University of Louisville
ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

8-2018

Solving school bus routing and student assignment problems with heuristic and column generation approach.

Di Zhang
University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

 Part of the Operational Research Commons

Recommended Citation

Zhang, Di, "Solving school bus routing and student assignment problems with heuristic and column generation approach." (2018). *Electronic Theses and Dissertations.* Paper 3051.
https://doi.org/10.18297/etd/3051

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

SOLVING SCHOOL BUS ROUTING AND STUDENT ASSIGNMENT PROBLEMS WITH HEURISTIC AND COLUMN GENERATION APPROACH

By

Di Zhang

A Dissertation
Submitted to the Faculty of the
J.B. Speed School of Engineering of
the University of Louisville
in Partial Fulfillment of the Requirements
for the degree of

Doctor of Philosophy in Industrial Engineering

Department of Industrial Engineering
University of Louisville
Louisville, Kentucky

August 2018

Copyright 2018 by Di Zhang

All rights reserved

SOLVING SCHOOL BUS ROUTING AND STUDENT ASSIGNMENT PROBLEMS WITH HEURISTIC AND COLUMN GENERATION APPROACH

By:

Di Zhang

A Dissertation Approved on

June 1, 2018

by the following Dissertation Committee:

Dissertation Director
Dr. Ki-Hwan G. Bae

Dr. Suraj M. Alexander

Dr. William E. Biles

Dr. Thomas Riedel

ACKNOWLEDGMENTS

I am deeply grateful to the help of my dissertation committee members (Dr. Ki-Hwan G. Bae, Dr. Suraj M. Alexander, Dr. William E. Biles and Dr. Thomas Riedel), whom gave me huge guidance with patience throughout my Ph.D. experience. I especially thank my advisor Dr. Ki-Hwan G. Bae, Dr. Sunderesh S. Heragu for their efforts to mentor me. I sincerely thank the Department of Industrial Engineering and Logistics and Distribution Institute (LoDI) at University of Louisville for supporting me financial assistance during my Ph.D. program. I also especially thank my wife, Bo Sun for her continuous support.

ABSTRACT
SOLVING SCHOOL BUS ROUTING AND STUDENT ASSIGNMENT PROBLEMS WITH HEURISTIC AND COLUMN GENERATION APPROCAH

Di Zhang

August 7, 2018

In this dissertation, we solve a school bus routing problem of transporting students including special education (handicapped) students and assigning them in Oldham county education district. The main contribution of this research is that we consider special education students (Type-2) along with other students (Type-1) and design a comprehensive school bus schedule to transport both kinds of students at the same time. Also, a student assignment mathematical model is presented to optimize the number of bus stops in use as well as one important measure of service quality, the total student walking distance. Comparing to the classic clustering methods, heuristic methods, or other methods from previous literatures, a mathematical optimization model is developed to solve a student assignment problem and to obtain the global optimal solution. The modeling constraints include budget limit, travel time limit, equity, school time window, and etc. Especially, the main difference between our model and other models is that it takes Type-2 students into consideration along with critical constraints accordingly, and solves the resulting more complex problem. Moreover, the school bus routing model in this work is one of the most general optimization models representing the school bus routing

problem. On the other hand, similar to all existing models, the developed model considers the total system cost as the objective function value to minimize, different bus capacities, and common vehicle routing constraints such as flow conservation on routes and subtour elimination. Furthermore, another main difference is that the bus scheduling and school time window is also considered and solved in the model. With two different types of students, both Type-1 and Type-2, the time restrictions are varying, resulting in more complexity and additional constraints. The results in this work present the difficulties of meeting the requirement of Type-2 student riding time limit and school time window simultaneously. Also, the constraints regarding service equity and quality are provided and they can be used by decision makers if necessary. Either densely populated urban areas or sparsely populated rural areas, the school bus routing problem is difficult to solve due to a large number of students or long travel distance. The school bus routing problem falls under vehicle routing problem (VRP) with additional requirements because each student represents one unit of capacity. In this dissertation, we present a modeling framework that solves a student assignment problem with bus stop selection, and subsequently a school bus routing problem with school time window constraints. We demonstrate the efficacy of heuristic methods as well as a column generation technique implemented to solve the problems using real data.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER I - INTRODUCTION

A. Motivation

According to the data published from (School Bus Safety Data, 2009), the public school transportation service in the USA is responsible for transporting approximately 26 million students twice a day. About 480,000 school buses provide service in the entire country, traveling approximately 4.4 billion miles each school year. “In 2004-05, the most recent year for which statistics were compiled, 55.3 percent of the 45,625,458 children enrolled in public K-12 schools were bused to schools at the taxpayers’ expense. The government spends 17.5 billion dollars per year on school bus transportation at an average cost of 692 dollars per student transported” (Pedroso, 2010).

However, the budgets have been cut in recent years and the inevitable effects on the public school transportation followed. The decreased budget forces researchers and industry workers to attempt to optimize the transporting system without sacrificing the service quality so that the quality and safety are ensured while saving the cost of operation.

B. Background

The school bus routing problem (SBRP) first appeared in the publication of Newton and Thomas (Newton, 1969). An SBRP can be defined as the problem determining the public school transportation plan for all students living

in an educational district area under the policy made by a regional board of education. Given (i) the information about the students such as residence address, school destination, and general (Type-1) or special-education, handicapped (Type-2), (ii) the information about the schools, such as school address, school bell time windows, number of students, (iii) the information about the school buses such as capacities, number of buses, and depot location. Using these input data, an optimal routing plan can be attained for students while meeting required constraints, reaching the best objective value.

C. Problem Statement

In this study, we investigate the Oldham County Schools District and solve the SBRP with a student assignment problem (SAP). “The Oldham County School District is located approximately 20 miles northeast of Louisville. Student enrollment has more than doubled in recent years, bringing the current number of students to 11,867, as of April 2013, in grades preschool through 12th. The district consists of one preschool, ten elementary schools, four middle schools, three high schools, one alternative school, one career center and one center for the arts and community education” (http://www.oldham.k12.ky.us/). So there are on average 600 to 700 students for each single school.

SBRP consists of five steps: “data preparation, bus stop selection, bus route generation, school bell time adjustment, and route scheduling” (Park, 2010).

The data preparation step assesses the required data. First, the information about students is needed before solving the problem, which includes the residence address of all students who require the service; the

number of students that share the same address; which school is the destination for each student; and the type of students (Type-1 or Type-2). Second, the information about schools is also needed, which includes the address of schools; the bell time window and the number of students for each school. Third, the information about buses needs to be obtained, which includes the number of buses available for each school or for all schools; the capacities of buses; the depots of buses; whether the buses can serve Type-2 students or Type-1 students only. Fourth, the origin-destination (OD) matrix must be constructed due to the residence addresses of students, bus stop locations and school addresses. It gives the data about the shortest travel time and distance between any pair of nodes (residence, bus stop, school, and bus depot). Fifth, the policies from regional board of education need to be clarified which decide the objective function(s) and constraint(s) in the model. The policies include the limit on student walking distance between home locations, bus stops or schools; the upper limit of number of students assigned to one bus stop; on travelling time and distance for each bus; on travelling time for students; the allowance of mixed loading; and the specific policies for transporting Type-2 students.

"The SAP seeks to select a set of bus stops and assign students to these stops" (Park, 2010). The condition of bus stop selection is always assumed to be given by many papers, and the only work left is to determine the routes so that every stop is visited at least once. In recent years, "the method of providing a group of potential bus stop locations and assigning students to the bus stops which are in use is introduced and it is more reasonable and practical" (Schittekat, 2006). In this study, the work of determining the set of bus stops in use is combined into the problem formulation. In this way, finding the set of bus

stops visited, assigning every student to a bus stop, and generating each bus route that visits the selected bus stops can be done step by step.

The bus route generation is an application of the classic vehicle routing problem (VRP) with some special constraints. For SBRP, the students are represented as the units of product. The bus stop selection and bus route generation are the important sub-problems in SBRP and we mainly focus on the joint problem of these two.

The school bell time window and route scheduling are about time issues. The school bell time window includes the starting and ending time of schools and different schools may have different bell time windows. Most papers assume the school bell time window is given as scheduling constraints in the model. It is reasonable because in most cases the school bell time window is decided by considering the condition of the school and class schedule instead of school bus transportation.

The route scheduling provides each route the starting and ending time so that each bus arrives at the school destination within the school bell time window. It also gives the schedule of each bus on each bus stop so that students and parents can follow.

Generally, SBRP is consists of many sub-problems as discussed above. To each specific SBRP study, many problem characteristics are involved and they decide the constraint and the model, and consequentially the solution. Based on the problem characteristics combined, the models from different studies could be very different and not general enough to be used to solve all the cases. This makes difficult to compare different studies directly. So it is necessary to introduce the problem with all the characteristics and assumptions

before developing the model.

First, different researches are aimed at varying objectives. It includes minimizing the number of buses in operation, minimizing the total travel distance of buses, minimizing the total travel time of buses, balancing a load of different mix, minimizing various types of cost to operate the school bus transportation system and others. Note that the number of buses, the total travel distance, and the total travel time contribute to the cost mainly and most studies consider part or all of them in the objective.

Second, the constraints represent the characteristics of the problem and the policy from the education district department. Some common constraints are shared by most of models, such as the upper limit of student walking distance; the upper limit of the number of students at one same bus stop; the lower limit of number of students to generate a route; the bus capacity constraint; the upper limit of the onboard time for one single student; the school bell time window and others. Besides, some logical constraints for SAP and SBRP are considered, including a bus traveling to a bus stop and at the same time, the bus travels to another stop from that bus stop; one bus stop being served by a bus or buses only when that bus stop is in use.

Third, one important characteristic is the number of schools being considered in the bus transportation system. Practically, it is decided by the policy from the educational district department. For example, in the Oldham County Schools District, different schools or school groups (several schools at the same location) have their own school bus service systems and they do not share the buses. The model for single-school problem is used for each school or school group. However, in the Jefferson County Public Schools District,

which is adjacent to the Oldham County, all schools share one school bus system and it becomes a multiple-school problem. “For a single-school problem, SBRP is similar to the Open VRP” (Fu, 2005). “One feature of Open VRP is that it is unnecessary for the vehicle to return to the depot after visiting the last customer on a route” (Li, 2007). For a multiple-school problem, the model can be more flexible but more complicated, and mixed loading is an additional consideration. Besides, schools may be treated as stops on the routes, and buses not only pick up students at stops but also drop off students at schools on the routes. In this way, the number of students on one bus may increase and decrease travelling along the route.

Fourth, as mentioned above, mixed loading means that one bus may serve students with different school destinations at the same route and at the same time. Therefore, it is usually considered only for a multiple-school problems. “It is true that mixed loading is considered only for a multiple-school problem and it increases flexibility and cost savings” (Braca, 1997). Also, mixed loading is widely used in the rural area because of scattered populated distribution.

Fifth, SBRP can be modeled in the context of bus routings in an urban or rural area. For either densely populated urban areas or sparsely populated rural areas, a school bus routing problem is difficult to solve due to a large number of students or long travel distance involved, respectively. The main difference between them is density of students. “In urban area with high density of students, it is common that bus capacity constraint is usually binding before the maximum riding time is reached” (Park, 2010). Hence, for SBRP in urban areas, maximum riding time constraints are not binding. Another difference is that in

urban areas, students have to walk from their homes location to the bus stops to wait for the bus, while in rural areas, students take their bus at front door. Thus, for SBRP in rural areas, mixed loading is desirable and the student assignment problem (SAP) becomes trivial.

Sixth, a bus is scheduled to run on routes in the morning and in the afternoon. In the morning, buses pick up students from bus stops and transport them to their schools. On the contrary, in the afternoon, buses pick up students from their schools and separate them to their bus stops. “Intensive school time windows and heavier traffic make the morning problem to be more difficult and urgent to solve” (Braca, 1997). Furthermore, “it is claimed that the afternoon problem can be converted into a morning problem with little modification” (Li, 2002).

Seventh, there may be several types of school buses in one school district system. This fact leads to two kinds of bus fleet groups, one is homogeneous and the other is heterogeneous. The heterogeneous fleet implies that the capacities of buses may not be the same and they have different fixed costs and travelling costs. The homogeneous fleet is simple with same capacity and cost for buses and it is widely used. Another consideration is the type of bus serving Type-2 students and the number of buses to accommodate them in one single route.

Eighth, most of studies have not considered Type-2 students. As mentioned above, they need service from a special type of buses with the proper equipment and assistance. “The problem of routing Type-2 students is fundamentally different in many respects from the problem of routing general students. First, Type-2 students are picked up and dropped off directly at their

homes and not at their bus stops. Second, there is a more rigid restriction on the maximum riding time for a Type-2 student in a bus than for a Type-1 student. Third, each student should be served differently depending on the severity of his/her own disability" (Park, 2010). When a Type-2 student's home location is selected as a bus stop, if mixed loading of Type-2 students and Type-1 students is allowed, other Type-1 students nearby may walk from their home to this bus stop within the upper bound of walking distance limit and take the ride together. The corresponding service time for a Type-2 student at the bus stop is longer and the upper limit of riding time is shorter. Due to this complexity, there may be only one Type-2 student on one route. In addition, based on the type of buses, a bus may have at most two Type-2 students on board at the same time. This characteristic depends on the policy and other practical situations. For this reason, often times modeling less number of Type-2 students per route improves solution tractability.

Ninth, SBRP is more complicated when taking transfer stations into account. In many cases nationwide, some students may be transported to arranged locations as transfer stations first, then be transferred to a different bus and finally be transported to their school destinations. In some extreme cases, multiple times of transfer are needed and more routes are possible because of route and schedule interaction. Note that the transfer stations could be schools or other places or even a bus stop. The advantage of applying a transfer station is that the number of buses in use and the total travel distance can be reduced. The disadvantage is that the complexity of the problem is increasing sharply and the service time at the transfer station should be accounted in the total travel time of students.

The problem characteristics and assumptions in this study are as follows. The objective is the total cost of the system which combines the number of buses with fixed cost and the total travelling cost. The constraints include the bus capacity with heterogeneous fleet, different riding time limits for Type-1 and Type-2 students, given a single school bell time window, the upper bound of walking distances for Type-1 students, and stops located at Type-2 student's homes. The SBRP in the Oldham County Schools District is a single-school problem due to the policy and it is located in a suburban area, so we construct the riding time limit constraint and a student assignment problem. The transfer stations are not considered. We focus on the morning problem. Transporting Type-2 students is considered with some specific constraints such as a bus stop being located at home, more critical riding time limits, and mixed loading of both types of students.

D. Dissertation Organization

To enhance understanding of SBRP, it is important and necessary to review all relevant literature. In chapter 2, a comprehensive literature review about SBRP is provided and it introduces scenarios of SBRP investigated, as well as methods and algorithms applied on SBRP along with corresponding performance and results.

Chapter 3 illustrates the models of SAP with different objectives. With the data about students, the SAP model searches for the optimal solution of assignment that meets all requirements and policies made by the educational department. This solution then will be used as the input data for the SBRP model. With the results from the SAP models, the SBRP model in Chapter 4

solves the routing problem with a time schedule and finds the final solution. It includes how many buses are required, the routes for buses, the time schedule of each bus.

Note that SAP and SBRP are $\mathcal{NP}$-hard problems and it may not be feasible to solve them directly with models developed in Chapter 3 and 4, especially for large size instances. Hence, it is necessary to apply heuristic algorithms on the SAP and SBRP. In Chapter 5, we propose an effective heuristic method of using Genetic Algorithm coupled with column generation approach applied to solve SBRP. The computational experiment results are analyzed and discussed in Chapter 6, and finally we provide summary and concluding remarks with future research in Chapter 7.

CHAPTER II - LITERATURE REVIEW

In this chapter, literatures about SBRP published in the early stages will be briefly introduced, and other SBRP papers published in recent 20 years will be introduced separately based on the single school problem or the multiple school problem.

A. Early SBRP Literatures

The SBRP falls into the Vehicle Routing and Scheduling Problems, and it is necessary to investigate the complexity first before we study the SBRP. A comprehensive research can be found in (Lenstra, 1981). In this paper, the problem class was defined and known NP-hardness results were reviewed.

Several papers published in early times will be briefly introduced first. (Bodin, 1979) solved a SBRP with the travel time limit for students and standard constraints such as the bus capacity limit using a specific heuristic approach. (Chapleau, 1985) used clustering technique to do the SAP with “compactness measure”. (Desrosiers, 1981) considered the case in urban area and the objective is to minimize the total number of routes. (Dulac, 1980) used location-allocation-routing strategy (LAR) to assign bus stops with students to school.

The first paper that considered transporting Type-2 students is proposed by (Russell, 1986). This paper pointed out that “in routing special-education students in an urban environment, the special-education students must be

picked up at home and delivered to selected schools which meet their specific educational needs such as special sequencing restrictions and route duration limitation".

Note that "the combined problem of bus stop selection and bus route generation falls into the class of location-routing problems" (Park, 2010). So the works about solving location-routing problems will help researchers better understand the SBRP. (Laporte, 1988) used "graph representation and then a graph extension" to convert the problem into "equivalent constrained assignment problems".

Another paper addressing location-routing problems is (Nagy, 2007). This paper is a survey of a location-routing problem and it investigated the variants of the problem with exact and heuristic algorithms.

B. Literatures on Single School Problem

The first comprehensive paper about SBRP is published by (Bowerman, 1995). In the paper, it introduced "a multi-objective approach to modeling the urban school bus routing problem and an algorithm for solution".

As mentioned before, SBRP can include multiple objectives. (Corberán, 2002). It investigated the case in a rural area, and explained the conflicts between cost and service.

Regarding multiple objectives, (Li, 2002) presented a "multi-objective combinational optimization problem" and a heuristic algorithm. The objectives include the total number of buses in use, the total travel time and equity measures.

As mentioned before, several steps are needed to solve SBRP, and the

most important one should be developing formulation. (Bektaş, 2007) believed that “the focus should be on developing better formulations for problems of moderate size, rather than devising solution algorithms that are problem-specific as the computer hardware and software technology is rapidly improving”.

(Martínez, 2011) presented a case for Lisbon. The problem is solved by two steps. The first step is to solve the SAP, and the second step is to generate the optimal routes with the results from the first step.

(Díaz-Parra, 2012) used genetic algorithm to solve SBRP. The algorithm includes “clusterization population pre-selection operator, tournament selection, crossover-k operator and an intelligent mutation operator”.

(Euchi, 2012) focused on the SBRP in urban areas. It proposed “a hybrid evolutionary computation based on an artificial ant colony with variable neighborhood local search algorithm”.

(Riera-Ledesma, 2012) solved SBRP as “multi-vehicle traveling purchaser problem”. A branch-and-cut algorithm is generated to solve randomly generated problems. The heuristic algorithms in this paper are “initial heuristic”, “constructive primal heuristic”, and “improvement phase”. The initial heuristic is a greedy approach assigning students to the nearest bus stop without breaking the capacity limit. Note that the initial heuristic may give no feasible solution. In this case the feasible solution may exist while the initial heuristic cannot find it. In the constructive primal heuristic, a feasible VRP solution can be reached from a fractional solution. “It consists two phases, the first phase constructs a set of feasible cycles with starting and ending points at the depot, and the second phase constructs a feasible assignment of students to the bus stops”

(Riera-Ledesma, 2012). Note that these two phases can be skipped if any one of them has no feasible solution. In the improvement phase, three procedures, cycle reduction procedure, cycle user interchange procedure, and cycle merging procedure were applied in order to refine the solution.

(Riera-Ledesma, 2013), which is a later work resulting from (Riera-Ledesma, 2012), presented a branch-and-price algorithm to solve SBRP as the Multiple Vehicle Traveling Purchaser Problem.

(Pacheco, 2013) addressed that SBRP in rural areas is the important practical applications of min-max VRP. A tabu search within the framework of Multiobjective Adaptive Programming was applied and compared to Non-dominated Sorting Genetic Algorithm.

(Addor, 2013) proposed an integer programming model and used an Ant Colony Based Metaheuristic to solve the problem.

(Arias-Rojas, 2013) considered a case in Bogotá, Colombia. It used a customized Ant Colony Optimization Algorithm to solve a real problem under cluster-first route second strategy.

Another complete paper on both SAP and SBRP is (Schittekat, 2013). It claims that “existing literature on routing of school buses focus mainly on building intricate models that attempt to capture as many real-life constraints and objectives as possible”. It presented an efficient parameter-free GRASP+VND metaheuristic to solve the SBRP first, and then checked the feasibility of SAP with SBRP solution.

Some works focus on service quality more than cost. (Song, 2013) focused on improving service quality rather than efficiency. (Manumbu, 2014) proposed a mathematical formulation and an exact optimal solution for SBRP.

Five procedures of mathematical formulation model are also given.

(Chalkia, 2014) introduced a Safe Map concept with safety parameters integrated to its links and its nodes. Several algorithms were also compared.

In (Kinable, 2014), column generation was applied to solve SBRP. 128 instances were tested and a better computation time results were reported. In the column generation approach, the mixed integer programming formulation of the SBRP was treated as the master problem. By replacing the integrality constraints, a linear programming relaxation was obtained. Then the pricing problem was presented with two algorithms, a local search heuristic and a labeling algorithm, to solve it. In the local search algorithm, the initial solution was randomly generated and it would be improved iteratively by local neighborhood searching. Three neighborhoods, insert neighborhood, remove neighborhood and swap neighborhood defined the set of feasible neighbor solutions. Note that in order to improve efficiency, only the neighborhood moves which are significantly different from the previous solutions are accepted. In the labeling algorithm, it is unnecessary to search for a column with the best objective value. The algorithm could return a negative reduced cost when it found one. After the pricing problem, two stabilization procedures, 3-piecewise proximal-type stabilization (PTS) and interior point stabilization (IPS) were applied and compared because the column generation algorithm would have slow convergence when the objective function gets closer to the optimal value. The authors claimed that “a major disadvantage of PTS methods is the extensive number of parameters that have to be tuned and the main argument against IPS is that it requires the restrict mater problem to be solved multiple times during each iteration to obtain the interior points” (Kinable, 2014). Then

the column pool manager (CPM) was applied to reduce the number of active columns in the restricted master problem because the time evaluating the master problem was too long when the total number of columns was too large. Finally, an exact branch-and-price framework was presented.

C. Literatures on Multiple Schools Problem

If an SBRP involves multiple schools instead of a single school problem, the complexity increases and it takes a longer time to solve the problem. The main difference is combining the mixed loads and the transfer stations into the problem, which in turn strictly limiting riding time for the Type-2 students. Some literatures investigate these aspects of modeling constraints.

One important model of routing Type-2 students was proposed by (Braca, 1997). In this paper, the case in five boroughs of New York City was investigated and the computerized system with solutions were used to transport Type-2 students.

When solving the SBRP, the best ideal solution needs to be decided by the practitioners after considering many aspects. So a decision-making technique can be effective to SBRP. (Spada, 2005) presented “a comprehensive methodology designed to support the decision of practitioners”.

(Fügenschuh, 2009) applied different methods on solving SBRP. It presented an integrated coordination of school time window and claimed that few buses are needed if the schools start at different times.

Scheduling is also a main part of SBRP. (Kim, 2012) investigated “VRP with time windows as a virtual stop” on given routes. The authors presented two assignment problem-based exact approaches and a heuristic algorithm. This

paper assumes that all buses are homogeneous and start at a fixed time point. In this way, “the problem can be modeled as an assignment problem.” Then, an assignment problem-based branch-and-bound algorithm was proposed. The initial upper bound and lower bound problem were solved using the Hungarian method which can solve the assignment problem easily. For the assumption of heterogeneous buses, an iterative AP based construction algorithm and an improvement algorithm were presented.

(Park, 2012) solved SBRP with mixed load assumption. This paper presented a mixed algorithm allowing students from different schools board on the same bus and there may be more than one school as stops on one bus route.

D. Literatures on Related Problems

Based on all works with respect to solving SBRP, heuristic or metaheuristic are shown to be the most effective method. Among various heuristic techniques, the revised branch and cut methods are widely chosen because they proved to be effective in solving hard combinatorial optimization problems. (Augerat, 1998) presented “a constructive algorithm, a randomized greedy algorithm and a tabu search procedure. It is the first time a metaheuristic is used for separation of the capacity constraints for the Capacitated Vehicle Routing Problem”.

SBRP is a practical problem even though many theories and techniques are applied to find the solution. To make the solution practical and useful to the real life problem, the real aspects involved when building the mathematical models are important to consider. For this purpose, a good survey is done by

(Howley, 2001). This paper presented “the nature and experience of riding the school bus in rural as compared to suburban locales in five states”.

Among the multiple objectives, one of the most important objectives is the cost. (Hanley, 2007) explained “the relationship between school district size and bus transportation costs”, and estimated “the change in costs if a statewide policy of consolidation is pursued”.

More generally, SBRP falls into the pickup and delivery problem. In some instances, the students are transferred from one bus to another at transfer stations. Considering this case, (Cortés, 2010) proposed a general formulation allowing students to transfer from one bus to another. The better solutions are obtained and corresponding heuristic methods are provided.

(Park, 2010) conducted a comprehensive review of the SBRP. It introduced all the previous literatures and clearly provided all the important basic modeling frameworks such as assumptions, constraints, and solution methods.

Note that some problems have in common with SBRP and the methods applied also can be used to solve SBRP. (Perugia, 2011) presented “a model and an algorithm for the design of a home-to-work bus service in a metropolitan area”. The authors used a tabu search algorithm to solve a multi-objective location-routing model.

Solving the similar case in Lisbon, (Eiró, 2011) presented the modeling process with four phases, “estimation of the potential demand, definition of stops, determination of the service O/D matrix and establishment of the most profitable routes”.

(Moccia, 2012) described “an incremental neighborhood tabu search

heuristic for the generalized vehicle routing problems with time windows".

Two papers, (Crainic, 2012) and (Nguyen, 2013), both described works on "Multi-Zone Multi-Trip Vehicle Routing Problem with Time Windows" which are also similar to SBRP, and proposed tabu search metaheuristic.

(Bock, 2011) presented a polynomial time 4-approximation algorithm and a constant factor approximation for the variant to solve SBRP.

One more paper that is worth mentioning here is (Yu, 2009). This paper presented an improved ant colony optimization with any-weight strategy.

E. Literatures Summary

In all, SBRP falls into the VRP and includes time schedule characteristics and assignment characteristics. Hence, literatures on related problems may provide insights to develop and methods in solving SBRP. Due to various problem characteristics of SBRP, it is necessary to summarize all literatures on SBRP.

Table 1 Comparison of Model Scopes of SBRP in Literature

References	Bus Stop Selection	Scheduling	Density of Population
(Bodin, 1979)	×	×	Urban
(Dulac, 1980)	×		Urban
(Desrosiers, 1981)	×	×	Both
(Chapleau, 1985)	×		Urban
(Russell, 1986)		×	Urban
(Bowerman, 1995)	×		Urban
(Braca, 1997)		×	Urban
(Corberán, 2002)			Rural
(Li, 2002)			Urban
(Spada, 2005)		×	Rural
(Bektaş, 2007)			Urban
(Fügenschuh, 2009)		×	Rural
(Martínez, 2011)	×	×	Urban
(Díaz-Parra, 2012)			Urban
(Kim, 2012)		×	Both
(Euchi, 2012)	×		Urban
(Riera-Ledesma, 2012)	×		Both
(Park, 2012)		×	Both
(Pacheco, 2013)			Rural
(Addor, 2013)			Both
(Riera-Ledesma, 2013)	×		Both
(Arias-Rojas, 2013)			Urban
(Schittekat, 2013)	×		Urban
(Song, 2013)		×	Urban
(Manumbu, 2014)			Urban
(Chalkia, 2014)			Both
(Kinable, 2014)			Both

Table 2 Comparison of Model Objectives of SBRP in Literature

References	Objectives							
	NB	TBD	TSD	SWD	LB	MRL	TL	LEB
(Bodin, 1979)	×							
(Dulac, 1980)	×	×						
(Desrosiers, 1981)	×	×						
(Chapleau, 1985)	×			×				
(Russell, 1986)		×						
(Bowerman, 1995)	×			×	×			
(Braca, 1997)	×							
(Corberán, 2002)	×					×		
(Li, 2002)	×	×	×		×			
(Spada, 2005)							×	
(Bektaş, 2007)	×	×						
(Fügenschuh, 2009)	×	×						
(Martínez, 2011)	×	×		×				
(Díaz-Parra, 2012)	×	×						
(Kim, 2012)	×	×						
(Euchi, 2012)	×	×						
(Riera-Ledesma, 2012)	×	×						×
(Park, 2012)	×							
(Pacheco, 2013)		×				×		
(Addor, 2013)	×	×						
(Riera-Ledesma, 2013)	×	×		×				
(Arias-Rojas, 2013)		×	×					
(Schittekat, 2013)		×						
(Song, 2013)			×					
(Manumbu, 2014)			×					
(Chalkia, 2014)			×					
(Kinable, 2014)		×						

Table 3 Comparison of Model Constraints of SBRP in Literature

References	Constraints					
	VC	MRT	TW	MWT	EPT	MSN
(Bodin, 1979)	×	×	×			
(Dulac, 1980)	×	×		×		
(Desrosiers, 1981)	×	×		×		
(Chapleau, 1985)	×	×		×		
(Russell, 1986)	×	×				
(Bowerman, 1995)	×			×		
(Braca, 1997)	×	×	×		×	×
(Corberán, 2002)	×					
(Li, 2002)	×					
(Spada, 2005)	×		×			
(Bektaş, 2007)	×	×				
(Fügenschuh, 2009)			×			
(Martínez, 2011)	×	×	×	×	×	
(Díaz-Parra, 2012)	×	×				
(Kim, 2012)	×		×		×	
(Euchi, 2012)	×			×		
(Riera-Ledesma, 2012)	×			×		
(Park, 2012)	×	×	×			
(Pacheco, 2013)	×					
(Addor, 2013)	×					
(Riera-Ledesma, 2013)	×	×	×			
(Arias-Rojas, 2013)	×					
(Schittekat, 2013)	×			×		
(Song, 2013)	×	×	×			
(Manumbu, 2014)	×					
(Chalkia, 2014)	×	×				×
(Kinable, 2014)	×					

In the table 2, the abbreviations are as follows:

$$NB = the\ \ number\ \ of\ \ buses\ \ used$$

$$TBD = total\ \ bus\ \ travel\ \ distance\ \ or\ \ time$$

$$TSD = total\ \ student\ \ riding\ \ distance\ \ or\ \ time$$

$$SWD = student\ \ walking\ \ distance$$

$$LB = load\ \ balancing$$

$$MRL = maximum\ \ route\ \ length$$

$$TL = child's\ \ time\ \ loss$$

$$LEB = length\ \ balancing$$

In the table 3, the abbreviations are as follows:

$$VC = vehicle\ \ capacity$$

$$MRT = maximum\ \ riding\ \ time$$

$$TW = school\ \ time\ \ window$$

$$MWT = maximum\ \ walking\ \ time\ \ or\ \ distance$$

$$EPT = earliest\ \ pick\ \ up\ \ time$$

$$MSN = minimum\ \ student\ \ number\ \ to\ \ create\ \ a\ \ route$$

In Table 1, the solution of bus stop selection was assumed to be given or predetermined in many studies, and the solution quality of routing and scheduling is directly subject to the bus stop selections. Not all works have taken into account school bus scheduling, and instead have mainly focused on minimizing the costs. The SBRP of urban areas has been studied more frequently than that of rural areas. The reasons are that the difficulty and variability of the urban area, which has a high density of students compared to the rural area, which has scatter population and longer traveling distance.

Most works in Table 2 focused on the minimization of total cost including fixed cost and routing cost. Fixed cost is directly associated with the number of

buses operating in system, whereas routing cost is mainly determined in proportion to total travel distance. Some literature also proposed methods to assess cost measures of time as well as service quality. Different costs associated with appropriate weights can be added such that the total routing cost is minimized. SWD is a common objective to minimize in a bus stop selection problem. NB and TBD can be used as objectives for cost, whereas TSD, SWD, LB, MRL, TL, and LEB serve as measures for service quality.

Table 3 presents a list of research works that considered various sets of constraints associated with SBRP. For example, VC was the most commonly used constraint in VRP and its variants were included in most of papers. MRT was used as an important restriction by many educational districts. TW was used when scheduling was considered. Some papers regarded MWT as constraints instead of optimizing it as an objective in a bus stop selection problem. EPT was more closely related to service quality just as MSN was to cost minimization.

Note that most SBRP studies minimize the total cost, which includes the fixed cost and the routing cost. The fixed cost is highly related to the number of buses in the system and the objective is to minimize it. The routing cost is mainly decided by the total travel distance. Some literatures also illustrate methods to calculate the measures of the time cost and the measures of the service quality. Adding different costs with corresponding weights which is determined by decision makers together so that the total routing cost can be minimized.

CHAPTER III - STUDENT ASSIGNMENT PROBLEM

When all the data about students and schools are available, it is possible to make school bus stops selection and solve the student assignment problem (SAP). The results from this part include that we know how many and which potential bus stop locations would be used as real bus stops and how many and which students are assigned to each real bus stop. As mentioned before, these results are assumed to be given as data in many papers. With all the potential bus stop location data and student data, we minimize the total walking distance for all students from their home to the bus stops they were assigned to, and minimize the number of bus stops in use in two separate models. The total walking distance represents the quality performance of our SAP model. Minimizing the total walking distance ensures that the students walk less and save time. The number of bus stops in use represents the number of demand points the buses should visit, and minimizing the number of bus stops reduces the total travelling time and distance for all buses which is also an important measure of the cost and the quality of service.

A. SAP Procedure

First, we introduce the data needed to start the SAP procedure. It includes the home location of each student, the potential bus stop locations and the walking distance between each pair of student's home and bus stop. Specifically, the home locations of Type-2 students are designated as potential

bus stops and the walking distance is zero between the bus stop and Type-2 student's home. Also, these bus stops would be used as constraints in the SAP model.

Then, we introduce the SAPD (Student Assignment Problem Distance) model that minimizes the total walking distance as the objective function as follows.

Notations:

$$z_{ij} = \begin{cases} 1 \ if \ student \ i \ is \ assigned \ to \ bus \ stop \ j \\ 0 \ otherwise \end{cases}$$

$$u_j = \begin{cases} 1 \ if \ potential \ bus \ stop \ point \ j \ is \ in \ use \\ 0 \ otherwise \end{cases}$$

$$w_{ij} = \ walking \ distance \ from \ student \ home \ i \ to \ bus \ stop \ j$$

$$MD = maximum \ walking \ distance \ for \ \ student$$

$$P = total \ number \ of \ potential \ bus \ stop \ points$$

$$Q = total \ number \ of \ students$$

$$UB = number \ of \ bus \ stops \ in \ use \ upper \ bound$$

(SAPD)

$$Minimize \ \sum_{i=1}^{Q} \sum_{j=1}^{P} w_{ij} z_{ij} \quad (1)$$

Subject to

$$\sum_{j=1}^{P} z_{ij} = 1 \qquad for \ i = 1, 2, \dots, Q \quad (2)$$

$$w_{ij} z_{ij} \leq MD \qquad for \ i = 1, 2, \dots, Q \quad j = 1, 2, \dots, P \quad (3)$$

$$z_{ij} \leq u_j \qquad for \ i = 1, 2, \dots, Q \quad j = 1, 2, \dots, P \quad (4)$$

$$\sum_{j=1}^{P} u_j \leq UB \quad (5)$$

$$z_{ij} = 0 \ or \ 1 \qquad for \ i = 1, 2, \dots, Q \quad j = 1, 2, \dots, P \quad (6)$$

$$u_j = 0 \ or \ 1 \qquad j = 1, 2, \dots, P \quad (7)$$

$$u_j = 1 \qquad for \ j = Type - 2 \ student \ home \ location \ bus \ stops \quad (8)$$

In this model, the objective function (1) is to minimize the total walking distance for all students between their homes and their bus stops. Constraint (2) ensures that every student is assigned to one and only one bus stop. In constraint (3), we set an upper limit of walking distance for each student and no one walks more than the limit. This limit could be decided due to the policy from an educational district. Constraint (4) ensures that every student is only assigned to the bus stop in use. In constraint (5), we set an upper bound of number of bus stops in use. In this way, we can reach the least number of bus stops in use with optimal solution by decreasing the upper bound till no feasible solution is found. Also, the number of bus stops in use could be the objective function because it may affect the travel time or complexity in latter work which is shown next. By decreasing the upper bound, we can get smaller number of bus stops in use while the optimal value increases. It is obvious that if some potential bus stops could not be used by decreasing the upper bound, the students would be assigned to the further distant bus stops which are in use within the upper limit of walking distance, which at the same time implies increasing the total walking distance. This is a tradeoff between these two objectives and it may be combined together in the objective function with weights, and the weights are also decided by the policy maker. Constraints (6) and (7) ensure the decision variables are binary, which represents the decisions about student assignments and bus stops picks, respectively. Constraint (8) forces the potential bus stops located at Type-2 students home locations to be used and other Type-1 students may walk to these bus stops as their bus stops.

Then, the SAPS (Student Assignment Problem Stop) model minimizing the number of bus stops in use is as follows.

Notations:

$$z_{ij} = \begin{cases} 1 \ if \ student \ i \ is \ assigned \ to \ bus \ stop \ j \\ 0 \ otherwise \end{cases}$$

$$u_j = \begin{cases} 1 \ if \ potential \ bus \ stop \ point \ j \ is \ in \ use \\ 0 \ otherwise \end{cases}$$

$$w_{ij} = \ walking \ distance \ from \ student \ home \ i \ to \ bus \ stop \ j$$

$$MD = maximum \ walking \ distance \ for \ \ student$$

$$P = total \ number \ of \ potential \ bus \ stop \ points$$

$$Q = total \ number \ of \ students$$

(SAPS)

$$Minimize \ \sum_{j=1}^{P} u_j \qquad (9)$$

$$\sum_{j=1}^{P} z_{ij} = 1 \qquad for \ i = 1, 2, \dots, Q \qquad (10)$$

$$w_{ij} z_{ij} \leq MD \qquad for \ i = 1, 2, \dots, Q \quad j = 1, 2, \dots, P \qquad (11)$$

$$z_{ij} \leq u_j \qquad for \ i = 1, 2, \dots, Q \quad j = 1, 2, \dots, P \qquad (12)$$

$$z_{ij} = 0 \ or \ 1 \qquad for \ i = 1, 2, \dots, Q \quad j = 1, 2, \dots, P \qquad (13)$$

$$u_j = 0 \ or \ 1 \qquad for \ j = 1, 2, \dots, P \qquad (14)$$

$$u_j = 1 \qquad for \ j = Type - 2 \ student \ home \ location \ bus \ stops \qquad (15)$$

Note that the differences between the SAPD and SAPS model introduced above are that the objective function and the upper bound of bus stops.

B. $\mathcal{NP}$-hard Problem

The size of the problem including the number of students, the number of potential bus stops and the number of Type-2 students that could be solved exactly and get global optimal solution is very limited, because SAP falls into a special case of the incapacitated facility location problem where only one customer point is school, and the set of potential facility points are the set of the

potential bus stops. Note the incapacitated facility location problem is $\mathcal{NP}$-hard, and here is the proof that SAP is $\mathcal{NP}$-hard.

The SAP is $\mathcal{NP}$-hard.

Proof. The $\mathcal{NP}$-hard proof of the SAP can be deduced from the node cover problem. Given a graph G and an integer n, does a subset of n nodes cover all the arcs of graph G? Note that if node k is an endpoint of arc a, then we say that node k covers arc a. The node cover problem is $\mathcal{NP}$-complete, and proof can be found in (Karp, 1972). Consider the set of potential bus stops S as the set of node, and the set of destination schools D as the set of arc. Then the SAP is defined by covering all the arcs D with minimum number of nodes S. In this way, the optimal solution of the SAP is equivalent to the node cover problem. Thus SAP is $\mathcal{NP}$-hard.□

C. SAP Instance

Here is the instance shows the different optimal solutions with two SAP models and their objectives.

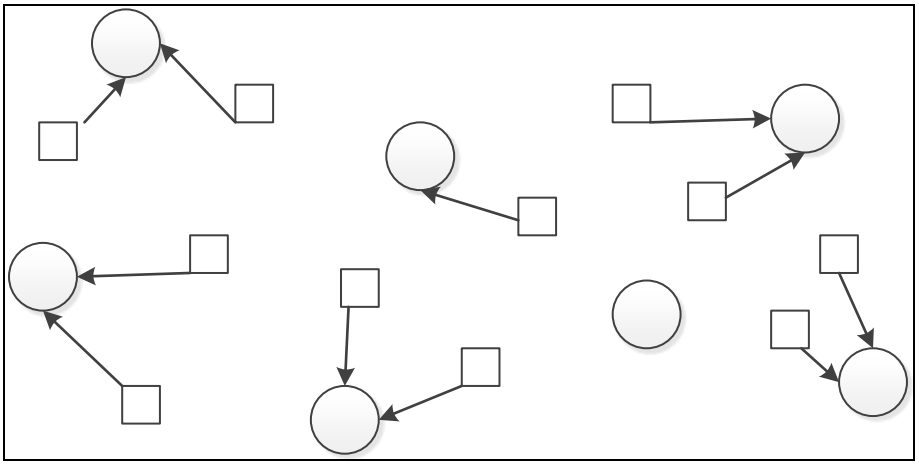

Figure 1. Minimize Walking Distance Solution

When the objective function is to minimize the total walking distance for all students, the solution shown in Figure 1 can be obtained. We can see that six of seven bus stops are in use and each single student is assigned to the nearest reachable bus stop.

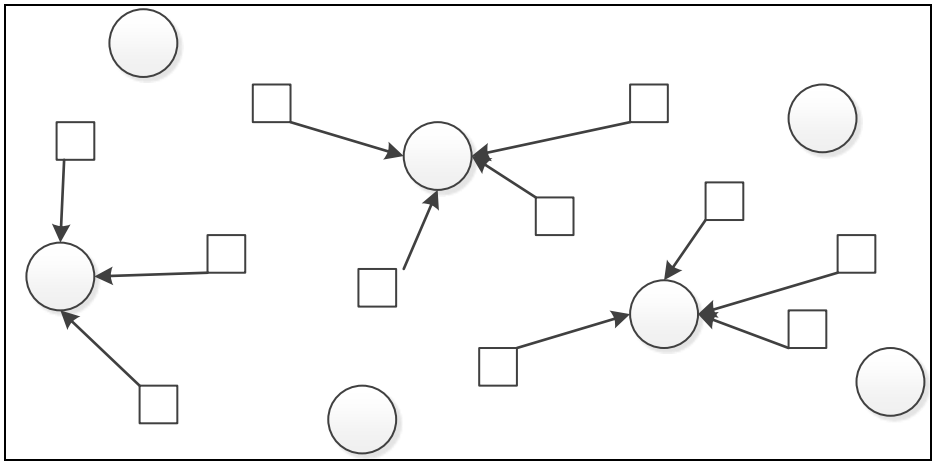

Figure 2. Minimize Number of Bus Stops Solution

When the objective function is to minimize the number of bus stops used,

the solution shown in Figure 2 can be obtained. Comparing to Figure 1, the number of bus stops is decreased to three with forcing students to the same bus stops. This is a more ideal solution for the later routing work because the total travel distance is shortened. Hence, the SAP model minimizing the number of bus stops in use may bring more benefits as long as the total walking distance, as a main measure of service quality, is acceptable.

D. Discussion

From the literatures discussed in Chapter 2, we can find that most works assume that the SAP solution is given and only several papers take SAP into consideration. The heuristic algorithms instead of mathematical model were widely used to construct and optimize solutions. All these heuristic algorithms can be classified into two strategies, the location-allocation-routing (LAR) strategy and allocation-routing-location (ARL) strategy (Park, 2010). For the LAR strategy, the determination of bus stops in use and assignment of students to these bus stops are decided before route generation. So the final optimal solution of SBRP is somehow decided by the condition SAP already provided. This is the drawback of LAR strategy. For the ARL strategy, clustering the students into groups within the bus capacity limit is the first step. Then SAP is done and SBRP starts for each cluster. The drawback of ARL strategy is how to make the policy of clustering. Simply using location gathering to reach the bus capacity as one cluster may lead to no feasible solution in SBRP. Especially, when the time restrictions such as travel time limit, school time window, are involved, there may be the situation that the bus serving the cluster that is furthermost from school destination cannot meet the schedule requirements.

So a general and effective cluster rule is required for the ARL strategy.

Finally in this chapter, we note that how previous work combines the SAP and SBRP, and why we divide them into two models and solve them separately. A combined model is proposed by the paper (Riera-Ledesma, 2012). The authors denote routing cost for buses and assigning cost for students. The objective function is to minimize the total routing cost plus the total assigning cost. Here the question is how to define the routing cost and, especially, the assigning cost. In our model, the student walking distance represents assigning cost and obviously it cannot be simply combined with the length of bus routes in the objective function without assigning weights to the distance students walk and the distance buses travel. It is a tradeoff between the student walking distance and the number of bus stops in use which affects the time spent at the bus stops by buses. Besides, it is possible to obtain better results for both SAP and SBRP if we apply different solving methods separately on SAP and SBRP with low complexity of model.

The results from SAP model decide whether a bus stop is selected or not and the assignment between the pair of each student, including Type-1 and Type-2, and the bus stop when the total students walking distance and the number of bus stops in use are minimized. These results would be used as the input data for the SBRP model we introduce in Chapter 4.

CHAPTER IV - SCHOOL BUS ROUTING PROBLEM

Using the solution from SAP as data, we can model the bus routing work with demand. At this stage of the modeling process of SBRP, the results would be the final optimal solution that includes the total cost, the number of buses, the routes, travelling distance, the time schedule for every bus, when and which bus stop each student should take the bus, and finally, the Type-2 student is served by the bus with handicapped facility. The objective in SBRP model is to minimize the total cost including three parts: the bus travelling cost, the bus fix cost and the additional cost for the bus with handicapped facility. The cost is one of the best measures of the SBRP model and it decides whether the theoretical optimal solution could be applied to practical use or for further decision making process.

A. SBRP Procedure

Before developing the SBRP model, we address data type requirements.

1. The parameters are generated from the SAP model, which are the number of bus stops in use and the number of students for each bus stop. They also identify the bus stop with Type-2 students and the bus with handicapped facility (denoted as Type-2 bus in the model).
2. Using the data regarding the bus stops in use and school locations, we could get the distance and bus travel time between each pair of locations and calculate the travelling cost matrix and the travelling time matrix.

3. We need the number of Type-1 buses and Type-2 buses available and some of them may not be used in the final optimal solution. Also, the capacities of buses are given and they can be a heterogeneous fleet.

4. We assume the fixed costs of Type-1 buses are proportional to the capacity and we predetermine a Type-1 bus fixed cost ratio. Also, we set an additional cost for Type-2 bus to calculate the fixed costs of Type-2 buses.

5. We apply different upper riding time limits for Type-1 and Type-2 students instead of the general maximum traveling distance limits.

6. We assume that the service time required at a bus stop for a Type-1 bus is positively correlated to the number of Type-1 students boarding at that stop. Similarly, more service time is added to the bus stop when picking Type-2 student up.

7. We assume the time window for school is given by the educational district, including the earliest and the latest arrival times for students at school to solve the morning problem. The time window is also referred to as the school bell time window and different schools may have different bell time windows. Most papers assume the school bell time window is given as scheduling constraints in the model. This is reasonable since in most cases the school bell time window is determined based on school policies and class schedules. In addition to the arrival schedule at every bus stop, route scheduling provides each route with the starting and ending times such that each bus reaches a school destination within the school bell time window.

Next, we introduce the SBRP model, and its mathematical formulation is developed as follows.

Notations:

$N = \ number\ of\ bus\ stops$

$P_1 = depot, all\ bus\ routes\ start\ here$

$P_2 \dots P_{n-1} = bus\ stops$

$P_n = \ school, all\ bus\ routes\ end\ here$

$NN = number\ of\ bus\ stops\ with\ Type - 1\ students$

$SN = number\ of\ bus\ stops\ with\ Type - 2\ students$

$B = number\ of\ buses$

$NB = number\ of\ Type - 1\ buses$

$SB = number\ of\ Type - 2\ buses$

$Ca_k = \ capacity\ of\ bus\ k$

$d_{ij} = \ travelling\ cost\ from\ stop\ i\ to\ j$

$t_{ij} = \ travelling\ time\ from\ stop\ i\ to\ j$

$y_{jk} = number\ of\ students\ picked\ up\ by\ bus\ k\ at\ stop\ j$

$S_j = total\ number\ of\ students\ assigned\ to\ stop\ j$

$A = arbitrary\ real\ numbers$

$TE = earliest\ arrival\ time\ for\ school$

$TL = latest\ arrival\ time\ for\ school$

$TA_{jk} = arrival\ time\ for\ bus\ k\ at\ stop\ j$

$TA_{nk} = arrival\ time\ for\ bus\ k\ at\ school\ n$

$TM = maximum\ riding\ time\ for\ Type - 1\ students$

$TMS = maximum\ riding\ time\ for\ Type - 2\ students$

$\alpha = Type - 1\ bus\ fixed\ cost\ ratio$

$\beta = Type - 1\ bus\ stop\ service\ time\ ratio$

$F = Type - 2\ bus\ additional\ fixed\ cost$

$L = Type - 2\ student\ additional\ service\ time$

$$x_{ijk} = \begin{cases} 1 & if\ bus\ k\ travels\ directly\ from\ stop\ i\ to\ j \\ 0 & otherwise \end{cases}$$

$$z_{jk} = \begin{cases} 1 \;\; if \;\; bus \;\; k \;\; picks \;\; up \;\; students \;\; at \;\; stop \;\; j \\ 0 \;\; otherwise \end{cases}$$

(SBRP)

$$Minimize \;\; \sum_{k=1}^{B}(\sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij}x_{ijk} + \alpha \times Ca_k) + F \times SB \qquad (16)$$

Subject to

$$\sum_{j=1}^{N} y_{jk} \leq Ca_k \qquad \forall k \qquad (17)$$

$$\sum_{k=1}^{B} y_{jk} = S_j \qquad \forall j \qquad (18)$$

$$\sum_{i=1}^{N-1} x_{ilk} = \sum_{j=2}^{N} x_{ljk} \qquad l = 2 \ldots (N-1) \qquad \forall k \qquad (19)$$

$$\sum_{i=1}^{N-1} x_{ijk} \geq z_{jk} \qquad j = 2 \ldots (N-1) \qquad \forall k \qquad (20)$$

$$\sum_{i=2}^{N} x_{jik} \geq z_{jk} \qquad j = 2 \ldots (N-1) \qquad \forall k \qquad (21)$$

$$y_{jk} \leq z_{jk} \times S_j \qquad \forall j \quad \forall k \qquad (22)$$

$$\sum_{k=1}^{SB} z_{jk} = 1 \qquad \forall j \in SN \qquad (23)$$

$$z_{1k} = 1 \qquad \forall k \qquad (24)$$

$$z_{Nk} = 1 \qquad \forall k \qquad (25)$$

$$x_{jjk} = 0 \qquad \forall j \quad \forall k \qquad (26)$$

$$\sum_{j=1}^{N-1} x_{jNk} = 1 \qquad \forall k \qquad (27)$$

$$\sum_{j=2}^{N} x_{1jk} = 1 \qquad \forall k \qquad (28)$$

$$A_i - A_j + N \times x_{ijk} \leq N - 1 \qquad for \;\; 2 \leq i \neq j \leq N \qquad \forall k \qquad (29)$$

$$TA_{ik} + \beta \times y_{ik} + t_{ij} = TA_{jk} + M\left(1 - x_{ijk}\right) \quad \forall i \in NN \quad \forall j \quad \forall k \epsilon NB \qquad (30)$$

$$TA_{ik} + \beta \times y_{ik} + L + t_{ij} = TA_{jk} + M\left(1 - x_{ijk}\right) \quad \forall i \in SN \;\; \forall j \quad \forall k \epsilon SB \qquad (31)$$

$$TA_{1k} = 0 \qquad \forall k \qquad (32)$$

$$TA_{jk} \leq M \times z_{jk} \qquad j = 2 \ldots (N-1) \qquad \forall k \qquad (33)$$

$$TA_{nk} - TA_{1k} \leq TM \qquad \forall k \in NB \qquad (34)$$

$$TA_{nk} - TA_{jk} \leq TMS \qquad \forall j \in SN \qquad \forall k \in SB \qquad (35)$$

$$TE \leq TA_{nk} \leq TL \qquad \forall k \qquad (36)$$

$$d_{1j} = 0 \qquad \forall j \qquad (37)$$

$$t_{1j} = 0 \qquad \forall \; j \qquad (38)$$

$$x_{ijk}, z_{jk} \ \ are \ \ binary \qquad \forall i, j, k \qquad (39)$$

$$y_{ik} \geq 0 \qquad \forall j, k \qquad (40)$$

In this model, the objective function (16) is to minimize the total cost including the total travelling cost for buses and the total fixed cost for both Type-1 buses and Type-2 buses. Note that the fixed cost of Type-1 buses is equal to Type-1 fixed cost ratio times the bus capacity. The fixed cost of Type-2 buses equals Type-1 fixed cost ratio times the bus capacity plus Type-2 bus additional fixed cost. The d_{ij} is decided by the bus travelling distance and we use the number of miles between two locations as d_{ij} directly. Note that the depots for all buses may not be the same locations and they could be the school bus parking lot or other parking lot. Hence we do not require all the buses to start their own route from the same depot or school and we set the distance and travelling cost between the depot and the first bus stop to visit as zero, and so is the travelling time. Constraint (17) is about the bus capacity limit and it is possible for different buses with different capacities. Constraint (18) ensures each student is picked up only once. Constraint (19) enforces the bus route because for bus stop l, any bus coming to visit bus stop l also leaves from bus stop l. Constraints (20) and (21) ensure that bus stop j is on the route of bus k (come and leave) if bus k picks up students at bus stop j. Constraint (22) guarantees two logical requirements. If the stop j is not served by bus k, the number of students picked up by bus k at stop j is zero. If stop j is served by bus k, the number of students picked up by bus k at stop j is less than or equal to the total number of students assigned to stop j. Constraint (23) asserts that the bus stop with a Type-2 student is served by the Type-2 bus. Constraints

(24) and (25) mean that depots and the school are on the all bus routes. Constraint (26) means that no bus travels any self-loop of a stop. Constraint (27) ensures that all bus routes end at school. Constraint (28) tells that all bus routes start at depots. Constraint (29) eliminates the subtour. Constraint (30) means that if bus k travels directly from stop i to j, the arrival time for bus k at stop j equals the sum of the arrival time for bus k at stop i, service time at stop i, and travelling time from stop i to j. Constraint (31) means that if a Type-2 student is served by Type-2 bus k from stop i to j, the service time for Type-2 bus k at stop i equals the sum of service time at stop i and the additional service time for Type-2 student. Constraint (32) makes the arrival time for any bus at depot is zero. Note that the travelling time from depot to any first bus stop is zero so the arrival time for the first bus stop following the depot is also zero. Constraint (33) means that if stop j is not visited by bus k, the arrival time is zero. Constraint (34) gives the maximum riding time limit for Type-1 students. Constraint (35) imposes the maximum riding time limit for Type-2 students. Constraint (36) promises the arrival time for any bus at school is within the school time window. Constraints (37) and (38) mean that the traveling distance and traveling time from depot to any first bus stop are zero. Constraint (39) is the binary constraint and constraint (40) makes sure that the number of students at every bus stop is greater than or equal to zero. Note that no integer constraint for y_{ik} because the constraint matrix is unimodular, so the optimal solution to the linear programming (LP) relaxation will assign all variables integer values and would therefore be the optimal solution to the integer programming (IP). The proof can be referred to (Shapiro, 1979).

B. $\mathcal{NP}$-hard Problem

For the SBRP model, the size of the problem that includes the number of students, the number of Type-2 students, the number of bus stops, and the number of buses, and can be solved exactly and reach global optimal solution is much more limited, because SBRP falls to Vehicle Routing Problem (VRP). Here is the simple proof that SBRP is $\mathcal{NP}$–hard.

The SBRP problem is $\mathcal{NP}$ -hard.

Since SBRP is a sub problem of VRP, and VRP is $\mathcal{NP}$ -hard (Lenstra, 1981), SBRP is also $\mathcal{NP}$ -hard. □

C. SBRP Instance

The following figures show the solution of the example with 45 students, 14 bus stops, three buses and two Type-2 students.

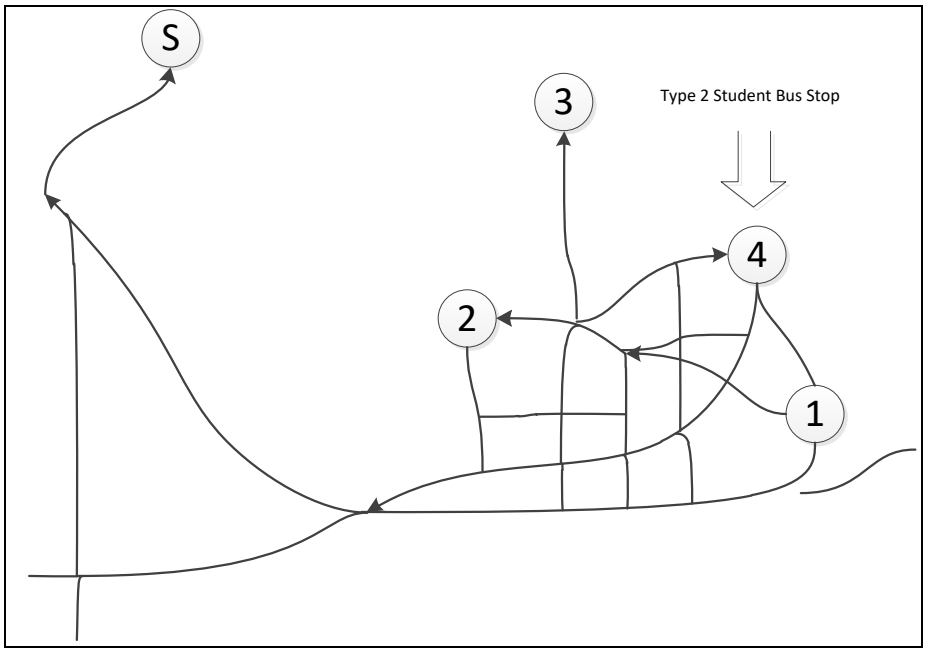

Figure 3. One Type-2 student four bus stops route

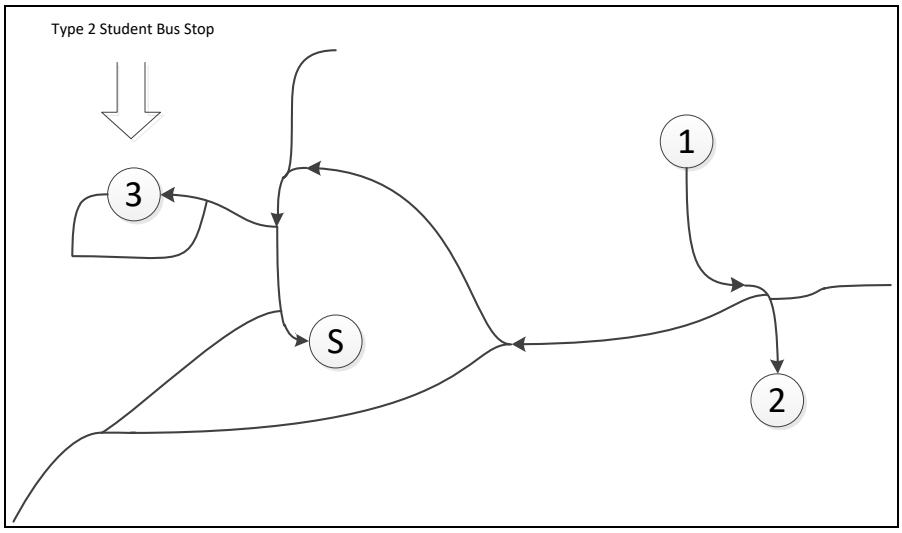

Figure 4. One Type-2 student three bus stops route

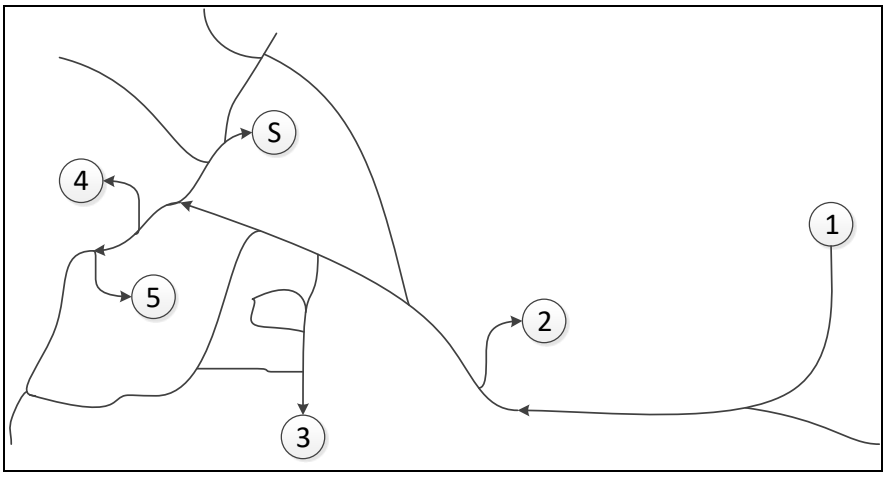

Figure 5. No Type-2 student five bus stops route

In this example, bus #1 and bus #2 are equipped with special handicapped facility (Type-2 buses), and the policy is that at most 1 Type-2 student can board on one bus. The arrows in Figure 3 and 4 point out the bus stops with Type-2 students. Thus, the bus stop with the first Type-2 student is assigned on the route of bus #1 and the bus stop with the second Type-2

student is assigned on the route of bus #2. We assume that no difference exists between two Type-2 students, bus #1 and #2. The only difference among all bus stops is that two of them are for the Type-2 students and the remaining bus stops are for the Type-1 students only. Note that the bus stops with Type-2 students are not only for the Type-2 students. That is, Type-1 students can walk from home locations to take their buses with the Type-2 students. The Figure 5 illustrates the route of bus #3. Observing the solution, we can easily find out two points. One is that bus #3 serves more bus stops than bus #1 and bus #2 separately. This is due to the time restriction. The service time at bus stop for the Type-2 student is much longer than for the Type-1 student, and the total travel time of each bus must be within the time window. It is obvious that bus #3 should serve more students and stops because the students onboard are all Type-1 students. Another important point is that the bus stops with Type-2 students are the last stops on the routes before school. The reason is still about the time restriction on the Type-2 student travel time. The upper travel time limit for the Type-2 students is much lower than for the Type-1 students, thus it is beneficial to shorten the distance between the bus stop with the Type-2 students and the school destination. In particular for the Figure 3, the locations of bus stop 2, 3 and 4 construct an approximate triangle shape, and bus #1 travels from bus stop 1 to 2 first, then to 3 second, and finally 4. In fact, the location of bus stop 4 is the nearest one to bus stop 1 compared to stop 2 and 3, and it decreases the travel distance and time if bus 1 travels to bus stop 4 first, 3 second, and finally 2. However, this route forces bus #1 to pickup Type-2 student first at bus stop 4, and transports him/her to 3 and 2. It violates the restriction on the travel time of Type-2 student, and the solution becomes

infeasible.

Figure 6 is another example with 30 students, 14 bus stops, 2 buses and 1 Type-2 student. It shows the effect of travel time limit clearly. In this example, one Type-2 student needs to take the bus at bus stop 5, which is on the way of bus travelling from stop 2 to 3 and 4. Because of the travel time limit, the bus has to turn back from the bus stop 4, pick up the Type-2 student at stop 5, and finally travel directly to the school destination. The dash line is the route of the other bus.

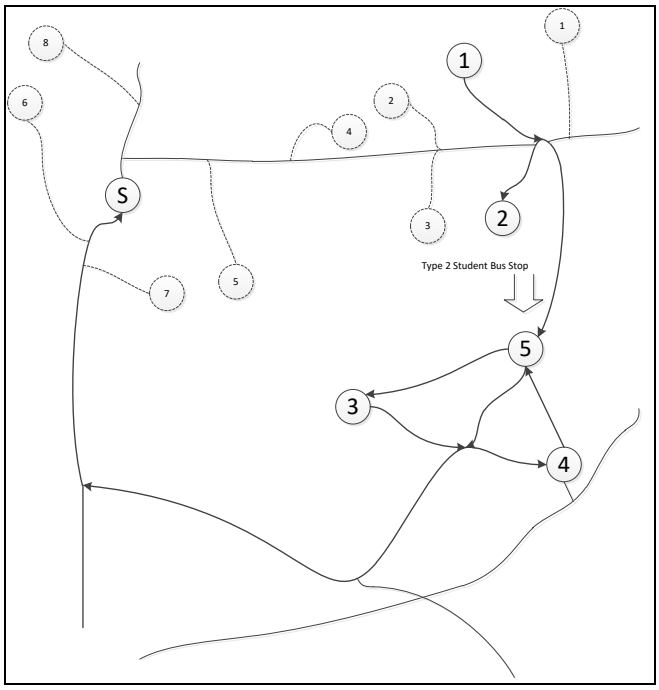

Figure 6. The turn back route of Bus

D. Multi-objective Model

From a practical standpoint, the service quality is a significant concern for public service, in particular for school bus transportation authorities. Other than the safety, equity is also another main performance measure of service quality. The equity consideration is about the fairness of all students, and it is usually an improvement and modification to the cost consideration. The load balance

and length balance are two most commonly used criteria and can be used to regard SBRP as a new multi-objective model. The idea of load balance is to balance the number of students on each single bus route so that all students are nearly evenly assigned to buses. It is a poor load balance situation if some buses are almost fully loaded, while other buses pick only few students. The idea of length balance is that the difference among all buses travel distance is small so that every student spends near equal time on board. It is a poor length balance if some buses travel much further with much longer time in comparison to other buses.

The SBRP model presented at the beginning of this chapter can be further enhanced by adding the equity objectives to a multi-objective model. The objective functions are as follows,

$$Minimize \quad \sum_{k=1}^{B}(\sum_{j=1}^{N} S_j \, z_{jk})^2 \tag{41}$$

$$Minimize \quad \sum_{k=1}^{B}(\sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij} \, x_{ijk} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij} x_{ijk}}{B})^2 \tag{42}$$

In the objective function (41), the variation in the number of students served on every bus route is minimized by adding the square of the number of students on each route. In the objective function (42), the variation in the route length of each bus is minimized by adding the square of the difference of each bus route length and the average bus route length.

E. Discussion

At the end of this chapter, we explain more details about the SBRP model. First, it is possible that multiple buses serve one bus stop which implicates that

the students at the same bus stop may not take the same ride. Second, in order to solve bigger size problem, for the instances tested above, we set the allowable maximum riding time for Type-1 student is 60 minutes, for Type-2 student is 30 minutes, the school time window is within 0 to 60 minutes, and the bus stop service time ratio is 0.5. Third, the instances are the morning cases about Goshen Elementary School in Oldham County School District. We pick bus stops from all the bus stops listed and measure the road distance and travel time between each pair of origin and destination.

CHAPTER V - SOLUTION APPROACH

From the computing results in Chapter 6, the sizes of the instances which can be solved directly with optimization software are very limited. From the literature, many works applied metaheuristics as tools to solve SBRP and got near optimal solutions within acceptable computing time. The computing time of solving SBRP may not be an important concern in practice because SBRP only needs to be solved one time for one whole academic year. However, if the quality of the solution is important and it is necessary to keep improving, solving SBRP in a short time brings benefits. For practical use purposes, the solutions from the computer programs should be revised manually under the policies from the education district department. The constraints may need to change over time until a good solution under the policies is obtained. In this way, how quickly SBRP can be solved and different solutions can be reached determine the efficiency of the whole decision making process.

In this study, we develop a Genetic Algorithm (GA) to achieve feasible solutions as initial columns, and then apply a Column Generation (CG) technique to improve the quality of the solutions. Next, we introduce each of them and present a combined strategy to solve a large practical problem within a reasonable time limit.

A. Genetic Algorithm

GA has been widely used as a robust heuristic method to find a feasible solution efficiently, and the basic steps are as follows. Basically, GA imitates the process of natural evolution, and by evaluating, selecting, breeding and filtering within randomly generated candidate solutions to obtain the optimal solution. In general, GA has the following basic steps.

1. Initialization: At the start of an algorithmic procedure, a certain number of individual solutions are randomly generated and they comprise the initial population.

2. Evaluation: A fitness score is assigned to rate each individual solution. Such a fitness score should reflect: a) if the solution satisfies the constraints of the optimization problem under consideration; b) how well the solution optimizes the objective function.

3. Selection: Existing individuals (solutions) are subsequently selected as parents to reproduce offspring, i.e., new solutions. Inspired by human evolution, one would make those solutions with good fitness scores have higher probability to be selected for reproduction. Individuals, or solutions, from one generation to the next, are likely to improve in terms of optimizing the objective value. Ideally, when enough generations pass over, the best solutions of the last generation should converge to the optimal solution. In short, GA often assigns those solutions with good fitness scores larger probability of being selected for reproduction, and those with poor fitness scores smaller probability. Note that in order to promote diversity for reaching

global optima, GA usually does not assign zero probability to individuals with poor fitness scores.

4. Reproduction: Once parents are selected, the next step is to generate offspring solutions to form the next generation, and two primary methods of creating next generations are crossover and mutation. The crossover process typically takes a set of alleles from one of the parents, and then switches them with the alleles of the other parent. It allows a new solution to share many characteristics of its parents and gathers them into one better solution as the offspring. Mutation on the other hand simply changes some alleles of one of the parents. The goal of mutation is to increase diversity of the population, which can effectively avoid trapping at local optima. The exchange mutation proceeds with probability. The times of mutation and the two points to exchange are also randomly selected. In many GA practices, it is not necessary to eliminate both parents once they reproduce offspring. The elitism strategy always keeps a certain percentage of the best solutions within current generation to be carried over to the next generation.

5. Termination: The process of Steps 2-4 repeats until the user-defined termination conditions are satisfied. Common terminating conditions are: a) a solution that satisfies some minimum criteria is found; b) a fixed number of generations is reached; c) a fixed amount of computation time passes; and d) successive iterations no longer produce better results.

B. Genetic Algorithm for SBRP

In this section, we introduce our customized GA in C# for the SBRP by going through one example with 30 students, 10 bus stops including depot

and school and 5 buses. In this example, two Type-2 students are assigned to the bus stop #2 and #6 from the SAP solution. Two Type-2 buses, bus #4 and #5, are required to serve Type-2 students and at most one Type-2 student can be onboard for one Type-2 bus.

B.1. Encoding

The SAP solution is encoded as an array with integer items. Total 30 students are encoded as $[0,1,2,3,4,5,6,5,4,0]$, which means that one student is assigned to stop #1, two students are assigned to stop #2, five students are assigned to stop #5, six students are assigned to stop #6, five students are assigned to stop #7 and so on. The first item is always depot and the last stop, the ninth item in this example, is always school. In these 30 students, 28 Type-1 students and two Type-2 students are encoded separately as $[0,1,1,3,4,5,5,5,4,0]$ and $[0,0,1,0,0,0,1,0,0,0]$. There are 10 items in the arrays indicating that 10 bus stops are in use. Also, the first item and the last item should always be zero because no student is assigned to the depot and the school. Comparing to the encoded SAP solution, 10 bus stops are numbered and encoded as $[0,1,2,3,4,5,6,7,8,9]$. The depot is numbered as 0 and the school is numbered as 9. The bus route in the population is encoded as $[0,1,2,9]$ for example. This means that the bus route starts travelling from depot to bus stop #1, then from bus stop #1 to #2, then from bus stop #2 to the school. Note that the first item should be depot, which is encoded as zero, and the last item should be school for every feasible route. The virtual bus depot is always numbered as stop #0 for all the routes and it is always the first item in each route. When numbering all the bus stops in order, the school is

numbered as the last one because school is always the last stop for all routes. In this way, the largest integer item in the array represents school, which is the last item in each array. The depot is set up for the modeling purpose, and the real depot locations for all bus routes may not be the same, and buses can start the route from anywhere as a virtual depot. Five buses are numbered and encoded as $[1,2,3,4,5]$, in which the Type-1 buses are encoded as $[1,1,1,0,0]$ and the Type-2 buses are encoded as $[0,0,0,1,1]$. This means that bus #1, #2 and #3 are Type-1 buses and bus #4 and #5 are Type-2 buses. The capacity of buses are encoded as $[8,8,8,7,7]$. Thereby, the capacities of three Type-1 buses, #1, #2 and #3, are eight students. The capacities of two Type-2 buses, #4 and #5, are seven students.

B.2. Initialization

When starting the initialization step, the number of available buses, and the bus stop numbers including depot and school are given. Also, the student assignment solution from SAP is given. and also encoded as an array with integer items. The integers provide the number of students assigned to the corresponding bus stops in order. Again, the first and the last integer items are always zero because no students can be assigned to the virtual depot or school. In addition, the SAP solution includes both the Type-1 and the Type-2 student assignment solutions and they are encoded in arrays separately. In short, the integer items in the array recording SAP solution are the number of students assigned, while the integer items in the array of route are the bus stop numbers. In the initialization step, the Type-2 bus routes serving Type-2 students are generated first as Figure 7 shows in the flowchart.

The approach starts at reading Type-2 stops from the Type-2 students SAP solution array $[0, 0, 1, 0, 0, 0, 1, 0, 0, 0]$. After this step, bus stop #2 and #6 are marked as Type-2 stops. Then the travelling time from stop #2 and #6 to the school (t_{29} and t_{69}) are calculated. If any travelling time is beyond the maximum riding time for Type-2 student (TMS), then no feasible solution exists because it is impossible to transport the corresponding Type-2 student to the school meeting the maximum riding time for Type-2 student constraint. In order to find a feasible solution, the maximum riding time for Type-2 student (TMS) needs to increase until both t_{29} and t_{69} are less than the new TMS.

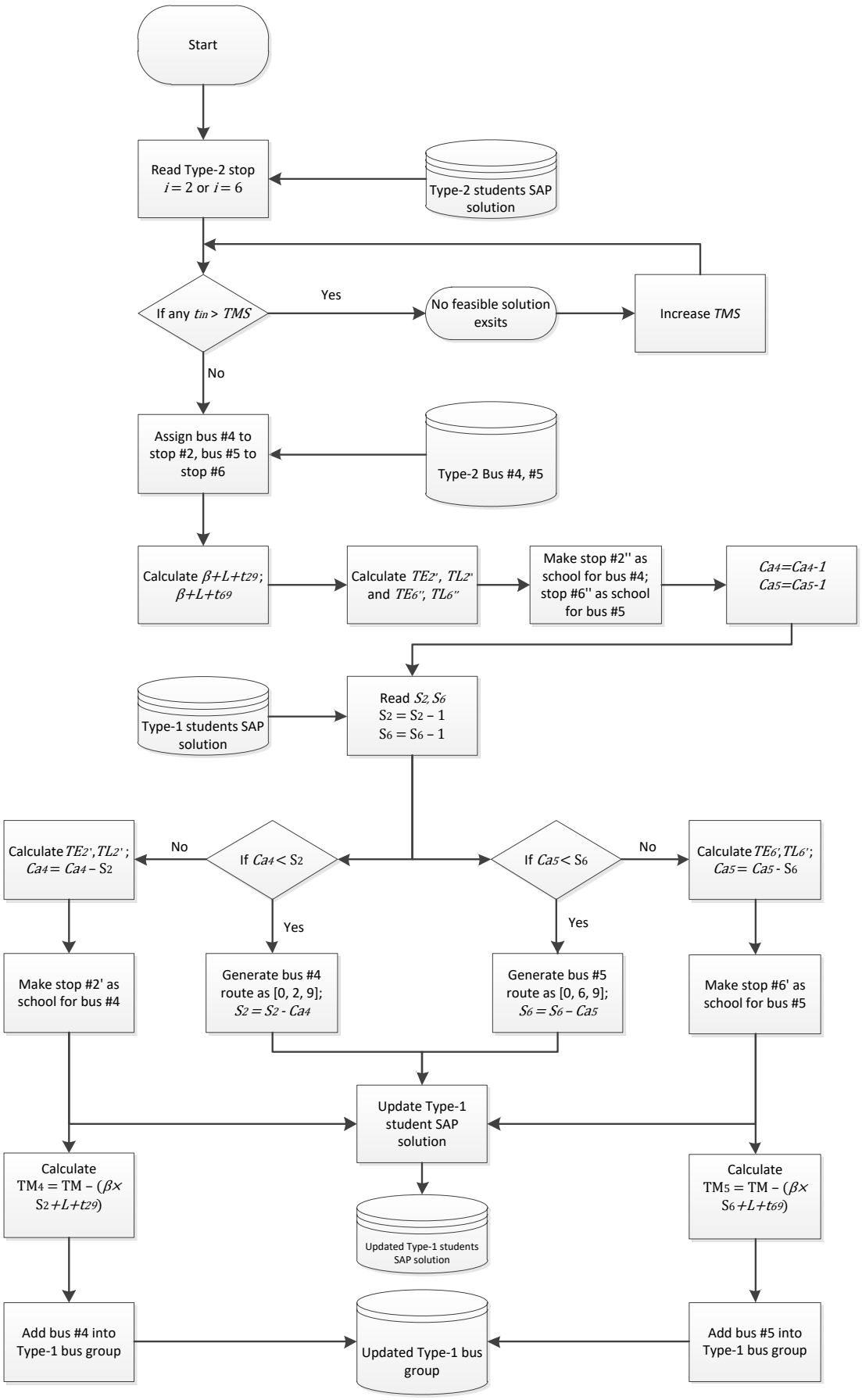

Figure 7. Type-2 student initial route generate procedure

When both t_{29} and t_{69} are less than the TMS, all Type-2 buses are read from the array $[0, 0, 0, 1, 1]$. Bus #4 is arbitrarily assigned to serve stop #2, and bus #5 is assigned to serve stop #6. Then we pick stop #2 to continue the procedure and the same approach will be used on stop #6 also. The sum of the service time for Type-2 student ($\beta \times 1 + L$) and the travelling time of transporting the Type-2 student to the school (t_{29}) is calculated. Note that if there are Type-1 students assigned to Type-2 stop #2 along with the Type-2 student, bus stop #2 is divided into two separate virtual stops #2' and #2''. The Type-2 student is assigned to stop #2'' and the Type-1 students are assigned to stop #2'. The travelling time between stop #2' and #2'' ($t_{2'2''}$) is zero. In this way, we can generate the route for bus #4 from back to front and the Type-2 student is picked up only before travelling directly to the school. This modeling technique is useful since the maximum riding time for a Type-2 student (TMS) is usually critical and the additional service time for a Type-2 student (L) is much longer, and in turn consumes a significant part of the maximum riding time limit for other Type-1 students on board. It also more likely to generate feasible routes in limited computational times. When the last two stops of bus #4's route are decided as $[2'', 9]$, we can treat stop 2'' as the new school destination for bus #4 with updated time windows ($TE_{2''}$ and $TL_{2''}$). Here $TE_{2''} = TE - t_{29} - (\beta \times 1 + L)$ and $TL_{2''} = TL - t_{29} - (\beta \times 1 + L)$. The capacity of bus #4 should also be updated as $Ca_4 = Ca_4 - 1$. As mentioned before, the time windows of Type-2 bus stop #6 ($TE_{6''}$ and $TL_{6''}$) and the capacity of bus #5 (Ca_5) are also updated. The last two stops of bus #5's route are also decided as $[6'', 9]$.

After both Type-2 students are served, it is reasonable to determine if bus #4 and #5 can serve other Type-1 students assigned to Type-1 bus stop #2' and #6'. Based on the total number of the Type-1 students assigned to stop #2 ($S_2 - 1$) and #6 ($S_6 - 1$) from Type-1 students SAP solution array $[0, 1, 1, 3, 4, 5, 5, 5, 4, 0]$, we pick stop #2' to continue the procedure and the same approach will be used on stop #6' also.

If the updated capacity of bus #4 (Ca_4) is less than the total number of Type-1 students assigned to stop #2' ($S_2 - 1$), bus #4 picks as many students as it can (Ca_4), and then travels directly to school. Bus #4's route is generated as $[0, 2, 9]$. The total number of students assigned to stop #2' ($S_2 - 1$) is updated as $S_2 = S_2 - 1 - Ca_4$.

If the updated capacity of bus #4 (Ca_4) is not less than the total number of Type-1 students assigned to stop #2' ($S_2 - 1$), we make Type-1 stop #2' as the school destination for bus #4 and calculate the time windows of stop #2' (TE_2, and TL_2,). The capacity of bus #4 (Ca_4) is updated as $Ca_4 = Ca_4 - (S_2 - 1)$. The maximum riding time for a Type-1 student on bus #4 (TM_4) is calculated as $TM_4 = TM - [\beta \times (S_2 - 1) + L + t_{29}]$. Finally, bus #4 is added into Type-1 bus group with the updated capacity (Ca_4) and the updated maximum riding time for a Type-1 student (TM_4).

After routing Type-2 students, we start routing all Type-1 students with updated information such as Type-1 students SAP solution, Type-1 bus group, school destinations, bus capacities, and maximum riding time for a Type-1 student. The flowchart in figure 8 shows the route generation procedure.

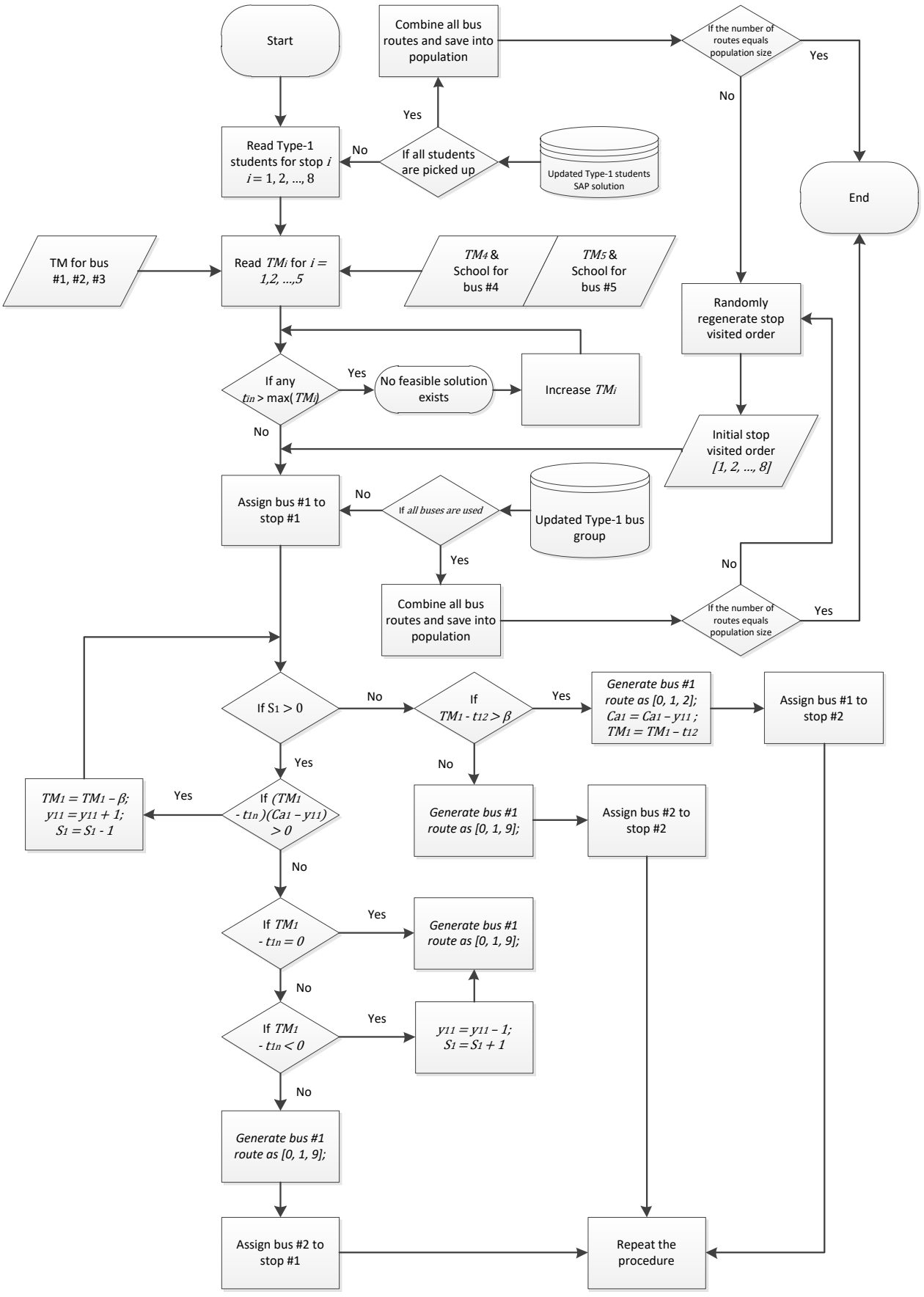

Figure 8. Type-1 student initial route generate procedure

First, we use Type-1 bus stops ($i = 1, 2, \ldots\ 8$) with the travelling time from every stop to the school (t_{in}), and Type-1 buses' ($j = 1, 2, \ldots\ 5$) maximum riding time for a Type-1 student (TM_j). If there exists any $t_{in} > \max(TM_j)$, which means that it is impossible for any bus to transport the Type-1 students from the corresponding bus stop to the school meeting the maximum riding time for a Type-1 student constraint, no feasible solution exists and the $\max(TM_j)$ must be increased. If any $t_{in} \leq \max(TM_j)$, we start the procedure by routing bus #1 from depot to the first stop in the initial visit order (stop #1). Note that the visit order will be regenerated randomly every time the procedure starts later.

Next, bus #1 starts to pick up the Type-1 students at stop #1. Once a student is picked up by bus #1, the maximum riding time (TM_1) is updated as ($TM_1 = TM_1 - \beta$), and the number of students picked up by bus #1 at stop #1 (y_{11}) is updated as ($y_{11} = y_{11} + 1$). The pick-up process repeats if the number of Type-1 students left on stop #1 $S_1 > 0$ and $(TM_1 - t_{1n})(Ca_1 - y_{11}) > 0$. If $(TM_1 - t_{1n})(Ca_1 - y_{11}) \leq 0$, it means that $TM_1 - t_{1n} \leq 0$ or $Ca_1 - y_{11} = 0$ because both Ca_1 and y_{11} are integers and y_{11} increments by one during the process. In the case of $TM_1 - t_{1n} = 0$, bus #1 is fully loaded and travels from stop #1 to the school meeting the maximum riding time (TM_1) exactly. Bus #1's route is generated as $[0, 1, 9]$. In the case of $TM_1 - t_{1n} < 0$, the maximum riding time (TM_1) limit is reached first and bus #1 has to travel from stop #1 to the school with empty seats. The number of students picked up by bus #1 is updated as $y_{11} = y_{11} - 1$ and the number of students left is updated as $S_1 = S_1 + 1$. Bus #1's route is generated as $[0, 1, 9]$.

In the case of $Ca_1 - y_{11} = 0$, bus #1 is fully loaded and there are more Type-1 students left at stop #1. Bus #1's route is generated as $[0, 1, 9]$ and we start the same procedure by routing bus #2 from depot to stop #1.

When $S_1 = 0$, which means that all Type-1 students at stop #1 are picked up by bus #1. The sum of the travelling time from stop #1 to #2 and the service time for one Type-1 student is calculated. If the updated TM_1 is more than the sum, we route bus #1 from stop #1 to #2 and the route is generated as $[0, 1, 2]$. The capacity of bus #1 is updated as $Ca_1 = Ca_1 - y_{11}$. The maximum riding time (TM_1) limit is updated as $TM_1 = TM_1 - t_{12}$. At bus stop #2, we start the same procedure again for bus #1 with the updated capacity Ca_1 and the updated maximum riding time (TM_1). If the updated TM_1 is not more than the sum, we generate bus #1's route as $[0, 1, 9]$ and route bus #2 from depot to stop #2. The same procedure is started on bus stop #2 for bus #2.

When the procedure starts again, the number of available buses and the number of Type-1 students left are checked. If all buses are used or all students are picked up, we combine all bus routes generated in one array and save it into the initial population. If there are still available buses and Type-1 students left, we repeat the procedure until no available bus or no student left.

When the routes generated are combined and saved into the initial population, the total number of solutions in the initial population is checked. If the total number of solutions reaches the population size we set as 1000, the initialization step ends. If the total number of solutions is less than the population size we set, the stop visiting order is randomly regenerated and the data including Type-1 students SAP solution, TM for bus #1, #2 and #3, TM_4 and TM_5, and school destinations for bus #4 and #5, is restored.

B.3. Evaluation

The pseudo-code for the GA is described in Algorithm 1. After generating the initial population with 1000 solutions, we calculate the total travelling distance for each solution based on the route generated and the travelling distance matrix. In our GA, The total traveling distance is represented as z. A binary variable $F = 0$ if the route is feasible, and $F = 1$ otherwise. The objective value (T) is calculated as $T = z + F \times P$, where P is a very large number of penalty. The fitness score is represented by $f = 1/T$, and the fitness scores of good feasible solutions are high, and the fitness scores of infeasible solutions are very low.

B.4. Selection

In our GA, we apply a rule of best keeping with the roulette wheel selection in order to choose solutions as parents to mate and produce offspring. For each generation, all N solutions are ranked from largest to smallest based on fitness score. The first solution with a largest fitness score is passed on from the parent generation to the offspring generation, and the rest $N - 1$ new offspring solutions are generated by reproduction. The roulette wheel method probabilistically selects individuals according to their fitness scores. The probability of an individual being selected (P_s) is given by the equation (43):

$$P_s = f_i/(\sum_{i=1}^{N} f_i) \quad (43)$$

B.5. Crossover

The probability of crossover (P_c) is set as the parameter of GA. For each selected pair of parents, if the random generated number (r) between 0 and 1 is less than or equal to P_c, we crossover the selected parent 1 and 2, and insert the offsprings to the next generation population $P(t+1)$.

Two-point crossover is applied with random two points (i and j) in GA. All alleles between i and j on the parent strings are swapped each other. Note that the range of the randomly generated two points (i and j) is $[1, N-1]$, where N is the number of bus stops. In this way, we ensure that the first allele is always the depot #0 and the last allele is always the school #N for all the solutions. Also, the Type-2 bus stop is the last stop before the school in each route at the initial generation step. However, this is not the case when the point j falls on the allele which represents the Type-2 bus stop because the next allele, which represents the school, is not swapped to the offspring. This design is beneficial and may lead to better solution because it brings the possibility that stops may be ordered between the Type-2 bus stops and the school. At the same time, this design limits the probability of inserting bus stops between the Type-2 bus stops and the school because the maximum travelling time limit for Type-2 students (TMS) is typically violated.

B.6. Mutation

The probability of mutation (P_m) is also set as the parameter of GA. For each selected pair of parents, if the randomly generated number (r) between 0 and 1 is more than the crossover probability (P_c), no crossover is applied between the selected pair of parents. Then two random number (x and y)

between 0 and 1 are applied to determine whether the mutation is applied. If x (or y) is less than or equal to the mutation probability (P_m), we mutate parent 1 (or 2) by exchanging the alleles on two random generated points (i and j), and insert the new solution to the next generation population $P(t+1)$. If x (or y) is more than the mutation probability (P_m), we do not mutate and copy parent 1 (or 2) directly to the next generation population $P(t+1)$.

Note that the crossover and the mutation will not be applied on the one pair of parents at the same time. Also, the selected pair of parents may be copied directly to the next generation population without any crossover or mutation. This facilitates maintaining good quality solutions with higher fitness scores in the population.

B.7. Termination

We limit the number of generations as 10,000 in the solving process. In the 10,000th generation population, the best solution with the highest fitness score is the final optimal solution.

After each run of the GA, if a feasible solution is found, it is then used as initial columns for the CG procedure as described in the next section.

Algorithm 1: GA Algorithm for SBRP

Genetic Algorithm

Data: *every generation population scale* (M); *fixed number of generations* (T); *crossover probability* (P_c); *mutation probability* (P_m); *number of bus stops* (N);
Result: *solution with highest fitness score*;
Initialization: *initial population* $P(0)$; *current generation* $t = 0$;
Calculate fitness score of each solution in $P(0)$;
While $t \leq T$ ***do***
    Copy the solution with highest fitness score to $P(t+1)$;
    Calculate the selection probability P_s *of each solution*;
    for $k = 0; k < M; k = k + 2$ ***do***
        Select parent 1 & 2 *from* $P(t)$;
        $r = random[0,1]$;
        if $r \leq P_c$ ***then***
        $i = random[1, N-1]$; $j = random[1, N-1]$;
            Crossover parent 1 & 2 *and insert offsprings to* $P\ (t+1)$;
        else
            $x = random[0,1]$;
            if $x \leq P_m$ ***then***
                $i = random[1, N-1]$; $j = random[1, N-1]$;
                mutate parent 1 *and insert new solution to* $P(t+1)$;
            else
                copy parent 1 *to* $P(t+1)$;
            end
            $y = random[0,1]$;
            if $y \leq P_m$ ***then***
                $i = random[1, N-1]$; $j = random[1, N-1]$;
                mutate parent 2 *and insert new solution to* $P(t+1)$;
            else
                copy parent 2 *to* $P(t+1)$;
            end
        end
    end
    Calculate fitness score of each solution;
end
Print the solution with highest fitness score in final generation;

C. Column Generation

The basic idea of Column Generation (CG) is the decomposition with the master model (or the master problem) and the submodel (or the pricing problem). The pseudo-code of the CG from (Stuber, M. 2010) is shown below.

Algorithm 2: Column generation approach to solve SBRP

Column Generation

Input: S *set of students,* Ca *be the capacity of buses and* r_{max} *be the maximum regret*

Output: γ' *the set of routes*

Step 0: *put in* γ' *all trivial routes*

Step 1: *solve master problem and get* $s_i = dual\ prices$
if $s_i = 0, for\ all\ i$ ***then***
break
else
go to step 2
end if

Step 2: *solve pricing problem and get* R *the new route*

Step 3: *add* R *to* γ', x_R *to the set of columns*

Step 4: *go to step* 1

In CG, all the available bus routes are treated as the variables and they are different from the variables in SBRP model. The main advantages of using CG is starting with only a subset of variables to solve the new problem and find the solution, and then determining one more part from the rest of the original problem to take into consideration. The CG algorithm keeps repeating this process in iterations until one (near) optimal solution is reached.

In the algorithm shown above, γ' denotes the set of bus routes to be considered, and the maximum regret r_{max} represents the most acceptable difference of the times between transported by school bus and transported by parents for students. The CG algorithm starts with all trivial routes with only one student each.

With γ', the master problem can be solved and the dual prices s_i can be obtained for all i constraints of the master problem. If all s_i are non-negative, the CG algorithm stops and the solution from the master problem is

optimal. If any negative dual price exists, the pricing problem is solved and a variable with negative dual price is found. This variable R representing the new route will be added to γ' and the master problem will be solved again. The CG algorithm is repeated with the set of bus routes to be considered γ', the master problem, negative dual price s_i, the pricing problem and new updated γ' until a solution having non-negative dual price.

D. Column Generation for SBRP

We utilize CG to effectively solve a large scale problem in the constraint matrix as in SBRP. All initial feasible columns are generated by GA and populated in an initial column pool for a CG procedure.

D.1 Master Problem

Next, we present a restricted master problem (RMP) and a pricing subproblem. Note that the decision variables in master problem are the routes not arcs, and also every student is considered instead of stop.

Notations:

$$x_r = \begin{cases} 1, & if\ route\ r\ is\ in\ the\ solution \\ 0, & otherwise \end{cases}$$

R: $the\ set\ of\ all\ feasible\ routes$

B: $the\ set\ of\ bus\ Type-1\ and\ Type-2$

H: $the\ set\ of\ bus\ Type-2\ only$

S: $the\ set\ of\ student\ Type-1\ and\ Type-2$

T: $the\ set\ of\ student\ Type-2\ only$

$c_r = cost\ of\ route\ r$

$$\delta_{sr} = \begin{cases} 1, & if\ student\ s\ is\ served\ by\ route\ r \\ 0, & otherwise \end{cases}$$

$$\mu_{br} = \begin{cases} 1, & if\ bus\ type\ b\ is\ used\ in\ route\ r \\ 0, & otherwise \end{cases}$$

(RMP)

$$Minimize \ \sum_{r \in R} c_r x_r \qquad (44)$$

Subject to

$$\sum_{r \in R} \delta_{sr}\, x_r = 1 \qquad \forall s \in S \qquad (45)$$

$$\sum_{r \in R} \mu_{br} x_r \leq 1 \qquad \forall b \in B \qquad (46)$$

$$\delta_{sr}(1 - \mu_{br}) = 0 \qquad \forall s \in T \quad \forall b \in H \quad \forall r \in R \qquad (47)$$

$$x_r \in \{0,1\} \qquad \forall v \in R \qquad (48)$$

The objective function (44) is to minimize the cost of route selection. The cost of route,c_r, is defined by the fixed cost of using bus and the traveling cost in proportion to distance. Constraints (45) assure that all students are served and visited exactly once. Constraints (46) restrict that one bus cannot serve more than one route. Constraints (47) enforce that a Type-2 student must be served by a Type-2 bus. Decision variables are binary in Constraints (48).

Remark 1. There are a couple of advantages of considering every student in lieu of bus stops in the decision variable δ_{sr}. First, the model allows that the solutions of a bus stop on multiple routes as well as a student riding different buses can be evaluated and compared. This rule enables potentially reaching a far better solution, in particular, when considering one stop with large number of students more than a bus capacity. Second, the quality of solution is not affected by any requirement resulting from grouping students. This obviates a need to investigate different grouping conditions leading to

different final solutions. However, the disadvantage is that it can increase the complexity of the problem. In our model, each student is counted as a stop and the traveling distance and time between two students, if they are located at the same bus stop, are effectively set as zeroes.

D.2 Sub problem

In the sub problem (SP), the decision variables are associated with arcs, and the constraints includes capacity, scheduling, and other restrictions.

Notations:

$$x_{ij} = \begin{cases} 1 & if\ arc\ (i,j), i,j \in N\ is\ used\ in\ the\ route \\ 0, & otherwise \end{cases}$$

d_{ij}: $travelling\ cost\ of\ arc\ (i,j), i,j \in N$

t_{ij}: $travelling\ time\ of\ arc\ (i,j), i,j \in N$

c_b: $capacity\ of\ bus\ type\ b, b \in B$

f_b: $fixed\ cost\ of\ bus\ type\ b, b \in B$

B: $the\ set\ of\ bus\ types$

S: $the\ set\ of\ students$

N: $S \cup \{n; school\ node\}$

a_i: $bus\ arrival\ time\ at\ student\ i, i \in S$

a_n: $bus\ arrival\ time\ at\ school$

e: $earliest\ arrival\ time\ for\ school$

l: $latest\ arrival\ time\ for\ school$

m_i: $maximum\ riding\ time\ limit\ for\ student\ i, i \in S$

s_i: $service\ time\ of\ student\ i, i \in S$

$\pi_i = dual\ variable\ corresponding\ to\ student - route\ assignment$

$constraint\ (45), i \in S$

$\gamma_b = dual\ variable\ corresponding\ to\ bus - route\ assignment$

$constraint\ (46), b \in B$

(SP)

$$Minimize\ y = \sum_{b \in B}(f_b - \gamma_b) + \sum_{i,j \in N}(d_{ij} - \pi_i)x_{ij} \quad (49)$$

Subject to

$$\sum_{i \in S}\sum_{j \in S} x_{ij} \leq c_b \qquad \forall b \in B \quad (50)$$

$$\sum_{i \in S} x_{il} - \sum_{j \in N} x_{lj} = 0 \qquad \forall l \in S \quad (51)$$

$$\sum_{i \in S} x_{in} = 1 \quad (52)$$

$$a_i + s_i + t_{ij} - M(1 - x_{ij}) \leq a_j \qquad \forall i, j \in N \quad (53)$$

$$e \leq a_n \leq l \quad (54)$$

$$a_n - a_i \leq m_i \qquad \forall i \in S \quad (55)$$

$$x_{ij}: binary \qquad \forall i, j \in N \quad (56)$$

The objective function (49) includes two terms: the fixed cost of using bus minus its dual cost, and the traveling cost associated with arcs minus its dual cost. The dual variables γ_b represent the negative cost of using additional bus and the dual variables π_i indicate the increment in the objective function by visiting the student i.

Constraints (50) limit the number of students on a bus to be less than or equal to the capacity of the bus type b. For any stop l, the flows entering l are the same as the flows leaving l in Constraints (51). Constraints (52) assert that the school n is on every route. Constraints (53) are related to time scheduling with respect to any pair of stops i and j. If the arc (i,j) is used in the route, the arrival time at j should be later than the arrival time at i (a_i) plus the service time at i (s_i) plus the traveling time from i to (t_{ij}). Constraints (54) specify the time window for each bus arrival time at school.

Constraints (55) check the upper time limit for each route, and decision variables x_{ij} are binary in Constraints (56).

CHAPTER VI - COMPUTATIONAL EXPERIMENTS

Three sets of computational experiments were performed in this section. First, in Section A, we used both LINGO and CPLEX to solve SAPS to compare the resulting number of bus stops selected as well as to measure the corresponding computational efficiency for each solver. Then, while limiting the computational run-time to 24 CPU hours in solving SBRP in Section B, we designed 36 test instances based on real data, and split them into small and large groups depending on the number of bus stops included in a data set. After identifying effective algorithmic strategies contingent upon the size of a test instance, we implemented and conducted two different sets of experiments accordingly. All runs were made using the commercial software IBM ILOG CPLEX Optimization Studio 12.6.3.0 on a ASUS U36JC computer having an Intel Core i5 2.67 GHz CPU, with 4.00 GB of RAM, and running Windows 7.

A. Analyses of SAPS Results

"There are no known polynomial time algorithms that can find an optimal solution for $\mathcal{NP}$–hard problem" (Lenstra, 1981). Table 4 shows the numerical results of solving SAPS exactly with Lingo and CPLEX Optimization Studio. Both the SAP model and SBRP are modeled as mixed integer linear programming. "The mixed integer solver of LINGO contains advanced solution techniques such as cut generation, tree reordering to reduce tree growth dynamically, and advanced heuristic and presolve strategies" (Cunningham, 1988). The mixed integer optimizers of CPLEX contains branch and cut, dynamic search, and MIP. LINGO and CPLEX apply different algorithms to solve the problem and they get different computation time for the same instances.

Table 4 SAPS Computation Time Results

Test Instance	No. of Students (Q)	No. of potential bus stop locations (P)	No. of selected Bus Stops	LINGO Run Time (sec.)	CPLEX Run Time (sec.)
T1	200	50	45	3	9
T2	200	70	61	2	10
T3	250	100	86	4	12
T4	300	150	126	10	17
T5	350	180	154	17	22
T6	400	200	170	23	27
T7	450	250	216	34	43
T8	500	280	240	54	60
T9	550	300	259	55	79
T10	600	350	301	111	116
T11	650	400	346	236	170
T12	700	450	390	441	239
T13	1000	500	436	604	563

To all the instances, the walking distance between student home locations and potential bus stops are generated randomly. The input data and the optimal solutions are checked manually to make sure there exists feasible

solutions, and the results of LINGO and CPLEX are the global optimal. As SAPS determines the number of selected bus stops, a stop is eliminated from feasible solutions if the walking distance to that stop from home location is more than the maximum walking distance limit. The number of students ranges from 200 to 1000 per test instance, with the number of potential bus stops from 50 to 500. There is no capacity limit for a single bus stop, and students are grouped into designated stops. As such, the number of selected bus stops is reduced and consequently facilitates the process of solving SBRP. For each student the model assigns one to the nearest bus stop no matter how many students are already assigned to the same bus stop. The main reason we do not set the capacity constraint for one bus stop is that it is possible for one single bus stop to be served by multiple buses. Also, the number of bus stops in use is minimized and the students are assigned accordingly. In addition, when the randomly generated Type-2 students were added, it did not affect the computational times substantially in solving SAPS. Based on the results in Table 4, for most of the practical size problems, both LINGO and CPLEX solve them directly and achieve the same optimal objective value along with the minimum number of potential bus stops locations within about 10 minutes of run time. For the largest instance T12, 1000 students were assigned to 500 potential bus stops and 436 of them were selected in the resulting solution. It took 604 seconds and 563 seconds for LINGO and CPLEX to solve, respectively. Overall, 84% to 90% of available bus stops are selected in the optimal solution.

B. Analyses of SBRP Results

Table 5 shows the computing results of solving SBRP exactly with Lingo and CPELX. We first introduce the (abbreviated) notations used for discussion of computational experiment results in Tables 5-8 as follows.

$stop$: $number\ of\ stops$

stu: $number\ of\ students$

$mt1$: $Type-1\ student\ maximum\ riding\ time$

$t2$: $the\ number\ of\ Type-2\ students$

$mt2$: $Type-2\ student\ maximum\ riding\ time$

bus: $number\ of\ available\ buses$

cap: $bus\ capacity$

$L-obj$: $objective\ value\ of\ LINGO\ solution$

$L-time$: $LINGO\ computation\ time\ in\ seconds$

$C-obj$: $objective\ value\ of\ CPLEX\ solution$

$C-time$: $CPLEX\ computation\ time\ in\ seconds$

GA: $objective\ value\ of\ GA\ only\ solution$

$GA-time$: $GA\ only\ computation\ time\ in\ seconds$

$GACG$: $objective\ value\ of\ GA+CG\ solution$

$GACG-time$: $GA+CG\ computation\ time\ in\ seconds$

Table 5 shows small size test instances having 7 to 20 stops, 10 to 98 students, one to three Type-2 students, two to four buses with capacity ranging from 5 to 25. Different maximum riding time limits (in seconds) are imposed to find feasible solutions. For the instance N17, 98 students are

Table 5 Small Size SPRB Test Instances

ID	stop	stu	mt1	t2	mt2	bus	cap
N1	7	10	3600	1	1800	2	5
N2	7	10	3600	2	1800	2	5
N3	7	30	3600	1	1800	3	10
N4	7	30	3600	2	1800	3	10
N5	10	30	4200	1	1800	2	15
N6	10	30	4200	2	1800	2	15
N7	14	30	5400	1	1800	2	15
N8	14	30	5400	2	1800	2	15
N9	14	45	5400	1	1800	3	15
N10	14	45	5400	2	1800	3	15
N11	15	67	3600	1	1800	3	30
N12	15	67	3600	2	1800	3	30
N13	15	67	3600	3	1800	3	30
N14	18	75	3600	1	1800	3	25
N15	18	75	3600	2	1800	3	25
N16	18	75	3600	3	1800	3	25
N17	20	98	3600	1	1800	4	25

located at 20 bus stops. The maximum riding time for Type-1 students is 3600 seconds and for Type-2 students is 1800 seconds. Only one Type-2 student needs service and four buses with capacity of 25 seats are available. If the number of available buses or the capacity are increased, some buses may not be used at all but the buses with handicapped facilities are utilized in accordance with the number of Type-2 students.

The objective values and computational run times required to directly solve SBRP using LINGO and CPLEX are presented in Table 6. Given the maximum runtime set as 24 hours, LINGO solved the instance N1 with ten students, seven bus stops, two buses, and one Type-2 student within 13 seconds and obtained the objective function value of 50.6. On the other hand, it took CPLEX two seconds to find the optimal objective value of 73.8. When increasing both the number of students to 98 and the number of bus stops to 20, it took LINGO about 4,749 seconds to find the optimal solution. For all instances with more Type-2 students, both LINGO and CPLEX cannot find feasible solution within 24 hours. For all instances with more than 100

students, no feasible solution is found from both LINGO and CPLEX within 24 hours. Overall, it takes from 5 seconds to 4749 seconds to achieve the final solution for LINGO. For CPLEX, the final solution can be found from 2 seconds up to 1,732 seconds. The objective value and computing times increase as the size of the problem and the number of Type-2 students are larger.

The run time is affected by not just the problem size but the (route) network structure from input data. Taking N13 and N14 for instance, the objective values and computing times from both LINGO and CPLEX for N13 are superior despite the size of N14 is larger. Also, the 15 stops and 67 students in N13 were randomly chosen and different from the 18 stops and 75 students in N14. Between the instances of N1 and N2, the only difference is that one Type-1 student in N1 is changed as a Type-2 student in N2, and similarly, the number of Type-2 students are in N11-N13 (see Table 5). The objective value of N11 using LINGO is 213.9, and when the number of Type-2 student increases from one to two as in N12, the objective value is 231.5 with an increase of 8.2%. Furthermore, when the number of Type-2 students increases from two to three as in N13, the objective value is 252.5 with an increase of 9.1%. Using CPLEX, the objective value of N11 is 224, and it increases 4.5% with the objective value of 234 when switching one Type-1 student to Type-2 in N12. Likewise, the objective value increases 9% resulting in the objective value of 255 when switching one more Type-1 student to Type-2 in N13. The corresponding objective values from both LINGO and CPLEX increase as more restriction to accommodate additional Type-2 students is imposed on time. As a result, N13 requires all three buses to

transport Type-2 students in order to satisfy the maximum riding time limit as well as the school time window constraints concurrently. For the number of buses in use and the bus capacity, we keep the total capacity close to the total number of students since the fixed cost of bus is usually much higher than the marginal traveling cost. Improved solutions and less computing times can be achieved if the maximum riding time limit is relaxed, especially for Type-2 students, and it is a tradeoff between the service quality and the cost. Moreover, additional fixed cost for the Type-2 bus is larger than the Type-1 bus. Hence, more Type-2 students demand additional Type-2 buses, and as a result these additional Type-2 buses contribute to the significant increase of the objective value.

Table 6 SBRP Results with LINGO and CPLEX

ID	L-obj	L-time	C-obj	C-time
N1	50.6	13	73.8	2
N2	65.3	15	90.6	3
N3	87.4	9	90.5	2
N4	116.5	10	115.3	2
N5	88.2	14	91.4	2
N6	100.5	15	102.1	2
N7	111.7	3400	125.7	2
N8	136.2	3446	144.8	2
N9	146.5	1400	148.2	2
N10	168	1364	171.2	2
N11	213.9	185	224	2
N12	231.5	2167	234	161
N13	252.5	4123	255	1732
N14	193.2	3161	198.1	5
N15	227.3	3024	229	182
N16	249.8	3582	254.3	193
N17	234.2	4749	256.3	5

For test instances of larger sizes in Table 7, heuristic methods are incorporated to solve the problems. The test instances are generated considering 20 to 55 stops, 98 to 700 students, one to five Type-2 student, 4 to 14 buses with capacity from 25 to 50. For the largest test instance D20, 55 bus stops are used along with 700 students for a single school problem. The maximum riding time for a Type-1 student is 3600 seconds, whereas that for a

Type-2 student is 1,800 seconds. A total of 14 buses with capacity of 50 seats are available and three of them are Type-2 buses.

Table 7 Large Size SBRP Test Instances

ID	stop	stu	mt1	t2	mt2	bus	cap
D1	20	98	3600	1	1800	4	25
D2	20	98	3600	2	1800	4	25
D3	20	98	3600	3	1800	4	25
D4	20	98	3600	4	1800	4	25
D5	28	200	3600	1	1800	4	50
D6	28	200	3600	2	1800	4	50
D7	28	200	3600	3	1800	4	50
D8	28	200	3600	4	1800	4	50
D9	40	400	3600	1	1800	8	50
D10	40	400	3600	2	1800	8	50
D11	40	400	3600	3	1800	8	50
D12	40	400	3600	4	1800	8	50
D13	45	500	3600	1	1800	10	50
D14	45	500	3600	2	1800	10	50
D15	45	500	3600	3	1800	10	50
D16	45	500	3600	4	1800	10	50
D17	45	500	3600	5	1800	10	50
D18	55	700	3600	1	1800	14	50
D19	55	700	3600	2	1800	14	50
D20	55	700	3600	3	1800	14	50

Table 8 presents the results of applying a heuristic method of GA only, and GA combined with CG. The instance D1 is the same instance as N17 (see Table 5 and Table 7), which indicates the limitation of solving the problem exactly using LINGO and CPLEX. Comparing the results from N17 in Table 6 and D1 in Table 8, both cases (GA only, GA with CG) achieved better solutions than LINGO and CPLEX, consuming 3,093 and 2,753 seconds of the run times, respectively. For a single school problem, the largest test instance includes 700 students, 55 bus stops, and 14 buses with 50 seats on each bus. All test instances D1-D20 were solved within a 24-hour run time limit using heuristic methods. Overall, the objective value and the computing time increase as the size of the problem and the number of Type-2 student increases. It took up to 80,415 seconds for GA to find a feasible solution. For GA with CG, the final solution was obtained in 2,139 to 79,253 seconds.

When comparing these two strategies, in compensation for longer run times, the solution quality of using CG coupled with GA is better than GA alone. This is because CG continues to improve the objective value on the initial column pool of solutions provided by GA. For instances D13-D17, the objective values increase as the number of Type-2 students increase, requiring more resources and associated cost. For example, if one Type-1 student is changed to become a Type-2 student, the original optimal routes for D13 are no longer feasible due to the maximum riding time limit of Type-2 students. However, the run time does not necessarily increase with more Type-2 students in D14 and D15 since the solution space of feasible region is reduced. On the other hand, when comparing D17 and D18, the computing time increases substantially when the problem size changes from 500 students with 45 stops to 700 students with 55 stops while decreasing the number of Type-2 students is decreased from five to one. Therefore, the size of the problem is a more important factor attributed to run times than the complexity associated with the number of Type-2 students.

Finally, we define $\%gap$ as $(\overline{GA} - \overline{GACG})/\overline{GACG} \times 100$, where $\overline{GA}$ and $\overline{GACG}$ denote objective function values obtained by using GA alone, and GA and CG combined, respectively. For the last ($\%gap$) column in Table 8, it measures approximately how much CG improves the optimal solution. For example, 17% gap from 169.3 to 144.7 reported for Instance D1 indicates the improved amount of the objective value resulting from applying GA in conjunction with CG, when compared to using GA alone. Among all instances, CG works best on Instance D5 in which the total cost decreased as much as 44.74%, whereas it improved only up to 5.22% from 565.45 to 537.39 in D19.

Table 8 SBRP results with GA and GACG

ID	GA	GA-time	GACG	GACG-time	%gap
D1	169.3	2753	144.7	3093	17.00%
D2	175.9	1677	153.3	3850	14.74%
D3	180.7	984	158.4	2139	14.08%
D4	193.8	2070	164.9	3199	17.53%
D5	126.5	1768	87.4	3957	44.74%
D6	129.9	889	104.6	8882	24.19%
D7	139.9	981	106.8	9871	30.99%
D8	153.6	4955	125.3	11843	22.59%
D9	339.8	5859	317.31	8704	7.09%
D10	352.81	5250	333.05	5864	5.93%
D11	397.78	5934	320.66	8839	24.05%
D12	426.21	5954	345.34	6462	23.42%
D13	354.54	26169	322.47	30281	29.66%
D14	418.13	27117	331.94	31930	6.81%
D15	462.48	17274	384.28	25863	20.35%
D16	522.94	17421	389.36	31021	34.31%
D17	525.73	21902	417.39	28465	25.96%
D18	556.8	80415	512.93	78394	8.55%
D19	565.45	78941	537.39	79253	5.22%
D20	625.36	55884	539	79157	16.02%

CHAPTER VII – CONCLUSIONS AND FUTURE RESEARCH

In this dissertation, we have proposed models of a student assignment problem and a school bus routing problem, and investigated the complexity of modeling and solution procedures by simultaneously considering Type-2 students and Type-1 students together. Using the available student data and potential bus stop locations, the bus stop selection and the student assignment were made while minimizing the number of stops in use or the total walking distance. In SBRP, we utilized solutions from SAP along with other information such as available bus capacity, time restrictions, student types, and distance and time matrices to minimize the total cost including fixed cost and traveling cost.

Both problems were formulated as a mixed-integer program and enhanced by a heuristic method in concert with Genetic Algorithm, and a Column Generation solution approach was designed to facilitate the solution process. Computational results were presented using real data obtained from a school district to demonstrate the efficacy of the modeling and algorithmic strategies. For all test instances of SAP (e.g., 200 students, 50 potential bus stops to 1000 students, 500 potential bus stops), the optimal solutions were found within 10 minutes of run time for LINGO and CPLEX. When the small instances (less than 100 bus stops) of SBRP we tested were solved exactly, LINGO performed slightly better than CPLEX while consuming more computational times. For any larger instances (e.g., 98 students, 20 stops to

700 students, 55 stops), no solutions were reported by both solvers within a 24-hour limit. However, our algorithmic approach by applying GA and subsequently CG was able to not only improve the final solutions by as much as 25% but achieve the tractable solutions within 24 hours. Overall, the combined strategy of GA and CG is a preferred option for solving relatively large-scale problems.

The SBRP is a challenging problem to solve considering a multitude of constraints and assumptions. Also, the tradeoff between cost optimization and service quality needs further investigation with more comprehensive and effective algorithmic solution approaches.

REFERENCES

http://www.schoolbusfacts.com

-----. (2008) Digest of Education Statistics, 2007. *U.S. Department of Education, National Center for Education Statistics.* http://nces.ed.gov/programs/digest/d07/tables/dt07_176.asp?referrer=list

-----. (2008) AASA Fuel and Energy Snapshot Survey. *American Association of School Administrators.* http://www.aasa.org/newsroom/pressdetail.cfm?ItemNumber=10637

-----. (2008) Facing Deficits, Most State are Imposing Cuts that Hurt Vulnerable Residents. *Center on Budget and Policy Priorities.* http://www.cbpp.org/3-13-08sfp.pdf

http://www.oldham.k12.ky.us/

-----. (2009) School Bus Safety Data. http://www.stnonline.com/

Addor, J. A., & Amponsah, S. K. (2013). School Bus Routing: A Case Study of Wood Bridge School Complex, Sekondi-Takoradi, Ghana. *International Journal of Business and Social Research*, 3(12), 26-36.

Arias-Rojas, J. S., Jiménez, J. F., & Montoya-Torres, J. R. (2013). Solving of school bus routing problem by ant colony optimzation (Resolución del problema de ruteo de buses escolares con optimización por colonia de hormigas). *Revista EIA*, 9(17), 193-208.

Augerat, P., Belenguer, J. M., Benavent, E., Corbéran, A., & Naddef, D. (1998). Separating capacity constraints in the CVRP using tabu search. *European Journal of Operational Research*, 106(2), 546-557.

Bektaş, T., & Elmastaş, S. (2007). Solving school bus routing problems through integer programming. *Journal of the Operational Research Society*, 58(12), 1599-1604.

Bock, A., Grant, E., Könemann, J., & Sanità, L. (2011). The school bus problem on trees. *Algorithmica* 67(1), 49–64.

Bodin, L. D., & Berman, L. (1979). Routing and scheduling of school buses by computer. *Transportation Science*, 13(2), 113-129.

Bowerman, R., Hall, B., & Calamai, P. (1995). A multi-objective optimization approach to urban school bus routing: Formulation and solution method. *Transportation Research Part A: Policy and Practice*, 29(2), 107-123.

Braca, J., Bramel, J., Posner, B., & Simchi-Levi, D. (1997). A computerized approach to the New York City school bus routing problem. *IIE Transactions*, 29(8), 693-702.

Chalkia, E., Grau, J. M. S., Bekiaris, E., Ayfandopoulou, G., Ferarini, C., & Mitsakis, E. (2014). Routing Algorithms for the Safe Transportation of Pupils to School Using School Buses. *Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment.*

Chapleau, L., Ferland, J. A., & Rousseau, J. M. (1985). Clustering for routing in densely populated areas. *European Journal of Operational Research*, 20(1), 48-57.

Clarke, G.,Wright,J.W., (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research* 12,568–581.

Corberán, A., Fernández, E., Laguna, M., & Marti, R. (2002). Heuristic solutions to the problem of routing school buses with multiple objectives. *Journal of the operational research society*, 53(4), 427-435.

Cortés, C. E., Matamala, M., & Contardo, C. (2010). The pickup and delivery problem with transfers: Formulation and a branch-and-cut solution method. *European Journal of Operational Research*, 200(3), 711-724.

Crainic, T., Gendreau, M., & Gajpal, Y. (2012). Multi-zone Multi-trip Vehicle Routing Problem with Time Windows. *CIRRELT.*

Cunningham, K., & Schrage, L. (1988). LINGO User's Manual. *LINDO System Inc., Chicago, IL.*

Desrosiers, J., Ferland, J.A., Rousseau, J.-M., Lapalme, G., Chapleau, L. (1981). An overview of a school busing system. *Scientific Management of Transport Systems*. North-Holland, Amsterdam, pp. 235–243.

Díaz-Parra, O., Ruiz-Vanoye, J. A., Buenabad-Arias, Á., & Cocón, F. (2012). A vertical transfer algorithm for the School Bus Routing Problem. *In Nature and Biologically Inspired Computing (NaBIC), 2012 Fourth World Congress* on (pp. 66-71). IEEE.

Dulac, G., Ferland, J. A., & Forgues, P. A. (1980). School bus routes generator in urban surroundings. *Computers & Operations Research*, 7(3), 199-213.

École des hautes études commerciales (Montréal, Quebec), Université de Montréal. Centre de recherche sur les transports, & Université de Montréal. Département d'informatique et de recherche opérationnelle. (1981). TRANSCOL: A multi-period school bus routing and scheduling system. *Montréal: Université de Montréal, Centre de recherche sur les tranports.*

Eiró, T., Viegas, J. M., & Martínez, L. M. (2011). A new optimization procedure to design minibus services: an application for the Lisbon Metropolitan Area. *Procedia-Social and Behavioral Sciences*, 20, 856-865.

Euchi, J., & Mraihi, R. (2012). The urban bus routing problem in the Tunisian case by the hybrid artificial ant colony algorithm. *Swarm and Evolutionary Computation*, 2, 15-24.

Fu, Z., Eglese, R., & Li, L. Y. (2005). A new tabu search heuristic for the open vehicle routing problem. *Journal of the operational Research Society*, 56(3), 267-274.

Fügenschuh, A. (2009). Solving a school bus scheduling problem with integer programming. *European Journal of Operational Research*, 193(3), 867-884.

Hanley, P. F. (2007). Transportation cost changes with statewide school district consolidation. *Socio-Economic Planning Sciences*, 41(2), 163-179.

Howley, C. B., Howley, A. A., & Shamblen, S. (2001). Riding the school bus: A comparison of the rural and suburban experience in five states. *Journal of Research in Rural Education*, 17(1), 41-63.

Karp, R. M. (1972). Reducibility among combinatorial problems (pp. 85-103). *springer US*.

Kim, B. I., Kim, S., & Park, J. (2012). A school bus scheduling problem. *European Journal of Operational Research*, 218(2), 577-585.

Kinable, J., Spieksma, F. C. R., & Vanden Berghe, G. (2014). School bus routing—a column generation approach. *International Transactions in Operational Research*, 21(3), 453-478.

Laporte, G., Nobert, Y., & Taillefer, S. (1988). Solving a family of multi-depot vehicle routing and location-routing problems. *Transportation science*, 22(3), 161-172.

Lenstra, J. K., & Kan, A. H. G. (1981). Complexity of vehicle routing and scheduling problems. *Networks*, 11(2), 221-227.

Li, F., Golden, B., & Wasil, E. (2007). The open vehicle routing problem: Algorithms, large-scale test problems, and computational results. *Computers & Operations Research*, 34(10), 2918-2930.

Li, L. Y. O., & Fu, Z. (2002). The school bus routing problem: a case study. *Journal of the Operational Research Society*, 552-558.

Manumbu, D. M., Mujuni, E., & Kuznetsov, D. (2014). Mathematical Formulation Model for a School Bus Routing Problem with Small Instance Data. *Mathematical Theory and Modeling*, 4(8), 121-132.

Martínez, L. M., & Viegas, J. M. (2011). Design and deployment of an innovative

school bus service in Lisbon. *Procedia-Social and Behavioral Sciences*, 20, 120-130.

Moccia, L., Cordeau, J. F., & Laporte, G. (2012). An incremental tabu search heuristic for the generalized vehicle routing problem with time windows. *Journal of the Operational Research Society*, 63(2), 232-244.

Nagy, G., & Salhi, S. (2007). Location-routing: Issues, models and methods. *European Journal of Operational Research*, 177(2), 649-672.

Newton, R. M., & Thomas, W. H. (1969). Design of school bus routes by computer. *Socio-Economic Planning Sciences*, 3(1), 75-85.

Nguyen, P. K., Crainic, T. G., & Toulouse, M. (2013). A tabu search for time-dependent multi-zone multi-trip vehicle routing problem with time windows. *European Journal of Operational Research*, 231(1), 43-56.

Pacheco, J., Caballero, R., Laguna, M., & Molina, J. (2013). Bi-objective bus routing: an application to school buses in rural areas. *Transportation Science*, 47(3), 397-411.

Park, J., & Kim, B. I. (2010). The school bus routing problem: A review. *European Journal of Operational Research*, 202(2), 311-319.

Park, J., Tae, H., & Kim, B. I. (2012). A post-improvement procedure for the mixed load school bus routing problem. *European Journal of Operational Research*, 217(1), 204-213.

Pedroso, M. (2010). National Statistics on School Transportation http://http://saferoutespartnership.org/

Perugia, A., Moccia, L., Cordeau, J. F., & Laporte, G. (2011). Designing a home-to-work bus service in a metropolitan area. *Transportation Research Part B: Methodological*, 45(10), 1710-1726.

Resende, M. G. (2009). Greedy randomized adaptive search procedures Greedy Randomized Adaptive Search Procedures. *Encyclopedia of Optimization*, 1460-1469.

Riera-Ledesma, J., & Salazar-González, J. J. (2012). Solving school bus routing using the multiple vehicle traveling purchaser problem: A branch-and-cut approach. *Computers & Operations Research*, 39(2), 391-404.

Riera-Ledesma, J., & Salazar-González, J. J. (2013). A column generation approach for a school bus routing problem with resource constraints. *Computers & Operations Research*, 40(2), 566-583.

Russell, R. A., & Morrel, R. B. (1986). Routing special-education school buses. *Interfaces*, 16(5), 56-64.

Schittekat, P., Sevaux, M., & Sorensen, K. (2006). A mathematical formulation for a school bus routing problem. *In Service Systems and Service Management, 2006 International Conference on (Vol. 2, pp. 1552-1557). IEEE.*

Schittekat, P., Kinable, J., Sörensen, K., Sevaux, M., Spieksma, F., & Springael, J. (2013). A metaheuristic for the school bus routing problem with bus stop selection. *European Journal of Operational Research*, 229(2), 518-528.

Shapiro, J. F., & Shapiro, J. F. (1979). Mathematical programming: structures and algorithms (p. 145). *New York: Wiley.*

Song, S. M., & Kim, T. (2013). Customer-oriented school bus operations for childcare centers in Korea. *Computers and Industrial Engineering*, 66(1), 116-124.

Spada, M., Bierlaire, M., & Liebling, T. M. (2005). Decision-aiding methodology for the school bus routing and scheduling problem. *Transportation Science*, 39(4), 477-490.

Stuber, M. (2010). A school bus route planning. École polytechnique fédérale de lausanne. Lausanne, Switzerland.

Yu, B., Yang, Z. Z., & Yao, B. (2009). An improved ant colony optimization for vehicle routing problem. *European journal of operational research*, 196(1), 171-176.

CURRICULUM VITA

NAME: Di Zhang

ADDRESS: Department of Industrial Engineering
J.B. Speed School of Engineering
University of Louisville
Louisville, Kentucky 40292

DOB: Beijing, China P.R. – April 30th, 1987

EDUCATION &TRAINING:

B.S., Automation Science and Electric Engineering
Beihang University
2005-09

M.S., Industrial Engineering
University of Louisville
2009-11