School Bus Routing with Stochastic Demand and Duration Constraints

Caceres, Hernan
Department of Industrial and Systems Engineering
University at Buffalo
hernanan@buffalo.edu

Batta, Rajan
Department of Industrial and Systems Engineering
University at Buffalo
batta@buffalo.edu

He, Qing
Department of Civil, Structural, and Environmental Engineering
Department of Industrial and Systems Engineering
University at Buffalo
qinghe@buffalo.edu

May 12, 2015

Abstract

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 Introduction

The school bus routing problem (SBRP) is crucial due to its impact on economic and social objectives (Delagado-Serna and Pacheco-Bonrostro, 2001). 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 (Bowerman et al., 1995).

SBRP has been intensively studied in the last few decades. One may refer to a fairly recent literature review of SBRP in Park and Kim (2010). 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 (Riera-Ledesma and Salazar-González, 2013; Kinable et al., 2014), tabu search (Pacheco et al., 2013), greedy randomized adaptive search procedure (Schittekat et al., 2013), branch-and-cut algorithm (Riera-Ledesma and Salazar-González, 2012), approximation algorithm (Bock et al., 2012) and genetic algorithm (Diaz-Parra et al., 2012). Additionally, the work of Park et al. (2012) and Kim et al. (2012) 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) (Christofides and Eilon, 1969; Braca et al., 1997). 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 (1983) 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 Laporte et al. (1989); the model uses fewer variables, but requires a homogeneous fleet of vehicles. Laporte et al. (1992) 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 et al. (2013) 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 in Park and Kim (2010). 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 Δt. Due to safety considerations, a limit on the amount of time students can spend on school buses is specified (Bodin and Berman, 1979; Desrosiers et al., 1981; Braca et al., 1997). 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 (Braca et al., 1997).

2. The probability of overcrowding a bus is less than α. 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 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.

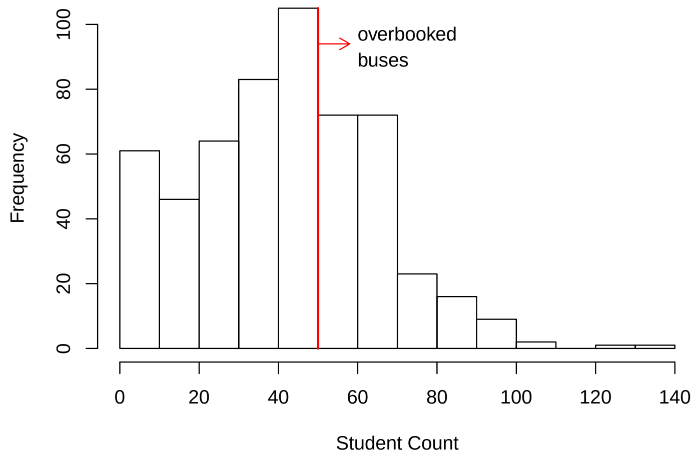

Figure 1: Histogram of student count.

3. The probability of being late to school is less than β. 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 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 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 student'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.

3 Model formulation

In this section we will formulate our problem based on the description provided in section 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 3 the focus will be in the formulating the bus routing problem, in Section 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.

3.1 General model

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

$$\begin{aligned}
\text{Min} \quad & \text{number of buses used} + \varepsilon\,(\text{total travel time}) && (1)\\
\text{s.t.} \quad & P\,(\text{overcrowding the bus}) \leq \alpha, \;\; \forall \text{ bus} && (2)\\
& P\,(\text{being late to school}) \leq \beta, \;\; \forall \text{ bus} && (3)\\
& E\,(\text{maximum ride time}) \leq \Delta t, \;\; \forall \text{ bus} && (4)
\end{aligned}$$

where the objective (1) is, first, to minimize the total number of buses needed and then the total travel time, provided that ε 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) provides an upper bound for the likelihood of overcrowding the bus, constraint (3) provides an upper bound for the likelihood of a bus being late to school, and constraint (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:

$$f_n^*\left(\{b_{ik}\}_n, \{t_{\text{avl}}^k\}_n\right) = \min_{\{x_{ijk}\}_n} \left\{v\left(P_n\right) + f_{n+1}^*\left[\Psi\left(\{b_{ik}\}_n, \{x_{ijk}\}_n\right), \Omega\left(\{t_{\text{avl}}^k\}_n, \{x_{ijk}\}_n\right)\right]\right\}, \quad n = 1, ..., N \quad (5)$$

where $f_{N+1}^* = 0$, $v\left(P_n\right)$ 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 $\{t_{\text{avl}}^k\}_n$ and the choice of routes $\{x_{ij}^k\}_n$ 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.

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 $(i, j) \in L^2$, μ_{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 κ_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:

$$\text{Min} \quad \sum_{k\in B}\sum_{i\in D}\sum_{j\in A}\kappa_i x_{ijk} + \varepsilon \sum_{k\in B}\sum_{i\in L}\sum_{j\in L}\left(\mu_{T_{ij}} + \mu_{T_i}\right) x_{ijk} \tag{6}$$

$$\text{s.t.} \quad \sum_{k\in B}\sum_{i\in D\cup A} x_{ijk} = 1, \;\; j \in A \tag{7}$$

$$\sum_{k\in B}\sum_{j\in A\cup S} x_{ijk} = 1, \;\; i \in A \tag{8}$$

$$\sum_{k\in B}\sum_{i\in L}\left(x_{iik} + \sum_{j\in D} x_{ijk} + \sum_{j\in S} x_{jik}\right) = 0 \tag{9}$$

$$\sum_{i\in D\cup A} x_{ijk} = \sum_{i\in A\cup S} x_{jik}, \;\; k \in B, j \in A \tag{10}$$

$$\sum_{i\in D\cup A} x_{ijk} \leq \sum_{g\in S}\sum_{i\in A} a_{jg} x_{igk}, \;\; k \in B, j \in A \tag{11}$$

$$\sum_{i\in D\cup A} x_{ijk} \leq \sum_{i\in D}\sum_{j\prime\in A} x_{ij\prime k}, \;\; k \in B, j \in A \tag{12}$$

$$\sum_{j\in L} x_{ijk} \leq b_{ik}, \;\; k \in B, i \in D \tag{13}$$

$$1 \leq u_{ik} \leq m + 2, \;\; k \in B, i \in L \tag{14}$$

$$u_{ik} - u_{jk} + (m+2)\, x_{ijk} \leq m + 1, \;\; k \in B, i \in L, j \in L \tag{15}$$

$$P\,(\text{overcrowding the bus}) \leq \alpha, \;\; k \in B \tag{16}$$

$$P\,(\text{being late to school}) \leq \beta, \;\; k \in B \tag{17}$$

$$E\,(\text{maximum ride time}) \leq \Delta t, \;\; k \in B \tag{18}$$

$$x_{ijk} \text{ binary} \tag{19}$$

where (6) minimizes the additional number of buses needed to run the corresponding bell time $\sum_{k\in B}\sum_{i\in D}\sum_{j\in A}\kappa_i x_{ijk}$ while maintaining the total length of the routes $\sum_{k\in B}\sum_{i\in L}\sum_{i\in L}\left(\mu_{T_{ij}} + \mu_{T_i}\right) x_{ijk}$ to a minimum, ε is set as the inverse of an upper bound for such length (the upper bound is found with the procedure described in section 4.2). The constraints ensure conditions as follow: (7) one and only one bus arrives to every stop, (8) one and only bus departures from every stop, (9) no bus stays at the same location nor arrives to a depot nor departures from a school, (10) same bus that arrives to a location departures from that location, (11) a bus only picks up students attending the same school, (12) a location can be visited by a bus only if that bus leaves the depot, (13) all buses start their route on their corresponding depot, (14) and (15) are the sub-tour elimination constraints where m is the maximum number of stops a bus can have, (16) to (18) are the stochastic constraints which will be developed in detail in the following section and (19) is the integrality condition.

3.3 Stochastic constraints

This section concentrates on the development of the stochastic constrains presented on the previous section that represent constraints (16) to (18).

3.3.1 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 $Y_k \sim Bin(\mathcal{Q}_k, p_j)$ where $\mathcal{Q}_k = \sum_{i\in A}\sum_{j\in L} w_i x_{ijk}$ 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 (16) is given by:

$$P(Y_k > c_k) \leq \alpha \ \ \forall k \in B \tag{20}$$

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

In order to introduce (20) into the MIP problem in section 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 α. Thus,

$$q_{jk} = max\left\{ q \,\middle|\, 1 - \sum_{v=0}^{c_k} \binom{q}{v} (p_j)^v (1-p_j)^{q-v} \leq \alpha \right\} \tag{21}$$

Proposition 1. *For all $k \in B$ the constraint*

$$\sum_{i\in L}\sum_{j\in L} w_i x_{ijk} \leq \sum_{i\in A}\sum_{j\in S} q_{jk} x_{ijk} \tag{22}$$

is an equivalent inequality for (20).

Proof. We know that the right hand side of (22) chooses q_{jk} according to the school that the bus is heading to. Thus, if (22) holds, then the following also holds

$$1 - \sum_{v=0}^{c_k} \binom{\mathcal{Q}_k}{v} (p_j)^v (1-p_j)^{\mathcal{Q}_k - v} \leq 1 - \sum_{v=0}^{c_k} \binom{q_{jk}}{v} (p_j)^v (1-p_j)^{q_{jk} - v}$$

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 (20) holds as well. □

3.3.2 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 τ_f and τ_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:*

$$\mu_{T_i} = \tau_f - \tau_f (1-p_j)^{w_i} + \tau_v w_i p_j \tag{23}$$

$$\sigma^2_{T_i} = \tau_f^2 (1-p_j)^{w_i} \left(1 - (1-p_j)^{w_i}\right) + 2\tau_f \tau_v w_i p_j (1-p_j)^{w_i} + \tau_v^2 w_i p_j (1-p_j) \tag{24}$$

Proof. We know that R_i, the actual number of students showing up at stop $i \in A$, is a r.v. such that $R_i \sim Bin\left(w_i, p_j\right)$. Then, let $T_i\left(R_i\right) = \begin{cases} \tau_f + \tau_v R_i & \text{if } R_i > 0 \\ 0 & \text{if } R_i = 0 \end{cases}$ 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 $p_{T_i}\left(t_i\right) = \begin{cases} p_{R_i}(0) & \text{if } t_i = 0 \\ p_{R_i}\left(r_i\right) & \text{if } t_i = \tau_f + \tau_v r_i \\ 0 & \text{otherwise} \end{cases}$ and the expected value and variance of T_i are then derived as follows:

$$
\begin{aligned}
\mu_{T_i} &= E\left[T_i\left(R_i\right)\right] = \sum_{r=0}^{w_i} T_i(r)\, p_{R_i}(r) = 0 \cdot p_{R_i}(0) + \sum_{r=1}^{w_i}\left(\tau_f + \tau_v r\right) p_{R_i}(r) \\
&= \tau_f \sum_{r=1}^{w_i} p_{R_i}(r) + \tau_v \sum_{r=1}^{w_i} r p_{R_i}(r) = \tau_f\left[\sum_{r=0}^{w_i} p_{R_i}(r) - p_{R_i}(0)\right] + \tau_v \sum_{r=0}^{w_i} r\, p_{R_i}(r) \\
&= \tau_f\left[1 - \left(1 - p_j\right)^{w_i}\right] + \tau_v E\left[R_i\right] = \tau_f - \tau_f\left(1 - p_j\right)^{w_i} + \tau_v w_i p_j \\
\sigma_{T_i}^2 &= V\left[T_i\left(R_i\right)\right] = E\left[T_i\left(R_i\right)^2\right] - \left[E\left[T_i\left(R_i\right)\right]\right]^2 = \sum_{r=0}^{w_i}\left[T_i(r)\right]^2 p_{R_i}(r) - \mu_{T_i}^2 \\
&= 0^2 \cdot p_{R_i}(0) + \sum_{r=1}^{w_i}\left(\tau_f + \tau_v r\right)^2 p_{R_i}(r) - \mu_{T_i}^2 \\
&= \tau_f^2 \sum_{r=1}^{w_i} p_{R_i}(r) + 2\tau_f\tau_v \sum_{r=1}^{w_i} r\, p_{R_i}(r) + \tau_v^2 \sum_{r=1}^{w_i} r^2\, p_{R_i}(r) - \mu_{T_i}^2 \\
&= \tau_f^2\left[\sum_{r=0}^{w_i} p_{R_i}(r) - p_{R_i}(0)\right] + 2\tau_f\tau_v \sum_{r=0}^{w_i} r\, p_{R_i}(r) + \tau_v^2 \sum_{r=0}^{w_i} r^2\, p_{R_i}(r) - \mu_{T_i}^2 \\
&= \tau_f^2\left[1 - \left(1 - p_j\right)^{w_i}\right] + 2\tau_f\tau_v E\left[R_i\right] + \tau_v^2 E\left[R_i^2\right] - \mu_{T_i}^2 \\
&= \tau_f^2\left[1 - \left(1 - p_j\right)^{w_i}\right] + 2\tau_f\tau_v w_i p_j + \tau_v^2\left[V\left[R_i\right] + E\left[R_i\right]^2\right] - \mu_{T_i}^2 \\
&= \tau_f^2\left[1 - \left(1 - p_j\right)^{w_i}\right] + 2\tau_f\tau_v w_i p_j + \tau_v^2\left[w_i p_j\left(1 - p_j\right) + w_i^2 p_j^2\right] - \left(\tau_f - \tau_f\left(1 - p_j\right)^{w_i} + \tau_v w_i p_j\right)^2 \\
&= \tau_f^2\left(1 - p_j\right)^{w_i}\left(1 - \left(1 - p_j\right)^{w_i}\right) + 2\tau_f\tau_v w_i p_j\left(1 - p_j\right)^{w_i} + \tau_v^2 w_i p_j\left(1 - p_j\right)
\end{aligned}
$$

□

An estimation for the fixed and variable time for picking up students can be found in Braca et al. (1997), 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 $\sigma_{T_{ij}}^2$ respectively.*

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

$$
\begin{aligned}
\mu_{\mathcal{T}_k} &= \sum_{i \in L}\sum_{j \in L}\left(\mu_{T_{ij}} + \mu_{T_i}\right) x_{ijk} \\
\sigma_{\mathcal{T}_k}^2 &= \sum_{i \in L}\sum_{j \in L}\left(\sigma_{T_{ij}}^2 + \sigma_{T_i}^2\right) x_{ijk}
\end{aligned}
$$

Then, the travel time constraint that represents (17) is given by:

$$
P\left(t_{\text{avl}}^k + \mathcal{T}_k > t_{\text{bell}}\right) \leq \beta \quad \forall k \in B \tag{25}
$$

where t_{avl}^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 β the given upper bound for the probability of bus $k \in B$ not making it on time to school. We now need to reformulate (25) such that it can be included in the single bell-time mix integer linear program.

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

$$\mathcal{T}_k = T_{0,1} + T_1 + T_{1,2} + ... + T_{m-1,m} + T_m + T_{m,m+1}$$

where m is the number of stops to be made by a bus. Then, $\mathcal{T}_k$ is a summation of $2m+1$ random variables.

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

Thus, we use the above conjecture in the following proposition in order to reformulate (25) into a set of linear inequalities.

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

$$t^k_{avl} + \mu_{\mathcal{T}_k} + \Phi^{-1}(1-\beta)\,\tilde{\sigma}_{\mathcal{T}_k} \le t_{bell} \tag{26}$$

$$\sum_{h=1}^{h^+} h^2 \gamma^k_h \ge \sigma^2_{\mathcal{T}_k} \tag{27}$$

$$\sum_{h=1}^{h^+} h \gamma^k_h = \tilde{\sigma}_{\mathcal{T}_k} \tag{28}$$

$$\sum_{h=1}^{h^+} \gamma^k_h = 1 \tag{29}$$

are valid inequalities for (25), where γ^k_h is a binary variable and h^+ is the maximum possible integer value for $\sigma_{\mathcal{T}_k}$.

Proof. From (25) it is obtained that

$$P\left(t^k_{\text{avl}} + \mathcal{T}_k < t_{\text{bell}}\right) \ge 1-\beta$$

which by standardizing becomes

$$\Phi\left(\frac{t_{\text{bell}} - \left(t^k_{\text{avl}} + \mu_{\mathcal{T}_k}\right)}{\sqrt{\sigma^2_{\mathcal{T}^k}}}\right) \ge 1-\beta$$

and by taking the inverse

$$t^k_{\text{avl}} + \mu_{\mathcal{T}_k} + \Phi^{-1}(1-\beta)\sqrt{\sigma^2_{\mathcal{T}^k}} \le t_{\text{bell}}$$

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

Since γ^k_h is a binary variable, the assignment constraints (28) and (29) ensure that the variable $\tilde{\sigma}_{\mathcal{T}_k}$ will only take an integer value between 1 and $h^+ = \lceil\left(t_{\text{bell}} - min\{t^k_{\text{avl}}\}\right)/2\rceil$ the maximum round up integer value the standard deviation can take. Then, the inequality in (27) constraints $\tilde{\sigma}_{\mathcal{T}_k}$ to be at least the round-up integer of $\sqrt{\sigma^2_{\mathcal{T}_k}}$.

Since now $\sqrt{\sigma^2_{\mathcal{T}_k}} \le \tilde{\sigma}_{\mathcal{T}_k}$, the following inequality holds:

$$t^k_{\text{avl}} + \mu_{\mathcal{T}_k} + \Phi^{-1}(1-\beta)\sqrt{\sigma^2_{\mathcal{T}_k}} \le t^k_{\text{avl}} + \mu_{\mathcal{T}_k} + \Phi^{-1}(1-\beta)\,\tilde{\sigma}_{\mathcal{T}_k}$$

Therefore, if (26) is satisfied then (25) will also be satisfied. □

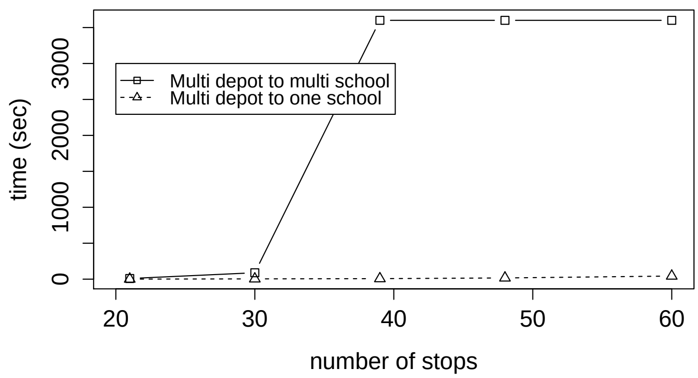

Figure 2: Solution time comparisons for difference sizes of the SBRP problem.

3.3.3 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 $\Delta t_{\max}$. 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 (18) reads as follows:

$$\mu_{\mathcal{T}_k} - \sum_{i \in D} \sum_{j \in A} \mu_{T_{ij}} x_{ijk} \leq \Delta t_{\max} \ \ \forall k \in B \tag{30}$$

where $\sum_{i \in D} \sum_{i \in A} \mu_{T_{ij}} x_{ijk}$ represents the expected time from the depot to the first stop.

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 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 Table 1). A 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.

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.

Table 1: Details for solution time comparison

	Multi depot to multi school							Multi depot to one school						
	# of	# of	bound at root		bound at end		CPU	# of	# of	bound at root		bound at end		CPU
size	variables	nodes	lower	upper	lower	upper	time*	variables	nodes	lower	upper	lower	upper	time*
21	2912	20201	3.69	4.91	4.25	4.75	9	2504	0	2.55	3.96	2.55	2.60	1
30	7716	52205	3.55	6.91	4.13	4.63	89	5740	318	2.35	4.84	2.38	2.56	4
39	16048	401480	3.51	7.88	3.56	6.61	3604	10992	388	3.38	5.88	3.40	3.55	7
48	28880	115779	3.51	6.64	3.49	8.93	3603	16068	2004	4.37	6.99	4.37	4.52	17
60	51792	38938	3.89	12.00	3.89	7.57	3613	32464	618	4.33	9.00	4.35	4.50	43

(*) CPU time in seconds

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* ($\frac{students \times ridership}{capacity}$).
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.

4.1 Set stops' location

As described in Section 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 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.

For simplification every student's resident represents a potential location of a stop. Then, the objective of the MIP problem is to minimize the total amount of stops subject to the maximum walking distance and the maximum number of students assigned to a single stop.

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 $j \in U$, δ the maximum walking distance and λ 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:

$$\text{Min} \quad \sum_{j\in U} z_j + \varepsilon \sum_{i\in U}\sum_{j\in U} d_{ij}y_{ij} \tag{31}$$

$$\text{s.t.} \quad \sum_{j\in U} y_{ij} = 1, \;\; i \in M \tag{32}$$

$$\sum_{i\in U} y_{ij} \leq \lambda z_j, \;\; j \in U \tag{33}$$

$$\sum_{j\in U} d_{ij}y_{ij} \leq \delta, \;\; i \in M \tag{34}$$

$$y_{ij}, z_i \text{ binary} \tag{35}$$

where (31) minimizes the total number of stops and $\varepsilon = (\delta|M|)^{-1}$, (32) ensures that every location gets assigned to one and only one stop-location, (33) that a maximum of λ students can be assigned to any stop-location, and (34) that no student walks more than the maximum walking distance.

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.

4.2 Finding an initial solution to the single school routing problem

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 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. (2002) and Alabas-Uslu (2008). However we will limit our work to implement a greedy algorithm combining Clarke and Wright saving algorithm (Clarke and Wright, 1964) and the Farthest First Heuristic proposed by Fu et al. (2005).

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 Fu et al. (2005), 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* $\mu_{T_{d^*i^*}} = \min_{d\in D}\{\mu_{T_{di^*}}\}$*, 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* $s_{ij} = \mu_{T_{di}} + \mu_{T_{is}} + \mu_{T_{j^-j}} - \mu_{T_{j^-i}} - \mu_{T_{ij}}$ *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* $s_i = \min_{d\in D}\{\mu_{T_{di}}\} + \mu_{T_{ij}}$ *where* j *is the corresponding school.*

7. Go to step 2 until $A = \emptyset$.

The results obtained with this algorithm are used to set ε in (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.

4.3 Solving the single school routing problem

The model formulation for this problem is identical to the one in section 3.2, with $|S| = 1$. Since there is only one school in the problem, constraint (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 Danna and Le Pape (2005) 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.

4.3.1 A column generation based approach

A closer look at the model in section 3.2 reveals that only constraints (7) and (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.

4.3.2 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_{ijk}^p be equal to 1 if edge $(i, j) \in L^2$ is covered by bus $k \in B$ when using path $p \in P^k$, $\theta_k^p = \sum_{i \in D} \sum_{j \in A} \kappa_k x_{ijk}^p + \varepsilon \sum_{i \in L} \sum_{j \in L} \left(\mu_{T_i} + \mu_{T_{ij}}\right) x_{ijk}^p$ 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 MP reads as follows:

$$\text{Min} \quad \sum_{k \in B} \sum_{p \in P_k} \theta_k^p y_k^p \tag{36}$$

$$\text{s.t.} \quad \sum_{k \in B} \sum_{p \in P_k} \nu_{ik}^p y_k^p = 1, \ \ i \in A \tag{37}$$

$$\sum_{p \in P_k} y_k^p \leq 1, \ \ k \in B \tag{38}$$

$$y_k^p \text{ binary} \tag{39}$$

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 considerable less variables reads as follows:

$$\text{Min} \quad \sum_{r\in R}\sum_{p\in P_r} \theta_r^p y_r^p \tag{40}$$

$$\text{s.t.} \quad \sum_{r\in R}\sum_{p\in P_r} \nu_{ir}^p y_r^p = 1, \;\; i \in A \tag{41}$$

$$\sum_{p\in P_r} y_r^p \leq K_r, \;\; r \in R \tag{42}$$

$$y_r^p \text{ binary} \tag{43}$$

4.3.3 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 π_i represent the dual variables associated with constraints (41) and ρ_r represent the dual variables associated with constraints (42). Then, for a given bus the SP minimizes the reduced cost $\theta_r^p - \left(\sum_{i\in A}\pi_i\nu_{ir}^p + \rho_r\right)$. Thus, the SP for class $r \in R$ reads as follows:

$$\text{Min} \quad \kappa - \rho + \sum_{i\in L}\sum_{j\in L}\left[\varepsilon\left(\mu_{T_{ij}} + \mu_{T_i}\right) - \pi_i\right] x_{ij} \tag{44}$$

$$\text{s.t.} \quad \sum_{j\in A\cup S} x_{ij} \leq 1, \;\; i \in A \tag{45}$$

$$\sum_{i\in L}\left(x_{ii} + \sum_{j\in D} x_{ij} + \sum_{j\in S} x_{ji}\right) = 0 \tag{46}$$

$$\sum_{j\in A} x_{ij} = 1, \;\; i \in D \tag{47}$$

$$\sum_{i\in D\cup A} x_{ij} = \sum_{i\in A\cup S} x_{ji}, \;\; i \in A \tag{48}$$

$$\sum_{i\in A} x_{ij} = 1, \;\; j \in S \tag{49}$$

$$1 \leq u_i \leq m + 2, \;\; i \in L \tag{50}$$

$$u_i - u_j + (m+2)\, x_{ij} \leq m + 1, \;\; i \in L, j \in L \tag{51}$$

$$\sum_{i\in L}\sum_{j\in L} w_i x_{ij} \leq q \tag{52}$$

$$t_{\text{avl}} + \mu_{\mathcal{T}} + \Phi^{-1}(1-\beta)\,\tilde{\sigma}_{\mathcal{T}} \leq t_{\text{bell}} \tag{53}$$

$$\sum_{h=1}^{h^+} h^2\gamma_h \geq \sigma_{\mathcal{T}}^2 \tag{54}$$

$$\sum_{h=1}^{h^+} h\gamma_h = \tilde{\sigma}_{\mathcal{T}} \tag{55}$$

$$\sum_{h=1}^{h^+} \gamma_h = 1 \tag{56}$$

$$\mu_{\mathcal{T}} - \sum_{j\in A} \mu_{T_{ij}} x_{ij} \leq \Delta t_{\max}, \;\; i \in D \tag{57}$$

$$x_{ij}, \gamma_h \text{ binary} \tag{58}$$

where $m = max\left\{|\mathcal{A}| : \sum_{i\in\mathcal{A}} w_i \leq q \wedge \mathcal{A} \subset A\right\}$ is the maximum number of stops a bus can visit.

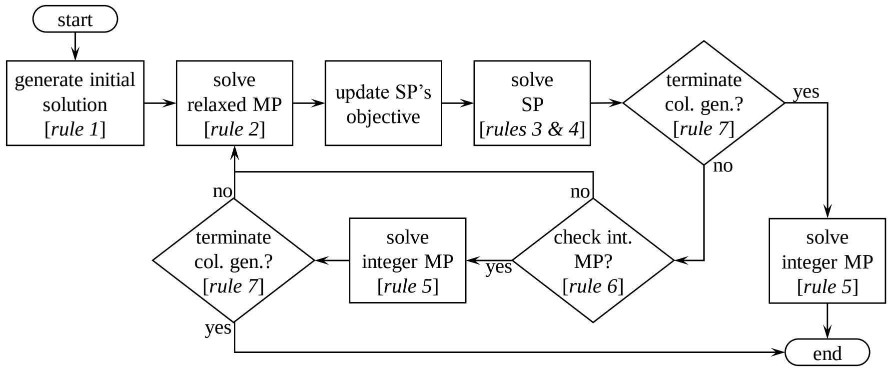

Figure 3: Column generation procedure.

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

$$l_i + \mu_{T_i} + \mu_{T_{ij}} \leq l_j + M\left(1 - x_{ij}\right), \;\; i \in L, j \in L \tag{59}$$

$$t_{\text{avl}} + \mu_{T_{di}} \leq l_i \leq t_{\text{bell}} - \mu_{T_i} - \mu_{T_{is}}, \;\; i \in L, d \in D, s \in S \tag{60}$$

where l_i is a decision variable representing the time at which a bus arrives to stop $i \in L$ and $M = \max\{t_{\text{bell}} - \mu_{T_i} - \mu_{T_{is}} + \mu_{T_{ij}} - t_{\text{avl}} - \mu_{T_{dj}}\}$, (59) establish the relation between the arrival time to one stop and its immediate successor and (60) 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.

4.3.4 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 et al. (1993), Barnhart et al. (2002), Patel et al. (2005) and Ceselli et al. (2009), 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 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 (41) and (42) with

$$\sum_{r \in R} \sum_{p \in P_r} \nu_{ir}^p y_r^p \geq \varphi, \;\; i \in A \tag{61}$$

$$\sum_{p \in P_r} y_r^p \leq \varphi K_r, \;\; r \in R \tag{62}$$

respectively, where φ 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 < -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 > 20 sec., or relative gap $< 10\%$.

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 Lübbecke and Desrosiers (2005)) and we replace (41) with

$$\sum_{r \in R} \sum_{p \in P_r} \nu_{ir}^p y_r^p \geq 1, \;\; i \in A \tag{63}$$

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 $\forall$ SP, or the last integer check shows no improvement.

4.3.5 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 3). Table 2 shows the treatment combinations and the results obtained and Figure 4 shows the iteration plot for the mean of both responses.

From both Table 2 and Figure 4 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 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.

Table 2: Factorial design for the single school routing problem

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

(*) CPU time in minutes

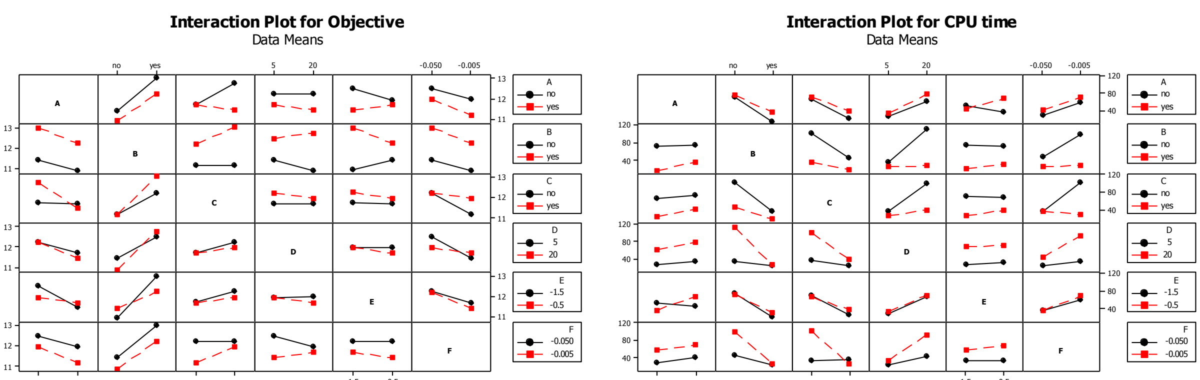

Figure 4: Results from factorial design.

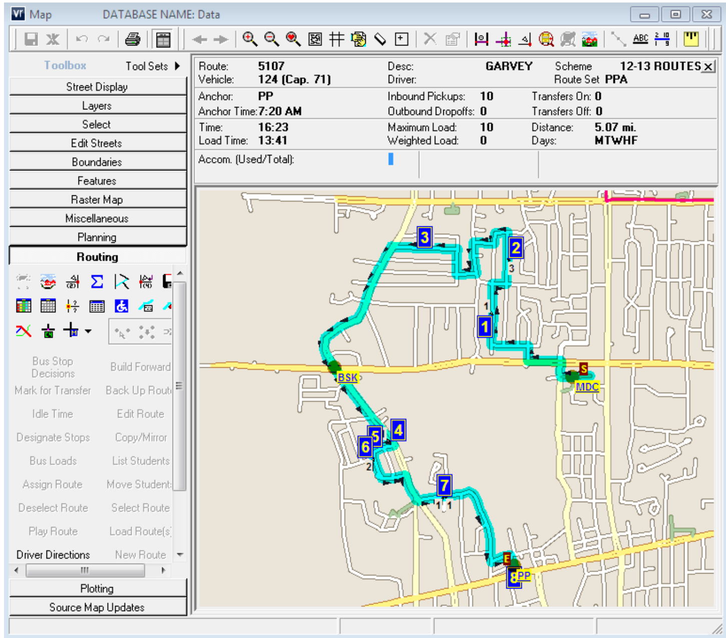

Figure 5: Student transportation management solution.

5 Model application to Williamsville Central School District

5.1 Data gathering

The Transportation Department at WCSD uses Versatrans (Tyler Technologies, 2014) 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 5 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 (23) and (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 (Braca et al., 1997).

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, 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. This data consists 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 3). Thus, overbooking the buses according to the school they go to is an appropriate strategy to a better utilization of the bus capacity.

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-depot to 1-school problems (single

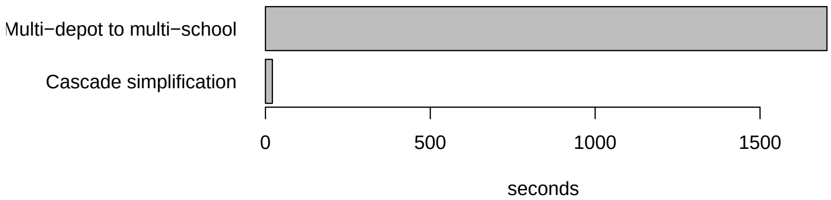

Figure 6: Solution time comparisons for the multi-depot to multi-school model and the cascade simplification.

Table 3: Results for the cascade simplification applied to WCSD

		drop				current	initial sol.		column generation			improved sol.		buses
school	grade	off time	students	stops	ridership	buses	buses	travel*	iterations	columns	time**	buses	travel*	saved
s01	HI	7:20	1237	172	36%	19	18	1002	246	524	17	18	1002	1
s02	HI	7:20	850	117	28%	13	12	653	338	753	24	10	391	3
s03	HI	7:20	1008	132	22%	12	12	695	310	635	21	11	476	1
Subtotal bell time						44	42	2350				39	1869	5
s04	EL	8:05	692	177	67%	16	15	800	505	623	33	15	800	1
s05	EL	8:05	622	169	72%	13	14	811	680	913	44	14	468	-
s06	EL	8:05	555	124	69%	12	11	579	840	1084	50	10	378	2
s07	EL	8:05	504	156	72%	12	13	660	400	492	25	13	663	-
Subtotal bell time						53	53	2850				52	2309	3
s08	MI	8:45	1009	135	69%	17	19	971	472	1627	29	19	964	-
s09	MI	8:45	883	127	53%	18	15	761	512	1460	35	15	755	3
s10	MI	8:45	679	90	62%	16	12	590	399	1180	28	12	590	4
s11	MI	8:45	644	83	52%	13	11	516	448	1412	32	11	240	2
s12	EL	8:45	663	154	62%	12	13	643	1092	1893	96	13	305	-
s13	EL	8:45	572	144	64%	10	13	697	1085	1553	96	11	410	-
Subtotal bell time						86	83	4178				81	3264	9
Total morning						86	83	9378				81	7442	9

(*) travel: total travel time (min). (**) time: CPU time in minutes

school).

Before applying any procedure to generate the routes, the first step is to set the stop locations using the model presented in section 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 6 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 (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 3 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 corresponding 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.

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 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: (i) find a route with potential for deletion, i.e., lowest capacity usage; (ii) find near by routes with the capacity to receive all or a fraction of the students from the route to be deleted; (iii) merge these routes and redistribute students following the corresponding overlapping routes found with the Cascade Simplification; (iv) check feasibility in both capacity and time; and (v) 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 5. Additionally, the set of new routes from Table 3 serves as reference of a final state of the routes in step (iii); 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.

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 β 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 β 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 $t_{\mathbf{avl}}$ 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.

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.

Acknowledgment

The authors are grateful to three anonymous reviewers whose comments on an earlier version of paper have helped improve it substantially.

References

Alabas-Uslu, C. (2008). A Self-Tuning Heuristic for a Multi-Objective Vehicle Routing Problem. *The Journal of the Operational Research Society*, 59(7):988–996.

Barnhart, C., Kniker, T. S., and Lohatepanont, M. (2002). Itinerary-Based Airline Fleet Assignment. *Transportation Science*, 36(2):199–217.

Bock, A., Grant, E., Könemann, J., and Sanità, L. (2012). The School Bus Problem on Trees. *Algorithmica*, 67(1):49–64.

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

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

Braca, J., Bramel, J., Posner, B., and Simchi-Levi, D. (1997). A computerized approach to the new york city school bus routing problem. *IIE Transactions*, 29(8):693–702.

Ceselli, A., Righini, G., and Salani, M. (2009). A Column Generation Algorithm for a Rich Vehicle-Routing Problem. *Transportation Science*, 43(1):56–69.

Christofides, N. and Eilon, S. (1969). An algorithm for the vehicle-dispatching problem. *OR*, 20(3):pp. 309–318.

Clarke, G. and Wright, J. W. (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12(4):568–581.

Corberan, A., Fernandez, E., Laguna, M., and Marti, R. (2002). Heuristic solutions to the problem of routing school buses with multiple objectives. *Journal of the Operational Research Society*, 53:427–435.

Danna, E. and Le Pape, C. (2005). Branch-and-price heuristics: A case study on the vehicle routing problem with time windows. In Desaulniers, G., Desrosiers, J., and Solomon, M., editors, *Column Generation*, pages 99–129. Springer US.

Delagado-Serna, C. R. and Pacheco-Bonrostro, J. (2001). Minmax vehicle routing problems: Application to school transport in the province of burgos. In Vo, S. and Daduna, J., editors, *Computer-Aided Scheduling of Public Transport*, volume 505 of *Lecture Notes in Economics and Mathematical Systems*, pages 297–317. Springer Berlin Heidelberg.

Desrosiers, J., Ferland, J., Rousseau, J.-M., Lapalme, G., and Chapleau, L. (1981). *An Overview of a School Busing System*, pages 235–243. North-Holland.

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

Fu, Z., Eglese, R., and Li, L. Y. O. (2005). A new tabu search heuristic for the open vehicle routing problem. *The Journal of the Operational Research Society*, 56(3):267–274.

Gounaris, C. E., Wiesemann, W., and Floudas, C. A. (2013). The robust capacitated vehicle routing problem under demand uncertainty. *Operations Research*, 61(3):677–693.

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

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

Krishnamurthy, N., Batta, R., and Karwan, M. (1993). Developing conflict-free routes for automated guided vehicles. *Operations Research*, 41(6):1077–1090.

Laporte, G., Louveaux, F., and Mercure, H. (1989). Models and exact solutions for a class of stochastic location-routing problems. *European Journal of Operational Research*, 39(1):71 – 78.

Laporte, G., Louveaux, F., and Mercure, H. (1992). The vehicle routing problem with stochastic travel times. *Transportation Science*, 26(3):161–170.

Lübbecke, M. E. and Desrosiers, J. (2005). Selected Topics in Column Generation. *Operations Research*, 53(6):1007–1023.

Pacheco, J., Caballero, R., Laguna, M., and Molina, J. (2013). Bi-Objective Bus Routing: An Application to School Buses in Rural Areas. *Transportation Science*, 47(3):397–411.

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

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

Patel, D. J., Batta, R., and Nagi, R. (2005). Clustering Sensors in Wireless Ad Hoc Networks Operating in a Threat Environment. *Operations Research*, 53(3):432–442.

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

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

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

Stewart Jr., W. R. and Golden, B. L. (1983). Stochastic vehicle routing: A comprehensive approach. *European Journal of Operational Research*, 14(4):371 – 385.

Tyler Technologies (2014). Versatrans School Transportation Software.