools. We assume a uniform distribution of students in a school district since the nature of land use in a school district tend to be the same. We randomly generate a certain number of uniformly distributed students in a two-dimensional plane with different side length measured by abscissa and ordinate range representing the school district of a certain area. Then the schools are located randomly within the central area of the school district. Suppose that there are four schools where each of them is at a different educational stage, and the number of students at each grade is identical, the ratio of the number of student attending each school is 1: 2: 1: 1.

Before the execution of the proposed algorithm, the eligibility to take school buses should be checked for all students. The criteria is that the students within their walking accessibility as stated in section *Walking accessibility* are not considered in the design of school bus system, which leads to students within the walking distance to his (her) school being not considered in the school bus demand, as indicated in column 2 and 7, where column 2 shows the number of students attending all schools in the school district, and column 7 shows the amount of students in need for the service of school bus after the eligibility certification based on walking accessibility.

B. COMPUTAIONAL RESULTS

Since the parameter sensitivity is analyzed later, referring to the successful application of ACO to the SBRP [27]–[31], the parameters of the algorithm are set as follows: $\alpha = 1$, $\beta = 2$, $\gamma = 2$, $\theta = 2$, $\rho = 0.9$, $q_0 = 0.9$, $v = 30km/h$, $\tau_0 = 1$, $m = n$, $I = 500$, where I represents the maximum number of iterations. The number of ants, m, in each iteration is set to be the number of bus stops generated at stop location stage, n. The stops and schools are taken into consideration separately when it comes to service duration at each node. The pickup duration at each stop (in seconds) is calculated by $t_i^s = 19.0 + 2.6 \times d_i, \forall i \in P^+$ while the formula $t_j^s = 29.0 + 1.9 \times d_j, \forall j \in P^0$ is adopted to determine the drop off duration at each school [11]. Notice that each school can be visited by many buses with different number of students d_j to drop off, leading to different drop off duration of different buses at a same school.

The experiment is conducted in an Intel (R) Core (TM) i5-4590 CPU @ 3.30 GHz desktop computer with 8 GB RAM. The proposed stop location method cluster students into a set of stops and students assigned to these stops. In order to calibrate the accuracy and robustness of the routing problem, the IACO routing algorithm was executed for 10 times for each instance as shown in Table 5. Columns 1 through 7 show the basic information of each instance as mentioned above. And Column 8 through 11 show the average result of multiple runs, including the number of stops in Column 8, the objective value, total commute time of all students, in Column 9, and run time in Column 11. Column 10 shows the average commute time of students which is the total commute time of all students divided by demand quantity, i.e. Column 9 divided by Column 7.

C. INFLUENCE FACTOR ANALYSIS OF TCT

From the computational results, we will analyze the influential factors in the total commute time of students in this section. Considering the overall process of taking school buses, the potential influence factors may include the bus capacity, the number of students, and the area of the district. The bus capacity is analyzed because it affects the number of bus stops according to the concentration degree of students and the number of stops to visit for each bus which could affect the ride time and stopover time of buses. As for the other potential factors, since the number of students directly affects the total commute time of all student, and there are major differences in the area of school district and the number of students among different instances, it appears improper to analyze the influence of the number of students or the area of the school district independently. Therefore, we transform the number of students and the area of the school district into a single factor, the student density, and analyze its effect on the average commute time of all students who taking school buses, i.e., the ACT.

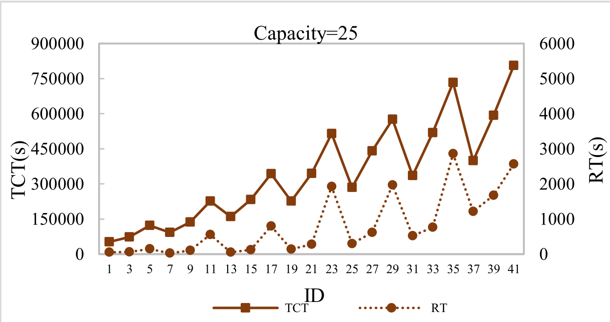
Fig.5 shows the comparison result of instances with only a difference of capacity. The horizontal axis indicates the ID of the corresponding instance. The primary and second vertical axes indicate the total commute time and run time at the routing stage of the instance. One can see that instances with only difference in bus capacity share similar trend with the increase of the number of students and the change of the area of the school district. The major difference resulting from the distinct capacity is that the total commute time of all students with capacity 50 is almost always larger than the TCT of instance with capacity 25, while there is little difference between RT. The reason is that buses with lower capacity could travel through fewer stops to arrive at schools leading to less time for all students. However, buses with lower capacity mean more buses for the same number of students, so it is the times for speed as opposite to instances with larger capacity. Since the number of students is identical,

TABLE 5. Computational results of SBSLRP-WA-ML for instances at different sizes.

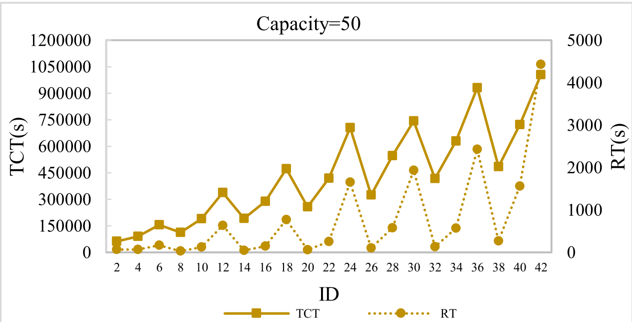
ID	N	District	StD (/km^2)	Capacity	n_k	Demand	n	TCT(s)	ACT (s)	RT (s)
1	75	3	8.33	25	4	67	36	52262.06	780.03	57.56
2	75	3	8.33	50	3	67	36	61876.08	923.52	67.24
3	75	5	3	25	4	72	36	73630.56	1022.65	58.75
4	75	5	3	50	3	72	36	90402.44	1255.59	67.31
5	75	7	1.53	25	4	75	56	122130.80	1628.41	145.67
6	75	7	1.53	50	3	75	56	155061.30	2067.48	165.15
7	150	3	16.67	25	7	142	23	92813.49	653.62	25.20
8	150	3	16.67	50	4	142	23	111699.80	786.62	26.19
9	150	5	6	25	7	144	51	136808.00	950.06	105.27
10	150	5	6	50	4	144	51	190006.20	1319.49	120.65
11	150	7	3.06	25	7	147	108	226343.10	1539.75	559.89
12	150	7	3.06	50	4	147	108	337497.40	2295.90	631.37
13	300	3	33.33	25	18	270	21	160254.50	593.54	51.20
14	300	3	33.33	50	7	270	21	191540.70	709.41	41.41
15	300	5	12	25	14	286	50	232646.30	813.45	121.02
16	300	5	12	50	7	286	50	287103.50	1003.86	143.05
17	300	7	6.12	25	13	295	107	343058.50	1162.91	796.18
18	300	7	6.12	50	7	295	107	471488.40	1598.27	771.44
19	450	3	50	25	24	409	27	226883.80	554.73	144.12
20	450	3	50	50	11	409	23	256855.60	628.01	55.49
21	450	5	18	25	23	436	54	344094.20	789.21	278.42
22	450	5	18	50	10	436	54	418182.30	959.13	253.93
23	450	7	9.18	25	19	444	165	514550.00	1158.90	1921.48
24	450	7	9.18	50	10	444	165	703935.00	1585.44	1652.87
25	600	3	66.67	25	32	551	36	285192.00	517.59	296.79
26	600	3	66.67	50	18	551	27	324927.00	589.70	97.87
27	600	5	24	25	32	584	69	440455.70	754.21	615.97
28	600	5	24	50	13	584	69	545533.30	934.13	573.97
29	600	7	12.24	25	27	593	111	576300.10	971.84	1972.07
30	600	7	12.24	50	13	593	111	743224.60	1253.33	1928.32
31	750	3	83.33	25	47	683	50	335885.40	491.78	519.18
32	750	3	83.33	50	20	683	22	416755.70	610.18	130.88
33	750	5	30	25	46	716	58	519079.90	724.97	763.44
34	750	5	30	50	16	716	58	628387.40	877.64	570.85
35	750	7	15.31	25	31	740	144	732413.60	989.75	2856.59
36	750	7	15.31	50	16	740	144	930262.70	1257.11	2428.68
37	900	3	100	25	54	825	63	398671.00	483.24	1217.66
38	900	3	100	50	24	825	26	484567.80	587.35	266.02
39	900	5	36	25	55	863	74	593248.20	687.43	1669.92
40	900	5	36	50	19	863	74	722214.30	836.86	1561.21
41	900	7	18.37	25	42	882	136	805494.80	913.26	2562.32
42	900	7	18.37	50	19	882	136	1003756.00	1148.46	4431.68

the result of the exchange of the speed and the times will arrive at a same level.

Fig.6 shows the trend of the ACT with the increase of student density. The gray dash-dot line with rectangle markers



(a)



(b)

FIGURE 5. Influence of bus capacity on TCT (a) trend of TCT when capacity = 25, (b) trend of TCT when capacity = 50.

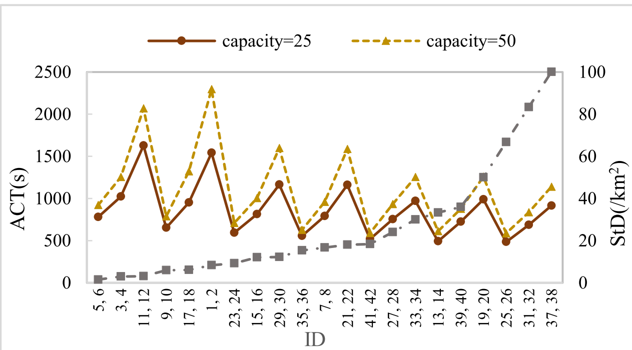


FIGURE 6. Influence of student density on TCT.

shows the corresponding student densities of the instances with ID number in the horizontal axis. The line charts show the change of the ACT with the increase of the student density in the instances with different bus capacities. Instances with capacity 25 is illustrated in the red line with the odd ID numbers shown in the horizontal axis corresponding to the markers, while the golden line represents the instances with capacity 50 with even instance ID numbers. From the figure, we can see that the average student commute time tends to decrease with the increase of the student density. The reason is that the dense distribution of students indicates that the bus stops, students' home locations, schools are all closer than dispersed distribution, which leads to the less travel time among stops and schools and less walking time from place of residence to schools. In spite of the overall downward trend of ACT, the decrease speed fluctuates when the density is no more than a certain value. Comparing to the neighborhood density in the solution section one can see that district with the student density at turning point has almost the same walking accessibility-neighborhood density. In other words, the walking accessibility neighborhood density of the school district at turning point in the line chart is almost as much as corresponding bus capacity. When the student distribution is less concentrated, more stops are needed to satisfy the walking accessibility constraint, which leads to increased bus tour time and stopover service time at stops, and resulting increased the average commute time. As the density grows, the number of students within walking accessibility increases, which leads to the decrease of the number of stops and the less travel time and stopover time at stops of buses. Consequently, the average commute time of students tends to decrease. However, the randomness of the student distribution and mutual influence among different factors may lead to the fluctuation of the decrease speed of ACT. When the neighborhood density is larger than bus capacity, the number of bus stops depends mainly on the number of students divided by the bus capacity, and the stops almost locate at the center of groups of students the number of which is capacity. Then buses go to one stop to pick up students up to capacity and directly ride to related schools, and this corresponds approximately to the direct service for students from accessible bus stops. Since the capacity and walking accessibility remain unchanged, the average commute time of students tends to stabilize under this situation.

D. COMPARISON TO DOOR-TO-DOOR SBRP

The proposed method first cluster students to several stop clusters and dispatch school buses to visit and pick up students and deliver to their corresponding schools. However, there exists another direct service, Door-to-Door School Bus, that is, school buses visit each student's place of residence to pick up and deliver to their attending schools. In other words, there is no school bus stop to bridge the students' homes and schools, which resembles the Dial-A-Ride Problem (DARP). Obviously, the Door-to-Door School Bus service is more convenient than the school bus service with stops. In order to obtain a precise comparison, we treat the generated instances as a Door-to-Door School Bus Routing Problem (Door-to-Door SBRP) and solve it using the proposed method as mentioned before. The minor difference is that the school bus stops in the SBSLRP-WA-ML serve several students and the location and demand of the stops need to be determined before the routing stage, while in the Door-to-Door SBRP, every student's location can be deemed as a school bus stop with the demand equals one. In addition to the difference in the school bus stops, the other main distinction lies in the constitution of the total commute time of the students. In the SBSLRP-WA-ML, the total commute time include the walking time from home to assigned stop, the in-vehicle ride time, and the service time at stops. However, in Door-to-Door SBRP, the walking time for all students is zero since all students are picked up at homes. In the following, we compare them with the same solution method on the same instances.

Table 6 shows the computational results of SBSLRP-WA-ML and Door-to-Door SBRP. The first two columns show the basic information of the instances. The third through sixth and the seventh through ninth columns show the results of the SBSLRP-WA-ML and Door-to-Door SBRP, respectively. The last column shows the difference of the two service modes in terms of the average commute time of students in the corresponding instance. A negative value means the Door-to-Door SBRP provides better service with less average commute time for students than the SBSLRP-WA-ML, vice versa. As we can see, the positive values make up the majority of the last column, which suggests that in most instances SBSLRP-WA-ML provides better service than Door-to-Door SBRP in terms of average commute time of students. The comparison result is opposite to the intuitive impression. In order to explain the phenomenon, taking the case of the school buses with capacity 50, we present Fig. 7 to show the change of the difference in ACT of two service modes with the student density of the school district in the corresponding instance.

In Fig. 7, the horizontal axis represents the student density of the instances. The primary axis represents the average commute time of students of the two service modes which are denoted by red line and red dash-dot line, while the secondary axis represents the difference between the two modes denoted by the gray dash-dot line. The labels next to the markers of the gray dash-dot line represent the corresponding instance ID. From the figure one can see that Door-to-Door SBRP dominates when the student density is very low. However, when the student density increases, the SBSLRP-WA-ML gradually takes it over and the comparative advantage tends to stabilize. The reason for the advantage shifting phenomenon between the two service modes lies in the change of the difference between walking time from home to stop in the SBSLRP-WA-ML and the additional detour and service time at stops in the Door-to-Door SBRP with the student density. At the first dominance period of Door-to-Door SBRP, the walking time is too large comparing to the detour and service time at stops. When the district becomes dense, the opposite relationship shows, leading to the advantage of SBSLRP-WA-ML in terms of the average commute time of students.

As to the safety issue of the SBSLRP-WA-ML and Door-to-Door SBRP, the latter mode is relatively safe when providing school bus service. However, as mentioned before, Door-to-Door SBRP may become less convenient in terms of student commute time when the school district becomes dense. Therefore, SBSLRP-WA-ML mode is more suitable for dense school district from the aspect of safety and student commute time. Since we consider individual walking accessibility in the stop location stage, the activity ability related safety issue can be neglected. In addition, the more student density of the school district usually indicates more population density and better safety security. In conclusion, the SBSLRP-WA-ML is a better mode for providing convenient school bus service in dense area, while students' security can be guaranteed by the individual walking accessibility and good public order.

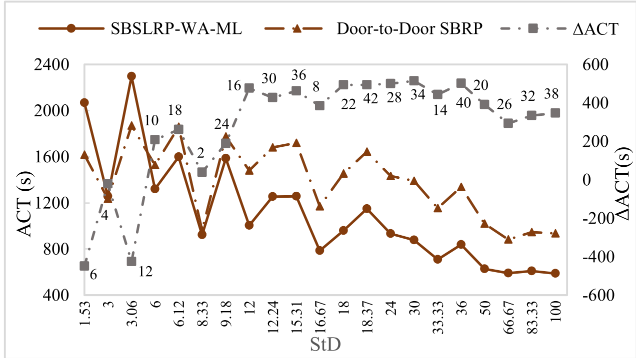


FIGURE 7. Change of difference between two service modes with student density.

E. ALGORITHM ANALYSIS

1) PERFORMANCE ANALYSIS

Considering the nonlinearity and complexity of the mathematical model of the problem in addition to its NP-hard characteristic, it is extraordinarily hard to solve it precisely, therefore we develop a two-stage heuristic solution method as mentioned before. In order to test the performance of the proposed method, we test and compare it with two other heuristic methods, Genetic Algorithm (GA), and Tabu Search (TS). It is worth noting that we only compare the methods in the second stage due to the inaccessibility to exact solution of the proposed routing method in the second stage and its similarity to GA and TS. In addition, GA and TS are the most frequently used methods in the school bus routing problem according to the latest review about SBRP [14].

We apply GA and TS to the same instances during the routing stage to obtain the results to analyze the algorithm performance. In order to enhance the comparability, we set the maximum number of iterations (generations) as the same number, 500. The initial solutions in GA and TS are generated randomly.

As shown in Table 7, the first column shows the instance ID. The second and third columns show the computational results of the proposed method including the objective value, TCT and the running time of the method, RT. The next four columns represent the results of GA in the routing stage while the last four columns represent the results of TS in the routing stage. Therein, Δ TCT represents the difference in total commute time of all students of different routing methods, which is calculated by the TCT obtained by GA or TS minus the TCT obtained by the proposed method. Δ TCT /TCT represents the proportion of the ΔTCT to the TCT obtained by the proposed method. As we can see, the proposed method dominates GA in terms of solution quality with acceptable increase in the run time. In some instances, the difference reaches more than 30 percent of the TCT of the proposed method. As to the comparison with TS, the solution quality

TABLE 6. Comparison of service quality of different service modes: SBSLRP-WA-ML and Door-to-Door SBRP.

ID	N	SBSLRP-WA-ML				Door-to-Door SBRP			
		Demand	TCT(s)	ACT (s)	RT (s)	TCT_{D2D}(s)	ACT_{D2D}(s)	RT_{D2D}(s)	ΔACT (s)
1	75	67	52262.06	780.03	57.56	57253.10	763.37	163.30	-16.66
2	75	67	61876.08	923.52	67.24	72223.06	962.97	183.52	39.45
3	75	72	73630.56	1022.65	58.75	75332.18	1004.43	143.63	-18.22
4	75	72	90402.44	1255.59	67.31	92604.83	1234.73	182.04	-20.86
5	75	75	122130.8	1628.41	145.67	98749.94	1316.67	194.91	-311.74
6	75	75	155061.3	2067.48	165.15	121348.20	1617.98	239.98	-449.50
7	150	142	92813.49	653.62	25.2	117480.78	783.21	238.21	129.59
8	150	142	111699.8	786.62	26.19	175581.61	1170.54	386.65	383.92
9	150	144	136808	950.06	105.27	152302.90	1015.35	434.83	65.29
10	150	144	190006.2	1319.49	120.65	229152.85	1527.69	584.40	208.20
11	150	147	226343.1	1539.75	559.89	187454.78	1249.70	841.78	-290.05
12	150	147	337497.4	2295.9	631.37	280655.23	1871.03	1020.02	-424.87
13	300	270	160254.5	593.54	51.2	184884.54	616.28	479.60	22.74
14	300	270	191540.7	709.41	41.41	345854.05	1152.85	717.41	443.44
15	300	286	232646.3	813.45	121.02	278198.58	927.33	1118.23	113.88
16	300	286	287103.5	1003.86	143.05	443900.56	1479.67	1369.42	475.81
17	300	295	343058.5	1162.91	796.18	358918.38	1196.39	2654.35	33.48
18	300	295	471488.4	1598.27	771.44	558080.03	1860.27	2760.21	262.00
19	450	409	226883.8	554.73	144.12	275081.09	611.29	1287.71	56.56
20	450	409	256855.6	628.01	55.49	458910.24	1019.80	1377.16	391.79
21	450	436	344094.2	789.21	278.42	383822.85	852.94	2571.16	63.73
22	450	436	418182.3	959.13	253.93	653779.46	1452.84	2386.28	493.71
23	450	444	514550	1158.9	1921.48	536263.28	1191.70	6404.55	32.80
24	450	444	703935	1585.44	1652.87	798715.25	1774.92	6697.30	189.48
25	600	551	285192	517.59	296.79	371409.05	619.02	2299.29	101.43
26	600	551	324927	589.7	97.87	529380.42	882.30	1870.66	292.60
27	600	584	440455.7	754.21	615.97	488361.30	813.94	4292.15	59.73
28	600	584	545533.3	934.13	573.97	860228.52	1433.71	4682.15	499.58
29	600	593	576300.1	971.84	1972.07	663596.94	1105.99	7367.46	134.15
30	600	593	743224.6	1253.33	1928.32	1008106.13	1680.18	9280.46	426.85
31	750	683	335885.4	491.78	519.18	424051.94	565.40	4360.02	73.62
32	750	683	416755.7	610.18	130.88	708836.17	945.11	2285.43	334.93
33	750	716	519079.9	724.97	763.44	555274.32	740.37	5812.58	15.40
34	750	716	628387.4	877.64	570.85	1044092.21	1392.12	6223.00	514.48
35	750	740	732413.6	989.75	2856.59	860792.54	1147.72	14267.92	157.97
36	750	740	930262.7	1257.11	2428.68	1289908.78	1719.88	14293.06	462.77
37	900	825	398671	483.24	1217.66	523186.55	581.32	7768.35	98.08
38	900	825	484567.8	587.35	266.02	840969.80	934.41	4068.72	347.06
39	900	863	593248.2	687.43	1669.92	653166.54	725.74	10618.12	38.31
40	900	863	722214.3	836.86	1561.21	1203932.16	1337.70	11204.96	500.84
41	900	882	805494.8	913.26	2562.32	930538.34	1033.93	19295.44	120.67
42	900	882	1003756	1148.46	4431.68	1478257.11	1642.51	19965.13	494.05

TABLE 7. Comparison of different routing methods in second stage.

ID	Proposed method		Stage 2: GA				Stage 2: TS			
	TCT(s)	RT (s)	TCT_{GA}(s)	RT_{GA}(s)	ΔTCT_{GA}(s)	$\Delta TCT_{GA}/TCT$	TCT_{TS}(s)	RT_{TS}(s)	ΔTCT_{TS}(s)	$\Delta TCT_{TS}/TCT$
1	52262.06	57.56	59018.92	68.41	6756.86	0.1293	53256.63	347.24	994.57	0.0190
2	61876.08	67.24	82430.1	57.51	20554.02	0.3322	76694.61	310.55	14818.53	0.2395
3	73630.56	58.75	87677.38	67.56	14046.82	0.1908	79898.15	373.81	6267.59	0.0851
4	90402.44	67.31	120231.4	56.55	29828.96	0.3300	111105.5	313.97	20703.06	0.2290
5	122130.8	145.67	140712.7	78.23	18581.9	0.1521	125229.6	1182.21	3098.8	0.0254
6	155061.3	165.15	202229.3	63.75	47168	0.3042	180143.2	920.23	25081.9	0.1618
7	92813.49	25.2	92158.61	93.75	-654.88	-0.0071	89104.47	160.26	-3709.02	-0.0400
8	111699.8	26.19	125684.3	68.33	13984.5	0.1252	117546.6	112.38	5846.8	0.0523
9	136808	105.27	151893.4	103.18	15085.4	0.1103	131714.7	1245.86	-5093.3	-0.0372
10	190006.2	120.65	233453.5	72.57	43447.3	0.2287	207319.7	868.16	17313.5	0.0911
11	226343.1	559.89	283101.5	124.16	56758.4	0.2508	225123.5	4436.04	-1219.6	-0.0054
12	337497.4	631.37	459018.5	96.35	121521.1	0.3601	387249.8	2923.9	49752.4	0.1474
13	160254.5	51.2	155084.9	185.45	-5169.6	-0.0323	155155.4	75.4	-5099.1	-0.0318
14	191540.7	41.41	189523.9	108.4	-2016.8	-0.0105	184400.8	69.76	-7139.9	-0.0373
15	232646.3	121.02	233406.1	179.96	759.8	0.0033	216200.8	854.28	-16445.5	-0.0707
16	287103.5	143.05	324566.7	120.5	37463.2	0.1305	293106.9	390.54	6003.4	0.0209
17	343058.5	796.18	410071.4	228.11	67012.9	0.1953	310163.9	4566.43	-32894.6	-0.0959
18	471488.4	771.44	639796.2	172.34	168307.8	0.3570	475898.4	5560.53	4410	0.0094
19	226883.8	144.12	224060.3	261.75	-2823.5	-0.0124	223948.9	171.56	-2934.9	-0.0129
20	256855.6	55.49	244082.3	159.48	-12773.3	-0.0497	241663.8	83.04	-15191.8	-0.0591
21	344094.2	278.42	336641.7	382.91	-7452.5	-0.0217	320071.9	1931.94	-24022.3	-0.0698
22	418182.3	253.93	447876.8	276.6	29694.5	0.0710	391370.1	1197.42	-26812.2	-0.0641
23	514550	1921.48	682296.1	469.84	167746.1	0.3260	444121.2	26869.92	-70428.8	-0.1369
24	703935	1652.87	1066138	403.34	362203	0.5145	753383.4	24393.65	49448.4	0.0702
25	285192	296.79	283390.7	433.25	-1801.3	-0.0063	283369.5	358.84	-1822.5	-0.0064
26	324927	97.87	319300.4	248.1	-5626.6	-0.0173	320106.3	186.13	-4820.7	-0.0148
27	440455.7	615.97	425630.6	689.85	-14825.1	-0.0337	409826.3	3065.37	-30629.4	-0.0695
28	545533.3	573.97	563698.7	540.61	18165.4	0.0333	490895.7	2897.13	-54637.6	-0.1002
29	576300.1	1972.07	639413.5	1649.61	63113.4	0.1095	525668	10885.42	-50632.1	-0.0879
30	743224.6	1928.32	926613.6	1548.63	183389	0.2467	704318.4	8351.93	-38906.2	-0.0523
31	335885.4	519.18	330709.9	759.31	-5175.5	-0.0154	329784.1	1018.89	-6101.3	-0.0182
32	416755.7	130.88	410569.1	272.45	-6186.6	-0.0148	410569.1	131.28	-6186.6	-0.0148
33	519079.9	763.44	503763	976.18	-15316.9	-0.0295	501264.9	2089.53	-17815	-0.0343
34	628387.4	570.85	620603.8	610.57	-7783.6	-0.0124	569771.3	1076.85	-58616.1	-0.0933
35	732413.6	2856.59	921049.1	1551.48	188635.5	0.2576	634723	31575.4	-97690.6	-0.1334
36	930262.7	2428.68	1213696	1623.07	283433.3	0.3047	829207.1	13176.35	-101055.6	-0.1086
37	398671	1217.66	386086.6	1455.06	-12584.4	-0.0316	384706.6	2249.54	-13964.4	-0.0350
38	484567.8	266.02	474388.4	441.04	-10179.4	-0.0210	474615	259.54	-9952.8	-0.0205
39	593248.2	1669.92	566619.3	1881.49	-26628.9	-0.0449	564192.9	3586.96	-29055.3	-0.0490
40	722214.3	1561.21	736251.8	1530.75	14037.5	0.0194	665747.5	2452.17	-56466.8	-0.0782
41	805494.8	2562.32	855397.5	2060.65	49902.7	0.0620	723066.3	24323.96	-82428.5	-0.1023
42	1003756	4431.68	1235134	3699.99	231378	0.2305	932262.7	19386.66	-71493.3	-0.0712

may improve by less than ten percent in exchange for multiple times of increase in running time. It is an unworthy trade-off between time and quality.

2) PARAMETER SENSITIVITY ANALYSIS

Considering there is no need to input manually decided parameters at the stop location stage of the algorithm and the stop location stage accounts for most time while running the algorithm, we will only apply sensitivity analysis on the parameters of the routing algorithm. Because there is a wide range of student density that has a significant impact on ACT among the instances, we conduct the sensitivity analysis on three instances with different student densities, including the minimum, maximum and median value of the densities. Thus, the instances number 5, 22, and 38 are selected and detailed in Table 5. A number of combinations of different parameters are tested, and for each set of parameters the IACO is executed for 10 times to get the average result to avoid the influence of the randomness.

Fig.8 shows the analysis result of each parameter. Since there are apparent difference of the total commute time of students when the number of students varies, the objective value is transformed to average commute time of students shown in the vertical axis to calibrate the impact of the parameters while eliminating the influence of other factors. The horizontal axis shows the values of the parameter being tested. We observe that instance being most sensitive to the parameters is the one with minimum student density. It can be easily understood that there exists need for more bus stops in the instance with minimum student density, i.e., the sparse area, to satisfy the walking accessibilities, which leads to the large search space for optimizing routing. Different values of parameters can significantly affect the search speed in the search space, resulting in different solution due to the nature of the solution approach, i.e., the approach iteratively searches for acceptable solutions over a period specified by certain stopping criteria.

Among all these parameters, most significant influence on ACT for all instances can be observed in Fig.8 (b), which shows sensitivity to *parameter* β, the factor indicating the importance of the heuristic information of stop and school location measured as the reciprocal of the riding time. This is because the tour time of buses constitutes one main part of total commute time of students, and the values larger than zero of *parameter* β means convergence to solutions with less tour time. However, as the value of *parameter* β continues increasing, the change of the convergence speed is trivial, leading to the slight decrease of ACT when the value of *parameter* β changes from one to two and leveling off when the value of *parameter* β is larger than two. Moreover, as for the difference of the extent of the variations of the ACT among different instances, the larger the student density is, the less the number of bus stops in need; the smaller the optimization space is, the more slightly the ACT decreases.

Another parameter significantly affecting the ACT of instances can be seen in Fig.8 (c), which indicates the sensitivity to *parameter* γ, the factor that representing the importance of stop assignment information denoted by the total walking time of students assigned to the stop. The consideration of stop assignment information contributes to the sharp increase of the ACT for instance with minimum student density and mild decrease for instance with median and maximum student density. The possible reason for the increase in the district with minimum student density is the low demand of stops. The variation of low demand has less effect on the ACT than the travel time between nodes. As for the other scenarios, the uniform distribution of students contributes greatly to the mild increase of the result.

The other parameters which indicate the corresponding relative importance of pheromone information, *parameter* α, and the relative importance of school attraction information, *parameter* θ, used in the route construction phase have no great impact on the ACT for all three instances except for the little fluctuation for the instance with minimum student density. The general trend for *parameter* α shown in Fig.8 (a) is due to the fact that the increased amount of pheromone is relatively small compared to the initial pheromone. When it comes to sensitivity to *parameter* θ, the much smaller number of schools compared with the number of bus stops, the uniform distribution of students and the constant student ratio for each school make contribution to the gentle trend indicated in Fig.8 (d).

Fig.8 (e) shows sensitivity to *parameter* ρ, which represents the evaporation rate of pheromone. The minimum value of ACT for instance occurs when *parameter* ρ takes the value of 0.9 with minimum student density. The value of ACT with *parameter* ρ taking the value of either 0.85 or 0.95 is larger than the one of 0.9. The possible reason is that there is need for proper rate of the evaporation of pheromone to lead to convergence to a better solution while avoiding falling into local minima.

Fig.8 (f) shows sensitivity to *parameter* q_0, the parameter directly affecting the choice of searching strategy between depth and breadth. One can observe a similar trend to *parameter* ρ. Since the two parameters both determine the preference on the diversification and intensification, the similar trends of them share the similar reason.

Fig.8 (i) shows sensitivity to *parameter* τ_0, the initial amount of pheromone at each link. The values of ACT of instances with median and maximum density are not sensitive to the change of the value of *parameter* τ_0, while there is little fluctuation for the instance with minimum density. This is due to the relatively small value of updated pheromone compared with initial pheromone and the relatively small value of pheromone compared with other information used in the route construction. However, the value of pheromone in instance with minimum density is relatively larger than the one with higher density, which leads

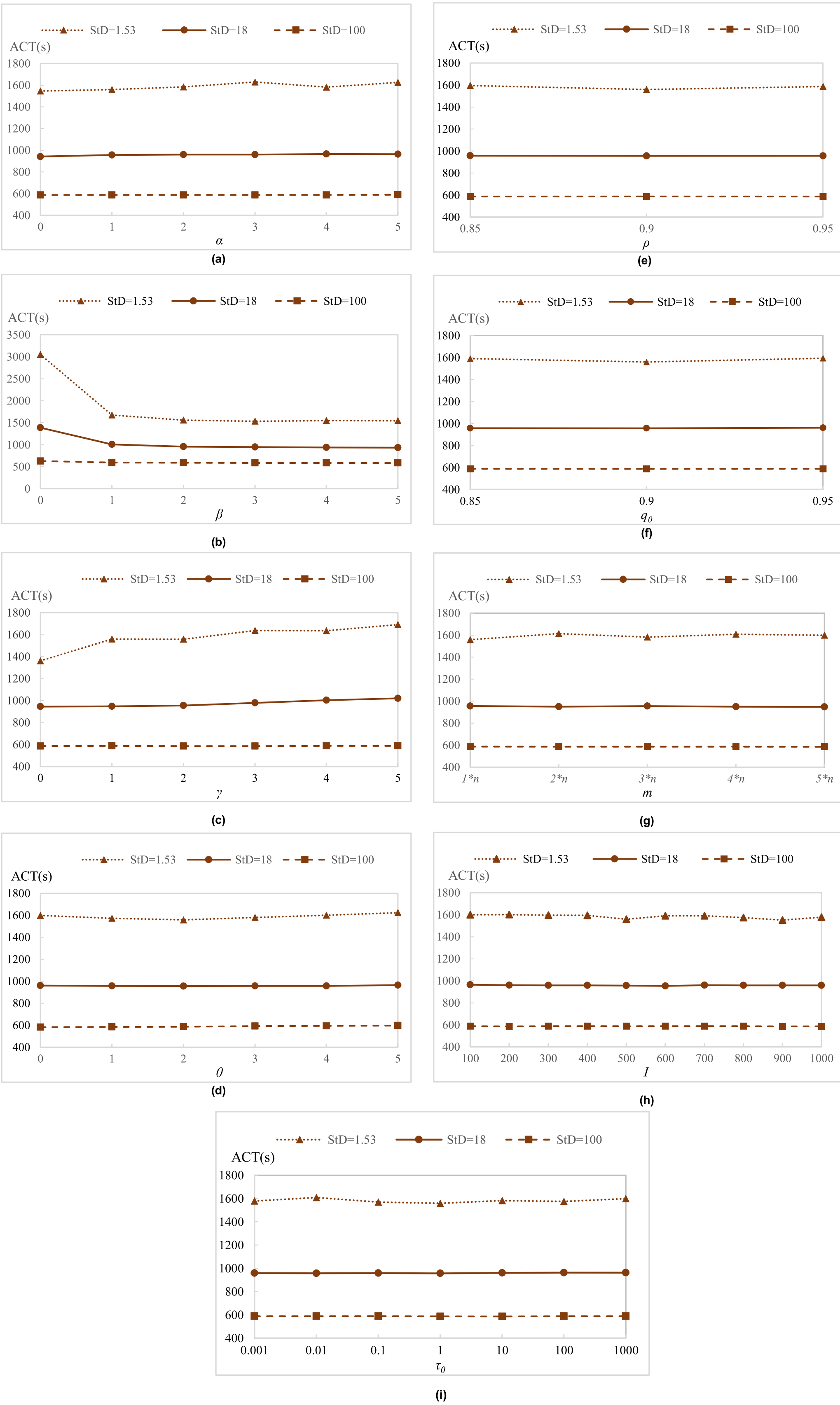
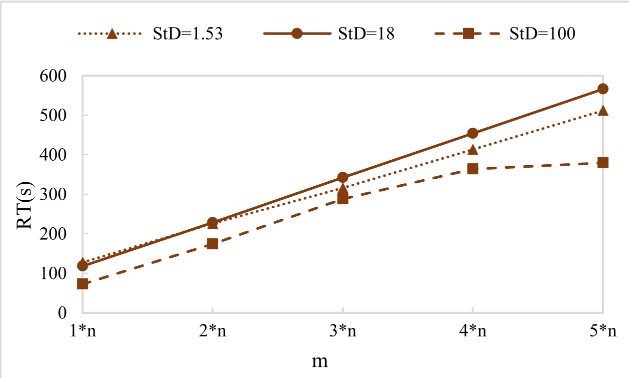
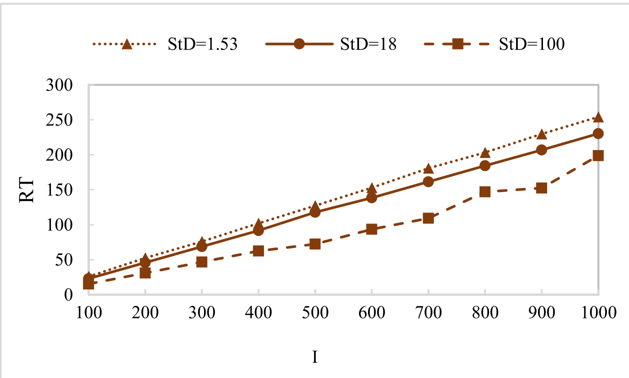


FIGURE 8. Sensitivity analysis of parameters (a) sensitivity to α, (b) sensitivity to β, (c) sensitivity to γ, (d) sensitivity to θ, (e) sensitivity to ρ, (f) sensitivity to q_0, (g) sensitivity to m, (h) sensitivity to I, (i) sensitivity to τ_0.



(a)



(b)

FIGURE 9. **Run time of routing process with various values of m and I (a) values of m, (b) values of I.**

to the more appreciable fluctuation in sparse area rather than the dense one.

Since the number of ants at each iteration, *parameter m*, and the number of iterations, *parameter I*, have a significant effect on the period of problem solving, the runtime of the routing process is analyzed for parameters m and I and shown in Fig.9 (a) and Fig.9 (b) in addition to the sensitivity of ACT shown in Fig.8 (g) and Fig.8 (h). One can see from the pictures that run time goes straight up with the increase of the value of parameters m and I for all instances. As for the change of the value of ACT, the change of the value of parameters m and I do not make change to the quality of the solutions for instances with median and maximum density, while the value of ACT fluctuates a little bit when the value of *parameter m* is less than 3 times the number of bus stops or the value of *parameter I* is less than 500 in instance with minimum student density. The possible reason lies in the relatively larger solution space in sparse area compared with dense ones and its vulnerability to randomness.

From all the analyses of sensitivity to various parameters, we conclude that the IACO routing method is robust and less vulnerable to the setting of parameters for dense areas while some of the parameters need to be adjusted properly to get better solution for sparse area.

VI. SUMMARY AND OUTLOOK

There is usually a great age difference among students from distinct schools at different educational stage, leading to the different walking accessibilities affecting the design of the school bus system, especially the school bus stop location phase. This study introduces the school bus stop location and routing problem with walking accessibility and mixed load with a mixed integer programming model. Considering we mainly study the problem of general students, the walking accessibility in the solution approach is described using the average walking speed of students attending schools at different stages of education and the same maximum walking time for all students. With the walking accessibility, the paper designs a two-stage solution method to minimize the total commute time of all students which can be used to evaluate the performance of the school bus system including walking time from home to stop, in-vehicle ride time, and service time at stops and schools. The solution method first clusters students and locates stops that can meet the requirements in terms of different walking accessibilities, and then determines a set of routes serving stops, while considering the interrelationship of the stages at each stage. To verify the solution approach, we generate a number of instances of different sizes to conduct the algorithm and analyze the factors influencing total commute time of students. In order to test the performance of the proposed solution for school bus problem, similar routing methods are applied to the same instances and compared to the proposed method. In addition, the door-to-door school bus system is compared with the same resources on the instances. The results show that the SBSLRP-WA-ML is a better mode for providing convenient and safe school bus service in dense area while door-to-door school bus service is preferable in sparse area. Several instances with different student densities are then selected to analyze the sensitivities to parameters and test the robustness and vulnerability of the solution approach. The robustness and less vulnerability of the IACO is validated while some of the parameters need to be adjusted properly to get better solution for sparse area.

The future direction of the research may exist in the formulation of a probabilistic walking accessibility model based on the real-world application instead of the average walking speed used in this paper. In addition, the walking accessibility can be represented as different walking velocity and maximum walking time for all students, which makes it easy to extend the model to consider special-education students for who maximum walking time can set to be zero or other values. The other possible improvement about the paper results from the limitations of two-stage heuristic methods. This can be studied further to design an effective solution algorithm which can locate stops and optimize routes considering walking accessibility simultaneously. In addition, the data used in the experiment is generated randomly according to the proposed problem. We are looking forward to the collaboration with realistic schools or related authorities to test and improve the method on realistic data.

REFERENCES

[1] R. M. Newton and W. H. Thomas, "Design of school bus routes by computer," *Socio-Econ. Planning Sci.*, vol. 3, no. 1, pp. 75–85, 1969. doi: 10.1016/0038-0121(69)90051-2.

[2] A. Ceder, *Public Transit Planning and Operation: Modeling, Practice and Behavior*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2016.

[3] W. Wu, R. Liu, and W. Jin, "Modelling bus bunching and holding control with vehicle overtaking and distributed passenger boarding behaviour," *Transp. Res. B, Methodol.*, vol. 104, pp. 175–197, Oct. 2017. doi: 10.1016/j.trb.2017.06.019.

[4] W. Wu, R. Liu, and W. Jin, "Designing robust schedule coordination scheme for transit networks with safety control margins," *Transp. Res. B, Methodol.*, vol. 93, pp. 495–519, Nov. 2016. doi: 10.1016/j.trb.2016.07.009.

[5] W. Wu, R. Liu, W. Jin, and C. Ma, "Stochastic bus schedule coordination considering demand assignment and rerouting of passengers," *Transp. Res. B, Methodol.*, vol. 121, pp. 275–303, Mar. 2019. doi: 10.1016/j.trb.2019.01.010.

[6] S. N. Parragh, K. F. Doerner, and R. F. Hartl, "A survey on pickup and delivery models Part II: Transportation between pickup and delivery locations," *J. für Betriebswirtschaft*, vol. 58, no. 2, pp. 81–117, 2008.

[7] M. W. P. Savelsbergh and M. Sol, "The general pickup and delivery problem," *Transp. Sci.*, vol. 29, no. 1, pp. 17–29, 1995. doi: 10.1287/trsc.29.1.17.

[8] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Manage. Sci.*, vol. 6, no. 1, pp. 80–91, 1959.

[9] G. Clarke and J. W. Wright, "Scheduling of vehicles from a central depot to a number of delivery points," *Oper. Res.*, vol. 12, no. 4, pp. 568–581, 1964. doi: 10.1287/opre.12.4.568.

[10] K. Braekers, K. Ramaekers, and I. Van Nieuwenhuyse, "The vehicle routing problem: State of the art classification and review," *Comput. Ind. Eng.*, vol. 99, pp. 300–313, Sep. 2016. doi: 10.1016/j.cie.2015.12.007.

[11] J. Braca, J. Bramel, B. Posner, and D. Simchi-Levi, "A computerized approach to the New York City school bus routing problem," *IIE Trans.*, vol. 29, no. 8, pp. 693–702, 1997. doi: 10.1023/A:1018526202990.

[12] J. Desrosiers, J. A. Ferland, J. M. Rousseau, G. Lapalme, and L. Chapleau, "An overview of a school busing system," in *Scientific Management of Transport Systems*, N. K. Jaiswal, Ed. Amsterdam, The Netherlands: North Holland, 1981, pp. 235–243.

[13] J. Park and B.-I. Kim, "The school bus routing problem: A review," *Eur. J. Oper. Res.*, vol. 202, no. 2, pp. 311–319, 2010. doi: 10.1016/j.ejor.2009.05.017.

[14] W. A. Ellegood, S. Solomon, J. North, and J. F. Campbell, "School bus routing problem: Contemporary trends and research directions," *Omega*, to be published. doi: 10.1016/j.omega.2019.03.014.

[15] R. Bowerman, B. Hall, and P. Calamai, "A multi-objective optimization approach to urban school bus routing: Formulation and solution method," *Transp. Res. A, Policy Pract.*, vol. 29, no. 2, pp. 107–123, 1995. doi: 10.1016/0965-8564(94)E0006-U.

[16] L. Y. O. Li and Z. Fu, "The school bus routing problem: A case study," *J. Oper. Res. Soc.*, vol. 53, no. 5, pp. 552–558, 2002. doi: 10.1057/palgrave.jors.2601341.

[17] J. Park, H. Tae, and B.-I. Kim, "A post-improvement procedure for the mixed load school bus routing problem," *Eur. J. Oper. Res.*, vol. 217, no. 1, pp. 204–213, 2012. doi: 10.1016/j.ejor.2011.08.022.

[18] J. F. M. Sarubbi, C. M. R. Mesquita, E. F. Wanner, V. F. Santos, and C. M. Silva, "A strategy for clustering students minimizing the number of bus stops for solving the school bus routing problem," in *Proc. IEEE IFIP Netw. Oper. Manage. Symp.*, S. O. Badonnel, Eds., Apr. 2016, pp. 1175–1180.

[19] M. Galdi and P. Thebpanya, "Optimizing school bus stop placement in Howard County, Maryland," *Int. J. Appl. Geospatial Res.*, vol. 7, no. 1, pp. 30–44, 2016.

[20] K. Worwa, "A case study in school transportation logistics," *Res. Logistics Prod.*, vol. 4, no. 1, pp. 45–54, 2014.

[21] M. Drexl and M. Schneider, "A survey of variants and extensions of the location-routing problem," *Eur. J. Oper. Res.*, vol. 241, no. 2, pp. 283–308, 2015. doi: 10.1016/j.ejor.2014.08.030.

[22] L. D. Bodin and L. Berman, "Routing and scheduling of school buses by computer," *Transp. Sci.*, vol. 13, no. 2, pp. 113–129, 1979. doi: 10.1287/trsc.13.2.113.

[23] P. Schittekat, J. Kinable, K. Sörensen, M. Sevaux, F. Spieksma, and J. Springael, "A metaheuristic for the school bus routing problem with bus stop selection," *Eur. J. Oper. Res.*, vol. 229, no. 2, pp. 518–528, 2013. doi: 10.1016/j.ejor.2013.02.025.

[24] M. F. Faraj, J. F. M. Sarubbi, C. M. Silva, M. F. Porto, and N. T. R. Nunes, "A real geographical application for the school bus routing problem," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2014, pp. 2762–2767.

[25] C. M. Silva, J. F. M. Sarubbi, D. F. Silva, M. F. Porto, and N. T. R. Nunes, "A mixed load solution for the rural school bus routing problem," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 1940–1945.

[26] S. P. Parvasi, M. Mahmoodjanloo, and M. Setak, "A bi-level school bus routing problem with bus stops selection and possibility of demand outsourcing," *Appl. Soft Comput.*, vol. 61, pp. 222–238, Dec. 2017. doi: 10.1016/j.asoc.2017.08.018.

[27] J. S. Arias-Rojas, J. F. Jiménez, and J. R. Montoya-Torres, "Solving of school bus routing problem by ant colony optimization," *Revista EIA*, vol. 9, no. 17, pp. 193–207, 2012.

[28] J. Euchi and R. Mraihi, "The urban bus routing problem in the Tunisian case by the hybrid artificial ant colony algorithm," *Swarm Evol. Comput.*, vol. 2, pp. 15–24, Feb. 2012. doi: 10.1016/j.swevo.2011.10.002.

[29] L. Huo, G. Yan, B. Fan, H. Wang, and W. Gao, "School bus routing problem based on ant colony optimization algorithm," in *Proc. IEEE Conf. Expo Transp. Electrific. Asia–Pacific*, Aug./Sep. 2014, pp. 1–5.

[30] B. Yao, Q. Cao, Z. Wang, P. Hu, M. Zhang, and B. Yu, "A two-stage heuristic algorithm for the school bus routing problem with mixed load plan," *Transp. Lett.*, vol. 8, no. 4, pp. 205–219, 2016. doi: 10.1080/19427867.2015.1110953.

[31] N.-A. Mokhtari and V. Ghezavati, "Integration of efficient multi-objective ant-colony and a heuristic method to solve a novel multi-objective mixed load school bus routing model," *Appl. Soft Comput.*, vol. 68, pp. 92–109, Jul. 2018. doi: 10.1016/j.asoc.2018.03.049.

[32] J. Müller, S. Müller, H. Baur, and F. Mayer, "Intra-individual gait speed variability in healthy children aged 1–15 years," *Gait Posture*, vol. 38, no. 4, pp. 631–636, 2013. doi: 10.1016/j.gaitpost.2013.02.011.

[33] M. H. Schwartz, A. Rozumalski, and J. P. Trost, "The effect of walking speed on the gait of typically developing children," *J. Biomech.*, vol. 41, no. 8, pp. 1639–1650, 2008. doi: 10.1016/j.jbiomech.2008.03.015.

[34] V. L. Chester, M. Tingley, and E. N. Biden, "A comparison of kinetic gait parameters for 3–13 year olds," *Clin. Biomech.*, vol. 21, no. 7, pp. 726–732, 2006. doi: 10.1016/j.clinbiomech.2006.02.007.

[35] T. Öberg, A. Karsznia, and K. Öberg, "Basic gait parameters: Reference data for normal subjects, 10-79 years of age," *J. Rehabil. Res. Develop.*, vol. 30, no. 2, pp. 210–223, 1993.

[36] B. W. Stansfield, S. J. Hillman, M. E. Hazlewood, A. A. Lawson, A. M. Mann, I. R. Loudon, and J. E. Robb, "Normalized speed, not age, characterizes ground reaction force patterns in 5-to 12-year-old children walking at self-selected speeds," *J. Pediatric Orthopaedics*, vol. 21, no. 3, pp. 395–402, 2001.

[37] D. Sutherland, "The development of mature gait," *Gait Posture*, vol. 6, no. 2, pp. 163–170, 1997. doi: 10.1016/S0966-6362(97)00029-5.

[38] *Traffic Engineering Manual*, (in Chinese), China Highway Transp. Soc., Beijing, China, 1998.

[39] M. Reimann, M. Stummer, and K. Doerner, "A savings based ant system for the vehicle routing problem," in *Proc. Genetic Evol. Comput. Conf.*, 2002, pp. 1317–1326.

[40] B. Yu, Z.-Z. Yang, and B. Yao, "An improved ant colony optimization for vehicle routing problem," *Eur. J. Oper. Res.*, vol. 196, no. 1, pp. 171–176, 2009.

[41] J. E. Bell and P. R. McMullen, "Ant colony optimization techniques for the vehicle routing problem," *Adv. Eng. Inform.*, vol. 18, no. 1, pp. 41–48, 2004.

[42] B. Bullnheimer, R. F. Hartl, and C. Strauss, "Applying the ANT system to the vehicle routing problem," in *Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization*, S. Voß *et al.*, Eds. Boston, MA, USA: Springer, 1999, pp. 285–296. doi: 10.1007/978-1-4615-5775-3_20.

[43] A. Colorni, M. Dorigo, and V. Maniezzo, "Distributed optimization by ant colonies," in *Proc. Eur. Conf. Artif. Life (ECAL)*. New York, NY, USA: Elsevier, 1991, pp. 134–142.

[44] M. Reimann, K. Doerner, and R. F. Hartl, "Insertion based ants for vehicle routing problems with backhauls and time windows," in *Proc. Int. Workshop Ant Algorithms*, 2002, pp. 135–148.

[45] M. Dorigo and T. Stützle, "The ant colony optimization metaheuristic: Algorithms, applications, and advances," in *Handbook of Metaheuristics*, F. Glover and G. A. Kochenberger, Eds. Boston, MA, USA: Springer, 2003, pp. 250–285.

JINGXUAN REN is currently pursuing the Ph.D. degree in transportation engineering with the South China University of Technology, China. Her main research interests include public transit system optimization and vehicle scheduling.

WENZHOU JIN received the B.S. degree in mathematics, the M.S. degree in probability theory and statistics, and the Ph.D. degree in transportation management engineering from Jilin University, China, in 1982, 1985, and 1994, respectively.

He is currently a Professor with the School of Civil Engineering and Transportation, South China University of Technology, China. His main research interests include road transport planning and development strategy, transport economy, transport policy, transport enterprise management, passenger and freight transport and logistics technology, urban transport planning and design, traffic safety management, and order management.

WEITIAO WU received the B.S. and Ph.D. degrees in transportation engineering from the South China University of Technology, China, in 2010, and 2015, respectively.

He is currently a Lecturer with the School of Civil Engineering and Transportation, South China University of Technology. His main research interests include public transit system optimization, vehicle scheduling, and traffic simulation.