A STUDY OF REALISTIC AND INNOVATIVE FEATURES OF THE SCHOOL BUS ROUTING PROBLEM

by

Hernan Andres Caceres Venegas

February 1, 2017

A dissertation submitted to the Faculty of the Graduate School of the University at Buffalo, State University of New York in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Industrial and Systems Engineering
To my parents.
# Contents

1 Introduction .................................................................................................................. 1
   1.1 School Bus Routing with Stochastic Demand and Duration Constraints . . . 1
   1.2 Special Need Students School Bus Routing, Consideration for Mixed Load
       and Heterogeneous Fleet ............................................................................................. 2
   1.3 Pricing tax return for student that opt-out using school bus . . . . . . . . . . . . 3

2 School Bus Routing with Stochastic Demand and Duration Constraints 4
   2.1 Introduction ............................................................................................................ 4
   2.2 Problem Description ............................................................................................... 8
   2.3 Model Formulation ................................................................................................ 10
       2.3.1 General model ............................................................................................... 11
       2.3.2 Single bell-time routing problem ................................................................... 12
       2.3.3 Stochastic constraints .................................................................................. 14
       2.3.4 Computational issues .................................................................................... 21
   2.4 Cascade Simplification .......................................................................................... 22
       2.4.1 Set stops’ location ....................................................................................... 23
       2.4.2 Finding an initial solution to the single school routing problem . . . . . . . . 25
       2.4.3 Solving the single school routing problem ...................................................... 26
   2.5 Model Application to Williamsville Central School District ................. 35
3 Special Need Students School Bus Routing, Consideration for Mixed Load and Heterogeneous Fleet

3.1 Introduction ........................................ 46
3.2 Problem description .................................. 49
3.3 Mathematical model .................................. 52
3.4 Solution strategy ..................................... 56
  3.4.1 Problem decomposition by column generation .......... 56
  3.4.2 Column generation procedure ....................... 59
  3.4.3 A lower bound .................................. 62
3.5 Computational experiments ........................... 65
  3.5.1 Simulated data .................................. 65
  3.5.2 Case study .................................... 67
3.6 Conclusion ........................................... 70
3.7 Algorithms ........................................... 71
  3.7.1 Generate initial solution for the master problem ........ 71
  3.7.2 Approximate solution for the subproblem ............ 73
  3.7.3 Try to add a stop to a route ....................... 74
  3.7.4 Selection of stop ................................ 75

4 Pricing tax return for student that opt-out using school bus 77

4.1 Introduction ......................................... 77
4.2 Literature review ......................................................... 79
4.3 Open offer problem ....................................................... 80
  4.3.1 Open offer for one school ......................................... 81
  4.3.2 Stop selection model ............................................... 82
  4.3.3 Routing model ...................................................... 83
  4.3.4 Numerical example ............................................... 84
4.4 Targeted offer problem ................................................ 88
  4.4.1 Illustrative example ............................................... 88
  4.4.2 Solution strategy .................................................. 91
  4.4.3 Numerical example ............................................... 92
4.5 Conclusion and further research ..................................... 93

Bibliography ................................................................. 94
# List of Figures

2.1 Histogram of student count. ................................................. 7
2.2 Solution time comparisons for difference sizes of the SBRP problem. .... 21
2.3 Column generation procedure. .............................................. 31
2.4 Results from factorial design. ............................................ 34
2.5 Results from factorial design. ............................................ 34
2.6 Student transportation management solution. ............................. 36
2.7 Overbooked capacity vs ridership. ....................................... 37
2.8 Solution time comparisons for the multi-depot to multi-school model and the cascade simplification. ............................................. 38

3.1 Bus layout for different seat type configurations. .......................... 50
3.2 Location of homes and schools. .......................................... 51
3.3 Column generation procedure. ............................................ 59
3.4 An illustrative example for bus-stop sequencing. .......................... 60
3.5 Numerical example with simulated data. .................................. 66
3.6 Average number of buses per Load Type and Maximum Ride Time. ...... 69

4.1 Average ridership per school and level. ................................... 78
4.2 Motivating illustration ..................................................... 79
4.3 Illustrative example. ....................................................... 85
# List of Tables

2.1 Details for solution time comparison .................................................. 22  
2.2 Details for solution time comparison .................................................. 22  
2.3 Factorial design for the single school routing problem .......................... 33  
2.4 Results for the cascade simplification applied to WCSD ........................ 39  
3.1 Previous work directly related to special education SBRP ....................... 48  
3.2 Computational results small size instances ........................................... 55  
3.3 Computational results for real instances of WCSD ............................... 68  
4.1 Computational results for real instances of WCSD ............................... 87
Abstract

The dissertation develops realistic and innovative features for the school bus routing problem (SBRP). The first part proposes a mathematical formulation responding to the “overbooking” policies applied at a real-world school district. According to our empirical studies, the probability of a student using the bus varies from 22% to 77%, opening the opportunity to overbook the buses in order to improve the utilization of their capacity. Due to the NP-hard nature of the problem, a cascade simplification algorithm is proposed to partition the multiple stage SBRP problems into multiple multi-depot and one-school subproblems that are solved sequentially, where the results for one are data inputs for the next. Furthermore, we develop column-generation-based algorithms to solve the scheduling problem, and different instances of the problem are examined. The second part focuses on the problem of routing special education students. We found the problem to be significantly different from that of routing regular students, the principle differences being the needs to pick up special education student from their home, to configure buses appropriately for special education students, and to provide a higher level of service. We developed a greedy heuristic coupled with a column generation approach to finding approximate solutions. Finally, the third part proposes two innovative policies for reducing the number of buses needed. Since ridership varies widely, many buses run with unused capacity over long routes. We explore the scenario where students are compensated for giving up the option to ride a bus, in an effort to reduce the overall cost of the system. Mathematical formulations for this problem are developed and analyzed. Results from a case study along with algorithmic computational results are presented.
Chapter 1

Introduction

This research is focused on the study of the School Bus Routing Problem (SBRP) from different perspectives, where we explore the impact of implementing various policies on the routing for both regular and especial education students. The three following chapters are the significant division of our work. The first part is concerned with SBRP for regular public schools, the second studies the particular complexities of SBRP for especial education students, and the third part examines creating an incentive for student to not use school transportation. These three chapters are written as independent papers. The first is accepted for the journal Transportation Science. The second has been submitted to the journal OMEGA. The third is a working paper that we will continue to work on for future publication. The following sections correspond to the abstract of these three pieces of research.

1.1 School Bus Routing with Stochastic Demand and Duration Constraints

The school bus routing problem (SBRP) is crucial due to its impact on economic and social objectives. A single bus is assigned to each route, picking up the students and arriving at their
school within a specified time window. SBRP aims to find the fewest buses needed to cover all the routes while minimizing the total travel distance and meeting required constraints. We propose a mathematical formulation responding to the “overbooking” policies applied at a real-world school district. According to our empirical studies, the probability of a student using the bus varies from 22% to 77%, opening the opportunity to overbook the buses in order to improve the utilization of their capacity. However, SBRP with “overbooking” has not attracted much attention in previous studies. In this work, “overbooking” is modeled via chance constrained programming. Additionally, to account for the uncertainty of the total travel time of the buses, a constraint limiting the probability of being late to school is also proposed in this paper. Due to the NP-hard nature of the problem, a cascade simplification algorithm is proposed to partition the multiple stage SBRP problems into multiple multi-depot and one-school subproblems that are solved sequentially, where the results for one are data inputs for the next. Furthermore, we develop column-generation-based algorithms to solve the scheduling problem, and different instances of the problem are examined. Our computational experiments on a real-world school district demonstrate desirable cost savings in terms of total number of buses used.

1.2 Special Need Students School Bus Routing, Consideration for Mixed Load and Heterogeneous Fleet

We consider the School Bus Routing Problem (SBRP) for routing special education students based on our experience at a large suburban school district in Western New York, United States. We found the problem to be significantly different from that of routing regular students. The principle differences include the need to pick up special education student from their home, the need to configure buses appropriately for special education students, and the
need to provide a higher level of service. Building upon prior work we developed a greedy heuristic coupled with a column generation approach to obtain approximate solutions for benchmark instances. Our findings demonstrated a 10\textendash20\% cost reduction, which is particularly significant since special education transportation account for 40\% of the transportation budget.

1.3 Pricing tax return for student that opt-out using school bus

School districts are often mandated to provide transportation but can encounter ridership that varies between 22\textendash72 percent. Consequently, buses run with unused capacity over long routes. We explore the scenario where students are compensated for giving up the option to ride a bus, in an effort to reduce the overall cost of the system. Mathematical formulations for this problem are developed and analyzed. Results from a case study along with algorithmic computational results will be presented.
Chapter 2

School Bus Routing with Stochastic Demand and Duration Constraints

2.1 Introduction

The school bus routing problem (SBRP) is crucial due to its impact on economic and social objectives [1]. In general, SBRP is the problem of finding a set of routes that optimizes specified objectives (e.g. total cost) for operating a fleet of school buses, which picks up students from bus stops near their homes and delivers them to their schools in the morning, and then does the opposite in the afternoon, while observing pre-specified physical and time limitations [2].

SBRP has been intensively studied in the last few decades. One may refer to a fairly recent literature review of SBRP in [3]. Later work has continued in tackling the computational complexity of the SBRP’s one-school instance by the design or adaptation of heuristics such as column generation [4, 5], tabu search [6], greedy randomized adaptive search procedure [7], branch-and-cut algorithm [8], approximation algorithm [9] and genetic algorithm [10]. Additionally, the work of [11] and [12] focuses on the potential gain of mixing students from
different schools in the same bus. The former designs an improvement algorithm and the
latter a particular branch and bound procedure.

Despite the work in SBRP, most of the previous studies focus on deterministic routing
problems with known student demand and fixed travel time. This paper formally defines
SBRP with Stochastic Demand and Duration Constraints, denoted as SBRP-SDDC, via
Chance Constrained Programming (CCP). The School Bus Routing and Scheduling Problem
is a generalization of the Vehicle Routing Problem (VRP) [13, 14]. Due to few studies
in SBRP-SDDC, we conduct a brief literature review on VRP with stochastic demand,
especially by the approach of CCP. In a CCP related problem, the decision maker selects a
here-and-now decision that satisfies all constraints with a pre-specified probability.

Stewart Jr. and Golden [15] proposed a chance-constrained model to identify minimum
cost tours subject to a threshold constraint on the probability of a tour failure. A similar
approach is proposed in [16]; the model uses fewer variables, but requires a homogeneous
fleet of vehicles. Laporte, Louveaux, and Mercure [17] developed a model to minimize a
linear combination of vehicle and routing costs while ensuring that the probability of the
duration of a route exceeding a set threshold is at most equal to a given value. Most
recently, Gounaris, Wiesemann, and Floudas [18] studied the robust capacitated vehicle
routing problem (CVRP), in which the decision maker selects minimum cost vehicle routes
that remain feasible for all realizations of uncertain customer demands. They established the
connection between the robust CVRP and a distributionally robust variant of the chance-
constrained CVRP.

This study develops a general solution framework to handle a multi-depot, multi-school
and multiple bell-time SBRP-SDDC. But in order to locate our model within the spectrum of
SBRP problems, we turn to the classification scheme used by Park and Kim [3]. Our problem
considers multiple schools in an urban area where the formulation can be used for both
morning and afternoon (however we limit the numerical example of the morning). No mixed
load are allowed and only general students are considered (as opposed of special-education students). The fleet considered is homogeneous, however we will show how the capacity of the bus will change depending on the school. Additionally, three chance constraints are considered in this paper.

1. Expectation of maximal travel times is less than $\Delta t$. Due to safety considerations, a limit on the amount of time students can spend on school buses is specified [19, 20, 14]. However, travel times are usually difficult to be accurately estimated because of many uncertain factors, such as weather conditions, traffic congestions, and student boarding/alighting times. The total travel time is decomposed into two parts: link travel time and bus stop time. It is assumed that link travel time follows a normal distribution and bus stop time is a linear function of number of students waiting at stops [14].

2. The probability of overcrowding a bus is less than $\alpha$. School districts, in order to better use their fleet, may overbook their buses. In other words, they can assign to a bus a number of students greater than the capacity, provided that is expected that not all students will actually ride the bus to school. However, a bus might end up being overcrowded if the overbooking level is too high. School bus overbooking has been largely ignored in the study of SBRP, even though it is crucial in practice. Although assigned to school buses, a large portion of students, especially high school students, tends to use other transportation modes. Figure 2.1 presents the histogram of the number of assigned students per bus of 50 student of capacity for all routes at Williamsville Central School District, Williamsville, NY. Twenty eight percent of buses are overbooked, with the number of assigned students exceeding bus capacity. We conclude that it is common practice for the transportation department of school districts to overbook buses on certain routes.
Figure 2.1: Histogram of student count.

3. The probability of being late to school is less than $\beta$. The starting time (bell time) of a school in the morning is one of the most important constraints in SBRP. Due to uncertainty in travel time, the chance of violating bell time exists.

Due to varying bell times for different schools, another new practical feature considered in this paper is multiple bell time school bus routing. In the morning or afternoon, each bus usually experiences multiple routes, denoted as a route set.

In this paper, a general multi-bell-time problem is decomposed into multiple single-bell-time problems. Schools are clustered into each time window according to their bell times. The first route is a typical depot-school route. The following route in a route set will originate from the schools in the current time window and end in the school whose bell time falls in the next time window.
2.2 Problem Description

This research began when Williamsville Central School District (WCSD), the largest suburban school district in Western New York, asked for assistance on their Transportation Operations Management Efficiency Program granted by the New York State Education Department. This paper focuses on one of the Program’s objectives, that is to increase the efficiency on bus routing.

WCSD encompasses 40 square miles including portions of the towns of Amherst, Clarence and Cheektowaga, enrolling over ten thousand K through 12 grade students in 13 public schools in 2012-13 school year. To successfully transport students WCSD uses a fleet of around a hundred buses, which is divided between their own fleet and a contractor’s fleet in a 2:3 ratio. The schools have different start and dismissal times; e.g. in the morning there are three schools starting at 7:45, four at 8:15 and six at 8:55.

Though WCSD provides transportation and assigns every student to a stop, not everyone uses the bus. School bus transportation in New York State is free and is required to be provided to all students by law. Parents and students have the ability to choose on any given day whether they want to access the bus transportation or not. Essentially, it is the same model that is in place for any public bus transportation system with the exception that our service takes students to their school and is free.

By studying the data gathered daily at the district, we found that the likelihood of a student using the bus is highly correlated to the school that he or she attends, and on whether it is a morning bus or an afternoon bus. Regardless if a student rides the buses, policy at WCSD states that all students are to have a stop assigned and all stops are to be visited regardless of the uncertainty of students not showing up. Thus, in order to have a better utilization of bus capacities, an overbooking policy is used resulting in having, for example, over a hundred students assigned to a 47-seat bus. In practice, such assignment situations
result in no more than 30 students actually riding the bus, indicating that formal study of
overbooking is worthwhile undertaking.

The overbooking policy, i.e. the limit on the number of students that can be assigned to
a bus, has been implemented gradually and only by observation. By reviewing the routes
on a yearly basis the district determines whether to update such limits or not, provided
that the information on actual ridership is available. Note that even though there are limits
for overbooking, many times these are not reached because of the length of route would
overpass the time windows provided for the runs. In other words, not all buses are assigned
the number of student that potentially could. In addition, the buses have a capacity of 71
students. In order to make it more comfortable for students, the capacity of 71 is only applied
to elementary students, and for the rest, middle and high school students, the capacity of the
buses is considered to be 47 (the bus’s capacity for adults). Finally, the number of assigned
students to a bus is determined in such a way that it is very unlikely to have an event of
overcrowding. However, if such event occurs, it would by the end of the route, in proximity
of the school and very close to the bell time, reasons why the students would simply be
squeezed into any seat.

Other routing related policies of WCSD are (i) walking distance restriction from a stu-
dent’s home to his or her designated stop; (ii) maximum riding time, and (iii) no mixed
loads on public schools routes. Saving opportunities for the school district come mainly
from reducing the maximum number of buses being used simultaneously at any given time
during the day, which provides the objective for this particular formulation of the SBRP.
The need for an additional bus implies either hiring a driver and purchasing a new bus, or
paying the contractor for another bus. Both these options are expensive.

For the 3 schools starting at 7:45 we have a 2-depot to 3-school SBRP. Say there are 50
buses at each depot, but we only use 20 of the first and 30 of the second for these 3 schools;
then, the problem for the set of schools starting at 8:15 is a 5-depot to 4-schools situation,
with the 5 depots being the 2 original that still have available buses and the 3 schools that have available buses from 7:45.

Since the objective is set to minimize the total number of buses used, there will be a set of constraints that will ensure a certain level of service that has to be met. When routing a bus, the risk of having a bus overcrowded or having a bus being late to school are not to be greater than a given threshold, and additionally there is an upper limit to the total time a student is expected to be riding the bus. Of course, all of WCSD’s policies must also be met.

2.3 Model Formulation

In this section we will formulate our problem based on the description provided in section 2.2. Our objective is to minimize the number of buses and secondary to minimize their length. As of the constraint considered we include bus capacity, maximum riding time and maximum walking distance.

Even though throughout Section 2.3 the focus will be in the formulating the bus routing problem, in Section 2.4 we will introduce a course of action that solves the problem sequentially for each school by first selecting the location of the stops and then solving the routing problem via column generation.
2.3.1 General model

In this section we represent the detailed formulation for the following conceptual model

\[
\begin{align*}
\text{Min} & \quad \text{number of buses used} + \varepsilon (\text{total travel time}) \\
\text{s.t.} & \quad P(\text{overcrowding the bus}) \leq \alpha, \forall \text{ bus} \\
& \quad P(\text{being late to school}) \leq \beta, \forall \text{ bus} \\
& \quad E(\text{maximum ride time}) \leq \Delta t, \forall \text{ bus}
\end{align*}
\]

where the objective (2.1) is, first, to minimize the total number of buses needed and then the total travel time, provided that \(\varepsilon\) is set as the inverse of an upper of such time. This would make \(\varepsilon (\text{total travel time}) \leq 1\), making the total number of buses the main objective. Thus, the weighted travel time encourage the generation of smoother routes.

Constraint (2.2) provides an upper bound for the likelihood of overcrowding the bus, constraint (2.3) provides an upper bound for the likelihood of a bus being late to school, and constraint (2.4) provides an upper bound for the expected maximum ride time of a student on any bus.

As it has been implied, the general model (routing for a whole morning or afternoon) is a succession of single bell-time multi-depot to multi-school routing problems, with the result of one being the input data for the next. A dynamic programming formulation captures the entire problem. An example of such formulation is as follows:

\[
f_n^* (\{b_{ik}\}_n, \{t_{av1}\}_n) = \\
\min_{\{x_{ijk}\}_n} \left\{ v (P_n) + f_{n+1}^* \left[ \Psi (\{b_{ik}\}_n, \{x_{ijk}\}_n), \Omega (\{t_{av1}\}_n, \{x_{ijk}\}_n) \right] \right\}, \quad n = 1, \ldots, N
\]

where \(f_{N+1}^* = 0\), \(v (P_n)\) represents the optimal value of a single bell-time stage multi-depot to multi-school routing problem, \(\{b_{ik}\}_n\) represents the initial position of the buses in stage \(n\),
Ψ operates \( \{b_{ik}\}_n \) and \( \{x_{ijk}\}_n \) to reposition the initial location of the buses for the following stage \( n + 1 \) and Ω operates the time at which the buses become available \( \{t_{avl}^k\}_n \) and the choice of routes \( \{x_{ijk}\}_n \) to reset the time at which buses are available for the following stage \( n + 1 \). Notice that if a bus is not used in a particular bell-time, \( t_{avl}^k \) remains the same in the next bell-time, making the model flexible enough to accommodate cases where a bus not used in a bell-time may be engaged in collecting student for future bell-times. A detailed definition of the parameter and variables is given in the following sections.

### 2.3.2 Single bell-time routing problem

Since the routing problem can be divided into separated periods of times, we define an MIP formulation for any given period.

Let us denote by \( D, A \) and \( S \) the set of depots, stops and schools, such that they are disjoint and \( D \cup A \cup S = L \) is the set of all locations. Let \( \mu_{T_{ij}} \) be the expected value of the travel time between locations \( i \) and \( j \) where \( (i, j) \in L^2 \), \( \mu_{T_i} \) the expected value of the waiting time or delay at location \( i \) where \( i \in A \), \( w_i \) the number of students assigned to stop \( i \in A \), \( a_{ij} \) equal to 1 if students at stop \( i \in A \) go to school \( j \in S \). And \( \kappa_i \) equal to 1 if depot \( i \in D \) is indeed a depot where buses are still idle and 0 if that depot represents in fact a school where there are buses ready to continue picking up students. Let \( B \) denote the set of buses and \( b_{ik} \) be equal to 1 if depot \( i \in D \) contains bus \( k \in B \) and 0 otherwise.

Let \( x_{ijk} \) be a binary decision variable that is equal to 1 when the edge \( (i, j) \in L^2 \) is covered by bus \( k \in B \) and 0 otherwise. Then, the single bell-time routing problem reads as follows:

\[
\begin{align*}
\text{Min} & \quad \sum_{k \in B} \sum_{i \in D} \sum_{j \in A} \kappa_i x_{ijk} + \varepsilon \sum_{k \in B} \sum_{i \in L} \sum_{j \in L} \left( \mu_{T_{ij}} + \mu_{T_i} \right) x_{ijk} \\
\text{s.t.} & \quad \sum_{k \in B} \sum_{i \in D \cup A} x_{ijk} = 1, \quad j \in A
\end{align*}
\]
The objective (2.6) minimizes the additional number of buses needed to run the corresponding bell time $\sum_{k \in B} \sum_{i \in D} \sum_{j \in A} \kappa_i x_{ijk}$ and as a secondary objective maintaining the total length of the routes $\sum_{k \in B} \sum_{i \in L} \sum_{j \in L} (\mu_{T_{ij}} + \mu_{T_{i}}) x_{ijk}$ to a minimum, $\varepsilon$ is set as the inverse of an upper bound for such length (the upper bound is found with the procedure described in section 2.4.2). The constraints ensure conditions as follow: (2.7) one and only one bus arrives to every stop, (2.8) one and only bus departures from every stop, (2.9) no bus stays at the same location nor arrives to a depot nor departures from a school, (2.10) same bus that arrives to a location departures from that location, (2.11) a bus only picks up...
students attending the same school, (2.12) a location can be visited by a bus only if that bus leaves the depot, (2.13) all buses start their route on their corresponding depot, (2.14) and (2.15) are the sub-tour elimination constraints where \( m \) is the maximum number of stops a bus can have, (2.16) to (2.18) are the stochastic constraints which will be developed in detail in the following section and (2.19) is the integrality condition.

2.3.3 Stochastic constraints

This section concentrates on the development of the stochastic constrains presented on the previous section that represent constraints (2.16) to (2.18).

Constraint on the likelihood of overcrowding the bus

On each route a bus will serve one and only one school. In practice, students do not always ride the bus and their decisions on whether to ride it or not is highly influenced by the grade, the school they attend and whether the route is done in the morning or in the afternoon. Also, a bus may have different capacity for different grades (e.g. a bus can hold up to 71 elementary students, whereas the capacity is set up to 47 with middle and high school students). Under such circumstances, though it is assumed to be using a homogeneous fleet, the bus capacity is dynamic and depends on the grade at which students attend and their choice on whether to ride the bus or not; the less willing the student are to ride the bus, the more students can be assigned to a bus, i.e., overbooking its capacity.

Definition 1. Let \( R_i \) be the actual number of students waiting at stop \( i \in A \). Then, \( R_i \) is a random variable following a Binomial distribution \( R_i \sim Bin(w_i, p_j) \) where \( w_i \) is the number of students assigned to stop \( i \in A \) and \( p_j \) the probability of any student attending school \( j \in S \) showing up at his or her stop, such that \( a_{ij} = 1 \). Note that the last condition requires that all students in any given stop must go to the same school.
Definition 2. Let \( Y_k = \sum_{i \in A} \sum_{j \in L} R_i x_{ijk} \) be the actual number of students riding bus \( k \in B \).

Then, \( Y_k \) is a random variable such that \( Y_k \sim \text{Bin}(Q_k, p_j) \) where \( Q_k = \sum_{i \in A} \sum_{j \in L} w_i x_{ijk} \) is the number of students assigned to bus \( k \in B \) and \( p_j \) the probability of any student attending school \( j \in S \) showing up at his or her stop.

Thus, the capacity constraint that represents (2.16) is given by:

\[
P(Y_k > c_k) \leq \alpha \quad \forall k \in B
\] (2.20)

where \( c_k \) is the capacity of bus \( k \), \( P(Y_k > c_k) = 1 - \sum_{v=0}^{c_k} \binom{Q_k}{v} (p_j)^v (1 - p_j)^{Q_k - v} \) for \( Q_k > c_k \) or 0 otherwise, is the probability of overcrowding the bus and \( \alpha \) is the upper bound on this probability.

In order to introduce (2.20) into the MIP problem in section 2.3.2, its representation needs to be transformed to a linear expression. Let \( q_{jk} \) be the maximum number of students that can be assigned to bus \( k \in B \) when going to school \( j \in S \). Then, the objective is to find how much overbooking is possible within a certain level of risk \( \alpha \). Thus,

\[
q_{jk} = \max \left\{ q \mid 1 - \sum_{v=0}^{c_k} \binom{Q_k}{v} (p_j)^v (1 - p_j)^{Q_k - v} \leq \alpha \right\}
\] (2.21)

Proposition 1. For all \( k \in B \) the constraint

\[
\sum_{i \in L} \sum_{j \in L} w_i x_{ijk} \leq \sum_{i \in A} \sum_{j \in S} q_{jk} x_{ijk}
\] (2.22)

is an equivalent inequality for (2.20).

Proof. We know that the right hand side of (2.22) chooses \( q_{jk} \) according to the school that
the bus is heading to. Thus, if (2.22) holds, then the following also holds

\[ 1 - \sum_{v=0}^{c_k} \binom{Q_k}{v} (p_j)^v (1 - p_j)^{Q_k - v} \leq 1 - \sum_{v=0}^{c_k} \binom{q_{jk}}{v} (p_j)^v (1 - p_j)^{q_{jk} - v} \]

and since for any \( q_{jk} \) the inequality \( 1 - \sum_{v=0}^{c_k} \binom{q_{jk}}{v} (p_j)^v (1 - p_j)^{q_{jk} - v} \leq \alpha \) holds, then (2.20) holds as well.

**Constraint on the likelihood of being late to school**

Since this SBRP considers transportation of students to their schools, the chance of arriving late to school must be assessed. At the same time, the buses are used to serve more than one school in different time spans; a bus picks up students from one school, drop them off and then starts a new route serving the second school and so on. The following definitions are made to account for these conditions.

**Proposition 2.** Let \( \tau_f \) and \( \tau_v \) represent the fixed and variable time when picking students up at each stop such that, if \( r \) students are to be picked up, it would take \( \tau_f + \tau_v r \) to do so. Then, given a stop location \( i \in A \) where there are \( w_i \) students assigned to go to school \( j \in S \) with a probability of showing up \( p_j \), the expected value and variance of the time required by a bus to pick them up are:

\[
\mu_{T_i} = \tau_f - \tau_f (1 - p_j)^{w_i} + \tau_v w_i p_j \tag{2.23}
\]

\[
\sigma_{T_i}^2 = \tau_f^2 (1 - p_j)^{w_i} (1 - (1 - p_j)^{w_i}) + 2\tau_f \tau_v w_i p_j (1 - p_j)^{w_i} + \tau_v^2 w_i p_j (1 - p_j) \tag{2.24}
\]

**Proof.** We know that \( R_i \), the actual number of students showing up at stop \( i \in A \), is a r.v. such that \( R_i \sim Bin (w_i, p_j) \). Then, let \( T_i (R_i) = \begin{cases} \tau_f + \tau_v R_i & \text{if } R_i > 0 \\ 0 & \text{if } R_i = 0 \end{cases} \) be the time that takes making a stop at node \( i \in A \). Then, the probability mass function (pmf) for \( T_i \) is
given by

\[
T_i(t_i) = \begin{cases} 
  p_{R_i}(0) & \text{if } t_i = 0 \\
  p_{R_i}(r_i) & \text{if } t_i = \tau_f + \tau_v r_i \\
  0 & \text{otherwise}
\end{cases}
\]

and the expected value and variance of \( T_i \) are then derived as follows:

\[
\begin{align*}
\mu_{T_i} &= E[T_i(R_i)] = \sum_{r=0}^{w_i} T_i(r) p_{R_i}(r) = 0 \cdot p_{R_i}(0) + \sum_{r=1}^{w_i} (\tau_f + \tau_v r) p_{R_i}(r) \\
&= \tau_f \sum_{r=1}^{w_i} p_{R_i}(r) + \tau_v \sum_{r=1}^{w_i} r p_{R_i}(r) = \tau_f \left[ \sum_{r=0}^{w_i} p_{R_i}(r) - p_{R_i}(0) \right] + \tau_v \sum_{r=1}^{w_i} r p_{R_i}(r) \\
&= \tau_f [1 - (1 - p_j)w_i] + \tau_v E[R_i] = \tau_f - \tau_f (1 - p_j)w_i + \tau_v w_i p_j
\end{align*}
\]

\[
\sigma^2_{T_i} = V[T_i(R_i)] = E[T_i(R_i)^2] - [E[T_i(R_i)]]^2 = \sum_{r=0}^{w_i} [T_i(r)^2] p_{R_i}(r) - \mu^2_{T_i}
\]

\[
= 0^2 \cdot p_{R_i}(0) + \sum_{r=1}^{w_i} (\tau_f + \tau_v r)^2 p_{R_i}(r) - \mu^2_{T_i}
\]

\[
= \tau_f^2 \sum_{r=1}^{w_i} p_{R_i}(r) + 2\tau_f \tau_v \sum_{r=1}^{w_i} r p_{R_i}(r) + \tau_v^2 \sum_{r=1}^{w_i} r^2 p_{R_i}(r) - \mu^2_{T_i}
\]

\[
= \tau_f^2 \left[ \sum_{r=0}^{w_i} p_{R_i}(r) - p_{R_i}(0) \right] + 2\tau_f \tau_v \sum_{r=0}^{w_i} r p_{R_i}(r) + \tau_v^2 \sum_{r=0}^{w_i} r^2 p_{R_i}(r) - \mu^2_{T_i}
\]

\[
= \tau_f^2 [1 - (1 - p_j)w_i] + 2\tau_f \tau_v E[R_i] + \tau_v^2 E[R_i^2] - \mu^2_{T_i}
\]

\[
= \tau_f^2 [1 - (1 - p_j)w_i] + 2\tau_f \tau_v w_i p_j + \tau_v^2 \left[ V[R_i] + E[R_i]^2 \right] - \mu^2_{T_i}
\]

\[
= \tau_f^2 [1 - (1 - p_j)w_i] + 2\tau_f \tau_v w_i p_j + \tau_v^2 \left[ w_i p_j (1 - p_j) + w_i^2 p_j^2 \right] - (\tau_f - \tau_f (1 - p_j)w_i + \tau_v w_i p_j)^2
\]

\[
= \tau_f^2 (1 - p_j)w_i (1 - (1 - p_j)w_i) + 2\tau_f \tau_v w_i p_j (1 - p_j) + \tau_v^2 w_i p_j (1 - p_j)
\]

\[
\square
\]

An estimation for the fixed and variable time for picking up students can be found in Braca et al. [14], where it was found \( \tau_f = 19 \) and \( \tau_v = 2.6 \) (both in seconds).

**Definition 3.** Let \( T_{ij} \) be the random travel time from location \( i \in L \) to location \( j \in L \) with
expected value and variance given by $\mu_{T_{ij}}$ and $\sigma^2_{T_{ij}}$ respectively.

**Definition 4.** Let $T_k = \sum_{i \in L} \sum_{j \in L} (T_{ij} + T_i) x_{ijk}$ be the total travel time for bus $k \in B$ with expected value and variance given by

$$\mu_{T_k} = \sum_{i \in L} \sum_{j \in L} (\mu_{T_{ij}} + \mu_{T_i}) x_{ijk} \tag{2.25}$$

$$\sigma^2_{T_k} = \sum_{i \in L} \sum_{j \in L} (\sigma^2_{T_{ij}} + \sigma^2_{T_i}) x_{ijk} \tag{2.26}$$

Then, the travel time constraint that represents (2.17) is given by:

$$P \left( t_{av1}^k + T_k > t_{bell} \right) \leq \beta \quad \forall k \in B \tag{2.27}$$

where $t_{av1}^k$ represents the time instant at which bus $k \in B$ becomes available, $t_{bell}$ the latest time instant at which the bus has to be at school and $\beta$ the given upper bound for the probability of bus $k \in B$ not making it on time to school. We now need to reformulate (2.27) such that it can be included in the single bell-time mixed integer linear program.

Given the previous definitions, $T_k$ represents the summation of the driving time $T_{ij}$ and the waiting time at stops $T_i$ of a particular bus. This is

$$T_k = T_{0,1} + T_1 + T_{1,2} + \ldots + T_{m-1,m} + T_m + T_{m,m+1} \tag{2.28}$$

where $m$ is the number of stops to be made by a bus. Then, $T_k$ is a summation of $2m + 1$ random variables.

**Conjecture 1.** The probability density function of $T_k$ can be approximated to a normal distribution with mean $\mu_{T_k}$ and variance $\sigma^2_{T_k}$ by means of the Central Limit Theorem.

Thus, we use the above conjecture in the following proposition in order to reformulate
(2.27) into a set of linear inequalities.

**Proposition 3.** For all \( k \in B \) the constraints

\[
t^k_{avl} + \mu T_k + \Phi^{-1} (1 - \beta) \tilde{\sigma} T_k \leq t_{bell}
\]  

(2.29)

\[
\sum_{h=1}^{h^+} h^2 \gamma^k_h \geq \sigma^2 T_k
\]  

(2.30)

\[
\sum_{h=1}^{h^+} h \gamma^k_h = \tilde{\sigma} T_k
\]  

(2.31)

\[
\sum_{h=1}^{h^+} \gamma^k_h = 1
\]  

(2.32)

are valid inequalities for (2.27), where \( \gamma^k_h \) is a binary variable and \( h^+ \) is the maximum possible integer value for \( \sigma T_k \).

**Proof.** From (2.27) it is obtained that

\[
P \left( t^k_{avl} + \tau_k < t_{bell} \right) \geq 1 - \beta
\]

which by standardizing becomes

\[
\Phi \left( \frac{t_{bell} - (t^k_{avl} + \mu T_k)}{\sqrt{\sigma^2 T_k}} \right) \geq 1 - \beta
\]

and by taking the inverse

\[
t^k_{avl} + \mu T_k + \Phi^{-1} (1 - \beta) \sqrt{\sigma^2 T_k} \leq t_{bell}
\]

where \( t_{bell} \), \( t^k_{avl} \) and \( \Phi^{-1} (1 - \beta) \) are constant numbers, and \( \mu T_k \) and \( \sigma^2 T_k \) are obtained as
stated in Definition 4. Notice that, as it is, the previous inequality is not linear. Then, the square root of \( \sigma_{T_k}^2 = \sum_{i \in L} \sum_{j \in L} \left( \sigma_{T_{ij}}^2 + \sigma_{T_i}^2 \right) x_{ijk} \) must be calculated while maintaining linearity.

Since \( \gamma_k^h \) is a binary variable, the assignment constraints (2.31) and (2.32) ensure that the variable \( \tilde{\sigma}_{T_k} \) will only take an integer value between 1 and \( h^+ = \lceil (t_{bell} - \min \{ t_{avl} \}) / 2 \rceil \) the maximum round up integer value the standard deviation can take. Then, the inequality in (2.30) constraints \( \tilde{\sigma}_{T_k} \) to be at least the round-up integer of \( \sqrt{\sigma_{T_k}^2} \).

Since now \( \sqrt{\sigma_{T_k}^2} \leq \tilde{\sigma}_{T_k} \), the following inequality holds:

\[
t_{avl}^k + \mu_{Tk} + \Phi^{-1} (1 - \beta) \sqrt{\sigma_{T_k}^2} \leq t_{avl}^k + \mu_{Tk} + \Phi^{-1} (1 - \beta) \tilde{\sigma}_{Tk}
\]

Therefore, if (2.29) is satisfied then (2.27) will also be satisfied.

Constraint on the expected maximum ride time

As part of the school district’s policy, it is expected that the average time a student spends on the bus should not be greater than a certain threshold \( \Delta t_{\text{max}} \). For this case, if we assure this condition to the first student who gets picked up, then the condition will apply to rest of the student in that bus as well.

Thus, the constraint that represents (2.18) reads as follows:

\[
\mu_{Tk} - \sum_{i \in D} \sum_{j \in A} \mu_{T_{ij}} x_{ijk} \leq \Delta t_{\text{max}} \quad \forall k \in B
\]

(2.33)

where \( \sum_{i \in D} \sum_{j \in A} \mu_{T_{ij}} x_{ijk} \) represents the expected time from the depot to the first stop.
2.3.4 Computational issues

The SBRP is a generalization of the VRP which is known to be NP-hard. As such, in this section we will show that it is essential to partition the general problem in a succession of sub problems (the following sections will show performance comparisons and the optimality gap).

For solving this problem the MIP formulation is programmed in Java 7 using the corresponding API of CPLEX 12.6 (64bit) in a computer running Windows 7 Enterprise (64bit) with a processor Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz and 15.90GB usable RAM. The customized settings for the branch and bound procedure are: node selection, best bound; variable selection, strong branching; branching direction, up branch selected first; relative MIP gap, 2%; absolute MIP gap, 0.5; and time limit of 1 hr. Also a priority order was issued to prioritize branching first on variable $x_{ijk} \forall i \in D, j \in A, k \in B$ which decides whether a bus leaves the depot or not, second on $x_{ijk} \forall i \in A, j \in S, k \in B$ which decides the destination of the bus, and then the rest.

Figure 2.2 shows, for a sample problem, the time needed to get to the optimal value for a single bell-time 2-depot to 3-schools routing problem (see details in Tables 2.1 and 2.2). A
Table 2.1: Details for solution time comparison

<table>
<thead>
<tr>
<th>Size</th>
<th>Variables</th>
<th>Nodes</th>
<th>Bound at root</th>
<th>Bound at end</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2912</td>
<td>20201</td>
<td>3.69</td>
<td>4.91</td>
<td>4.25</td>
</tr>
<tr>
<td>30</td>
<td>7716</td>
<td>52205</td>
<td>3.55</td>
<td>6.91</td>
<td>4.13</td>
</tr>
<tr>
<td>39</td>
<td>16048</td>
<td>401480</td>
<td>3.51</td>
<td>7.88</td>
<td>3.56</td>
</tr>
<tr>
<td>48</td>
<td>28880</td>
<td>115779</td>
<td>3.51</td>
<td>6.64</td>
<td>3.49</td>
</tr>
<tr>
<td>60</td>
<td>51792</td>
<td>38938</td>
<td>3.89</td>
<td>12.00</td>
<td>3.89</td>
</tr>
</tbody>
</table>

Table 2.2: Details for solution time comparison

<table>
<thead>
<tr>
<th>Size</th>
<th>Variables</th>
<th>Nodes</th>
<th>Bound at root</th>
<th>Bound at end</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2504</td>
<td>0</td>
<td>2.55</td>
<td>3.96</td>
<td>2.55</td>
</tr>
<tr>
<td>30</td>
<td>5740</td>
<td>318</td>
<td>2.35</td>
<td>4.84</td>
<td>2.38</td>
</tr>
<tr>
<td>39</td>
<td>10992</td>
<td>388</td>
<td>3.38</td>
<td>5.88</td>
<td>3.40</td>
</tr>
<tr>
<td>48</td>
<td>16068</td>
<td>2004</td>
<td>4.37</td>
<td>6.99</td>
<td>4.37</td>
</tr>
<tr>
<td>60</td>
<td>32464</td>
<td>618</td>
<td>4.33</td>
<td>9.00</td>
<td>4.35</td>
</tr>
</tbody>
</table>

size over 30 stop locations produces unreasonable running times. Clearly it is too expensive to try to solve the single bell-time problem to optimality. Moreover, it is hardly possible to solve the dynamic program with the multi bell-time routing problem at each stage to find the optimal solution for the real world entire morning problem. We conclude that a decomposition of the problem into a sequence of multi-depot to 1-school problems is needed.

2.4 Cascade Simplification

The concept behind this simplification is straightforward: the routing problem is solved for one school at a time. Consequently, we will have a succession of multi-depot to 1-school routing problems, where the solutions for one will be input data for the next. The following algorithm states the steps of this procedure.

Algorithm 1. Cascade simplification

1. Order schools by increasing bell time and then, for each school with same bell time, order
by decreasing ratio between expected riders and capacity of the bus \( \frac{\text{students} \times \text{ridership}}{\text{capacity}} \).

2. Select the first bell time

3. Position available buses at each depot

4. Select the first school in selected bell time

5. Set stops’ location, calculate the mean and variance of the waiting time at each stop, calculate the maximum number of students to be assigned to each bus, find starting solution, set routes and update the availability of buses at each depot.

6. Select the next school in selected bell time and go to step 5 if not all schools in selected bell time have been processed, otherwise go to step 7.

7. If selected bell time is not the last, let all schools in this bell time be depots

8. Select the next bell time and go to step 3 if not all bell times have been processed, otherwise go to step 9.

9. Send all buses to the original depots.

Notice that step 1 sets the order in which each school is selected within the bell times. Since the ratio between expected riders and capacity of the bus represents a lower bound on the amount of buses to be utilized for each school, this ordering criteria prioritizes schools with greater demand for buses.

### 2.4.1 Set stops’ location

As described in Section 2.2, our work is motivated by the WCSD’s Efficiency Program. One of the tasks that the district instructed us was to study the location of the bus-stops as part of the main objective of increase the efficiency on bus routing. The aforementioned give reasons for the inclusion of a stop selection procedure as part of our work.

The goal here is to minimize the total number of stops, subject to policy requirements established by WCSD. These are: (1) the maximum walking distance, defined as the distance...
from a student’s home to his or hers designated bus stop, and (2) the maximum number of
students assigned to a single bus-stop.

We now describe an IP formulation for the stop selection problem. Let $U$ be the set of all
potential bus-stop locations and $M$ the set of students, namely their home’s locations. Let $d_{ij}$ be the distance from student $i \in M$ to location $j \in U$, $\delta$ the maximum walking distance
and $\lambda$ the maximum number of students that can be assigned to a stop. Let $y_{ij}$ be the binary
decision variables that are equal to 1 if student $i \in M$ gets assigned to stop-location $j \in U$
and 0 otherwise; $z_i$ is equal to 1 if location $i \in U$ is set to be a stop and 0 otherwise. Then,
the stop location selection problem can be stated as follows:

$$\text{Min} \quad \sum_{j \in U} z_j + \epsilon \sum_{i \in U} \sum_{j \in U} d_{ij} y_{ij}$$

s.t.

$$\sum_{j \in U} y_{ij} = 1, \quad i \in M$$

$$\sum_{i \in U} y_{ij} \leq \lambda z_j, \quad j \in U$$

$$\sum_{j \in U} d_{ij} y_{ij} \leq \delta, \quad i \in M$$

$y_{ij}, z_i$ binary

where (2.34) minimizes the total number of stops and its second member homogenizes the
average walked distance as a secondary objective by setting $\epsilon = (\delta |M|)^{-1}$, (2.35) ensures that
every student gets assigned to one and only one bus-stop location, (2.36) that a maximum of
$\lambda$ students can be assigned to any stop-location, and (2.37) that no student walks more than
the maximum walking distance towards his or hers designated bus-stop. For simplification
in our application (see Section 2.5) every student’s resident represents a potential location
for a bus-stop.

The stops’ location definition has a direct effect in the total number of buses needed:
a bigger number of stops will increase the length of the route, hence the need for more
buses. Additionally, the effect of more stops will increase the computational effort, given
the complexity of the routing problem. Thus, an appropriate choice of the stops’ location is
worthy of attention in our study.

2.4.2 Finding an initial solution to the single school routing prob-
lem

In this section we implement a simple heuristic with the sole purpose of generating an initial
solution that will be used in the procedure described in section 2.4.3. In the literature one
can find several works that focuses on the development of heuristics to solve the school
bus routing problem, such that of Corberan et al. [21] and Alabas-Uslu [22]. However we
will limit our work to implement a greedy algorithm combining Clarke and Wright saving
algorithm Clarke and Wright [23] and the Farthest First Heuristic proposed by Fu, Eglese,
and Li [24].

The idea of Algorithm 2 is to initiate the heuristic by creating a new route and assigning
the farthest stop to it (steps 1 and 2); then, while complying with the capacity and time
constraints, stops should be added to the route prioritizing those that add the less time to
the route and are the farthest (steps 3 and 4). As indicating in [24], the strategy of not
starting a new route until the vehicle can’t hold any more stops due to the capacity or time
constraints, aims for a solution that keeps the number of buses needed to a minimum.

Algorithm 2. The initial solution for the single school routing problem

1. For each stop \( i \in A \) set \( s_i = \min_{d \in D} \{ \mu_{T_{di}} \} + \mu_{T_{ij}} \) where \( s_i \) is the travel time of serving
   stop \( i \) with one bus exclusively and \( j \) is its corresponding school.
2. Find a stop \( i^* \in A \) such that \( s_{i^*} = \max \{ s_i \} \), set NewRoute as a new route, set depot
   at \( d^* \in D \) such that \( \mu_{T_{di^*}} = \min_{d \in D} \{ \mu_{T_{di^*}} \} \), set first stop at \( i^* \) and remove it from set
A and set second stop at corresponding school.

3. For each \( i \in A \) that attend new route’s school and for each stop \( j \) in NewRoute set \( j^- \) as the stop before \( j \) and \( s_{ij} = \mu_{T_i} + \mu_{T_{j^-}} + \mu_{T_{j^-}} - \mu_{T_{j^-}} - \mu_{T_{j^+}} \) where \( s_{ij} \) is the saving in travel time of pulling students in stop \( i \) from their exclusive (hypothetical) bus into NewRoute. If inserting \( i \) before \( j \) is not feasible, set \( s_{ij} = -\infty \).

4. Add into NewRoute stop \( i^* \) before stop \( j^* \) such that \( s_{i^*,j^*} = \max\{s_{ij}\} \) and remove \( i^* \) from \( A \).

5. Go to step 3 until \( s_{ij} = -\infty \ \forall (i,j) \).

6. Update availability of buses at each depot. If any depot has no bus, remove this depot from set \( D \) and recalculate \( s_i = \min_{d \in D} \{\mu_{T_d}\} + \mu_{T_{ij}} \) where \( j \) is the corresponding school.

7. Go to step 2 until \( A = \emptyset \).

The results obtained with this algorithm are used to set \( \varepsilon \) in (2.6) as the inverse of the summations of the travel time of all buses. Further, the initial solution can reduce the given total number of buses available in the problem with the purpose of reducing the amount of decision variables.

### 2.4.3 Solving the single school routing problem

The model formulation for this problem is identical to the one in section 2.3.2, with \(|S| = 1\). Since there is only one school in the problem, constraint (2.11) is dropped. While the problem remains NP-hard, this size is far smaller and more likely to allow for a close to optimal solution in a reasonable time.

In order to take advantage of the formulation’s structure, in the following sections we turn to column generation as mean of finding good solutions to the single school routing problem. Our implementation is based on the standard procedure used in the literature.
for column generation (see [25] for a related implementation), but in addition we included
features as part of our acceleration strategy. Hereunder we first present the decomposition
of our formulation followed by the implementation of the column generation procedure and
a set of computational experiment.

A column generation based approach

A closer look at the model in section 2.3.2 reveals that only constraints (2.7) and (2.8)
combine the vehicles while the rest deal with each vehicle separately. This strongly suggests
the use of decomposition to break up the overall problem into a master problem (MP) and
a subproblem (SP) for each vehicle.

The master problem

Let $P_k$ be the set of feasible paths for bus $k \in B$, where $p \in P_k$ is an elementary path. Let
$x^p_{ijk}$ be equal to 1 if edge $(i, j) \in L^2$ is covered by bus $k \in B$ when using path $p \in P^k$,
$\theta^p_k = \sum_{i \in D} \sum_{j \in A} k_i x^p_{ijk} + \varepsilon \sum_{i \in L} \sum_{j \in L} (\mu_{Ti} + \mu_{Tij}) x^p_{ijk}$ be the cost of using path $p \in P^k$
with vehicle $k \in B$ and $\nu^p_{ik} = \sum_{j \in A \cup S} x^p_{ijk}$ be equal to 1 if stop $i \in A$ is visited by bus $k \in B$
when using path $p \in P_k$ and 0 otherwise. Let $y^p_k$ be the binary decision variables that are
equal to 1 if path $p \in P_k$ is used by bus $k \in B$ and 0 otherwise. Then, the MP reads as
follows:

Min $\sum_{k \in B} \sum_{p \in P_k} \theta^p_k y^p_k \quad (2.39)$

s.t. $\sum_{k \in B} \sum_{p \in P_k} \nu^p_{ik} y^p_k = 1, \; i \in A \quad (2.40)$

$\sum_{p \in P_k} y^p_k \leq 1, \; k \in B \quad (2.41)$

$y^p_k$ binary \quad (2.42)
Notice that because the fleet of buses is homogeneous in regard to their capacity, we could potentially drop $k$ in our formulation in order to break down the symmetry. However, the buses may have different times of availability and also be positioned at different depots. Therefore, let $R$ define the set of unique bus classes, where each element $r \in R$ represents a bus class with distinct pairs of time of availability and depot, and let $K_r$ be the number of available buses for each class. Then, a new MP formulation with considerably less variables reads as follows:

$$\text{Min} \sum_{r \in R} \sum_{p \in P_r} \theta_p^r y_p^r$$

s.t. $\sum_{r \in R} \sum_{p \in P_r} \nu_{ir}^p y_p^r = 1, \quad i \in A$ \hspace{1cm} (2.44)

$\sum_{p \in P_r} y_p^r \leq K_r, \quad r \in R$ \hspace{1cm} (2.45)

$y_p^r \text{ binary}$ \hspace{1cm} (2.46)

The subproblem

Since the buses are based at different depots and have different time of availability, one SP must be solved for each bus class. Thus, there will be $|R|$ SPs to solve separately, each one with $|D| = |S| = 1$.

Let $\pi_i$ represent the dual variables associated with constraints (3.21) and $\rho_r$ represent the dual variables associated with constraints (3.22). Then, for a given bus the SP minimizes the reduced cost $\theta_p^r - (\sum_{i \in A} \pi_i \nu_{ir}^p + \rho_r)$. Thus, the SP for class $r \in R$ reads as follows:

$$\text{Min} \quad \kappa - \rho + \sum_{i \in L} \sum_{j \in L} \left[ \varepsilon (\mu_{T_{ij}} + \mu_{T_{ji}}) - \pi_i \right] x_{ij}$$

s.t. $\sum_{j \in A \cup S} x_{ij} \leq 1, \quad i \in A$ \hspace{1cm} (2.48)
\[
\sum_{i \in L} \left( x_{ii} + \sum_{j \in D} x_{ij} + \sum_{j \in S} x_{ji} \right) = 0 \quad (2.49)
\]

\[
\sum_{j \in A} x_{ij} = 1, \quad i \in D \quad (2.50)
\]

\[
\sum_{i \in D \cup A} x_{ij} = \sum_{i \in A \cup S} x_{ji}, \quad i \in A \quad (2.51)
\]

\[
\sum_{i \in A} x_{ij} = 1, \quad j \in S \quad (2.52)
\]

\[
1 \leq u_i \leq m + 2, \quad i \in L \quad (2.53)
\]

\[
u_i - u_j + (m + 2)x_{ij} \leq m + 1, \quad i \in L, j \in L \quad (2.54)
\]

\[
\sum_{i \in L} \sum_{j \in L} w_i x_{ij} \leq q \quad (2.55)
\]

\[
ts_{avl} + \mu_T + \Phi^{-1} (1 - \beta) \bar{\sigma}_T \leq t_{bell} \quad (2.56)
\]

\[
\sum_{h=1}^{h^+} h^2 \gamma_h \geq \sigma_T^2 \quad (2.57)
\]

\[
\sum_{h=1}^{h^+} h \gamma_h = \bar{\sigma}_T \quad (2.58)
\]

\[
\sum_{h=1}^{h^+} \gamma_h = 1 \quad (2.59)
\]

\[
\mu_T - \sum_{j \in A} \mu_{T_j} x_{ij} \leq \Delta t_{\max}, \quad i \in D \quad (2.60)
\]

\[
x_{ij}, \gamma_h \text{ binary} \quad (2.61)
\]

where \( m = \max \{|A| : \sum_{i \in A} w_i \leq q \land A \subset A \} \) is the maximum number of stops a bus can visit.

In order to decrease the size of the solution space of the SP we added the following constraints to the formulation:

\[
l_i + \mu_{T_i} + \mu_{T_{ij}} \leq l_j + M (1 - x_{ij}), \quad i \in L, j \in L \quad (2.62)
\]
\[ t_{\text{avl}} + \mu T_{di} \leq l_i \leq t_{\text{bell}} - \mu T_i - \mu T_{is}, \quad i \in L, d \in D, s \in S \] (2.63)

where \( l_i \) is a decision variable representing the time at which a bus arrives to stop \( i \in L \) and \( M = \max\{t_{\text{bell}} - \mu T_i - \mu T_{is} + \mu T_{ij} - t_{\text{avl}} - \mu T_{dj}\} \), (2.62) establish the relation between the arrival time to one stop and its immediate successor and (2.63) define the time windows for the bus arrival to each stop. By adding this set of constraints we aim to attain stronger lower bounds when solving the relaxation of the problem within the branch and bound procedure.

Solution strategy

A special column generation procedure is designed which includes several rules that aim to obtain good quality solutions in a reasonable time. Much like the work of Krishnamurthy, Batta, and Karwan [26], Barnhart, Kniker, and Lohatepanont [27], Patel, Batta, and Nagi [28] and Ceselli, Righini, and Salani [29], our strategy is heuristic in nature as the generation of columns will be only allowed in the root of the branch and bound tree of the Master Problem. Thus, we sacrifice optimality over computational time, which is reduced significantly given that the column generation procedure is used only once at the root node of the math program.

The implementation of the column generation procedure is based on the standard practice available in the literature. However, as part of our own acceleration strategy, we introduced a series of rules within the procedure as depicted in Figure 2.3. The following elaborates in such rules:

**Rule 1.** The MP is in fact restricted since it only deals with the generated set of routes or columns. Then, an initial start for the restricted master problem (RMP) is provided by the set routes obtained from Algorithm 2.
Rule 2. When solving the relaxed RMP we replace (3.21) and (3.22) with

\[
\sum_{r \in R} \sum_{p \in P_r} \nu_{ir}^p y_{ir}^p \geq \varphi, \quad i \in A
\]

(2.64)

\[
\sum_{p \in P_r} y_{ir}^p \leq \varphi K_r, \quad r \in R
\]

(2.65)

respectively, where \(\varphi\) is an integer greater than 1 (by default we set \(\varphi = 2\)). Such modification intends to amplify the values of the dual variables that later will be included in the corresponding SP.

Rule 3. When solving the SP the branch and bound procedure is terminated if: best integer known solution, the incumbent in less than \(-1.5\) (a threshold implying the solution’s potential of reducing the number of buses in at least one unit in the RMP), or elapsed time is greater than 20 sec., or relative gap is lower than 10%.
Rule 4. All feasible solutions found when solving the SP are added into the RMP as new columns if their objective value is negative.

Rule 5. When solving the integer completion of the RMP we do not consider those variables with reduced cost higher than average nor those variables that are dominated (their stops are covered by other less expensive variable, see [30]) and we replace (3.21) with

$$
\sum_{r \in R} \sum_{p \in P_r} \nu^p_{i} y^p_{i} \geq 1, \quad i \in A
$$

(2.66)

where such modification allows the existence of repeated stops. Additionally, all known solutions are included as the algorithm starts, and the branch and bound procedure is terminated if: incumbent – best bound ≤ 0.1, or incumbent is better than last known solution and elapsed time > 2 min., or elapsed time > 10 min.

Rule 6. Every 50 iterations the integer completion of the RMP is checked. If the solution of this check contains repeated stops, the solution is modified to only contain unique stops. Such modified solution is added to the RMP.

Rule 7. The column generation procedure is terminated if: objective > −0.05 ∀ SP, or the last integer check shows no improvement.

Computational experience

In order to obtain good quality solutions in a reasonable time we need to establish a suitable configuration of the different rules within the column generation procedure. Therefore, we perform a $2^{6-2}$ fractional factorial design, where in the expression of form $l^{k-p}$ the parameter $l$ is the number of levels of each factor investigated, $k$ is the number of factors investigated, and $p$ describes the size of the fraction of the full factorial used. For the design we considered
The following factors: (A) whether to apply Rule 2 or not, (B) whether to apply Rule 6 or not, (C) whether to apply Rule 4 or not, (D) the time in seconds for terminating a SP in Rule 3, (E) value of the incumbent for terminating a SP in Rule 3, and (F) the objective value of SPs for terminating the column generation procedure in Rule 7. In addition, we consider two responses in the design: the value objective function and the CPU time needed to obtain such value. For every run we solve the same single school routing problem with 117 stops, the third of the real instances at WCSD (results for all instances are later in Table 2.4). Table 2.3 shows the treatment combinations and the results obtained. Figures 2.4 and 2.5 shows the iteration plot for the mean of both responses.

From both Table 2.3 and Figures 2.4 and 2.5 we conclude the following. The application of Rule 2 improves the objective value significantly without causing a considerable increase in the running time; applying Rule 6 saves a great amount of running time but with a notable

### Table 2.3: Factorial design for the single school routing problem

<table>
<thead>
<tr>
<th>Run</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>CPU*</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>5</td>
<td>-1.5</td>
<td>-0.05</td>
<td>36.0</td>
<td>11.712</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>5</td>
<td>-0.5</td>
<td>-0.005</td>
<td>164.4</td>
<td>10.652</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>5</td>
<td>-0.5</td>
<td>-0.005</td>
<td>31.1</td>
<td>11.670</td>
</tr>
<tr>
<td>4</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>20</td>
<td>-0.5</td>
<td>-0.05</td>
<td>53.3</td>
<td>11.629</td>
</tr>
<tr>
<td>5</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>5</td>
<td>-0.5</td>
<td>-0.005</td>
<td>36.2</td>
<td>11.696</td>
</tr>
<tr>
<td>6</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>20</td>
<td>-0.5</td>
<td>-0.05</td>
<td>23.4</td>
<td>12.751</td>
</tr>
<tr>
<td>7</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>5</td>
<td>-1.5</td>
<td>-0.05</td>
<td>2.3</td>
<td>13.837</td>
</tr>
<tr>
<td>8</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>20</td>
<td>-1.5</td>
<td>-0.005</td>
<td>2.2</td>
<td>13.837</td>
</tr>
<tr>
<td>9</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>5</td>
<td>-0.5</td>
<td>-0.05</td>
<td>35.8</td>
<td>11.669</td>
</tr>
<tr>
<td>10</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>5</td>
<td>-0.5</td>
<td>-0.005</td>
<td>170.2</td>
<td>10.625</td>
</tr>
<tr>
<td>11</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>5</td>
<td>-1.5</td>
<td>-0.005</td>
<td>36.6</td>
<td>10.623</td>
</tr>
<tr>
<td>12</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>20</td>
<td>-1.5</td>
<td>-0.05</td>
<td>60.1</td>
<td>10.667</td>
</tr>
<tr>
<td>13</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>5</td>
<td>-1.5</td>
<td>-0.005</td>
<td>35.8</td>
<td>11.684</td>
</tr>
<tr>
<td>14</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>20</td>
<td>-1.5</td>
<td>-0.05</td>
<td>43.2</td>
<td>12.762</td>
</tr>
<tr>
<td>15</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>5</td>
<td>-0.5</td>
<td>-0.05</td>
<td>27.4</td>
<td>12.769</td>
</tr>
<tr>
<td>16</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>20</td>
<td>-0.5</td>
<td>-0.005</td>
<td>37.4</td>
<td>11.734</td>
</tr>
</tbody>
</table>

(*) CPU time in minutes.
Figure 2.4: Results from factorial design.

Figure 2.5: Results from factorial design.
degradation of the objective value, whereas applying Rule 4 shows savings in running time with no significant degradation of the objective value. In regards to Rule 3, setting a shorter time to terminate the SP, saves time with no significant impact on the objective; moreover, the value of the incumbent as criteria for terminating the SP shows insignificant impact in both the running time and the objective value. The application of Rule 7, seeking a value closer to zero for the SPs, improves the objective value with an increase on the running time depending on the settings for Rule 4 and Rule 6. Finally, an appropriate setting that yields a good solution in reasonable time is the one for run 11 in Table 2.3.

2.5 Model Application to Williamsville Central School District

2.5.1 Data Gathering

The Transportation Department at WCSD uses Versatrans [31] as their student transportation management solution, where all information related to students, routes and buses is handled. However, the software only allows to manually build routes and allocate bus stops, i.e., neither is computer generated (Figure 2.6 shows the route editing tool of Versatrans with which routes are manually constructed and students are assigned to stops). From this database we obtained the students’ addresses which were revised and corrected so that they would be identified by mapping engines. Then, the distance and time matrices were obtained using the Open Directions API offered by MapQuest in a process that took several days.

As for the waiting time at each stop, in (2.23) and (2.24) we use $\tau_f = 19$ sec. and $\tau_v = 2.6$ sec. as the estimation for the fixed and variable time for picking up students [14].

WCSD utilizes up to a hundred buses (self-owned and from a contractor) to meet routing requirements. The fleet for the regular students can be said to be homogeneous; in general,
Figure 2.6: Student transportation management solution.
all buses can handle 47 middle and high school students or 71 elementary students. For the set of students involving public schools, WCSD currently uses a fleet of 86 buses. This is the instance in which we developed this study. The routes start right after 6:00 AM with the first stage being for the high school students. For the high school set of routes, overbooking has been frequently used; high school students are the ones who use the buses the least. The probability of a student not showing up to his or her designated stop (or ridership) was estimated over daily data collected throughout two weeks in January, 2013. These data consist of the head count for each bus in the morning and afternoon. Results depending on the schools at which a student attends varies from 22% to 72% (see Table 2.4). Thus, overbooking the buses according to the school they go to is an appropriate strategy to a better utilization of the bus capacity. Figure 2.7 shows the overbooked capacity as a function of the ridership, and display the ranges of the ridership for each group of students.

2.5.2 Results

In this section, in order to measure performance, different sample instances of the WCSD problem are solved as both the multi-depot to multi-school (single bell time) and the multi-
Before applying any procedure to generate the routes, the first step is to set the stop locations using the model presented in section 2.4.1. The model is applied to each school separately and for the WCSD case the parameters were set as follow: \( \lambda = 15 \) (students), \( \delta = 0.1 \) (miles) for elementary school students and \( \delta = 0.2 \) for middle school and high school students.

Figure 2.8 shows the solution time for the same instance (with 50 stop locations), a two-depot to three-school problem, while comparing the performance when applying the multi-depot to multi-school model and its decomposition using the cascade simplification. Savings on computational time are significant due to partitioning the problem for each school that implies a reduction on the buses needed and stops visited in the SP. Though there exists a degradation in optimality of the objective (2.6), for this sample problem the total amount of buses used remains the same for both procedures (at 7 buses), where the difference is the total travel time of all buses (152 min for the multi-depot to multi-school and 174 min for the Cascade Simplification).

Table 2.4 shows the solution found with the cascade simplification applied to all 13 school in the morning run. For each school we present the group of the grade of its students (EL: elementary, MI: middle, and HI: high), the drop off time (which is precedes the correspond-
Table 2.4: Results for the cascade simplification applied to WCSD

<table>
<thead>
<tr>
<th>School</th>
<th>Grade</th>
<th>Drop off time</th>
<th>Students</th>
<th>Stops</th>
<th>Current sol.</th>
<th>Initial sol.</th>
<th>Improved sol.</th>
<th>Buses</th>
<th>Travel*</th>
<th>Iterations</th>
<th>CPU time**</th>
<th>Saved</th>
<th>Buses travel</th>
<th>Total morning bell time</th>
</tr>
</thead>
<tbody>
<tr>
<td>s01</td>
<td>HI</td>
<td>7:20</td>
<td>1237</td>
<td>172</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>1</td>
<td>1002</td>
<td>1</td>
<td>10</td>
<td>391</td>
<td>10</td>
<td>390</td>
</tr>
<tr>
<td>s02</td>
<td>HI</td>
<td>7:20</td>
<td>850</td>
<td>117</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>3</td>
<td>653</td>
<td>24</td>
<td>24</td>
<td>310</td>
<td>635</td>
<td>476</td>
</tr>
<tr>
<td>s03</td>
<td>HI</td>
<td>7:20</td>
<td>1008</td>
<td>132</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>2</td>
<td>695</td>
<td>21</td>
<td>21</td>
<td>310</td>
<td>635</td>
<td>476</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s04</td>
<td>EL</td>
<td>8:05</td>
<td>692</td>
<td>177</td>
<td>16</td>
<td>15</td>
<td>15</td>
<td>1</td>
<td>800</td>
<td>17</td>
<td>17</td>
<td>33</td>
<td>15</td>
<td>800</td>
</tr>
<tr>
<td>s05</td>
<td>EL</td>
<td>8:05</td>
<td>622</td>
<td>169</td>
<td>13</td>
<td>14</td>
<td>14</td>
<td>7</td>
<td>680</td>
<td>28</td>
<td>28</td>
<td>44</td>
<td>14</td>
<td>468</td>
</tr>
<tr>
<td>s06</td>
<td>EL</td>
<td>8:05</td>
<td>555</td>
<td>124</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>579</td>
<td>32</td>
<td>32</td>
<td>50</td>
<td>10</td>
<td>378</td>
</tr>
<tr>
<td>s07</td>
<td>EL</td>
<td>8:05</td>
<td>504</td>
<td>156</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>7</td>
<td>690</td>
<td>25</td>
<td>25</td>
<td>492</td>
<td>13</td>
<td>663</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s08</td>
<td>MI</td>
<td>8:45</td>
<td>1009</td>
<td>135</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td>1</td>
<td>971</td>
<td>99</td>
<td>99</td>
<td>472</td>
<td>1627</td>
<td>964</td>
</tr>
<tr>
<td>s09</td>
<td>MI</td>
<td>8:45</td>
<td>883</td>
<td>127</td>
<td>18</td>
<td>15</td>
<td>15</td>
<td>4</td>
<td>761</td>
<td>52</td>
<td>52</td>
<td>312</td>
<td>1460</td>
<td>755</td>
</tr>
<tr>
<td>s10</td>
<td>MI</td>
<td>8:45</td>
<td>679</td>
<td>90</td>
<td>16</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>590</td>
<td>90</td>
<td>90</td>
<td>399</td>
<td>1180</td>
<td>590</td>
</tr>
<tr>
<td>s11</td>
<td>MI</td>
<td>8:45</td>
<td>644</td>
<td>88</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>6</td>
<td>516</td>
<td>52</td>
<td>52</td>
<td>448</td>
<td>1412</td>
<td>240</td>
</tr>
<tr>
<td>s12</td>
<td>EL</td>
<td>8:45</td>
<td>663</td>
<td>154</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>643</td>
<td>149</td>
<td>149</td>
<td>1092</td>
<td>1893</td>
<td>305</td>
</tr>
<tr>
<td>s13</td>
<td>EL</td>
<td>8:45</td>
<td>572</td>
<td>144</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td>10</td>
<td>697</td>
<td>139</td>
<td>139</td>
<td>1085</td>
<td>1553</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal bell time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s01</td>
<td>HI</td>
<td>7:20</td>
<td>1237</td>
<td>172</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>1</td>
<td>1002</td>
<td>1</td>
<td>10</td>
<td>391</td>
<td>10</td>
<td>390</td>
</tr>
<tr>
<td>s02</td>
<td>HI</td>
<td>7:20</td>
<td>850</td>
<td>117</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>3</td>
<td>653</td>
<td>24</td>
<td>24</td>
<td>310</td>
<td>635</td>
<td>476</td>
</tr>
<tr>
<td>s03</td>
<td>HI</td>
<td>7:20</td>
<td>1008</td>
<td>132</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>2</td>
<td>695</td>
<td>21</td>
<td>21</td>
<td>310</td>
<td>635</td>
<td>476</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s04</td>
<td>EL</td>
<td>8:05</td>
<td>692</td>
<td>177</td>
<td>16</td>
<td>15</td>
<td>15</td>
<td>1</td>
<td>800</td>
<td>17</td>
<td>17</td>
<td>33</td>
<td>15</td>
<td>800</td>
</tr>
<tr>
<td>s05</td>
<td>EL</td>
<td>8:05</td>
<td>622</td>
<td>169</td>
<td>13</td>
<td>14</td>
<td>14</td>
<td>7</td>
<td>680</td>
<td>28</td>
<td>28</td>
<td>44</td>
<td>14</td>
<td>468</td>
</tr>
<tr>
<td>s06</td>
<td>EL</td>
<td>8:05</td>
<td>555</td>
<td>124</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>579</td>
<td>32</td>
<td>32</td>
<td>50</td>
<td>10</td>
<td>378</td>
</tr>
<tr>
<td>s07</td>
<td>EL</td>
<td>8:05</td>
<td>504</td>
<td>156</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>7</td>
<td>690</td>
<td>25</td>
<td>25</td>
<td>492</td>
<td>13</td>
<td>663</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s08</td>
<td>MI</td>
<td>8:45</td>
<td>1009</td>
<td>135</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td>1</td>
<td>971</td>
<td>99</td>
<td>99</td>
<td>472</td>
<td>1627</td>
<td>964</td>
</tr>
<tr>
<td>s09</td>
<td>MI</td>
<td>8:45</td>
<td>883</td>
<td>127</td>
<td>18</td>
<td>15</td>
<td>15</td>
<td>4</td>
<td>761</td>
<td>52</td>
<td>52</td>
<td>312</td>
<td>1460</td>
<td>755</td>
</tr>
<tr>
<td>s10</td>
<td>MI</td>
<td>8:45</td>
<td>679</td>
<td>90</td>
<td>16</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>590</td>
<td>90</td>
<td>90</td>
<td>399</td>
<td>1180</td>
<td>590</td>
</tr>
<tr>
<td>s11</td>
<td>MI</td>
<td>8:45</td>
<td>644</td>
<td>88</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>6</td>
<td>516</td>
<td>52</td>
<td>52</td>
<td>448</td>
<td>1412</td>
<td>240</td>
</tr>
<tr>
<td>s12</td>
<td>EL</td>
<td>8:45</td>
<td>663</td>
<td>154</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>643</td>
<td>149</td>
<td>149</td>
<td>1092</td>
<td>1893</td>
<td>305</td>
</tr>
<tr>
<td>s13</td>
<td>EL</td>
<td>8:45</td>
<td>572</td>
<td>144</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td>10</td>
<td>697</td>
<td>139</td>
<td>139</td>
<td>1085</td>
<td>1553</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal bell time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total morning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*) travel: total travel time (min).
(**) time: CPU time in minutes.
ing bell time), the number of students and stop locations, and the ridership (the average percentage of student that actually ride the bus). The current number of buses represents the practice of WCSD. The column “initial sol.” is the solution used to start the column generation procedure and is found with Algorithm 2. Basic information of the column generation procedure is shown. The column “improved sol.” corresponds to the solution provided by the column generation procedure, and the number of buses saved shows the potential of savings for the district.

We observe savings in the number of buses at each bell time, but not for every school. For those schools where no improvement is found we keep the current routes. When looking at the reason as to why our approach yields inferior solutions, we find a number factors which may contribute the most to this effect. First, the overbooking for middle school is often set above the threshold used in our experiment. This is done with no major concern since the buses can in fact hold up to 71 students whereas the capacity considered to assign the students is 47 (recall that a bus is set to hold 47 high school or middle school students, or 71 elementary students). And provided that students in middle school are not all grown up, it is not of big concern to have some route with more than 47 students that actually ride. Second, our work considers a limit in the probability of getting late to school, however in the current practice only the expected value of the travel time is considered to set the length of the routes. This makes the routes in our work inherently shorter than the potential in the current practice.

The last bell time contains the highest amount of concurrent students to be transported to school, hence producing a spike in the number of buses needed to a total of 86 in the current state, where the Cascade Simplification saves 9 buses. Since the last bell time needs the highest number of buses, it determines the total amount of buses needed throughout the morning. Therefore, the Cascade Simplification reduced the number of buses to a total of 77, observing a 10% reduction for the entire fleet from the current practice.
2.5.3 Implementation

We now relate the results of this research to some implementation issues encountered in the School District.

The decision making process of school bus routing is not based only on efficiency criteria. School bus routing is highly sensitive to the public’s opinion, particularly to the families of students that utilize this service and that have become accustomed to a certain schedule. At the same time, route changes affects the drivers that become concerned about reducing their hours or possible firings. Then, not only costs are to be consider on the implementation process, but also the students and drivers need to taken into account.

An additional implementation issue is the change of the contractor after the latest bidding process, which will start operation at the beginning of the 2014-2015 school year. We need to consider that the previous contractor worked with the district for over 2 decades and that they carried out about two-thirds of the transportation operation of the district. Therefore, measures to ensure a smooth transition need to be in place, including maintaining the contractor’s routes for the beginning of the 2014-2015 school year as they were by the end of the previous year.

The results found in section 2.5.2 show potential for savings in the route set of 8 schools. Implementation of new routes for these schools would reach the schedule of over 6,000 students and their families, and at the same time all of the drivers, both in the district and the contractor side. Therefore, the District decides to make a gradual transition that aims to close the gap between the current situation and the ideal potential in the long term while acknowledging all of the issues previously presented.

A simple procedure of route merging and student re allocation to nearby routes was introduced. The basic steps are:

1. find a route with potential for deletion, i.e., lowest capacity usage;
2. find near by routes with the capacity to receive all or a fraction of the students from the route to be deleted;

3. merge these routes and redistribute students following the corresponding overlapping routes found with the Cascade Simplification;

4. check feasibility in both capacity and time; and

5. repeat until feasibility cannot be found or some non efficiency related criteria is meet.

This procedure attains a very localized number of routes, reducing significantly the effect of the change on non efficiency related factors. Carrying out this procedure is fairly easy given the tools provided in Versatrans, their student transportation management solution shown in Figure 2.6. Additionally, the set of new routes from Table 2.4 serves as reference of a final state of the routes in step 3; therefore, the closer the resulting routes are to the reference set, the faster the projected savings will be attained.

A driving factor for savings in the number of buses is the definition of the overbooked capacity for all schools based on their ridership. The introduction of a formal procedure to obtain this number for each of the schools revealed that a significant number of routes had far more free room than previously thought, making it easier to check for feasibility of capacity in the latter procedure.

Our joined venture with WCSD will continue in a monitoring phase that the Transportation Operations Management Efficiency Program considers until the end of 2015. By spring of 2014 6 routes have been removed by a partial implementation of this research’s findings. Finally, it is our belief that the full implementation of the proposed policies throughout this paper will produce significant and sustainable savings with minimum effect on service level.
2.6 Discussion

We hereby discuss two limitations of our work, namely concerning assumptions embedded in the stochastic component of the model.

One main assumption is that of all students from the same school have the same probability of showing up at their designated bus stop. On the one hand, this facilitates the formulation of the problem by allowing to model the number of students in a bus as a binary random variable. On the other hand, the data available at the school district could not provide enough information to model the problem any differently.

Ideally every student should have their own probability $p_i$ of showing up at their respective stop. Then, for those students who never use the bus $p_i = 0$, for those who always use it $p_i = 1$, and so on. In this way we can truly represent the stochastic behavior of the number of student on a given route.

Among others, some of the factors that may predict such probability are the age of the pupil, whether they can drive or not, how close they live from the school and whether their neighborhood is rich or poor. Be that as it may, the data available at the school district limited the scope of our model. The only data available is the headcount for the current routes, where no distinction of individual students exists. Additionally, we know that the buses only collect student of the same school on a given bell-time. Therefore, we choose to generalize such probability to the school level.

Notice that schools are of three types: elementary, middle and high. Thus, grouping by school also somewhat distinguish age, a factor that one can easily presume as relevant. Remember as well that the schools considered in this study are public, and therefore their students live (almost 100% of them) within the school’s boundary. Thus, the same association by school somewhat distinguish the neighborhood, hence the average level of household income of the students.
A second assumption that we discussed is that if a bus is used in a particular bell-time, it is assumed to be available right after the end of that period, to potentially continue on collecting students for the next bell-time. As formulated in our model, there exists a $\beta$ probability that a bus will be late for school. In such case the bus would have a later time of availability for the next bell-time and, if no idle time is planned before the start of such next route, the probability to be late to the next school would increase, even above $\beta$ depending on the case.

However unlikely, the situation described could potentially snowball the bus to be late for the rest of the morning. A way of tackling this drawback could be treating the time of availability $t_{\text{avl}}$ of a bus as a random variable, namely a normal random variable with parameters mean and variance of the sum between the travel time and time of availability of the previous route of such bus. That said, even though this feature can be implemented within the framework of our work, further analysis is needed to make any definitive conclusion regarding the approach previously suggested.

2.7 Conclusion and Further Work

Most of previous SBRP studies focus on deterministic routing problems with known student demand and fixed travel times. This paper formally defines SBRP with Stochastic Demand and Duration Constraints, denoted as SBRP-SDDC, via Chance Constrained Programming (CCP). This formulation allows for a considerable increase of the capacity of the buses by permitting overbooking, hence reducing the need of buses due to capacity constraints. Overbooking the buses induces the generation of longer route; therefore, the travel time constraints will be binding more frequently making more likely the occurrence of late arrivals to school given the random nature of the travel time. Thus, chance constraints for the travel time lessen the likelihood of late arrivals to an acceptable level.
Due to different bell times for different schools a dynamic programming formulation is proposed to model the multi-bell time problem, each stage representing a multi-depot to multi-school MIP problem. This formulation responds to the characteristics of the operation observed at Williamsville Central School District, Williamsville, NY. Given the NP-hard nature of the problem, a cascade simplification is proposed to partition the entire SBRP problem into multiple multi-depot to 1-school sub-problems that are solved sequentially using column generation based algorithms. This framework allows the generation of good solutions in a reasonable time. In addition, the numerical experience shows that the solution time and solution quality are very sensitive to different configurations of the proposed column generation procedure; therefore, it is significant to define a specific set of rules by experimentation. Finally, the application of the Cascade Simplification reduces the total number of used buses from 86 (in the current practice) to 77 for the whole morning operation for Williamsville Central School District.

Understanding the variability of the ridership is relevant, whether it changes in the morning and afternoon or its dependency on the school and grade of the students, this understanding allows the implementation of a proper overbooking policy that aims to improve the utilization of the capacity of the buses. This is fairly easy to do for a school district that has in place a process to record and monitor the daily capacity usage of every single bus; a simple calculation would provide them with key information on how to better use the capacity of their fleet.

Given that the operation of the nation’s school districts are very similar, we can easily see the replicability of our approach. Consequently, the introduction of uncertainty, specially for the demand, in the school bus routing problem opens the opportunity to attain significant savings in the total number of buses needed, allowing any school district to move part of their cash flow from transportation towards the classroom while maintaining service level.
Chapter 3

Special Need Students School Bus Routing, Consideration for Mixed Load and Heterogeneous Fleet

3.1 Introduction

The Individuals with Disabilities Education Act (IDEA) was originally enacted by the United States Congress in 1975 to ensure that children with disabilities have the opportunity to receive a free appropriate public education, just like other children. In IDEA, the term transportation embodies travel to and from school, between schools, and around school buildings. Specialized equipment, such as adapted buses, lifts, and ramps, is required when needed to provide the transportation. Also, transportation includes transit from house to bus; hence, when students are unable to get to school without assistance, door-to-door transportation is required. In addition to providing specialized equipment that is required to transport students safely, school boards even have to provide nurses or aides on vehicles if needed [32, 33].
The additional requirements needed for transportation of special education students make this operation costly. During the 1999-2000 school year the U.S. total expenditure on special transportation services is estimated to be about $3.7 billion. This represents around 28% of the total transportation expenditures ($13.1 billion) in the nation and approximately 7% of the total spending on special education services ($50 billion) [34]. The transportation cost per regular education student is approximately $200 to $400, whereas the same for special education ranges from $4,000 to $6,000 per student; note that the actual cost depends on school schedules and district geography [32]. Since the transportation cost of special education students is significantly higher than the cost for the regular case, even small improvements can benefit school boards. In national studies, computer-generated routes have proved to be significantly (32%) more efficient and cost-effective than hand-developed routes [32].

A significant difference between routing special education students as opposed to regular students is the diversity of the students and the programs they attend. In addition to special restrictions in travel time and equipment, special education students do not necessarily live close to their programs, whereas the most common School Bus Routing Problem (SBRP) for regular programs have a significant number of students living relatively close to their school. Moreover, the number of students per program is dramatically lower than for regular schools. In regards to the programs, these are geographically dispersed and have different start and end times. Their location is also particularly troubling since it is often beyond the limits of the school district, making the routes very long, which then allows the assignment of only a few students per bus. Thus, school districts often allow a mixed load configuration for these buses, where a bus can serve students from different schools. However, there is a lack of work in studying the mixed ride scenario for both special education and general students [3].

The most recent review of the SBRP states its research is extensive; however, the review
Table 3.1: Previous work directly related to special education SBRP

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schools</td>
<td>Russel 1986</td>
</tr>
<tr>
<td>Location type</td>
<td>Ripplinger 2005</td>
</tr>
<tr>
<td>Mixed load</td>
<td>Kamali 2013</td>
</tr>
<tr>
<td>Fleet</td>
<td>Multiple</td>
</tr>
<tr>
<td>Objective</td>
<td>Single</td>
</tr>
<tr>
<td>Constraints</td>
<td>Urban</td>
</tr>
<tr>
<td>Problem size</td>
<td>Ripplinger</td>
</tr>
<tr>
<td>Math model</td>
<td>Multiple</td>
</tr>
<tr>
<td>Solution method</td>
<td>Homogenous</td>
</tr>
<tr>
<td></td>
<td>Minimize total travel time</td>
</tr>
<tr>
<td></td>
<td>Minimize cost and travel time</td>
</tr>
<tr>
<td></td>
<td>Minimize total distance</td>
</tr>
<tr>
<td></td>
<td>Bus capacity</td>
</tr>
<tr>
<td></td>
<td>Bus capacity, maximum travel time per student</td>
</tr>
<tr>
<td></td>
<td>Bus capacity</td>
</tr>
<tr>
<td></td>
<td>140 students</td>
</tr>
<tr>
<td></td>
<td>131 students</td>
</tr>
<tr>
<td></td>
<td>111 students</td>
</tr>
<tr>
<td></td>
<td>Greedy heuristic</td>
</tr>
<tr>
<td></td>
<td>Greedy heuristic</td>
</tr>
<tr>
<td></td>
<td>Greedy heuristic</td>
</tr>
</tbody>
</table>

clarifies only a few papers consider the problem of routing special education students [3]. To the best of our knowledge, only three papers have directly studied this problem [35, 36, 37]. Table 3.1 shows the main characteristics of each work, where the common theme is to focus on developing a heuristic solution method for a real-world inspired problem. Only in two of these papers the problem was modeled with a mathematical formulation, having that in neither case such formulation was used as part of the solution procedure.

The SBRP mainly consists of morning and afternoon problems. Many studies are dedicated to the morning problem, whereas the afternoon problem is only mentioned briefly [3], despite the fact they are different in their formulation, and are equally complex. For example, the drop-off location for students in the afternoon may differ. Nowadays, many students do not return home after school; instead, they go to after school caretakers. Students with disabilities are entitled to be transported to caretakers even when those caretakers live out of a district’s attendance boundaries [33]. Also, schools do not necessarily have their time windows in the reverse sequence as in the morning. Thus, routes in the afternoon cannot be simply obtained as the reverse of the morning case.

In this paper, we study the SBRP for special education students inspired by our work with the Williamsville Central School District (WCSD), in Western New York, United States. The main contributions of this work are as follows:
• present a unified mathematical formulation for morning and afternoon routing problem
  with consideration for mixed loading, school time windows, and student maximum
  travel time;
• introduce a heterogeneous fleet, where the difference between buses is not only capacity
  but also the seat type configuration, combination of regular seat and wheelchair space;
  and
• present a numerical example along with real word instances.

3.2 Problem description

The problem studied in this research is based on the operation of the Transportation De-
partment at Williamsville Central School District (WCSD) in New York, and focuses on the
transportation of students enrolled in Special Education programs. Though similar to the
typical SBRP, there are significant differences that need careful consideration.

As in many school districts, transportation for special education students in WCSD is
outsourced. However, the school district designs the routes and decides the number of
buses to be utilized. This is done at least a month prior to the beginning of each school year.
Because the District does not own the buses, it has the liberty of choosing a fleet combination
that fits the need for each year. A bus can be configured to hold up to three wheelchairs,
having the remainder space setup with regular seats. Figure 3.1 shows the layout of different
seat type configurations, where that of 3.1a has only regular seats; 3.1b can hold only one
wheelchair (in the back of the bus), 3.1c two wheelchairs, and 3.1d up to three. It is worth
mentioning that for each configuration the bus capacity changes. Thus, the District not only
decides the routes and the number of buses, but also the seat type configuration for those
buses.

There are significantly fewer students in special education, about 4% of the total student
population. However, the number of schools involved is much greater than for regular programs. They have different bell times and their locations are more disperse in comparison to the other schools, having several out of the district’s boundaries. From Figures 3.2a and 3.2b we can see the significant difference between the locations of homes (in gray circles) and schools (in black squares) for the regular and special education cases for WCSD. These particular features tend to make routes significantly longer [35] and with less use of bus capacity. Consequently, the operation of special education students is very expensive, reaching up to 40% of the annual transportation budget of WCSD. Thus, careful planning of this operation is needed to identify saving opportunities while guaranteeing the quality of service that these students require.

At WCSD, special education students are picked up at their doorstep whereas regular students are required to walk to a particular stop location, to which multiple students may be assigned. Additionally, every bus is obliged to have an aide on board to care for the students and assist those who require more attention; a nurse may also be required depending on the student’s needs. If needed, the buses have to be specially equipped, for example, they need to be able to handle safe transportation for students in wheelchairs, which reduces the capacity of the bus. We say a student has a demand for each seat type, equivalent to that
of multicommodity problems. A bus may also carry students from different schools at the same time; hence, we modeled our problem considering mixed load.

Constraints associated with time, on one hand have a time windows for morning drop-offs and afternoon pick-ups. In the other hand, students have a default maximum ride time set to one hour that can be changed upon request. There is no restrictions for the hours that a bus can operate. Finally, each of the schools and students have a delay, i.e. service time, at drop-offs and pick-ups. This time is about 10 minutes for schools, and for pick-ups and drop-offs at the students’ home the delay varies from 1 to 5 minutes depending on their needs.
3.3 Mathematical model

The following formulation considers the characteristics described in the previous section and supports school bus routing for both morning and afternoon runs. The objective of the model is to minimize the total number of buses used in the morning and afternoon runs. Later, we provide a solution strategy based on column generation. Even though our methods and testing are based on the Williamsville School District scenario, we believe that the methods can be readily modified for other school districts, and the results are widely applicable.

To support both AM and PM runs, we define $\theta$ as a binary parameter equal to 1 if routes are for the AM runs and 0 for the PM runs. In addition we define $\phi = 2 (\theta - \frac{1}{2})$, i.e., equal to 1 and -1 for AM and PM runs respectively. These parameters enable us to formulate a unique model for both cases.

Consider the set $A$ of all stops for students, the set $S$ of all schools and the sets $D_1$ and $D_2$ corresponding to the start and end depots respectively. The set of all locations is $L$, with $L = D_1 \cup A \cup S \cup D_2$. Let the function $\delta(i)$ represent the school of the students in stop $i \in A$ (this means $\delta(i) \in S$). Let $t_i$ be the fixed waiting time at node $i \in A \cup S$ (time per stop), $\dot{t}_i$ be the variable waiting time at node $i \in A \cup S$ (time per student), $t_{ij}$ be the travel time from node $i$ to node $j$, and $\tau_i$ the maximum travel time for student $i \in A$. Let $a_i$ and $b_i$ represent the earliest and latest time of arrival to location $i \in L$ (time window). For $i \in S$ the time windows of arrival $[a_i, b_i]$ for the rest of the assignment is as follows:

$$a_i = \left\{ \begin{array}{ll}
  a_j - \theta \max\{t_i + \tau_i, t_i + t_{ij}\} + (1 - \theta) (t_j + t_{ji}) & : j = \delta(i) \quad \text{if } i \in A \\
  \theta \min_{j \in A} \{a_j - t_{ij}\} + (1 - \theta) \min_{j \in S} \{a_j - t_{ij}\} & : i \in D_1 \\
  \theta \min_{j \in S} \{a_j + t_j + t_{ji}\} + (1 - \theta) \min_{j \in A} \{a_j + t_j + t_{ji}\} & : i \in D_2
\end{array} \right.$$
Let \( G = (L, E) \) be a directed graph. The set of edges is \( E = E_{DA} \cup E_{AA} \cup E_{AS} \cup E_{SA} \cup E_{SS} \cup E_{SD} \) where \( E_{DA} = \{(i, j) \in D_1 \times A\} \) is the set of edges connecting the depot to the students, \( E_{AA} = \{(i, j) \in A^2 \mid i \neq j, t_{ij} + t_j + t_{j,\delta(i)} \leq \max\{\tau_i, t_{i,\delta(i)}\}, a_i + t_i + t_{ij} \leq b_j\} \) is the set of feasible links between students, \( E_{AS} = \{(i, j) \in A \times S \mid j \neq \delta(i), t_{ij} + t_j + t_{j,\delta(i)} \leq \max\{\tau_i, t_{i,\delta(i)}\}, a_i + t_i + t_{ij} \leq b_j\} \cup \{(i, \delta(i)) \mid i \in A\} \) is the set of feasible links from students to schools, \( E_{SA} = \{(i, j) \in S \times A \mid i \neq \delta(j), a_i + t_i + t_{ij} \leq b_j\} \) is the set of feasible links from schools to students, \( E_{SS} = \{(i, j) \in S \times A \mid i \neq j, a_i + t_i + t_{ij} \leq b_j\} \) is the set of feasible links between schools, and \( E_{SD} = \{(i, j) \in S \times D\} \) is the set of edges connecting schools to the depot.

Additionally, let \( A_j = \{i \in A : j = \delta(i)\} \) be the set of students attending school \( j \in S \).

Notice that the sets \( A_j \) are mutually exclusive.

Regarding the attributes of the vehicles we consider in our model, we define \( B \) to be the set of buses and \( Q \) to be the set of seat types that the buses have (e.g. regular seats, wheelchair spaces). Let \( d_{ijk} \) be a binary parameter equal to 1 if bus \( k \in B \) starts in depot \( i \in D_1 \) and finished \( j \in D_2 \), \( s_i^q \) be the number of seat types \( q \in Q \) used by student \( i \in A \), and \( c_k^q \) capacity of bus \( k \in B \) in regards of seat-type \( q \in Q \).

We now define the decision variables of our model. Let \( z_k \) be a binary variable indicating if bus \( k \in B \) is used, and \( x_{ijk} \) be a binary variable indicating if bus \( k \in B \) goes from node \( i \in L \) to node \( j \in L \). Let \( u_{ik} \) be the time of arrival of bus \( k \in B \) at node \( i \in L \) and \( v_{ik} \) the load of bus \( k \in B \) upon arrival at node \( i \in L \). Thus, the formulation of our problem reads
as follows:

\[ P_1: \min \sum_{k \in B} z_k + \varepsilon \sum_{k \in B, (i,j) \in E} t_{ij} x_{ijk} \]  

s.t. \[ \sum_{k \in B, j \in L, j \neq i} x_{ijk} = 1 \quad i \in A \]  

\[ \sum_{(i,j) \in E} x_{ijk} \leq z_k \quad k \in B \]  

\[ \theta \sum_{j \in A} x_{ijk} + (1 - \theta) \sum_{j \in S} x_{ijk} \leq \sum_{i \in D_1} d_{ijk} \quad i \in D_1, k \in B \]  

\[ \theta \sum_{i \in S} x_{ijk} + (1 - \theta) \sum_{i \in D_1} x_{ijk} \leq \sum_{j \in D_2} d_{ijk} \quad j \in D_2, k \in B \]  

\[ \sum_{i \in L} x_{ijk} = \sum_{j \in L} x_{jik} \quad j \in A \cup S, k \in B \]  

\[ \sum_{g \in L} x_{gik} \leq \sum_{g \in L} x_{g,\delta(i),k} \quad i \in A, k \in B \]  

\[ u_{ik} + t_i + t_i \sum_{q \in Q} s^q_i + t_{ij} \leq u_{jk} + M_1 (1 - x_{ijk}) \quad (i,j) \in E : i \in L \setminus S, k \in B \]  

\[ u_{ik} + t_i + t_i \sum_{g \in L} s^q_{ig} x_{gek} + t_{ij} \leq u_{jk} + M_1 (1 - x_{ijk}) \quad (i,j) \in E : i \in S, k \in B \]  

\[ 0 \leq \phi (u_{\delta(j),k} - u_{ik}) \leq \max \{ \tau_{i,\delta(j)}, \theta t_i, \delta(j) + (1 - \theta) t_{\delta(j),i} \} \quad i \in A, k \in B \]  

\[ a_i \leq u_{ik} \leq b_i \quad i \in L, k \in B \]  

\[ v^q_{ik} + \phi s^q_i \leq v^q_{jk} + M_2 (1 - x_{ijk}) \quad (i,j) \in E : i \in L \setminus S, k \in B, q \in Q \]  

\[ v^q_{ik} - \phi \sum_{e \in A_i} s^q_{gek} x_{gek} \leq v^q_{jk} + M_2 (1 - x_{ijk}) \quad (i,j) \in E : i \in S, k \in B, q \in Q \]
Table 3.2: Computational results small size instances

<table>
<thead>
<tr>
<th>Stops</th>
<th>Schools</th>
<th>CPU time</th>
<th>Initial Solution</th>
<th>Final Solution</th>
<th>Lower bound</th>
<th>Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
<td>1</td>
<td>2.5</td>
<td>2.43</td>
<td>2.43</td>
<td>0%</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>11</td>
<td>3.5</td>
<td>3.48</td>
<td>3.48</td>
<td>0%</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>43</td>
<td>5.5</td>
<td>4.56</td>
<td>4.56</td>
<td>0%</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>3603</td>
<td>6.5</td>
<td>5.52</td>
<td>5.45</td>
<td>1%</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>3607</td>
<td>6.5</td>
<td>6.48</td>
<td>3.58</td>
<td>45%</td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td>3608</td>
<td>7.5</td>
<td>4.69</td>
<td>4.39</td>
<td>32%</td>
</tr>
<tr>
<td>35</td>
<td>6</td>
<td>3614</td>
<td>7.5</td>
<td>7.50</td>
<td>4.36</td>
<td>42%</td>
</tr>
<tr>
<td>40</td>
<td>6</td>
<td>3617</td>
<td>9.5</td>
<td>9.50</td>
<td>3.42</td>
<td>64%</td>
</tr>
<tr>
<td>45</td>
<td>6</td>
<td>3626</td>
<td>10.5</td>
<td>10.50</td>
<td>3.29</td>
<td>69%</td>
</tr>
<tr>
<td>50</td>
<td>6</td>
<td>3633</td>
<td>11.5</td>
<td>11.50</td>
<td>3.30</td>
<td>71%</td>
</tr>
</tbody>
</table>

\[ 0 \leq w_{qik}^{t} \leq c_{qk}^{t} \quad i \in L, k \in B, q \in Q \quad (3.14) \]

\[ x_{ijk} \in \{0,1\} \quad (i,j) \in E, k \in B \quad (3.15) \]

where (3.1) minimizes the total number of buses \( \sum z_{k} \) while maintaining the routes’ total length \( \sum t_{ij} x_{ijk} \) to a minimum, \( \varepsilon \) is set as half the inverse of an upper bound for such length. The constraints enforce conditions as follow: (3.2) ensures that all students are picked up, (3.3) each bus leaves the depot at most once, (3.4) and (3.5) verify that all buses start and return to their respective depot, (3.6) is the flow conservation, (3.7) ensures the bus visits the corresponding schools (notice that a bus can visit a school one time at the most), (3.8) and (3.9) define the time of arrival at each stop, (3.10) ensures a school is visited after picking up the corresponding students for the AM run or the opposite in the PM run and the travel time is lower than the maximum allowed for each student, (3.11) is the time windows for the arrival to school, (3.12) and (3.13) define the variation of the number of used seats in the bus, and (3.14) verifies the capacity of the bus is not violated.

In Table 3.2 the model is used to solve different size problems. We can see how rapidly the solving time increases. In particular for this test, we set a limit of 1 hr for every instance. For the cases that reach the time limit, the optimality gap is presented.
3.4 Solution strategy

Because of the complexity of the routing problem $P_{77}$ from the previous section, optimization packages like CPLEX are not able to optimally solve realistic size instances. Thus, we elect to solve the problem approximately. Our approach is based on a column generation decomposition and a procedure tailored for our problem’s special characteristics. In the work of Caceres, Batta, and He [38] a similar decomposition was successfully applied for the regular school bus routing problem. In this work we also use the concept of bus class and establish a subproblem for each one of them, and the decomposition is used heuristically as well. In addition, our subproblem is more complex, and we adopt a different approach to finding good solutions. The following elaborates on the aspects of this decomposition.

3.4.1 Problem decomposition by column generation

Master problem

Let $P_k$ be the set of feasible paths for bus $k \in B$, where $p \in P_k$ is an elementary path. Let $x_{ijk}^p$ be equal to 1 if edge $(i, j) \in L^2$ is covered by bus $k \in B$ when using path $p \in P_k$, $\psi_k^p = \sum_{i \in D} \sum_{j \in A} x_{ijk}^p + \varepsilon \sum_{i \in L} t_{ij} x_{ijk}^p$ be the cost of using path $p \in P_k$ with vehicle $k \in B$ and $\nu_{ik}^p = \sum_{j \in A \cup S} x_{ijk}^p$ be equal to 1 if stop $i \in A$ is visited by bus $k \in B$ when using path $p \in P_k$ and 0 otherwise. Let $y_k^p$ be the binary decision variables that are equal to 1 if path $p \in P_k$ is used by bus $k \in B$ and 0 otherwise. Then, the master problem reads as follows:

$$P_2 : \min \sum_{k \in B} \sum_{p \in P_k} \psi_k^p y_k^p$$

s.t. $\sum_{k \in B} \sum_{p \in P_k} \nu_{ik}^p y_k^p = 1, \quad i \in A$
\[
\sum_{p \in P_k} y^p_k \leq 1, \; k \in B \tag{3.18}
\]
\[
y^p_k \text{ binary} \tag{3.19}
\]

Given that there are different seat configurations for the buses, we group the buses that share the same configuration into \textit{bus classes}. Therefore, let \( W \) define the set of bus classes, where each element \( w \in W \) represents a set of indistinguishable buses, and let \( K_w \) be the number of available buses for each class. Then, the new master problem reads as follows:

\[
P_3 : \min \sum_{w \in W} \sum_{p \in P_w} \psi^p_w y^p_w \tag{3.20}
\]
\[
\text{s.t. } \sum_{w \in W} \sum_{p \in P_w} v^p_{iw} y^p_w = 1, \; i \in A \tag{3.21}
\]
\[
\sum_{p \in P_w} y^p_w \leq K_w, \; w \in W \tag{3.22}
\]
\[
y^p_w \text{ binary} \tag{3.23}
\]

\textbf{Sub problem}

Since the variables or columns of the master problem represent paths for each of the bus classes, one subproblem must be solved for each bus class. Thus, there will be \(|W|\) subproblems to solve separately.

Let \( \pi_i \) represent the dual variables associated with constraints (3.21) and \( \rho_w \) represent the dual variables associated with constraints (3.22). Then, for a given bus the subproblem minimizes the reduced cost \( \psi^p_w - (\sum_{i \in A} \pi_i v^p_{iw} + \rho_w) \). Thus, the subproblem for class \( w \in W \) reads as follows:

\[
P_4 : \min 1 - \rho_w + \sum_{(i,j) \in E} [\varepsilon_{t_{ij}} - \pi_i] x_{ij} \tag{3.24}
\]
\[
\text{s.t. } \sum_{j \in L, j \neq i} x_{ij} \leq 1 \quad i \in A \tag{3.25}
\]

\[
\theta \sum_{j \in A} x_{ij} + (1 - \theta) \sum_{j \in S} x_{ij} = \sum_{j \in D_2} d_{ij} \quad i \in D_1 \tag{3.26}
\]

\[
\sum_{i \in S} x_{ij} + (1 - \theta) \sum_{i \in A} x_{ij} = \sum_{i \in D_1} d_{ij} \quad j \in D_2 \tag{3.27}
\]

\[
\sum_{i \in L} x_{ij} = \sum_{i \in L} x_{ji} \quad j \in A \cup S \tag{3.28}
\]

\[
\sum_{g \in L} x_{gi} \leq \sum_{g \in L} x_{g,\delta(i)} \quad i \in A \tag{3.29}
\]

\[
u_i + t_i + \dot{t}_i \sum_{q \in Q} s^q_i + t_{ij} \leq \nu_j + M_1 (1 - x_{ij}) \quad (i,j) \in E : i \in L \setminus S \tag{3.30}
\]

\[
u_i + t_i + \dot{t}_i \sum_{q \in Q} s^q_i x_{ge} + t_{ij} \leq \nu_j + M_1 (1 - x_{ij}) \quad (i,j) \in E : i \in S \tag{3.31}
\]

\[
0 \leq \phi (u_{\delta(j)} - u_i) \leq \max \{ \tau_i, \theta t_{i,\delta(j)} + (1 - \theta) t_{\delta(j),i} \} \quad i \in A \tag{3.32}
\]

\[
a_i \leq u_i \leq b_i \quad i \in L \tag{3.33}
\]

\[
u^q_i + \phi s^q_i \leq v^q_j + M_2 (1 - x_{ij}) \quad (i,j) \in E : i \in L \setminus S, q \in Q \tag{3.34}
\]

\[
u^q_i - \phi \sum_{e \in A} s^q_e x_{ge} \leq v^q_j + M_2 (1 - x_{ij}) \quad (i,j) \in E : i \in S, q \in Q \tag{3.35}
\]

\[
0 \leq v^q_i \leq c^q \quad i \in L, q \in Q \tag{3.36}
\]

\[
x_{ij} \in \{0,1\} \quad (i,j) \in E \tag{3.37}
\]

where (3.24) minimizes the reduced cost of the new variables for the master problem. Constraint (3.25) ensures that every student is considered at most once on any new path; recall that in the subproblem we are interested in finding single routes that do not necessarily need to contain all students. Constraints (3.26)–(3.37) are the single bus case of (3.4)–(3.15).
3.4.2 Column generation procedure

Our implementation of the column generation procedure follows the framework that Figure 3.3 shows. In a general sense, the method contemplates the following four steps:

*Step 1.* Using a saving heuristic (see Section 3.4.2 for a description and Appendix 3.7.1 for details), we construct a set of initial solutions that is used to start the column generation process. We use the best solution from the initial set to establish the value of $\varepsilon$ as one tenth of the inverse of its total travel time; recall that $\varepsilon$ is a parameter in the objective functions (3.1) and (3.24).

*Step 2.* Solve the linear relaxation of the master problem with all the available columns. Update the coefficient of the objective function of the subproblems with the new values of the dual variables.

*Step 3.* Using a second saving heuristic (see Section 3.4.2 for a description and Appendix 3.7.2 for details), we generate new columns. If at least one of the generated columns in any of the subproblems has a negative cost, then go to *Step 2.*

*Step 4.* Solve the master problem as an integer program using regular branch and bound. In the following sections, we present further explanation on how the subroutines used here work.
Generating an initial solution

The procedure to generate the initial solution can be summarized as follows (in the Appendix, Algorithm 3.7.1 shows the detail):

1. Create a new route $R$.
2. For every unassigned stop find, if feasible, the cheapest potential insertion into $R$.
3. If at least one stop can be inserted into $R$, assign to $R$ one stop based on the saving cost that such insertion yields. If not, go to Step 1.
4. If there are unassigned stops go to Step 2.

The algorithm creates a new route in Step 1 by first selecting a bus with the most availability of the seat type with least demand among the students that have not yet been assigned a route. Next, we filter the stops that have a positive demand for the selected seat type.

To choose which stop to add into the new route, for each stop, we calculate the resulting new cost of the route from adding such a stop. We will use Figure 3.4 to illustrate Steps
2 and 3 of the procedure. Let us assume that we begin Step 2 with the provided current route, where two students have been assigned so far, and they both go to the same school. Three other students have not yet been assigned a route, and for each, we then evaluate their possible insertions. Student 3 has two possible options to be inserted into the route. Similarly, Student 4 has four possible options but note that this student goes to a different school, for which the algorithm needs to determine a position as well. We keep the cheapest option for each student as their potential insertion to the route. In the example, options 2, 3 and 1 are the most economical for students 3, 4 and 5 respectively.

In Step 3, the algorithm chooses one of the students based on the saving cost of their potential insertion. There are two ways to perform this operation: (1) deterministically, by choosing the students with the maximum saving cost, and (2) randomly, by selecting the student based on a probability proportional to $\left(\frac{S[i] - S[1]}{S[n] - S[1]}\right)^\gamma$ where $i$ represents the student, $\gamma = 2$ and $S[i]$ is the student’s saving cost previously calculated. In the example, we chose student 4, and we then repeat the process for the modified route. In the case where the algorithm cannot find any feasible insertion for any of the unassigned students, we stop working with the current route and create a new one in Step 1.

This algorithm returns a feasible solution for our problem: a set of routes where every student has been assigned to one and only one bus. To create diversity in the set of initial solutions, we run our implementation of the algorithm once in deterministic mode and fifty times in random mode (recall Step 3, where the next insertion can be randomly selected). Then, we pass the set of initial solutions to the master problem to continue with the column generation procedure.

**Approximate solution for the subproblem**

The subproblem is NP-hard. Therefore, finding an optimal solution for it is computationally intensive. Thus, we chose to solve the subproblems approximately by means of a heuristic
described in detail in the Appendix (Algorithm 3.7.2) and summarized as follows.

Step 1. Create a new route $R$ for the given bus class.

Step 2. For every unassigned stop find, if feasible, the cheapest potential insertion into $R$.

Step 3. If at least one stop can be inserted into $R$, assign to $R$ one stop based on the saving cost that such insertion yields. If not, terminate.

Step 4. If there are unassigned stops go to Step 2.

The algorithm creates an empty route using a bus from the given class (recall that there is a subproblem for each bus class). Steps 2 and 3 of this algorithm are fundamentally the same as in the procedure explained in the previous section. Thus, the example from Figure 3.4 also illustrates the subroutine in this algorithm.

This algorithm returns a single route for the given bus class. To accelerate the column generation, our implementation runs the algorithm once in deterministic mode and twenty times in random mode for each subproblem. All the solutions generated with negative objective value are then passed to the master problem to continue with the column generation procedure.

3.4.3 A lower bound

A simple way to find a lower bound of problem $P_{3.1}$ is to solve its linear relaxation, i.e., exchange constraint (3.15) for the inequality $0 \leq x_{ijk} \leq 1$. However, this bound is weak and is computationally costly to obtain. Alternatively, a column generation procedure can be used for the same purpose. In the preceding Section we described a decomposition by column generation where the subproblem is solved approximately due to its complexity; hence, it does not yield a valid lower bound. Then, we hereby present an alternate formulation of $P_{3.1}$ that allows for an easy and quick way of obtaining a lower bound.
Let $z_k$ be a binary variable indicating if bus $k \in B$ is used, and $x_{ik}$ be a binary variable indicating if node $i \in A \cup S$ is visited by bus $k \in B$. For any non-empty set of students $G \subset A$ let $r(G) = |G|$ if a feasible route containing all nodes in $G$ exists, and let $r(G) = |G| - 1$ otherwise. Consider the following formulation:

$$P_5: \min \sum_{k \in B} z_k$$

$$\text{s.t. } \sum_{i \in A \cup S} x_{ik} \leq M_1 z_k \quad k \in B$$

$$\sum_{i \in A_j} x_{ik} \leq M_2 x_{ik} \quad k \in B, j \in S$$

$$\sum_{k \in B} x_{ik} = 1 \quad i \in A$$

$$\sum_{i \in G} x_{ik} \leq r(G) \quad k \in B, G \subset A$$

$$x_{ik} \in \{0, 1\} \quad k \in B, i \in A \cup S$$

$$z_k \in \{0, 1\} \quad k \in B$$

where the objective (3.38) is to minimize the number of buses, constraints (3.39) ensures only buses being used can visit any node, (3.40) only allow assignment of students to buses that visit their corresponding school, (3.41) ensures every students is assigned a bus, and (3.42) cut infeasible solutions.

Notice $P_5$ is a compact formulation of $P_1$. Hence, we now focus on finding a lower bound for $P_5$. Also, notice the number of constraints represented by (3.42) increases exponentially with the number of students. To handle this, we leave out all constraints in (3.42) gives a relaxation of $P_5$ and then we proceed to solve with cutting-planes, where at each iteration we add the corresponding cut in (3.42) if an infeasible solution is found. Since this method yields a lower bound at every iteration, we can stop at any time. However, at the beginning of the
algorithm the value of the lower bound is very weak, as can be expected since constraints (3.39)–(3.41) contain almost no information about the feasibility of the routes.

To strengthen the lower bound we now modify these constraints. Let $Z = \{(i, j) \in N^2 : i < j, (i, j) \notin E, (j, i) \notin E\}$ be the set of conflicting rides, where any element in this set represents two nodes that cannot be visited by the same vehicle. Let $\hat{t}_i = t_i + \min_j \{t_{ij}\}$ be the minimum amount of time spent when visiting node $i \in A \cup S$. Let $b(i)$ be the set of buses that student $i \in A$ can ride. Consider the following formulation

$$
P_6 : \min \sum_{k \in B} z_k \quad (3.45)
$$

s.t. 

$$
x_{ik} + x_{jk} \leq z_k \quad (i, j) \in Z, k \in B \quad (3.46)
$$

$$
\sum_{i \in A \cup S} \hat{t}_i x_{ik} \leq \left( \max_i \{b_i + \hat{t}_i\} - \min_i \{a_i + \hat{t}_i\} \right) z_k \quad k \in B \quad (3.47)
$$

$$
\sum_{i \in A_j} s^q_i x_{ik} \leq c_{kq} x_{jk} \quad k \in B, q \in Q, j \in S \quad (3.48)
$$

$$
\sum_{k \in b(i)} x_{ik} = 1 \quad i \in A \quad (3.49)
$$

$$
\sum_{k \notin b(i)} x_{ik} = 0 \quad i \in A \quad (3.50)
$$

$$
\sum_{i \in G} x_{ik} \leq r(G) \quad k \in B, G \subset A \quad (3.51)
$$

$$
x_{ik} \in \{0, 1\} \quad k \in B, i \in A \cup S \quad (3.52)
$$

$$
z_k \in \{0, 1\} \quad k \in B \quad (3.53)
$$

where the objective (3.45) minimizes the number of buses, (3.46) prohibit having in the same bus two students that cannot ride together, (3.47) is a proxy of travel time constraints, (3.48) only allow assignment of students to buses that visit their corresponding school and also act as a proxy of capacity constraints. (3.49) and (3.50) ensures that students are assigned in
buses that they are allow to ride. Finally, (3.51) is equivalent to (3.42).

Constraints (3.47) and (3.48) are the modified version of (3.39) and (3.40) respectively, and (3.49) and (3.50) modify constraints (3.41). Notice that the feasible space of $P_6$ in more constrained than that for $P_5$. As a result we can obtain a stronger lower bound for $P_{3,1}$.

3.5 Computational experiments

3.5.1 Simulated data

To understand the influence of some characteristics of the problem on its objective function, we designed a factorial experiment with five factors, where the response is the number of buses needed. The first factor is “Bell Time Offset”, defined as the difference in time of the start time of any two groups of schools. The four levels for this factor are: 0, 10, 20 and 30 minutes. The second factor is “Load Type”, with two levels: mixed and single. Recall that the problem allows for mixed ride, i.e., having students of different school in the same bus. The single load type represents the case where only students of the same school can share a bus. The third factor is the maximum distance of any school relative to an arbitrary center, with five levels: 2.5, 5, 7.5, 10 and 12.5 miles. The fourth factor is the bus capacity, which was simply set to two levels: 5 and 10 seats. Finally, the fifth factor is the maximum ride time with two levels: 45 and 60 minutes.

In our experiment we randomly generated 100 student locations and 20 schools around a single point. The students are randomly located within 5 miles of the center point, and the schools are located within different levels of distance (2.5, 5, 7.5, 10 and 12.5 miles) from the center depending on the experiment.

Figure 3.5 shows the result of the numerical experiment. We can see that all the five factors influence the number of buses needed, having the bell time offset with the highest.
Figure 3.5: Numerical example with simulated data.
influence. The greater the offset, the fewer buses are needed. Having wider spacing between
start time of the schools allows better re-utilization of the buses. In this experiment we
observed an average reduction of 52% when varying the bell time offset from zero to half an
hour.

Regarding capacity, as expected, the more seats available the less buses are needed. The
school location also affects fleet size, the further the schools are allowed to be located the
more buses are needed.

The load type, more often than not, influences the number of buses. A mixed ride policy
allows for the need of fewer buses in most cases. This is specially true for rather narrow bell
time offsets; we can see the biggest difference of the policy is for an offset of 10 minutes.
Let us assume that there are only 10 minutes between the start time of two schools. A bus
wouldn’t be able to pick up students after the first school to be able to drop them off at the
second school. Therefore, to be able to reuse that bus for a second school, students for both
schools would need to be picked up prior to visiting the first school, i.e., a mixed ride policy
is appropriate. Now, if more time exits between start times, there will be sufficient time
to pick up students of the second school after visiting the first one. Therefore, a combined
policy of mixed ride and wider bell time offsets is desirable for minimizing the need of buses.

3.5.2 Case study

We hereby consider four real-world instances from WCSD, corresponding to AM and PM
runs of the 2013-2014 and 2014-2015 school years. Recall that the geographical dispersion
of these instances follows the nature seen in Figure 3.2b. We run each of these instances
for each combination of the factors “Maximum Ride Time” and “Load Type”, resulting in
a total of 24 numerical experiments. For the maximum ride time we considered three levels:
40, 50 and 60 minutes. For the load type there are two levels: mixed and single. The results
Table 3.3: Computational results for real instances of WCSD

<table>
<thead>
<tr>
<th>Instance</th>
<th>Stops</th>
<th>Schools</th>
<th>Bus classes</th>
<th>Max ride</th>
<th>Load type</th>
<th>Edges</th>
<th>Valid</th>
<th>PC time</th>
<th>Buses</th>
<th>Current</th>
<th>Buses</th>
<th>Current</th>
<th>Buses</th>
<th>Current</th>
<th>Buses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AM 227</td>
<td>39</td>
<td>2</td>
<td></td>
<td>60</td>
<td>Mixed</td>
<td>71,824</td>
<td>40,345</td>
<td>37</td>
<td>95</td>
<td>14</td>
<td>34</td>
<td>31</td>
<td>41</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>71,824</td>
<td>5,765</td>
<td>5</td>
<td>15</td>
<td>14</td>
<td>35</td>
<td>32</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AM 227</td>
<td>39</td>
<td>2</td>
<td></td>
<td>50</td>
<td>Mixed</td>
<td>71,824</td>
<td>35,536</td>
<td>30</td>
<td>76</td>
<td>14</td>
<td>35</td>
<td>34</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>71,824</td>
<td>5,765</td>
<td>6</td>
<td>20</td>
<td>14</td>
<td>36</td>
<td>34</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AM 227</td>
<td>39</td>
<td>2</td>
<td></td>
<td>40</td>
<td>Mixed</td>
<td>71,824</td>
<td>29,386</td>
<td>20</td>
<td>28</td>
<td>17</td>
<td>39</td>
<td>36</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>71,824</td>
<td>5,762</td>
<td>6</td>
<td>12</td>
<td>17</td>
<td>38</td>
<td>37</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PM 225</td>
<td>39</td>
<td>2</td>
<td></td>
<td>60</td>
<td>Mixed</td>
<td>70,756</td>
<td>38,960</td>
<td>16</td>
<td>72</td>
<td>14</td>
<td>29</td>
<td>29</td>
<td>36</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>70,756</td>
<td>5,938</td>
<td>3</td>
<td>11</td>
<td>14</td>
<td>34</td>
<td>31</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PM 225</td>
<td>39</td>
<td>2</td>
<td></td>
<td>50</td>
<td>Mixed</td>
<td>70,756</td>
<td>35,578</td>
<td>15</td>
<td>52</td>
<td>14</td>
<td>34</td>
<td>32</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>70,756</td>
<td>5,938</td>
<td>4</td>
<td>16</td>
<td>14</td>
<td>36</td>
<td>33</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PM 225</td>
<td>39</td>
<td>2</td>
<td></td>
<td>40</td>
<td>Mixed</td>
<td>70,756</td>
<td>29,351</td>
<td>8</td>
<td>31</td>
<td>16</td>
<td>38</td>
<td>36</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>70,756</td>
<td>5,935</td>
<td>3</td>
<td>11</td>
<td>16</td>
<td>38</td>
<td>36</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AM 189</td>
<td>33</td>
<td>2</td>
<td></td>
<td>60</td>
<td>Mixed</td>
<td>50,176</td>
<td>28,933</td>
<td>26</td>
<td>448</td>
<td>13</td>
<td>24</td>
<td>23</td>
<td>32</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>50,176</td>
<td>4,724</td>
<td>5</td>
<td>44</td>
<td>13</td>
<td>27</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AM 189</td>
<td>33</td>
<td>2</td>
<td></td>
<td>50</td>
<td>Mixed</td>
<td>50,176</td>
<td>24,840</td>
<td>19</td>
<td>184</td>
<td>14</td>
<td>27</td>
<td>26</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>50,176</td>
<td>4,724</td>
<td>6</td>
<td>40</td>
<td>14</td>
<td>29</td>
<td>26</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AM 189</td>
<td>33</td>
<td>2</td>
<td></td>
<td>40</td>
<td>Mixed</td>
<td>50,176</td>
<td>20,793</td>
<td>11</td>
<td>145</td>
<td>15</td>
<td>30</td>
<td>28</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>50,176</td>
<td>4,724</td>
<td>5</td>
<td>60</td>
<td>15</td>
<td>31</td>
<td>28</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PM 200</td>
<td>34</td>
<td>2</td>
<td></td>
<td>60</td>
<td>Mixed</td>
<td>55,696</td>
<td>31,670</td>
<td>16</td>
<td>251</td>
<td>15</td>
<td>25</td>
<td>25</td>
<td>29</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>55,696</td>
<td>4,781</td>
<td>4</td>
<td>14</td>
<td>15</td>
<td>30</td>
<td>28</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PM 200</td>
<td>34</td>
<td>2</td>
<td></td>
<td>50</td>
<td>Mixed</td>
<td>55,696</td>
<td>29,055</td>
<td>12</td>
<td>270</td>
<td>15</td>
<td>29</td>
<td>27</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>55,696</td>
<td>4,781</td>
<td>3</td>
<td>23</td>
<td>15</td>
<td>30</td>
<td>29</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PM 200</td>
<td>34</td>
<td>2</td>
<td></td>
<td>40</td>
<td>Mixed</td>
<td>55,696</td>
<td>23,026</td>
<td>9</td>
<td>102</td>
<td>16</td>
<td>31</td>
<td>29</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>55,696</td>
<td>4,781</td>
<td>4</td>
<td>14</td>
<td>16</td>
<td>32</td>
<td>31</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

are shown in Table 3.3.

For each instance we show the name, the number of stops (students), schools, and maximum ride time allowed in minutes. Under column Edges we show the total number of edges and valid edges after discarding those that do not conform to the set defined as a function of the time windows with the procedure described in Section 3.3. Under CPU time we show the computational time needed to find the initial solution and the time needed to run the column generation procedure (both times are in seconds). Under Buses we show the number of vehicles in the solution for each one of the methods utilized. The lower bound for each instance is shown under LB and it was obtained following the method described in Section 68.
3.4.3. Column Initial presents the number of vehicles found with Algorithm 3.7.1 described in Section 3.4.2 that finds a set of initial solutions for the column generation procedure. Under column CG, we show the final result for our procedure, i.e., the result obtained at the end of the column generation. Column Current buses shows the real number used at WCSD. This amount is only available for instances with 60 minutes of maximum ride time. Finally, column Buses saved shows the difference between the current situation and the solution found after using column generation. Notice that the current situation is available for four instances with 1 hr of maximum ride time and a mixed ride policy.

By comparing the load type we can see the benefit of implementing the “Mixed Ride” over the “Single Ride” policy. The former requires, on average, 4% fewer buses than the latter. Also, in Figure 3.6 we can see some interaction between the maximum ride time and the load type. The effect of the load type policy increases when the maximum ride is 60 minutes; if students can ride for a longer time in a bus, then there is more time to go around picking students up from different schools. Regarding computational complexity, the single ride policy results in a smaller problem, evidenced by the lower number of edges (under valid edges), requiring less time to solve.
In addition, we compare our results with the current situation at the school district for the corresponding school year. We can see that in all of the instances our solution outperforms the current practice, saving an average 20% of buses in a reasonable computational time. Recall that our approach takes advantage of the mixed load strategy for building the initial solution and further improves it via column generation. We can see that the column generation procedure improves the initial solution in almost every instance in an average of 6%.

3.6 Conclusion

In this work, we modeled and presented a solution scheme for the mixed load routing problem for special education students. Given the nature of dispersed locations of the stops and schools for this class of problem, we took advantage of the mixed ride strategy that allows a bus to carries students attending different schools and modeled the problem accordingly.

Special education students may require specially equipped buses for wheelchairs. Therefore, our model supports the use of a heterogeneous bus fleet. To find the appropriate number of regular and specially equipped buses to be leased, we start with a overestimated number of buses available to solve the problem and let the optimization choose the best number for each class of buses.

In addition, our formulation responds to both morning and afternoon problems in a uniform way. Provided that for this problem we allow mixed ride and that the schools have different start and end times, it is not appropriate to use the same sequence of stops for the morning in the afternoon as suggested in previous work. See Park and Kim [3] for a discussion in the morning versus afternoon problems.

Our numerical experiment demonstrates the benefits of using a combined policy of mixed ride and wider bell time offsets: fewer buses are needed when students from different school
can ride the same bus simultaneously and when schools have significantly different start
times. In this regards, two other findings are worth noting. First, by only looking at the
bell time offset we observe a reduction by half of the fleet size when varying from zero to 30
minutes. Second, the biggest reduction of the fleet size explained by the mixed ride policy is
observed when the bell time offset is 10 minutes, a rather modest level of such factor; mixed
load allows having students from different schools picked up before visiting two schools 10
minutes apart.

By using a customized column generation procedure, we approximately solved a set of
instances from a real school district in Western New York. We found that our approach
responds well to cases with high dispersion of student and school locations. One reason for
this is that our approach builds the routes allowing mixed ride from the beginning, whereas
Park, Tae, and Kim [11]’s approach is to improve a set of routes initially created with a
homogeneous load. Validated by instances in a real school district, our proposed approach
could save an average of 20% buses for special education students as compared to the existing
bus operations.

### 3.7 Algorithms

#### 3.7.1 Generate initial solution for the master problem

1: function GETINITIALSOLUTION($L, B$) \[\triangleright L: \text{set of stops}, B: \text{set of buses.}\]
2: \hspace{1em} for $i = 1$ to $n$ do \[\triangleright n$ is the number of student stops.\]
3: \hspace{2em} $t \leftarrow$ travel time from stop $i$ to its corresponding school
4: \hspace{2em} $T[i] \leftarrow 1 + \varepsilon \times t$
5: \hspace{1em} end for
6: \hspace{1em} $q \leftarrow$ index of seat type with least demand

71
7: \( w \leftarrow \) index of the bus-class with the most availability of seat type \( q \)
8: \( U \leftarrow \) set of stops with demand of seat type \( q \)
9: \( r \leftarrow \) new route with start and end depot using bus-class \( w \)
10: \( n_b[w] = n_b[w] - 1 \quad \triangleright \) \( n_b[w] \) is the number of available buses in class \( w \).
11: \textbf{while} \( U.size > 0 \) and \( n_b[w] \geq 0 \) \textbf{do}
12: \hspace{1em} \( c \leftarrow r.cost \)
13: \hspace{1em} \textbf{for each} student \( i \) in \( U \) \textbf{do}
14: \hspace{2em} \( c' \leftarrow \text{TRYADDSTOP}(r, i) \) \( \triangleright \) Get new cost of \( r \) if stop \( i \) is added. See Algorithm 3.7.3.
15: \hspace{2em} \( m \leftarrow \) number of seats requested in stop \( i \)
16: \hspace{2em} \( S[i] \leftarrow m \times (T[i] - (c' - c)) \) \( \triangleright \) Calculate the saving.
17: \hspace{2em} \textbf{end for}
18: \hspace{1em} \textbf{if} \( \text{count}(S[i] > -\infty) > 0 \) \textbf{then}
19: \hspace{2em} \textbf{if} \( \text{mode} = \text{deterministic} \) \textbf{then}
20: \hspace{3em} \( i^* \leftarrow \arg \max \{S[i]\} \)
21: \hspace{2em} \textbf{else if} \( \text{mode} = \text{random} \) \textbf{then}
22: \hspace{3em} \( i^* \leftarrow \text{SELECTSTOP}(S, L) \) \( \triangleright \) See Algorithm 3.7.4.
23: \hspace{2em} \textbf{end if}
24: \hspace{2em} \( r.addStop(i^*) \)
25: \hspace{2em} \( U.remove(i^*) \)
26: \hspace{2em} \textbf{else}
27: \hspace{3em} \text{update} \( q, w, U \) \( \triangleright \) See lines 6 to 8.
28: \hspace{2em} \( r \leftarrow \) new route with start and end depot using bus-class \( w \)
29: \hspace{3em} \( n_b[w] = n_b[w] - 1 \)
30: \hspace{2em} \textbf{end if}
31: \textbf{end while}
3.7.2 Approximate solution for the subproblem

1: function GetShortestPath(w,\rho,\pi) \triangleright w: bus-class; \rho, \pi: dual values of master problem.

2: \quad U \leftarrow \text{set of students that can be accommodated in bus-class } w

3: \quad \textbf{for } i = 1 \textbf{ to } n \textbf{ do}

4: \quad \quad t \leftarrow \text{travel time from stop } i \text{ to its corresponding school}

5: \quad \quad T[i] \leftarrow 1 - \rho[w] + \varepsilon \times t - \pi[i]

6: \quad \textbf{end for}

7: \quad r \leftarrow \text{new route containing start and end depot}

8: \quad \textbf{while } U.\text{size} > 0 \textbf{ do}

9: \quad \quad c \leftarrow r.\text{cost}

10: \quad \quad \textbf{for each student } i \text{ in } U \textbf{ do}

11: \quad \quad \quad c' \leftarrow \text{TRYADDSTOP}(r, i) \triangleright \text{Get new cost of } r \text{ if stop } i \text{ is added. See Algorithm 3.7.3.}

12: \quad \quad \quad m \leftarrow \text{number of seats requested in stop } i

13: \quad \quad \quad S[i] \leftarrow m \times (T[i] - (c' - c)) \triangleright \text{Calculate the saving.}

14: \quad \quad \textbf{end for}

15: \quad \quad \textbf{if } mode = \text{deterministic} \textbf{ then}

16: \quad \quad \quad i^* \leftarrow \text{arg max}\{S[i]\}

17: \quad \quad \textbf{else if } mode = \text{random} \textbf{ then}

18: \quad \quad \quad i^* \leftarrow \text{SELECTSTOP}(S, L) \triangleright \text{See Algorithm 3.7.4.}

19: \quad \textbf{end if}
20: \( r.add(i^*) \)
21: \( U.remove(i^*) \)
22: if \( \text{count}(S[i] > -\infty) = 0 \) then
23: break
24: end if
25: end while
26: return \( r \)
\( \triangleright \) Returns the corresponding shortest path.
27: end function

3.7.3 Try to add a stop to a route

1: function TRYADDSTOP\((r, s)\) \( \triangleright r \) is a route and \( s \) a stop.
2: \( c \leftarrow \infty \)
3: \( n \leftarrow \text{number of stops in route } r \)
4: if \( r \) contains the school of student \( s \) then
5: \( j \leftarrow \text{position in } r \) of the school of stop \( i \)
6: \( i \leftarrow j \)
7: while \( i > 1 \) do
8: insert stop \( s \) in position \( i \) of route \( r \)
9: if ROUTEISFEASIBLE\((r)\) then \( \triangleright \) Runs a feasibility check.
10: \( c \leftarrow \min\{c, r.cost\} \)
11: end if
12: remove stop \( s \) from route \( r \)
13: \( i \leftarrow i - 1 \)
14: end while
15: else
for $j = 2$ to $n$ do

insert the school of stop $s$ in position $j$ of $r$

$i \leftarrow j$

while $i > 1$ do

insert stop $s$ in position $i$ of route $r$

if $\text{ROUTEISFEASIBLE}(r)$ then

$c \leftarrow \min\{c, r.cost\}$

end if

remove stop $s$ from route $r$

$i \leftarrow i - 1$

end while

remove the school stop $s$ from route $r$

end for

end if

return $c$

end function

3.7.4 Selection of stop

function $\text{SELECTSTOP}(S, L)$

$n \leftarrow \text{length of array } S$

$\gamma \leftarrow 0.1$

sort($S, L$)

$F[1] \leftarrow 0$

for $i = 2$ to $n$ do

$F[i] \leftarrow \left(\frac{S[i] - S[1]}{S[n] - S[1]}\right)^\gamma + F[i - 1]$

end function
8: \textbf{end for}

9: \( r \leftarrow \text{random value from uniform distribution } U[0, 1] \)

10: \( r \leftarrow r \times F[n] \)

11: \( i \leftarrow 1 \)

12: \textbf{while } F(i) < r \textbf{ do}

13: \( i \leftarrow i + 1 \)

14: \textbf{end while}

15: \textbf{return } L[i] \quad \triangleright \text{Return the corresponding student.}

16: \textbf{end function}
Chapter 4

Pricing tax return for student that opt-out using school bus

4.1 Introduction

The literature on school bus routing has focus on minimizing the length or cost of routes, using always the constraint that all students need to be assign to a bus [3]. However, depending on their age, students often don’t ride the bus. Figure 4.1 shows the ridership, the ratio between actual number of students riding a bus and the total number originally assigned to it, for each school, separating the AM and PM cases. We can see how no more than 80% of the students actually use the transportation provided by the school district. The fact that these students are still assigned to a bus accounts for longer routes and low capacity usage.

In our previous work we tackled this problem by implementing overbooking, i.e., assign to a bus more students than the actual capacity [38]. In this paper we explore encouraging students to opt-out from using school buses. Figure 4.2 illustrates the motivation of our work, on the left 10 students are being covered by 3 routes, but when 2 of the students opt-
out from using the buses, and are given an incentive $\tau^*$ to do so, we see that the 8 remaining
students can be covered by only two buses. By removing 2 students from the set we are able
to reduce the fleet size in one unit. Then, we can see that this scenario is convenient for the
school district only if the savings derived from using fewer buses are greater than the sum
of the incentives paid to the students that accepted the deal.

In the following sections we analyze two policies that will try to reduce the number of
students that use the school buses to ultimately reduce the fleet size. The first policy is the
“Open Offer”, that aims to openly offer a certain amount of money (or tax return) to any
student willing to opt-out from using the bus. In this policy the school district does not
control which students would accept the offer, but it only controls the value of the incentive.
This incentive is determined before the students decide whether to accept the deal or not.
Thus, the incentive value needs to be determined to reduce the possibility of having to pay
more for incentives than what would be saved by using fewer buses.

The second policy corresponds to a “Targeted Offer”, where a school district aims to
find the smallest set of students that would need to opt-out in order to achieve a certain
Figure 4.2: Motivating illustration

goal on savings, e.g., in terms of number of buses. Then, the incentive offered would be
determined once the set of opt-out students is known. In his case the school district controls
how many students could potentially receive an incentive. Hence, the incentive offered can
be significantly larger than the open offer.

4.2 Literature review

The first policy, Open Offer, is a school bus routing problem (SBRP) where not all the
students (customers) need to be routed since a fraction of them opt-out from the service.
However, the routing problem exists after the students have decided whether or not to opt-
out, and hence, the routing problem is simply a SBRP with fewer students. The SBRP is
part of the family of Vehicle Routing Problems (VRP). The VRP has been widely studied
for over half a century; its practical relevance and economical impact is demonstrated by the
large number of real-world applications that have yielded substantial savings in the global
transportation cost [39, 40, 41, 42]. Many exact and approximate algorithms are available
[43, 41]. The SBRP can be formulated as a vehicle routing problem with time windows
(VRPTW), a very popular member of the VRP family [44, 45, 46]. Park and Kim [3] review
the most typical approaches to solve the SBRP. More recent work has widened the study of the SBRP [7, 47, 48, 49, 50]. Also, there has been recent effort on considering mixed load, i.e., allowing students from different schools to ride the same bus [51, 52, 53]. Finally, the stop selection problem is many times solved separately from the routing problem but is also considered in the SBRP family [54, 55, 56].

The second policy, Targeted Offer, is a composite of two problems. The first is a full size school bus routing problem [3, 38], where the objective is to minimize the fleet size when all the students have to be picked up. The second, consists on finding the smallest set of students to achieve a reduction of one unit in the bus fleet. Though this second problem has not been explored in the context of school transportation, we can see it as a routing problem where we want to drop the most unprofitable customers [57] or one where we select only the customers that increases the total profit [58]. The latter problem, the capacitated team orienteering problem (CTOP), is mostly solved heuristically by adapting methods for the VRP family [59, 60].

4.3 Open offer problem

In the open offer problem the school district offers a predefined incentive to any student willing to opt out from using the bus. We will focus then on finding a value for the incentive \( \tau \) such that the school district is less likely to lose money. Notice that in this problem the number of students that need transportation is unknown, i.e., is a random variable. Let \( Y_i = 1 \) if student \( i \in A \) opts-out from riding the bus, with \( A \) being the set of all students. Then, we aim to minimize the total cost:

\[
P_1 : \min_{\tau} \quad \mathbb{E} \left[ \sum_{i \in A} \tau Y_i + R \left( i \in A : Y_i = 0 \right) \right] \quad (4.1)
\]
where $\tau$ is the incentive (tax return) to students that opt-out, $Y_i$ is a Bernoulli random variable with parameter $r_i$, the probability of opting-out, and $\mathcal{R}(i \in A : Y_i = 0)$ is the cost associated with routing the remaining students. Notice that $\mathcal{R}$ is also a random variable since is a function of the set $Y_1, ..., Y_n$.

We assume that each student decides whether to use the school bus at the beginning of the school year. The decision can be modeled as a function of the incentive and the distance to school:

$$r_i = f(\tau, d_i)$$  \hspace{1cm} (4.2)

where $r_i$ is the probability of opting-out from using the bus, $\tau$ is the tax return offer to each student and $d_i$ is the distance from the students home to his or her school. Let us assume that

$$\frac{\delta r_i}{\delta \tau} < 0 \quad \text{and} \quad \frac{\delta r_i}{\delta d_i} > 0$$  \hspace{1cm} (4.3)

### 4.3.1 Open offer for one school

We now present the procedure to determine $\tau$ for the case of one school with $n$ students. A student can choose to opt-out from riding the bus by accepting a tax return (a price represented with $\tau$) from the district. Those students that don’t accept the price, will need to be assigned a bus. Let $r_i = u_i^{-1}(\tau)$ represent the probability that student $i$ opts-out of riding the bus. For simplicity, let us assume that $r_i = r \ \forall i \in A$. Then, the expected number of students is $n \times r$. In order to find the cost function, we proceed as follows:

1. Choose $p$
2. Set $r_i = u_i^{-1}(\tau)$
3. Sample the $n$ student using $r_i$
4. Choose stop location based on the sampled students
5. Create routes for the chosen stops
If students opt-out from using the school bus, the fleet can be reduced. We explore offering an incentive such that a fleet reduction can be attained.

4.3.2 Stop selection model

After students decide whether or not to opt-out from using the bus, we need to assign bus stops to the students that choose to keep using school transportation. Let $U$ be the set of all potential stop locations and $M$ the set of students. Let $d_{ij}$ be the distance from student $i \in M$ to location $j \in U$, $\delta$ the maximum allowed walking distance and $\lambda$ the maximum number of students that can be assigned to a stop. Let $y_{ij}$ be the binary decision variables that are equal to 1 if student $i \in M$ gets assigned to stop-location $j \in U$ and 0 otherwise; $z_i$ is equal to 1 if location $i \in U$ is set to be a stop and 0 otherwise. Then, the stop location selection problem can be stated as follows:

$$
P_2 : \min \sum_{j \in U} z_j + \varepsilon \sum_{i \in M} \sum_{j \in U} d_{ij} y_{ij} \quad (4.4)$$

s.t.  
$$
\sum_{j \in U} y_{ij} = 1, \ i \in M \quad (4.5)$$
$$
\sum_{i \in M} y_{ij} \leq \lambda z_j, \ j \in U \quad (4.6)$$
$$
\sum_{j \in U} d_{ij} y_{ij} \leq \delta, \ i \in M \quad (4.7)$$
$$
y_{ij}, z_i \text{ binary} \quad (4.8)$$

where (4.4) minimizes the total number of stops and $\varepsilon = (\delta |M|)^{-1}$, (4.5) ensures that every location gets assigned to one and only one stop-location, (4.6) that a maximum of $\lambda$ students can be assigned to any stop-location, and (4.7) that no student walks more than the maximum walking distance.
For simplification, in our application we have that every student’s residence represents a potential location of a stop. Thus, we have that the sets $U$ and $M$ are equal.

### 4.3.3 Routing model

In the following we provide a mixed integer linear program.

**Sets:**
- $B$ set of buses.
- $A$ set of stops.
- $Q$ set of seat types that the buses have (e.g. regular seats, wheelchair).

**Parameters:**
- $t_{ij}$ travel time from node $i \in L$ to node $j \in L$.
- $d_i$ number of students in stop $i \in A$.
- $q$ capacity per bus.
- $a_i$ and $b_i$ are the earliest and latest time of arrival to location $i \in L$ (time window).

**Variables:**
- $x_{ijk}$ binary variable indicating if bus $k \in B$ goes from node $i \in L$ to node $j \in L$.
- $u_{ik}$ time of arrival of bus $k \in B$ at node $i \in L$.

Then, the formulation reads as follows:

$$
\mathcal{P}_3 : \min \sum_{k \in B} \sum_{i \in L} \sum_{j \in L} c_{ij} x_{ijk} \quad (4.9)
$$

$$
\text{s.t.} \quad \sum_{k \in B} \sum_{j \in L} x_{ijk} = 1 \quad i \in A \quad (4.10)
$$

$$
\sum_{i \in A} \sum_{j \in L} d_i x_{ijk} \leq q \quad k \in B \quad (4.11)
$$

$$
\sum_{j \in L} x_{0jk} = 1 \quad k \in B \quad (4.12)
$$
\[ \sum_{i \in L} x_{ihk} - \sum_{j \in L} x_{hjk} = 0 \quad h \in A, k \in B \quad (4.13) \]

\[ \sum_{i \in L} x_{i,n+1,k} = 1 \quad k \in B \quad (4.14) \]

\[ u_{ik} + t_{ij} - M_{ij} (1 - x_{ijk}) \leq u_{jk} \quad i \in L, j \in L, k \in B \quad (4.15) \]

\[ a_i \leq u_{ik} \leq b_i \quad i \in L, k \in B \quad (4.16) \]

\[ x_{ijk} \in \{0, 1\} \quad (4.17) \]

where \( c_{ij} = 1 \) for \( i = 0 \) and \( j \in A \), and \( c_{ij} = 0 \) otherwise. Thus, the objective function represents the number of buses needed to visit all the bus stops.

### 4.3.4 Numerical example

Our numerical example begins with exploring the behavior in one High School. Figure 4.3 shows the joint results of 200 replicates for random values of \( r \), the probability of opting-out, following a uniform random distribution between 0 and 1. In the horizontal axis the plot shows the values for the probability \( r \) and the vertical axis the fleet size needed to cover the students that did not opt out on each of the 200 realizations. Because of the stochastic selection of students (their choice of opting out or not) the same total number of students may result in different fleet size. Depending on which students opt out, namely their location, the need for buses varies. That is the reason why in the graph we see an area or band where the number of buses lies instead of a single line. The area above the band represents the savings on the number of buses for each realization of \( r \).

Let us assume that the annual cost of a bus is fixed and equal to 35,000 US dollars. We now are interested in finding \( \tau \) such that the savings from reducing the fleet size equals the sum of the incentive handed to the students. We will call this amount “maximum \( \tau \)” or \( \tau^+ \). This amount quantifies the maximum incentive per student that the school district is willing
Figure 4.3: Illustrative example.

Figure 4.4: Illustrative example.
to give such that it does not lose money. In Figure 4.4 we show the same data used in Figure 4.3 but here the vertical axis is $\tau^+$. We can see how the resulting values can be supported by a so called safe maximum $\tau$ which represents the maximum value of $\tau$ that can be offered such that the district does not lose money regardless of the level of $r$.

Figure 4.5 illustrates selected cases for four levels of $r$. The maps displays three high schools represented with red squares. The bus stops are represented in blue and the black dots represents students’ homes. As $r$ increases we see how there are fewer stops covered by routes.

Table 4.1 shows summarized results for each of eleven levels of $r$, the probability of opting-out from riding the bus. For each level 20 replicates were run, so the rest of the columns of the table show the range for the indicated parameters. The last column shows how $\tau^+$ remains over 278 dollars for North High School and over 160 dollars for South High School.
Table 4.1: Computational results for real instances of WCSD

<table>
<thead>
<tr>
<th>School</th>
<th>Probability of opting-out</th>
<th>Number of opting-out students</th>
<th>Number of stops</th>
<th>Number of buses</th>
<th>Savings</th>
<th>Max r</th>
</tr>
</thead>
<tbody>
<tr>
<td>North High School</td>
<td>0</td>
<td>0</td>
<td>133-134</td>
<td>15-16</td>
<td>0-35,000</td>
<td>-</td>
</tr>
<tr>
<td>Students: 1,235</td>
<td>0.1</td>
<td>108-135</td>
<td>123-130</td>
<td>14-15</td>
<td>35,000-70,000</td>
<td>282-648</td>
</tr>
<tr>
<td>Ridership: 36.3%</td>
<td>0.2</td>
<td>227-262</td>
<td>112-121</td>
<td>13</td>
<td>105,000</td>
<td>401-463</td>
</tr>
<tr>
<td>Maximum ride time: 50 min.</td>
<td>0.3</td>
<td>341-399</td>
<td>104-115</td>
<td>11-12</td>
<td>140,000-175,000</td>
<td>358-482</td>
</tr>
<tr>
<td>Maximum walking dist: 0.2 miles</td>
<td>0.4</td>
<td>459-516</td>
<td>95-107</td>
<td>10-11</td>
<td>175,000-210,000</td>
<td>339-426</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>600-649</td>
<td>87-99</td>
<td>9-10</td>
<td>210,000-245,000</td>
<td>324-407</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>710-777</td>
<td>77-90</td>
<td>8-10</td>
<td>210,000-280,000</td>
<td>285-391</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>854-881</td>
<td>69-81</td>
<td>8-9</td>
<td>245,000-280,000</td>
<td>286-328</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>960-1,010</td>
<td>60-72</td>
<td>7-8</td>
<td>280,000-315,000</td>
<td>278-328</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>1,091-1,132</td>
<td>39-53</td>
<td>5-6</td>
<td>350,000-385,000</td>
<td>309-349</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1,235</td>
<td>0</td>
<td>0</td>
<td>560,000</td>
<td>453</td>
</tr>
<tr>
<td>South High School</td>
<td>0</td>
<td>0</td>
<td>77</td>
<td>8</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Students: 850</td>
<td>0.1</td>
<td>74-108</td>
<td>68-74</td>
<td>7-8</td>
<td>0-35,000</td>
<td>0-461</td>
</tr>
<tr>
<td>Ridership: 28.4%</td>
<td>0.2</td>
<td>141-197</td>
<td>65-69</td>
<td>6-7</td>
<td>35,000-70,000</td>
<td>178-400</td>
</tr>
<tr>
<td>Maximum ride time: 50 min.</td>
<td>0.3</td>
<td>229-274</td>
<td>59-64</td>
<td>6</td>
<td>70,000</td>
<td>255-306</td>
</tr>
<tr>
<td>Maximum walking dist: 0.2 miles</td>
<td>0.4</td>
<td>320-380</td>
<td>56-60</td>
<td>5-6</td>
<td>70,000-105,000</td>
<td>200-316</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>396-455</td>
<td>51-57</td>
<td>5-6</td>
<td>70,000-105,000</td>
<td>161-265</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>484-532</td>
<td>48-54</td>
<td>5</td>
<td>105,000</td>
<td>197-217</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>571-629</td>
<td>43-50</td>
<td>4-5</td>
<td>105,000-140,000</td>
<td>167-241</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>664-702</td>
<td>37-44</td>
<td>4</td>
<td>140,000</td>
<td>199-211</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>754-782</td>
<td>25-34</td>
<td>3-4</td>
<td>140,000-175,000</td>
<td>185-232</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>850</td>
<td>0</td>
<td>0</td>
<td>280,000</td>
<td>329</td>
</tr>
</tbody>
</table>

Each row contains 20 replications.
4.4 Targeted offer problem

In the open offer policy the school district decides an incentive prior to knowing who would take the deal (the set of student opting out from riding the bus). In the targeted offer, the district first decides the number of buses that it desires to save and then finds the smallest set of students who would need to opt-out from using a bus in order to accomplish such savings. The objective is to target the smallest number of students that would realize the desired savings.

The targeted offer is a twofold problem, each being in the family of VRPs. What distinguishes these two problems is the transportation request. In the first it is require to serve all stops, a regular collection and delivery problem. In the second the stops are selected upon convenience, which classifies the problem to routing with profits and service selection [41].

4.4.1 Illustrative example

Given a solution of a VRPTW problem, we want to find the smallest set of customer (bus stops) to drop so that the number of vehicles used is reduced in one unit. Let $m^*$ be the minimum number of vehicles needed to serve the set of customers $C$. The following is the model for the regular VRPTW:

\[
P_4 : \quad m^* = \min \sum_{k \in V} \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ijk} \tag{4.18}
\]

s.t.

\[
\sum_{k \in V} \sum_{j \in N} x_{ijk} = 1 \quad i \in C \tag{4.19}
\]

\[
\sum_{i \in C} \sum_{j \in N} d_{ij} x_{ijk} \leq q \quad k \in V \tag{4.20}
\]

\[
\sum_{j \in N} x_{0jk} = 1 \quad k \in V \tag{4.21}
\]
\[
\sum_{i \in N} x_{ihk} - \sum_{j \in L} x_{hjk} = 0 \quad h \in C, k \in V \tag{4.22}
\]

\[
\sum_{i \in N} x_{i,n+1,k} = 1 \quad k \in V \tag{4.23}
\]

\[
u_{ik} + t_{ij} - M_{ij} (1 - x_{ijk}) \leq u_{jk} \quad i \in N, j \in N, k \in V \tag{4.24}
\]

\[
a_i \leq u_{ik} \leq b_i \quad i \in N, k \in V \tag{4.25}
\]

\[
x_{ijk} \in \{0,1\} \quad i \in N, j \in N, k \in V \tag{4.26}
\]

where, for this problem, \(c_{0j} = 1\), for \(j \in C\) and \(c_{ij} = 0\), otherwise. The VRPTW as been widely studied and there are numerous approaches to solve it [44, 45, 46].

Using the result from the previous model, let \(m^*\) be the minimum number of vehicles. Let \(V'\) be a set of vehicles with \(m^* - 1\) elements. Then, we are now concerned with finding the maximum number of customers or bus stops that can be visited using only \(m^* - 1\) vehicles. The mixed integer formulation of the problem reads as follow:

\[
\mathcal{P}_5: \quad \text{max} \quad \sum_{k \in V'} \sum_{i \in C} \sum_{j \in N} p_i x_{ijk} \tag{4.27}
\]

s.t. \[
\sum_{k \in V} \sum_{j \in N} x_{ijk} \leq 1 \quad i \in C \tag{4.28}
\]

\[
\sum_{i \in C} \sum_{j \in N} d_i x_{ijk} \leq q \quad k \in V' \tag{4.29}
\]

\[
\sum_{j \in N} x_{0jk} = 1 \quad k \in V' \tag{4.30}
\]

\[
\sum_{i \in N} x_{ihk} - \sum_{j \in L} x_{hjk} = 0 \quad h \in C, k \in V' \tag{4.31}
\]

\[
\sum_{i \in N} x_{i,n+1,k} = 1 \quad k \in V' \tag{4.32}
\]

\[
u_{ik} + t_{ij} - M_{ij} (1 - x_{ijk}) \leq u_{jk} \quad i \in N, j \in N, k \in V' \tag{4.33}
\]

\[
a_i \leq u_{ik} \leq b_i \quad i \in N, k \in V' \tag{4.34}
\]
Figure 4.6: Three iterations for the targeted offer problem.

\[ x_{ijk} \in \{0, 1\} \quad i \in N, j \in N, k \in V' \]  

This model corresponds to a special case of the capacitated team orienteering problem (CTOP) introduced by Archetti et al. [58]. The CTOP is a variant of the well-known team orienteering problem (TOP). It is defined in a graph \( G = (V, E) \), where \( V = \{1, ..., n\} \) is the set of vertices and \( E \) is the set of edges. Vertex 1 represents the depot and the remainder are the customers \( C = \{2, ..., n\} \). Each customer \( i \in C \) has a demand \( d_i \) and a profit \( p_i \). Each edge \((i, j)\) has a travel time \( t_{ij} \). It is assumed that the distances satisfy the triangle inequality. In addition, the fleet consists of \( m \) homogeneous vehicles each with capacity \( Q \) and maximum travel distance \( T_{\text{max}} \). Chu [57] applies the CTOP to service outsourcing of unprofitable customers, a closer interpretation of what the targeted offer problem tries to do. Luo et al. [59] and Tarantilis, Stavropoulou, and Repoussis [60] offer metaheuristics to solve this problem.

Figure 4.6 shows an illustration where the previous two models are solved sequentially.
In this example the capacity $Q = 47$, the profit $p_i = d_i$ and the demand $d_i$ represents the number of students at each stop (vertex). The first map from the left is the solution for a VRPTW where 4 vehicles are needed to visit 80 stops. The second map from the left shows the routes that can cover the maximum number of students with 3 vehicles (one unit fewer than that from the VRPTW solution). We can observe how only 7 stops need to be left out in order to save one bus. The following two maps show the same logic, but for a reduction in two and three units respectively. Notice that the first reduction is the one where the fewest number of stops need to be left out.

4.4.2 Solution strategy

Prior to choosing what students to target, we need to know what is the smallest fleet of buses needed to serve all students. This routing problem (SBRP) is equivalent to that on Caceres, Batta, and He [38]. Given the solution of a full size SBRP, we want to find the smallest set of students that would reduce the number of buses needed by one. Let $m^*$ be the minimum number of vehicles needed to serve the set of all students $A$. Let $B'$ be a set of buses with $m^* - 1$ elements.

The Targeted Offer problem is a variation of the CTOP. By starting with the full size SBRP we change to components. First, we modify the objective (2.6) to maximize the number of students covered:

$$\max \sum_{k \in B} \sum_{i \in D} \sum_{j \in A} w_i x_{ijk} \quad (4.36)$$

Second, constraints (2.7) and (2.8) are respectively modify as follows

$$\sum_{k \in B} \sum_{i \in D \cup A} x_{ijk} \leq 1, \quad j \in A \quad (4.37)$$
\[ \sum_{k \in B} \sum_{j \in A \cup S} x_{ijk} \leq 1, \ i \in A \] (4.38)

where they now don’t require visiting every bus stop.

The students that are not visited after solving the CTOP version of the full size SBRP are those that would need to be targeted and offered an incentive in order to achieve a reduction of the bus fleet by one unit.

4.4.3 Numerical example

In the numerical example we run the targeted offer problem to three high schools in WCSD. For each we calculate the minimum number buses needed considering all students. After we reduce one bus for each of the schools, we determine the largest set of students in each school that can be picked up with the remaining buses.

Figure 4.7 shows on the left the routes that cover all the students. The blue dots are the bus stops and the black dots are the students residences. The red squares are the three schools. On the right we show the routes, considering one route less for each school. Consequently, here there are students not covered by any route, which are shown with a red dot.

For each school we show the number of students covered before and after the fleet reduction. The difference between these quantities are the students that need to opt-out from riding a bus in order to be able to reduce the fleet size. Thus, the maximum incentive each student could receive \( \tau^+ \) is calculated by dividing the cost of one bus among the students opting-out. The difference between the school is very significant, varying from 700 to 11,667 dollars per student.
Schools districts design bus routes in order to pick up all students. However, students often choose not to use the busing system, as seen in our case study. Other work has looked into better utilization of capacity through overbooking. In this work we explored two policy options to reduce the fleet size by reducing the set of students. Offering an incentive, in the form of a tax return may influence students to opt out from using the bus system and hence would reduce the fleet size. Naturally, the sum of the incentives should not be greater than saving in the routing cost.

With the Open Offer Policy we found a valuation of tax return for students opting out from using school bus that protects a school district from losing money. We determined a procedure to find an incentive value that would put the district in a safe position regardless of the amount of student that decide to opt out from the system. The numerical example assumes that all student have the same likelihood of opting out. Therefore, future work should include modeling the probability of opting out for each student as a function of the
distance to school, and should also consider the influence of the intrinsic average ridership of the corresponding school.

With the Targeted Offer Policy we showed how significant savings can be attained by targeting a small set of the students and offering them an incentive. Since the number of students that are offered the incentive is limited to a small set, the magnitude of the incentive can be quite large and hence more persuading. Future work in this line should focus on finding alternative sets of students that can be targeted in order to obtain the desired savings. Note that when targeting particular students they don’t necessarily accept to opt out, and therefore we need to have a strategy to account for this possibility.

Additionally, for both policies all schools from a district should be considered simultaneously in order to assess global savings for the district.
Bibliography


[38] H. Caceres, R. Batta, and Q. He. “School Bus Routing with Stochastic Demand and Duration Constraints”. In: Accepted for Transportation Science (2016).


