SHAREABILITY NETWORK BASED DECOMPOSITION APPROACH FOR SOLVING THE LARGE-SCALE SCHOOL BUS ROUTING PROBLEM

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Xiaotong Guo

May 2019

SHAREABILITY NETWORK BASED DECOMPOSITION APPROACH FOR SOLVING THE LARGE-SCALE SCHOOL BUS ROUTING PROBLEM

Xiaotong Guo, M.S.

Cornell University 2019

We consider the classic School Bus Routing Problem (SBRP) combined with alternate modes, where students are either picked up by a fleet of school buses subject to some constraints or transported by alternate transportation modes to a common destination (school). The constraints that are typically imposed for school buses are a maximum fleet size, a maximum walking distance to a pickup point and a maximum commute time for each student. This is a special case of the Vehicle Routing Problem (VRP) with a common destination. We propose a decomposition approach for solving this problem based on the existing notion of a shareability network, which has been used recently in the context of dynamic ridepooling problems. Furthermore, we build a connection between the weighted set covering problem and SBRP after decomposition via a shareability network. To scale this method to large-scale problem instances, we propose i) a node compression method of the shareability network based decomposition approach, and ii) heuristic-based edge compression techniques that works well in practice. We show that the compressed problem leads to an Integer Linear Programming (ILP) of reduced dimensionality that can be solved very efficiently using off-the-shelf ILP solvers. Numerical experiments on small-scale, large-scale and benchmark networks are used to evaluate the performance of our approach and compare it to existing large-scale SBRP solving techniques.

BIOGRAPHICAL SKETCH

Xiaotong Guo is currently in his 2th year of study in the Transportation Systems Engineering Program, Department of Civil and Environmental Engineering at Cornell University. In May 2019, he will graduate with an Master of Science degree, with a focus in transportation systems. Starting from August 2019, Xiaotong will pursue his Ph.D. degree in Interdepartmental Transportation Program at MIT.

During the past two years at Cornell University, Xiaotong worked with Prof. Samitha Samaranayake on optimization for school bus scheduling. During the summer of 2018, Xiaotong cooperated with Prof. Jinhua Zhao, who is currently an associate professor at Department of Urban Studies and Planning, MIT, on the project about introducing individuals' time flexibility in ridesharing systems. In a short term, Xiaotong will continue to work on solving problems in the urban mobility systems, ridesharing systems for instance. He tries to combine the methodology, algorithm and social impact together to gain more insights an ideas about urban mobility systems.

Outside of academics, Xiaotong enjoys playing basketball and power-lifting. He is one of the members of Cornell Chinese basketball team.

ACKNOWLEDGEMENTS

This research is founded by U.S. Department of Transportation. The author would like to thank Prof. Samitha Samaranayake for being a wonderful advisor for two years and Ph.D. candidate Yang Liu for his comments on this paper.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

According to the American School Bus Council, nearly 480,000 school buses transported 25 million children to and from school and school-related activities every school day in 2010 [3]. Meanwhile, based on a recent report from Nation Center for Education Statistics, 23 billion dollars were spent on public school transportation during the academic year 2013-2014, which is nearly 5 percent of the total expenditures for public schools [20]. Every dollar spent on transporting students is a dollar lost for direct spending to improve the education of students. Therefore, the efficient and economical operation of school bus systems is of significant importance to school districts that are trying to make the most of their limited education budgets.

The major costs associated with operating a school bus service are the capital and operational cost of the buses and the wages of the drivers. Thus, an efficient solution will serve the students by traveling the shortest distance and using the fewest buses. This needs to be done subject to getting everyone to school on time and not making some students spend a very long time sitting on a bus (e.g. one hour maximum in Boston). This leads to the so-called school bus routing problem (SBRP). Furthermore, we incorporated alternate transportation modes in SBRP by allowing students of senior years to take other transportation modes (dedicated vehicles for instance) to school, which could be beneficial for the whole bus routing scheduling by reducing the number of buses.

The SBRP is a generalization of the metric Traveling Salesman Problem (TSP) and a special case of Vehicle Routing Problem (VRP), both of which are NP-hard problems [19]. While the metric TSP has a number of good approximation techniques for obtaining provable guarantees on the solution accuracy, the VRP and SBRP problems are harder to approximate and typically solved using heuristic techniques. Therefore, the state-of-art methods for solving SBRP can only solve small-scale problems optimally. To solve the SBRP at scale, the problem is typically formulated as an Integer Linear Programming (ILP) and solved using different heuristics techniques [13, 15, 17]. One limitation of these approaches is that they lead to very high dimensional ILP problems that have a very large decision space, and are hard to solve well at-scale even with very sophisticated heuristic techniques.

This paper proposes a new approach for solving the SBRP considering alternates modes at-scale with a decomposition approach via a shareability network. Compared to classical approaches, our decomposition method consists of a multi-step approach that leads to a much simpler ILP problem compared to the traditional ILP formulations. Our decomposition approach utilizes the following steps:

- Decoupling the bus routing and student matching problems via the construction of a shareability network and student-trip assignment graph.
- Using a node compression technique for the shareability network by assigning students to bus stops subject to maximum walking constraints.
- Using a set of heuristic-based edge compression techniques for the shareability network to delete edges and compress the feasible set.

The steps described above lead to a much simpler and smaller ILP. For extreme large-scale problems, node and edge compression techniques for the shareability network can be combined with the traditional large-scale ILP heuristics to obtain solutions more efficiently (column generation for instance).

The contributions of this article include the following three components:

1. Considering to offer students alternates modes in the SBRP for the first time, which could benefit the whole bus scheduling.
2. Modeling the SBRP using the shareability network framework (used in high-capacity ridepooling), defining the corresponding student-trip graph and formulating the corresponding ILP problem.
3. Showing that techniques used in high-capacity ridepooling using a shareability network can not be applied to the SBRP directly, due to the density of the resulting shareability network, and developing network compression techniques to improve the tractability of the problem.
4. Connecting SBRP with weighted set covering problem with the decomposition approach via a shareability network.
5. Numerical results that validate the performance of our approach in solving large-scale SBRPs efficiently. Conducting benchmark testing with two different approach for solving SBRPs and showing the improvement on the objective function.

The rest of the article is organized as follows. Chapter 2 reviews the related literature. Chapter 3 provides a basic definitions for the SBRP with alternate modes. The model formulation for our decomposition approach via a compressed shareability network is shown in Chapter 4. Chapter 5 states the numerical experiments, benchmark testing and sensitivity analyses for our approach. Finally, Chapter 6 recaps the main ideas of this thesis and lists future directions for this research.

CHAPTER 2

LITERATURE REVIEW

The SBRP has been studied since 1969 when Newton and Thomas first proposed a method to generate school bus routes and schedules [10]. A comprehensive review of the SBRP can be found in Park and Kim [12], where SBRP is decomposed into five steps including data preparation, bus stop selection, bus route generation, school bell time adjustment and route scheduling. This paper focuses on solving the bus stop selection and bus route generation aspects of the SBRP, which we refer to as the SBRP.

Bekta's et al. [4] proposed an ILP model based on the open vehicle routing problem (OVRP), in which vehicles do not return to the depot after serving the last demand, to solve the real-life SBRP for transporting the students of an elementary school throughout central Ankara, Turkey. They considered a capacity constraint for the vehicles and a maximum travel distance constraint for each student, and an objective of minimizing the bus operating cost. This paper provides a basic mathematical formulation of the SBRP.

Different constraints and objectives for the SBRP have been considered in the literatures. Park et al. [13] developed a mixed load algorithm for the SBRP, where students from different schools can be served using the same bus. The problem is modeled using an ILP and solved by a post-improvement algorithm applied to a single load solution. The algorithm they proposed is an improvement on the mixed load algorithm given by Braca et al. [7], which addressed the New York City school bus routing problem. Shafahi et al. [18] proposed a new formulation of the SBRP with a homogeneous fleet that maximized trip compatibility (two trips are compatible if they can be served by the same bus) while minimizing the total travel time, and generated eight mid-size data sets to test the performance of the model.

The literatures on solving large-scale SBRPs are dominated by heuristic approaches. Riera-Ledesma and Salazar-Gonzalez [15] solved the large-scale SBRPs by modeling it as an ILP of the multi-vehicle traveling purchaser problem, which is a generalization of the VRP. The LP-relaxation method was used to efficiently solve the high dimensional ILP and a heuristic algorithm was proposed to round the fractional results. This approach was tested by using synthetic data and shown to solve instances with up to 125 students.

Schittekat et al. [17] proposed a sophisticated ILP considering both the bus stop selection and the bus routing generation simultaneously and used a metaheuristic approach to solve the problem. The metaheuristic approach contains two steps i) a construction phase that uses a greedy randomized adaptive search procedure to compute sub-optimal starting solutions for improvement phase and ii) an improvement phase where a variable neighborhood descent method is applied, it uses different neighborhood structures and ensures a local optimum in all neighborhoods. The method can produce satisfying solutions within one hour for problems of up to 80 stops and 800 students, and therefore we use the generated instances from this paper as the benchmark for testing our method.

More recently, Bertsimas et al. [5] proposed the first optimization model for the School Time Selection Problem(STSP), which is a generalization of the school bus routing problem. A state-of-art bus routing algorithm, named BiRD (Bi-objective routing decomposition), was proposed by them. The BiRD algorithm consists of generating single-school bus routes as sub-problems and combining sub-problems via mixed-integer optimization to identify a trip-by-trip itinerary for each bus in the fleet. The implementation of their approach led to a $5 million annually saving in Boston. The BiRD algorithm will serve as a benchmark to test our shareability network based decomposition approach in the experiments section.

In summary, most of the recent papers use ILP as a basic approach and concentrate on proposing heuristic techniques to improve efficiency for solving the ILP. To improve the efficiency and accuracy of current approaches, this paper will propose a shareability network based decomposition approach to solve large-scale SBRPs and conduct real-world experiments of Boston public schools. Our approach modeling the SBRP is used in dynamic high-capacity ridepooling problem [2], which is a special case of dynamic open capacitated VRP with time windows. They proposed a sequential approach via the shareability network to get a low dimensional ILP for solving the real-time high-capacity ride-sharing systems, and this approach is adapted to the SBRP. The notion of the shareability network is firstly described by Santi et al. [16], which helps them efficiently compute optimal sharing strategies on a massive dataset.

CHAPTER 3

PROBLEM FORMULATION

In this chapter, we will give a formal definition of the school bus routing problem (SBRP) with alternate modes. The problem description will be consistent throughout the paper.

Let $G_r(V_r, E_r)$ denote the road network, for any pair of nodes $i, j \in V_r$, d_{ij} represents the shortest path distance and t_{ij} stands for the corresponding travel time. Consider a set of students S need to be transported to a single destination(school) $v_d \in V_r$ with a fleet of school buses V, in which the fleet is homogeneous with capacity C and it consists of at most K^{max} buses. Without loss of generality, we define that each student $s \in S$ is located at some location $v_s \in V_r$ and the set D indicates students pickup locations, $D \subseteq V_r$. Moreover, we let M represent the set of potential bus stops, where $M \subseteq V_r$ [1].

Giving a set of alternate modes A, each student $s \in S$ could either take school buses or an alternate transportation mode $a \in A$ to the school. We define $\alpha^b_{C_1}$ as the cost for possessing a bus per day, $\alpha^b_{C_2}$ as the cost for operating a bus per mile and α^a_C as the cost for taking an alternate mode a per mile [2]. The objective of this problem is to minimize the total cost for the school bus scheduling. Figure 3.1 illustrates a simple instance of this problem.

We enforce the following constraints in the SBRP formulation:

1. The maximum travel time for any student s staying on the school bus is

[1] The graph can be augmented to add vertexes corresponding to each pickup or bus stop location.

[2] For the simplicity, we assume the cost for the alternate mode a has a linear relation with the trip distance, which can be replaced by a cost function in the future.

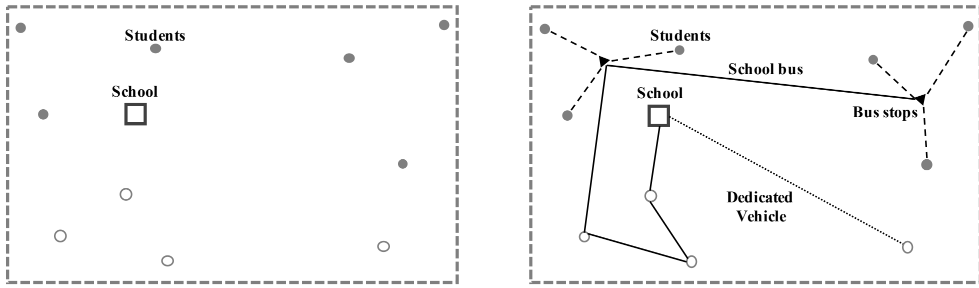


Figure 3.1: Instance of SBRP with alternate modes

$t^{\max}$.

2. Each student $s \in S$ has a maximum walking distance $d_s^{\max}$ from their residence to bus stops. This distance can be student specific and equal to zero if the students need a door-to-door pickup. Furthermore, we let N_s represent the set of reachable stops for student s, i.e., $N_s = \{m \in M | d_{v_s m} \leq d_s^{max}\}$.

3. All school buses start at a single pre-specified location $v_0 \in V_r$.

We let $N = D \cup M \cup \{v_0, v_d\}$ denote the set of pickup locations combined with potential bus stops, bus depot and school location. The decision variables for this problem are x_{ijk}, y_{ik}, z_{isk} and u_s, where $x_{ijk} = 1$ if bus k travels from vertex i to j through the shortest path, $y_{ik} = 1$ if bus k visits vertex i, $z_{isk} = 1$ if student s is picked up by bus k at vertex i and $u_{sa} = 1$ if student s takes an alternate mode a to the destination. Assuming each bus stop or student residence can be visited by at most one bus, the ILP formulation for the SBRP considering alternate modes is described in the following:

$$\min \quad \alpha^b_{C_1} \cdot K + \alpha^b_{C_2} \cdot \sum_{i\in N}\sum_{j\in N} d_{ij} \sum_{j\in V} x_{ijk} + \sum_{a\in A} \alpha^a_C \cdot \sum_{s\in S} u_{sa} d_{v_s v_d} \tag{3.1}$$

$$\text{s.t.} \quad \sum_{j\in N} x_{ijk} = \sum_{j\in N} x_{jik} = y_{ik} \quad \forall i \in N, \forall k \in V \tag{3.2}$$

$$\sum_{j\in N\setminus\{v_0,v_d\}} x_{v_0 jk} = \sum_{i\in N\setminus\{v_0,v_d\}} x_{iv_d k} \quad \forall k \in V \tag{3.3}$$

$$\sum_{i,j\in Q} x_{ijk} \leq |Q| - 1 \quad \forall Q \subseteq N, \forall k \in V \tag{3.4}$$

$$\sum_{i,j\in N\setminus\{v_0\}} t_{ij} \cdot x_{ijk} \leq t^{max} \quad \forall k \in V \tag{3.5}$$

$$\sum_{i\in N}\sum_{s\in S} z_{isk} \leq C \quad \forall k \in V \tag{3.6}$$

$$\sum_{k\in V} y_{ik} \leq 1 \quad \forall i \in N \setminus \{v_0, v_d\} \tag{3.7}$$

$$z_{isk} \leq y_{ik} \quad \forall i \in N, \forall s \in S, \forall k \in V \tag{3.8}$$

$$\sum_{a\in A} u_{sa} + \sum_{i\in N_s\cup\{v_s\}} \sum_{k\in V} z_{isk} = 1 \quad \forall s \in S \tag{3.9}$$

$$\sum_{k\in V}\sum_{i\in N} x_{iv_d k} = K \leq K^{max} \tag{3.10}$$

$$x_{ijk} \in \{0, 1\} \quad \forall i, j \in N, \forall k \in V \tag{3.11}$$

$$y_{ik} \in \{0, 1\} \quad \forall i \in N, \forall k \in V \tag{3.12}$$

$$z_{isk} \in \{0, 1\} \quad \forall i \in N, \forall s \in S, \forall k \in V \tag{3.13}$$

$$u_{sa} \in \{0, 1\} \quad \forall s \in S, \forall a \in A \tag{3.14}$$

The objective function (3.1) minimizes the overall school bus scheduling cost considering the number of buses, vehicle miles travel and alternate modes cost. Constraints (3.2) ensure that if bus k visits pickup location i, then there will be a flow entering i and a flow leaving out of i for bus k. Constraints (3.3) impose that the bus entering the destination should also have left the depot. Constraints

(3.4) keep the connectivity for bus k, which eliminates the potential sub-tours. Constraints (3.5) consider the maximum travel time for each student by restricting the total travel time for each bus route. Constraints (3.6) enforce that the number of students in bus k never exceed the capacity C. Constraints (3.7) guarantee that each vertex should be visited at most once. Constraints (3.8) ensure that student s will not be picked up by bus k at vertex i if bus k never visit the vertex i. Constraints (3.9) impose that student s either takes an alternate mode or be picked up by a school bus to the destination. Constraint (3.10) enumerates the number of non-idle buses and enforces the maximum number of available buses K^{max}. Constraints (3.11) - (3.14) make sure that the decision variables are binary.

The ILP is constructed for solving the optimal school bus routing schedules for a single school. The single school setting reduces the problem complexity, and moreover, it has real-world meanings. Different schools may obtain school bus services from different providers and potential conflicts could be induced by sharing buses among schools. The large-scale SBRP considering alternate modes is intractable for the state-of-art ILP solvers. Therefore, solving this problem at scale in a computationally tractable manner requires some decomposition and heuristic methods. The following section describe our new approach and heuristics for the SBRP with alternate modes that scales better than existing techniques, while still preserving the same quality of solutions.

CHAPTER 4

METHODOLOGY

In this chapter, we will propose a decomposition method through the shareability network framework, which is recently used in the shared mobility literature [2, 16], to solve the SBRP with alternate modes. Furthermore, we will show the connection between the weighted set covering problem and decomposed SBRP. This theoretical connection provides a different perspective with underlying approach to deal with the SBRP. In the last part, we will increase the tractability for large-scale instances by introducing network compression techniques based on the shareability network.

4.1 Decomposition through the shareability network

In order to reduce the complexity and dimensionality of the ILP for solving the SBRP with alternate modes, we propose a decomposition method which consists of several steps leading to an assignment problem via the shareability network, which yields a much-simplified ILP.

The shareability network [16] is an undirected graph $G = (V, E)$, where V corresponds to the set of trips and each edge $(i, j) \in E$ indicates that trip i can share a vehicle with trip j under some compatibility constraints. The shareability network under the SBRP setting is constructed as follows. The vertex set V designates to be the set of students and each edge $(s_i, s_j) \in E$ reflects that students s_i and s_j can share the same school bus. The students s_i and s_j can share a same bus if both of them could be picked up by the same bus at reachable bus stops or their residences and arrive at the destination v_d within the maximum travel time t^{max} using the same bus.

Figure 4.1(a) shows an instance of a shareability network for four students.

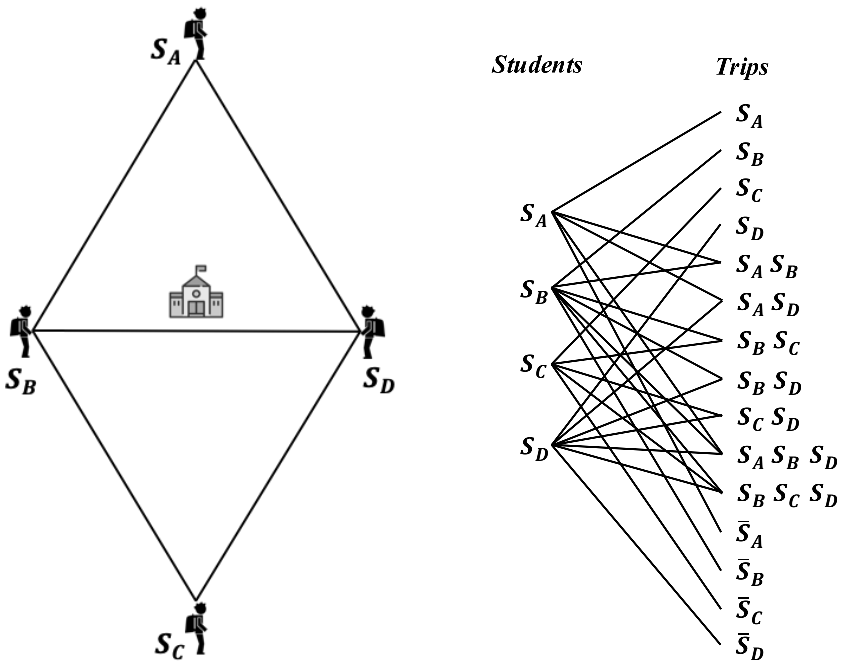


Figure 4.1: Instance of the shareability network and ST-graph

Next, we establish a bipartite graph $G_{ST} = (V_{ST}, E_{ST})$ with a set of students and a set of all possible trips allocations (school bus or alternate modes assignment) based on the shareability network. This bipartite graph is referred as the student-trip graph (ST-graph). The set of feasible trips T includes bus trips T_b and trips T_a for an alternate mode a , i.e., $T = T_b \bigcup_{a \in A} T_a$. Let $S(\tau_b)$ denote the set of students who participate in a feasible bus trip $\tau_b \in T_b$. A bus trip $\tau_b \in T_b$ is feasible if the travel time $t_s \leq t^{max}$, $\forall s \in S(\tau_b)$ and the total number of students $s \in S(\tau_b)$ is smaller than bus capacity C. For each student $s \in S$, $\tau_a^s \in T_a$ represents the trip that student s directly takes the alternate mode a to school. The node set V_{ST} is the union of the set of student and the set of feasible trips, i.e. $V_{ST} = S \cup T$, and there will be an edge $e(s, \tau_b) \in E_{ST}$ if $s \in S, \tau_b \in T_b$, $s \in S(\tau_b)$, and an edge $e(s, \tau_a^s) \in E_{ST}$ for every students $s \in S, \tau_a^s \in T_a$. Figure 4.1(b) shows an instance of ST-graph corresponding to the shareability network in Figure 4.1(a).

The set of feasible bus trip T_b is generated using the shareability network. The following observation is typically made to efficiently compute the feasible bus trips in T_b based on the shareability network G [2].

Lemma 1 *(Lemma 1 in [2]) A bus trip τ_b can be feasible only if all students s in the bus trip τ_b forms a clique in the shareability network G (i.e. $\forall s_i, s_j \in S(\tau), e(s_i, s_j)$ exists).*

Given this observation, if any pair of students in a bus trip does not have an edge in the shareability graph G, this bus trip will be infeasible. Thus, if a set of n students can not form a feasible bus trip, we know that they can not form any feasible bus trips that include another additional student s_{n+1}. We construct the set of feasible bus trip T_b by first considering trips which consist of one student and progressively considering larger sets only when the smaller set is feasible. Algorithm 1 describes the details for generating the feasible trip list T_b. The input function PATHTSP$(\cdot)$ is a blackbox for solving the path traveling salesman problem (path-TSP), which is NP-hard. We will give a decent heuristic approach for solving the path-TSP in the experiments section. In general, this algorithm provides an efficient pruning mechanism that eliminates the consideration of infeasible bus trips.

The last step of our approach is to compute the optimal student-trip assignment given the ST-graph G_{ST}, and it is formalized as an ILP. From the ST-graph G_{ST}, we can calculate the travel cost C_τ for each trip $\tau \in T$. Therefore, we can formulate an assignment problem based on G_{ST} which assigns all students to

Algorithm 1 Generating the set of feasible bus trips. Input: the shareability network G, the set of students S, maximum travel time t^{max}, bus capacity C, path-TSP solver for any trip τ with optimal travel time t^* as the output, i.e., $t^* = \textsf{PathTsp}(\tau)$

function BusTripGeneration($G = (V, E), S, t^{max}, C, \textsf{PathTsp}(\cdot)$)
 $T_b, T_b^1 \leftarrow \emptyset$
 for $s \in S$ **do** ▷ Generate the trip list with one student
  $\tau \leftarrow \{s\}$
  $T_b^1 \leftarrow T_b^1 \cup \{\tau\}$
 $T_b \leftarrow T_b \cup T_b^1$
 $k \leftarrow 2$ ▷ Iterate from trips with two students
 while true **do**
  $T_b^k \leftarrow \emptyset$ ▷ Initialize the trip list with k students
  for $\tau \in T_b^{k-1}$ **do**
   for $s \in S$ and $s \notin \tau$ **do**
    $\tau' \leftarrow \tau \cup \{s\}$ ▷ Add one more student to the trip with $k - 1$ students
    if $\textsf{CliqueCheck}(\tau, s, G(V, E)) = true$ **then**
     if $\textsf{FeasibilityCheck}(\tau', t^{max}, C, \textsf{PathTsp}(\cdot)) = true$ **then**
      $T_b^k \leftarrow T_b^k \cup \{\tau'\}$ ▷ Add feasible trips with k students into the list
  if $|T_b^k| = 0$ **then** ▷ Break when there are no feasible trips with k students
   break
  $T_b \leftarrow T_b \cup T_b^k$; $k \leftarrow k + 1$
 return T_b
function CliqueCheck($\tau, s, G(V, E)$)
 for $s' \in S(\tau)$ **do**
  if $e(s, s') \notin E$ **then**
   return false
 return true
function FeasibilityCheck($\tau, t^{max}, C, \textsf{PathTsp}(\cdot)$)
 $t^* \leftarrow \textsf{PathTsp}(\tau)$
 if $t^* \leq t^{max}$ and $|S(\tau)| + 1 \leq C$ **then**
  return true
 else
  return false

trips while minimizing the overall school bus scheduling cost which consists of number of buses, vehicle miles travel and the cost for alternate modes.

Decision variables are $x_{s\tau}$ and y_τ, where $x_{s\tau} = 1$ represents that student s chooses trip τ and $y_\tau = 1$ if trip τ is chosen in optimal trip set, for all $s \in S$ and $\tau \in T$. Let $L(\tau)$ denote the number of students in trip τ, $L(\tau_a^s) = 1$ and $L(\tau_b) = |S(\tau_b)|$, and we have the following ILP for student-trip assignment:

$$\min \quad \sum_{\tau_b \in T_b} [\alpha^b_{C_1} + \alpha^b_{C_2} \cdot C_{\tau_b}] \cdot y_{\tau_b} + \sum_{a \in A} \sum_{\tau_a \in T_a} \alpha^a_C \cdot C_{\tau_a} \cdot y_{\tau_a} \tag{4.1}$$

$$\text{s.t.} \quad \sum_{s \in S} x_{s\tau} = y_\tau \cdot L(\tau) \quad \forall \tau \in T \tag{4.2}$$

$$\sum_{\tau \in T} x_{s\tau} = 1 \quad \forall s \in S \tag{4.3}$$

$$x_{s\tau} \in \{0, 1\} \quad \forall s \in S, \forall \tau \in T \tag{4.4}$$

$$y_\tau \in \{0, 1\} \quad \forall \tau \in T \tag{4.5}$$

The objective function (4.1) minimizes the overall school bus scheduling cost. Constraints (4.2) ensure that if a trip τ is selected, then all students participated in this trip τ won't be considered by other trips. Constraints (4.3) impose that the student $s \in S$ are considered exactly once in all selected trips. Constraints (4.4) and (4.5) make sure that decision variables are binary.

We simplify the ILP (3.1) - (3.14) via the shareability network and ST-graph to the ILP (4.1) - (4.5). However, this simplified ILP is still intractable when considering the large-scale SBRP since the size of the feasible bus trip list $|T_b|$ will increase exponentially when $|S|$ increases. The shareability network is dense because of the loose constraint, which is the maximum travel time t^{max} (usually

1 hour) for the SBRP. The loose constraints induce many large-sized cliques in the shareability network compared to the dynamic ridepooling problem where stricter quality of service constraints lead to a sparser shareability network. In order to address this issue, we will propose some network compression techniques to induce sparsity in the shareability network that arises from a SBRP and make the problem tractable to solve.

4.2 Connection between SBRP and weighted set covering problem

After decomposing SBRP through the shareability network, it offers us a chance to build the connection between the weighted set covering problem and SBRP.

Definition 1 (Weighted set covering problem) *Giving a set of n elements $\mathcal{U} = \{e_1, e_2, ..., e_n\}$ and m subsets of $\mathcal{U}$, $\mathcal{S} = \{S_1, S_2, ..., S_m\}$ with a cost function $c : \mathcal{S} \longrightarrow \mathbb{R}^+$, $c(S_j)$ denote the cost of subset S_j. The objective is to find a set $\mathcal{A} \subseteq \mathcal{S}$ such that:*

- *All elements are covered by the set $\mathcal{A}$, and*
- *The sum of the cost of the subsets in $\mathcal{A}$ is minimized.*

Let x_S be the binary variable for whether choosing subset $S \in \mathcal{S}$ in the set $\mathcal{A}$, and the weighted set covering problem can be formulated as the following ILP:

$$\min \quad \sum_{S \in \mathcal{S}} c(S) \cdot x_S \tag{4.6}$$

$$\text{s.t.} \quad \sum_{S : e \in S} x_S \geq 1 \quad \forall e \in \mathcal{U} \tag{4.7}$$

$$x_S \in \{0, 1\} \quad \forall S \in \mathcal{S} \tag{4.8}$$

In the weighted set covering problem, if we consider the non-overlapping constraints

$$\forall S_i, S_j \in \mathcal{A}, S_i \cap S_j = \emptyset,$$

which means the set $\mathcal{A}$ is a collection of disjoint subsets in $\mathcal{S}$, the problem becomes the weighted set partitioning problem. For the ILP above, constraints (4.7) become

$$\sum_{S : e \in S} x_S = 1 \quad \forall e \in \mathcal{U}, \tag{4.9}$$

which imply that each element in $\mathcal{U}$ will be covered by $\mathcal{A}$ exactly once.

In order to build the connection between the SBRP and the weighted set covering problem, we first make the following claims:

Claim 1 *The SBRP (4.1)-(4.5) with $\alpha^b_{C_2} = 1, \alpha^b_{C_1} = \alpha^a_C = 0$ and $T = T_b$, which represents the feasible bus trip list, is a special case of the weighted set partitioning problem with the following constraints:*

$$\forall S \in \mathcal{S}, \{S' : S' \subseteq S\} \in \mathcal{S} \tag{4.10}$$

$$\forall S' \subseteq S, c(S) \geq c(S'). \tag{4.11}$$

Proof. For the ST-graph of the SBRP, we let each student $s \in S$ as an element. The set of elements will be $\mathcal{U} = S$ and each bus trip τ_b is the subset of S with trip cost C_{τ_b}. The feasible bus trip list T_b is the collection of subsets $\mathcal{S}$, and the SBRP is equivalent to the weighted set partitioning problem if we only minimize the overall trip cost for the SBRP ($\alpha^b_{C_2} = 1, \alpha^b_{C_1} = \alpha^a_C = 0$). However, the SBRP is a special case of the weighted set partitioning problem since it has more constraints on the collection of subsets $\mathcal{S}$.

When generating the feasible bus trips $\tau_b \in T_b$, let $\tau_b \in T_b$ be a feasible bus trip and $T_b^{sub} = \{\tau'_b : S(\tau'_b) \subseteq S(\tau_b)\}$ be the collection of sub-trips for τ_b, which $S(\tau_b)$ represents the set of all students in the bus trip τ_b. Because the students in the trip τ_b should form a clique in the shareability network, the students in the trip τ'_b will also form a clique. Also, the trip cost for τ'_b is smaller than the cost for τ_b. Thus, $\forall \tau'_b \in T_b^{sub}$, $\tau'_b \in T_b$. And we have constrains (4.10) and (4.11) for the collection of subsets T_b (or $\mathcal{S}$). □

Claim 2 *The weighted set partitioning problem with constraints (4.10) and (4.11) can be solved by the algorithm for the weighted set covering problem.*

Proof. We prove this claim by the contradiction. Let $\mathcal{A}$ be the optimal solution for the weighted set covering problem, and suppose there exits two subsets $S_1, S_2 \in \mathcal{A}$, $S_1 \cap S_2 \neq \emptyset$.

Let $S' = S_1 \cap S_2$, $S'_1 = S_1 \setminus S'$ and $S'_2 = S_2 \setminus S'$. According to the constraints (4.10), $S'_1, S'_2 \in \mathcal{S}$ since $S'_1 \subseteq S_1$ and $S'_2 \subseteq S_2$. By the constraints (4.11), we have $c(S'_1) \leq c(S_1)$ and $c(S'_2) \leq c(S_2)$. We can reduce the total cost for $\mathcal{A}$ by replace either S_1 with S'_1 or S_2 with S'_2 in the optimal set $\mathcal{A}$ while still covering all elements. Thus, the optimal set $\mathcal{A}$ should be a collection of disjoint subsets in $\mathcal{S}$, and the optimal set $\mathcal{A}$ is also optimal for the set partitioning problem with same $\mathcal{U}$ and $\mathcal{S}$. □

Corollary 1 *The SBRP with the objective function for minimizing the overall bus trip cost ($\alpha^b_{C_2} = 1, \alpha^b_{C_1} = \alpha^a_C = 0$) can be solved by the algorithm for the weighted set covering problem.*

Proof. Combining Claim 1 and Claim 2, the SBRP that only minimizes overall bus trip cost ($\alpha^b_{C_2} = 1, \alpha^b_{C_1} = \alpha^a_C = 0$) can be solved by the algorithm of the weighted set covering problem. □

And problem (4.1)-(4.5) with $\alpha^b_{C_2} = 1, \alpha^b_{C_1} = \alpha^a_C = 0$ and $T = T_b$ can be solved by the following ILP:

$$\min \quad \sum_{\tau \in T} C_\tau \cdot y_\tau \tag{4.12}$$

$$\text{s.t.} \quad \sum_{\tau : s \in \tau} y_\tau \geq 1 \quad \forall s \in S \tag{4.13}$$

$$y_\tau \in \{0, 1\} \quad \forall \tau \in T \tag{4.14}$$

Thus, the SBRP benefits from approaches for solving the weighted set covering problems. However, the weighted set covering problem is NP-hard and the state-of-art methods for solving the weighted set covering problems are metaheuristics. The large-scale problems are intractable and hard to reach the optimality. In the experiment section, we will compare the decomposition approach with a state-of-art metaheuristic method for solving the SBRP instead of the weighted set covering problem. The main take-away for this section is to offer a new perspective for solving the SBRP, which could lead to potential theoretical analyses.

4.3 Network compression techniques

Our approach considering the shareability network is intractable for large-scale instances as the size of the feasible bus trip list T_b is enormous (over billions). Generating the feasible trip list T is a time-consuming job, and T as the input for the student-trip assignment ILP (4.1) - (4.5) make this program intractable for off-the-shelf ILP solvers.

To eliminate the computation bottleneck of our decomposition approach, we present the network compression techniques aiming to induce sparsity in the shareability network while retaining all (or most of the) useful information that is embedded in the network. We present compression techniques from two perspectives that work by either pruning the nodes or the edges in the shareability network. The sparse shareability network will lead to a shorter feasible trip list, which makes the large-scale SBRPs tractable to solve.

4.3.1 Node compression technique

For the node compression technique, we reduce the number of nodes in the shareability network by generating bus stops and allowing students to walk to bus stops within their maximum walking distance. The school buses will pick up students at bus stops instead of students' residence, and the shareability network will be constructed based on bus stops.

With the set of potential bus stops M, We formulate the following ILP to generate the minimum number of bus stops with the binary decision variable x_m, where $x_m = 1$ represents picking the potential bus stops $m \in M$.

$$\min \quad \sum_{m \in M} x_m \tag{4.15}$$

$$\text{s.t.} \quad \sum_{m \in N_s} x_m \geq 1 \quad \forall s \in S \tag{4.16}$$

$$x_m \in \{0, 1\} \quad \forall m \in M \tag{4.17}$$

The objective function (4.15) minimizes the total number of bus stops which have been chosen. Constraints (4.16) ensure that each student could find at least one bus stops within their maximum walking distance. Constraints (4.17) impose that decision variables are binary. This ILP generates the minimum number of bus stops while covering all students within their maximum walking distance.

For the student who need door-to-door pickup ($d_s^{max} = 0$ and $N_s = \emptyset$), the node compression technique can be used by assigning a vitual walking distance to the student and quantifying the penalty brought by this assumption. The penalty is bounded by the distance of a round-trip between the assigned bus stop and door-to-door students' residence. This penalty will be incorporated in the input function $\textsc{PathTsp}(\cdot)$ when considering the feasibility of trips. Figure 4.2 shows an example of the shareability network before and after applying the node compression technique with vitual distance 0.5 miles for students with door-to-door pickups.

The node compression technique decreases the maximum number of effective students for any trip $\tau_b \in T_b$, since each bus stop is now considered as a single request with a larger capacity. Nonetheless, even with reduced number of nodes in the shareability network in this manner, the density of the shareability network still does not lead to an adequate level of sparsity for large-scale problem

Figure 4.2: Instance of applying node compression technique

instances. Therefore, we also adopt a heuristic-based edge compression technique to delete edges that are unlikely to induce shared trips. This compression can lead to a sub-optimal solution in theory because we are eliminating feasible sharing possibilities, but our aim is to generate a set of rules that only eliminate pairings that are very unlikely to occur in practice.

4.3.2 Edge compression technique

For the edge compression technique, we reduce edges in the shareability network following some mechanisms. The main idea behind pruning edges in the shareability network is to consider sharing trips with nodes that are relatively close to each other (e.g. it is unlikely for a student 5 miles north of the school to share a ride with a student 5 miles south of the school if there are enough students to fill a bus from north of the school). Meanwhile, the travel time between two nodes is not the only factor corresponding to the likelihood of sharing trips, the relative position of the school and nodes also matters (e.g. two students could share the same trip if they are at the same side of the school even if they are far away from each other).

To incorporate both factors, we define the adjusted travel time $\bar{t}_{ij} = \delta_{ij} \cdot t_{ij}$ between any two nodes i and j in the shareability network, where δ_{ij} represents the detour factor and $\delta_{ij} = \frac{t_{ij}+t_{jv_d}}{t_{iv_d}}$. The detour factor $\delta_{ij} \geq 1$ reflects how much detour for node i to share the trip with node j comparing to node i directly go to the destination.

With the definition of the adjusted travel time, we apply a mechanism that a node only share trips with *nearby nodes* in the shareability network. More formally, the *nearby nodes* for any node i are generated by calculating the adjusted travel time between node i and all other nodes $V \setminus \{i\}$, sorting the nodes in ascending order by the adjusted travel time, and choosing the k closest nodes such that $k \leq \beta \cdot C$. If nodes correspond to bus stops with multiple students as a result of the node compression, we consider the k closest nodes such that the sum of students is less than $\beta \cdot C$ (i.e. $\sum_{j=1}^{k} n(m_j) \leq \beta \cdot C$, where m_j represents the bus stop and $n(m_j)$ is the number of students at stop m_j).

β is a control parameter and it should be greater than 1 (at least considering C students nearby). The number of edges decreases while β decreasing, so the number of feasible bus trips $|T_b|$ decreases. We choose β to keep an appropriate size of feasible bus trip list $|T_b|$ while giving satisfying bus routes. If β is too small, some feasible trips belong to the optimal solution will be eliminated. On the other hand, if β is too large and only a few edges are eliminated in the shareability network, the large-scale SBRPs will keep intractable. As a conclusion, we should choose a β to keep a balance between computability and optimality, the value of β can be different with respect to different numbers and distributions of the students. Figure 4.3 shows an example of the shareability network before and after applying the edge compression technique.

The edge compression technique is a heuristic approach without guarantee for

Figure 4.3: Instance of applying edge compression technique

the loss of optimality. When the size of SBRPs becomes larger, in order to make problems tractable, we have to use a small value of β to restrict the size of feasible bus trip list $|T_b|$. By the criterion of finding a feasible trip in Lemma 1, all students in the feasible trip form a clique in the shareability network. Therefore, the edge compression technique could provide a terrible results when β is small. In order to compensate for this side-effect, we propose a γ-quasi-clique process based on the Algorithm 1 to find groups of students who form quasi-cliques in the shareability network. More precisely, the quasi-clique designates to be a structure which is similar to the clique. The γ-quasi-clique process replaces the function CLIQUECHECK$(\tau, G(V, E))$ in the Algorithm 1 and details are shown in the Algorithm 2.

Algorithm 2 γ-quasi-clique process. Input: the shareability network G, a feasible bus trip $\tau \in T_b$, a student s, heuristic parameter γ.

function γ-QUASICLIQUEPROCESS$(\tau, s, G(V, E), \gamma)$
    $x \leftarrow 0$
    for $s' \in S(\tau)$ **do**
        if $e(s, s') \notin E$ **then**
            $x \leftarrow x + 1$
    if $x > \gamma \cdot |\tau|$ **then**
        return false
    return true

The γ-quasi-clique process is determined by a heuristic parameter γ, which indicates the difference between quasi-cliques and cliques. γ is upper bounded by 1 since a student s is not connected with at most $|\tau|$ students in the trip τ. By introducing the γ-quasi-clique process in the Algorithm 1, we can maintain a satisfying solution by having a large γ while decreasing β to reduce the computation time.

CHAPTER 5

EXPERIMENTS

5.1 Real-world experimental results

To test the applicability of our proposed algorithm to at-scale SBRPs, we conducted a number of numerical experiments using some publicly accessible benchmark problem instances. All the experiments are run on a 2.7 GHz Intel Core i7 Processor with 16 GB Memory using Python 3.6. The first data set we use is from Transportation Challenge held by Boston Public Schools (BPS), which contains 22420 simulated students addresses (to protect student privacy) from 134 public schools with the same aggregate pickup location distributions as in the real-world. The simulated dataset can be downloaded from `https://www.bostonpublicschools.org/transportationchallenge`. The spatial distribution of this dataset is shown in the Figure 5.1.

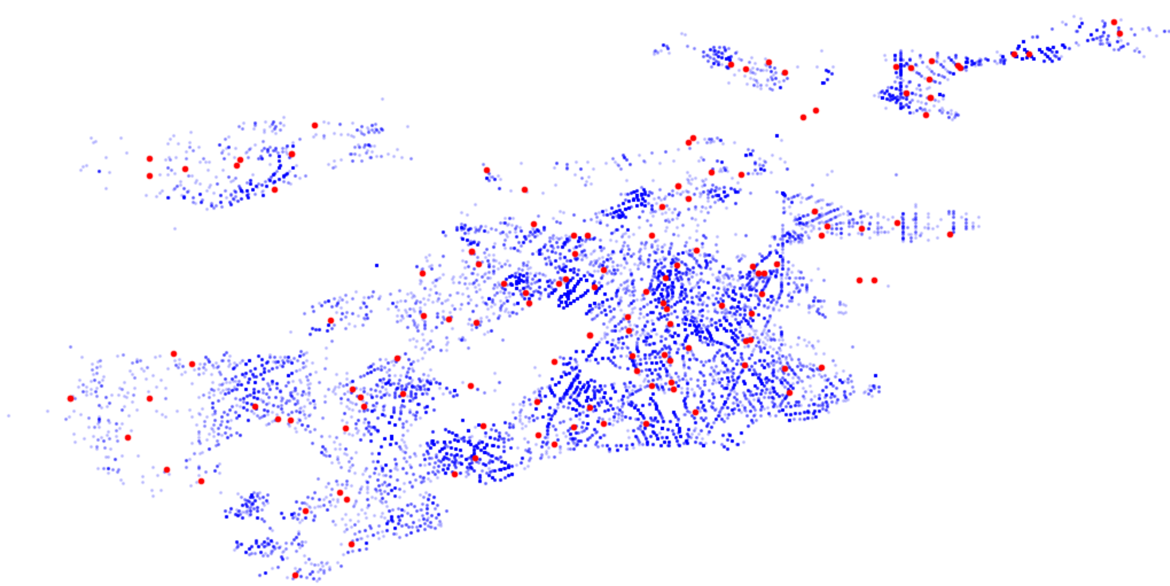


Figure 5.1: Simulated data from Boston Public School (BPS)

The Boston road network $G_r(V_r, E_r)$ is obtained from Open Street Maps (OSM) [11] using the open source Python library OSMnx [6]. All the visualizations of the results are generated using the Python library NetworkX [9]. The implementation for the function PATHTSP($\cdot$) is shown in A. We used the state-of-art ILP solver Gurobi [8] in the experiments with a 3600 seconds maximum computation time. We make the following assumptions based on the BPS data and requirements:

- Assume that school buses can start at any location in the network.
- The delay time d for buses at each bus stop m follows a function $d = 15 + 5 \cdot n(m)$, where $n(m)$ is the number of students at the bus stop m.
- Each door-to-door student $s \in S$ has the same vitual maximum walking distance $d_s^{max} = 0.5$ miles.
- The maximum number of school buses $K^{max} = |S|$, i.e. there is no restriction on the bus fleet size.
- The school bus capacity $C = 72$.
- The maximum travel time for students is $t^{max} = 3600$ seconds.
- The set of potential bus stops M is same to the set of nodes in the road network V_r.
- The set of alternate modes A contains only the dedicated vehicles.
- The cost for owning a bus is $\alpha_1 = 200$ dollars per day and the cost for operating a bus is $\alpha_2 = 1$ dollar per mile [1], which implies that the bus capital cost is much larger than the bus operational cost. This leads to solutions that minimize the number of buses needed. The cost for taking a dedicated vehicle is $\alpha_3 = 2$ dollars per mile [14].

With all the parameters and assumptions for running the experiments on the Boston road network with simulated data from Boston Public School (BPS), we picked ten schools and computation results are shown in Table 5.1 [1]. N_S, $N_{S_{d2d}}$ represent the number of students and door-to-door students, N_{M^*} denotes the number of bus stops, N_{T_b} designates for the length of the feasible bus trip list, SND indicates the objective value for the shareability network based decomposition approach, N_B stands for the number of buses, N_U means the number of students who take dedicated vehicles and T speaks for the overall computation time (seconds). The optimal school bus schedules can be found in B.

Table 5.1: Computational results for real-world experiments

School	N_S	$N_{S_{d2d}}$	N_{M^*}	N_{T_b}	β	γ	SND	N_B	N_U	T
Tommy H.	51	7	19	62632	-	-	342.00	1	24	13.76
Craig K.	71	11	20	107253	-	-	461.39	2	2	26.77
Deven M.	91	14	22	100021	-	-	527.03	2	13	21.76
Frank M.	160	30	30	333994	1.2	0.3	760.93	3	11	84.58
Dick W.	183	28	22	71672	-	-	622.07	2	42	18.36
Dick B.	208	40	35	760551	1.3	0.4	790.47	3	23	191.25
Dutch L.	243	42	42	1951866	2	0.4	942.85	3	36	515.85
Christian V.	344	66	35	553380	3	0.4	1109.87	5	0	3730.93
Dennis E.	403	75	45	322566	2	0.4	1312.87	5	44	4194.26
Rick F.	573	109	55	546319	2.5	0.4	1781.92	8	2	151.26

5.2 Benchmark testing

To further evaluate the performance and scalability of this approach, we compared our approach with two state-of-art methods for solving the SBRP [17, 5].

For the BiRD (Bi-objective Routing Decomposition) algorithm for SBRP [5], we implemented the single-school routing algorithm and compared on the BPS

[1]"-" indicates that we didn't use β or γ heuristic in edge compression techniques.

simulated data. In the comparison, we set $\alpha_1 = 500, \alpha_2 = 1$ and $\alpha_3 = \infty$ (consider the school bus as the only transportation mode). In the BiRD algorithm, N determines the optimality of solutions, which represents the number of feasible trips covering each bus stop. In the application in [5], N was set to be 400. Because of the randomness in BiRD algorithm, we set N equal to 1000 and run 10 times to get the best results. Comparison results are shown in the Table 5.2 and the optimal school bus schedules can be found in C. N_S represents the number of students, N_{M^*} denotes the number of bus stops, β and γ are parameters for the edge compression techniques, SND indicates the objective value for the shareability network based decomposition approach, *BiRD* stands for the objective value for the Bi-objective Routing Decomposition algorithm, *Improv′* shows the improvement of routing cost for SND compared to *BiRD*, *Improv* implies the improvement of objective value for SND compared to *BiRD*.

Table 5.2: Comparison results with BiRD algorithm [5]

School	N_S	N_{M^*}	N_{T_b}	β	γ	SND	BiRD	Improv′	Improv
Tommy H.	51	19	62632	-	-	1549.16	1553.44	8.01%	0.28%
Craig K.	71	20	107253	-	-	1548.97	1554.28	9.78%	0.34%
Deven M.	91	22	100021	-	-	1553.94	1558.34	7.54%	0.28%
Frank M.	160	30	333994	1.2	0.3	2080.18	2081.9	2.10%	0.08%
Dick W.	183	22	71672	-	-	1559.98	1560.32	0.56%	0.02%
Dick B.	208	35	760551	1.3	0.4	2083.13	2083.91	0.93%	0.04%
Dutch L.	243	42	1951866	2	0.4	2614.7	3120.28	4.64%	16.20%
Christian V.	344	35	553380	3	0.4	2609.87	3113.03	2.80%	16.16%
Dennis E.	403	45	548316	2.3	0.4	3127.26	3632.52	3.97%	13.91%
Rick F.	573	55	124257	2	0.4	4669.67	4673.46	2.18%	0.08%

Schittekat et al. [17] used a metaheuristic approach combining a greedy search procedure with a neighborhood search and solved the bus stop selection and school bus routing problem together while we solved these two problems separately. We used the generated instances in Euclidean space given in [17] and adopted our approach under their settings by assigning students to bus stops first and solving the optimal bus route. The comparison results are shown in Table 5.3. *ID* corresponds to the instance number, *stop* denotes the number of bus stops, *stud* represents the number of students, *cap* indicates the bus capacity, *wd* is the walking distance for students in Euclidean space, *beta* and γ are parameters for the edge compression techniques, N_{T_b} stands for the number of feasible bus trips, *MH* is the objective value for metaheuristic method [17], SND indicates the objective value for the shareability network based decomposition approach, *Improv* implies the improvement of objective value for SND compared to *MH*. We picked instances with minimum students walking distance 5 because larger walking distance led to fewer selected bus stops and made problem not interesting any more with respect to SND approach.

Table 5.3: Comparison results with metaheuristic method [17]

ID	stop	stud	cap	β	γ	N_{T_b}	MH	SND	Improv
73	40	200	25	3	0.3	2776	831.94	804.4	+3.31%
74	40	200	50	1.5	0.3	127211	593.35	585.08	+1.39%
81	40	400	25	-	-	575	1407.05	1428.2	−1.5%
82	40	400	50	3	0.3	6514	858.80	848.8	+1.16%
89	40	800	25	-	-	40	2900.14	3085.11	−6.38%
90	40	800	50	-	-	726	1345.70	1404.61	−4.38%
97	80	400	25	3	0.4	5928	1546.23	1494.38	+3.35%
98	80	400	50	1.5	0.4	40583	1048.56	1025.62	+2.19%
105	80	800	25	-	-	1410	2527.96	2623.77	−3.79%
106	80	800	50	5	0.4	45252	1530.58	1499.88	+2.01%

For instance 81, 89, 90 and 105, our approach yields worse solutions because we split the bus stops selection and bus routing problem while applying the node compression technique. Our approach has optimality limitations on these instances. However, SND approach gets better solutions for instances without such limitations while reducing a significant amount of time. From implementation perspective, it is intractable to solve both bus stops selection and bus routing problem together for large instances. Solving them separately is a reasonable decomposition which increases the problem tractability and yields satisfying school bus schedules.

5.3 Sensitivity analyses

Finally, we conduct a sensitivity analyses for network compression techniques using the large-scale BPS instance of Christian Vazquez School with 344 students, where 66 students need door-to-door pickup. We set $\alpha_1 = 500, \alpha_2 = 1$ and $\alpha_3 = \infty$ (consider the school bus as the only transportation mode).

For measuring the sensitivity to the control parameter β for the edge compression technique, we choose β values ranging from 1 to 4 with a step size of 0.2 and keep the γ equal to 0.4. The sensitivity analyses results are shown in Figure 5.2. The objective value and number of buses required decrease in unison as β increased, since a larger set of sharing options are made available via the shareability network and can be potentially fit into a smaller number of buses. As β decreases, fewer options are available and more buses maybe needed leading to a higher objective function value. In contrast, both the running time and length of the feasible trip list increase exponentially with respect to β, since the shareability graph grows exponentially in size with respect to β. The fluctuation of the running time is induced by the uncertainty for solving the ILP.

For measuring the sensitivity to the control parameter γ for the edge compression technique, we consider a range from 0 to 0.9 with a step size of 0.1 meters and keep the beta parameter fixed at $\beta = 2$. The sensitivity analyses results are shown in Figure 5.3. Unsurprisingly, we notice that the objective value will decrease when γ increases. The running time and the number of the feasible

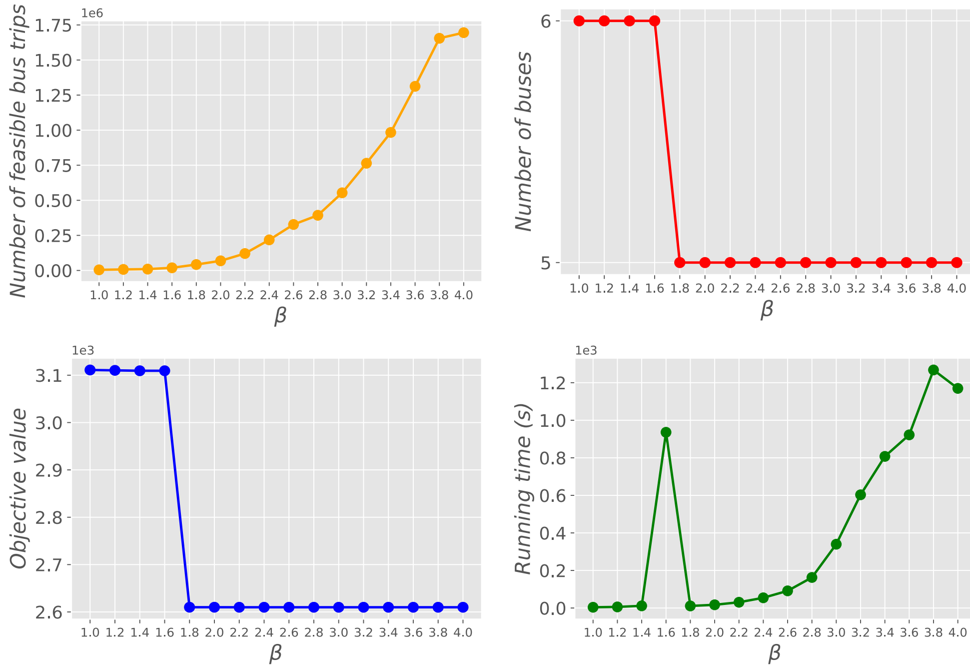


Figure 5.2: Sensitivity analysis for β

bus trips increases when γ increases. The number of buses required does not change because we can get 5 buses even without applying $\gamma - quasi - clique$ process, which is the minimum number of buses for this instance.

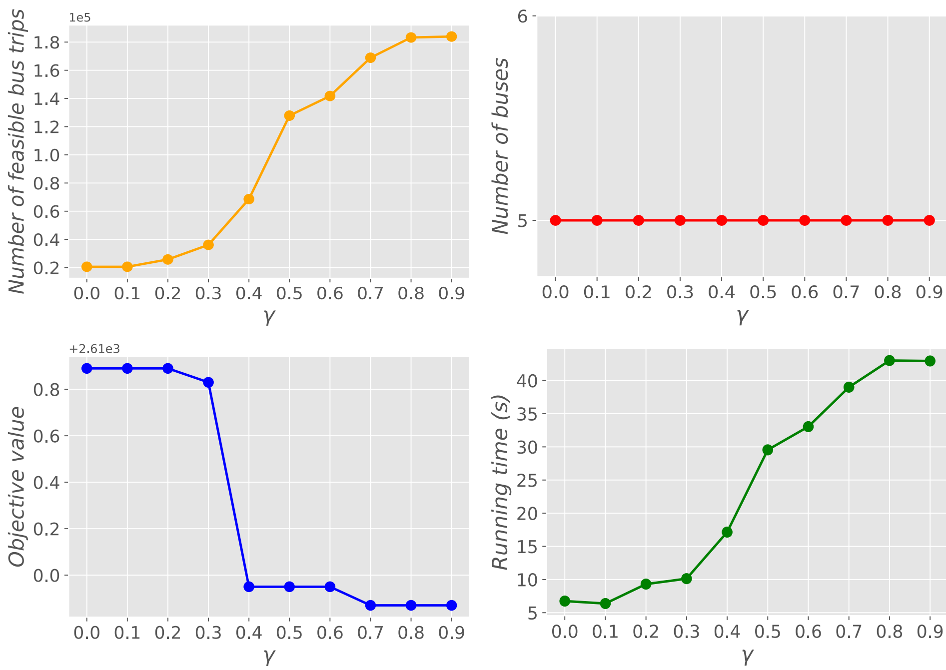


Figure 5.3: Sensitivity analysis for γ

CHAPTER 6

DISCUSSION

This thesis proposes a shareability network based approach for solving large-scale SBRP considering alternate modes. We build the theoretical connections between SBRP and the weighted set covering problem by decomposition through the shareability network. Moreover, we present a node compression technique and heuristic edge compression techniques to obtain a simplified ILP and enable tractable computation of the SBRP at-scale. The node compression technique uses an ILP to generate bus stops while satisfying maximum walking constraints for all students and decreases the number of nodes in the shareability network. In contrast, the edge compression techniques are heuristics that are applied to reduce the density of the shareability network, and work well in practice. We evaluate our solution using synthetic data-sets from BPS and show that our approach can compute decent solutions for large-scale problems. To further evaluate the performance for our approach, two benchmark testings are conducted with state-of-art SBRP sovling techniques. Finally, a sensitivity analysis of the parameters used for network compression techniques is provided to get insights into how they influence the trade-off between solution quality and computation time.

This work is the first attempt to adapt techniques from the ridepooling literature on the shareability networks to the SBRP, consider alternate modes in SBRP and connect SBRP with the weighted set covering problem. The key extension for enabling these techniques to work in practice for large-scale instances for the SBRP is the compression of the shareability network. One important future direction is to develop more sophisticated and nuanced edge compression schemes that more precisely target the edges that are unlikely to be relevant to the optimal solution. Moreover, the simplified ILP we present can be combined with state-of-art ILP solving techniques (e.g. column generation) to solve the extreme large-scale SBRP instances or allow for larger β values. This is also an area to be explored.

APPENDIX A

HEURISTIC INSERTION PATH-TSP SOLVER

One of the most time-consuming parts in our decomposition method through the shareability network is generating the bus trip list T_b. In Algorithm 1, we need to solve a path-TSP every time we find a clique in the shareability network to determine the feasibility of this trip. The path-TSP itself is NP-hard and it takes an off-the-shelf solver seconds to solve even with small amount of students as the input. The problem becomes intractable if we have to check millions of cliques in the Algorithm 1. Therefore, we propose a heuristic insertion path-TSP solving technique to decrease the computation time while outputting the satisfying feasible bus trip list T_b.

The essential idea of this heuristic insertion path-TSP solving technique is that if we know the optimal path p^* for a set of k students S_k, we generate a sub-optimal path for $k+1$ students $S_k \cup \{s\}$ by fixing p^* and inserting the student s into the order of p^* where yields a path with the minimal travel time. To be specific, we modified the function FEASIBILITYCHECK(τ, t_max, C, PATHTSP($\cdot$)) to Algorithm 3. For the input of Algorithm 3, we need the optimal path and travel time for a feasible trip τ. We can store the optimal routes and travel time once we generate a feasible trip τ with k students, which will be used when considering trips with $k+1$ students.

Algorithm 3 yields a sub-optimal route and travel time within linear time. The experiment results in the chapter 5 show the routes computed by Algorithm 3 are decent.

Algorithm 3 Trips feasibility check with the heuristic inserting path-TSP solving technique. Input: a feasible bus trip $\tau \in T_b$, the optimal route p^*_τ for the trip τ, the optimal travel time t^*_τ for the trip τ, a student s, maximum travel time t^{max}, bus capacity C, travel time function between any two vertices in the road network $\{t_{ij} | \forall i, j \in V_r\}$

.

```
function HeuristicFeasibilityCheck(τ, p*_τ, t*_τ, s, t^max, C, {t_ij | ∀i, j ∈ V_r})
  n ← |S(τ)|
  p*_τ := [s_1, s_2, ..., s_n]
  t* ← ∞, p* ← ∅
  for i in [0, 1, ..., n] do
    if i = 0 then
      t' ← t_{v_s v_{s_1}} + t*_τ
      if t' < t* then
        t* ← t'; p* ← [s, s_1, s_2, ..., s_n]
    else if i = n then
      t' ← t_{v_{s_n} v_s} + t_{v_s v_d} − t_{v_{s_n} v_d} + t*_τ
      if t' < t* then
        t* ← t'; p* ← [s_1, s_2, ..., s_n, s]
    else
      t' ← t_{v_{s_i} v_s} + t_{v_s v_{s_{i+1}}} − t_{v_{s_i} v_{s_{i+1}}} + t*_τ
      if t' < t* then
        t* ← t'; p* ← [s_1, ..., s_i, s, s_{i+1}, ..., s_n]
  if t* ≤ t^max and |S(τ)| + 1 ≤ C then
    return true
  else
    return false
```

APPENDIX B

REAL-WORLD EXPERIMENT RESULTS

This section shows the optimal school bus schedules for the real-world experiments with simulated data from Boston Public School (BPS). For figures in the following, red stars denote the school location, blue dots represent student locations and red crosses indicate students who take dedicated vehicles.

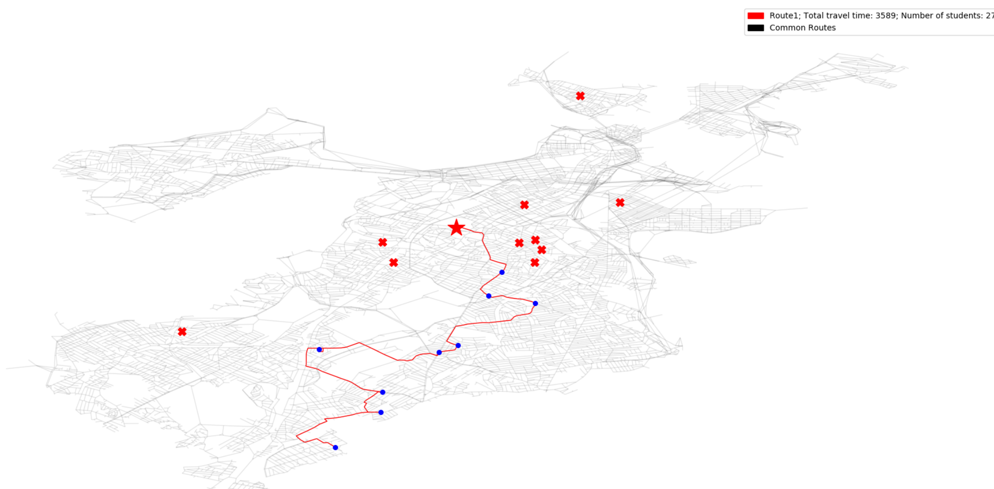


Figure B.1: Optimal school bus schedules for Tommy Harper

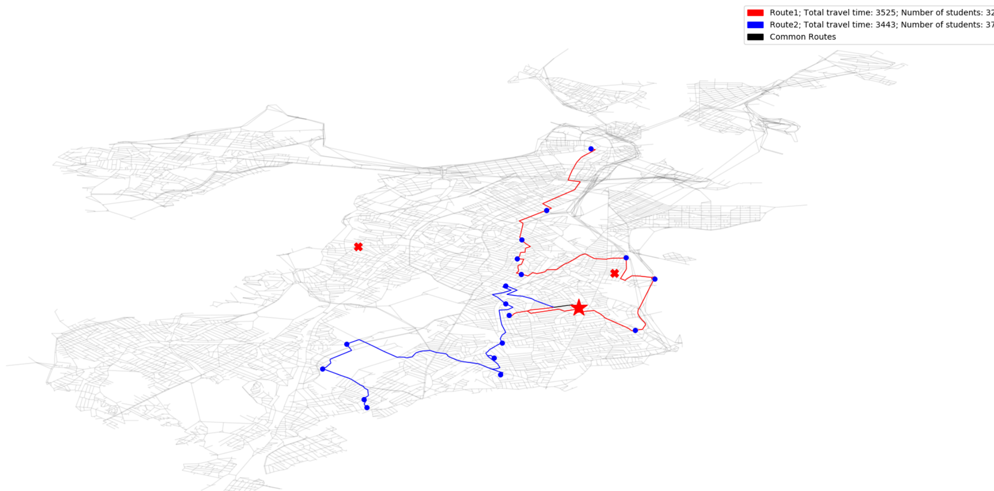


Figure B.2: Optimal school bus schedules for Craig Kimbrel

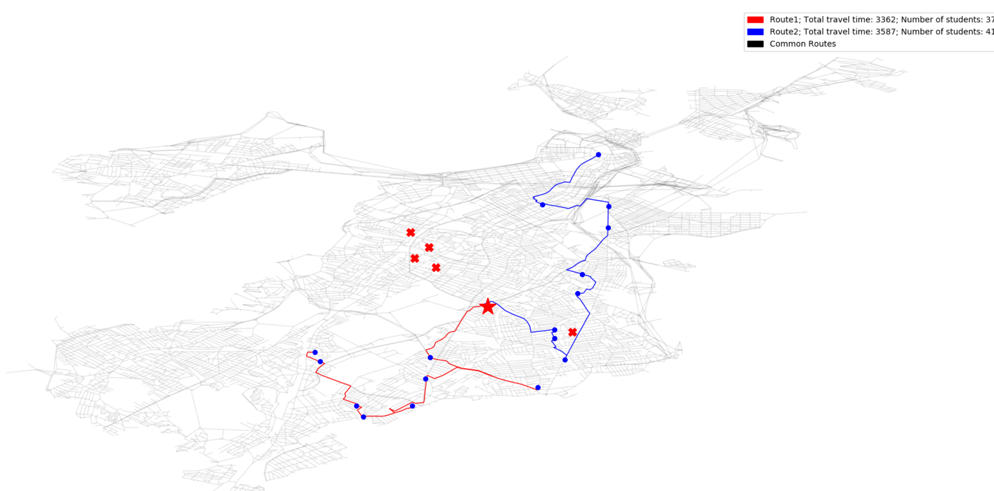


Figure B.3: Optimal school bus schedules for Deven Marrero

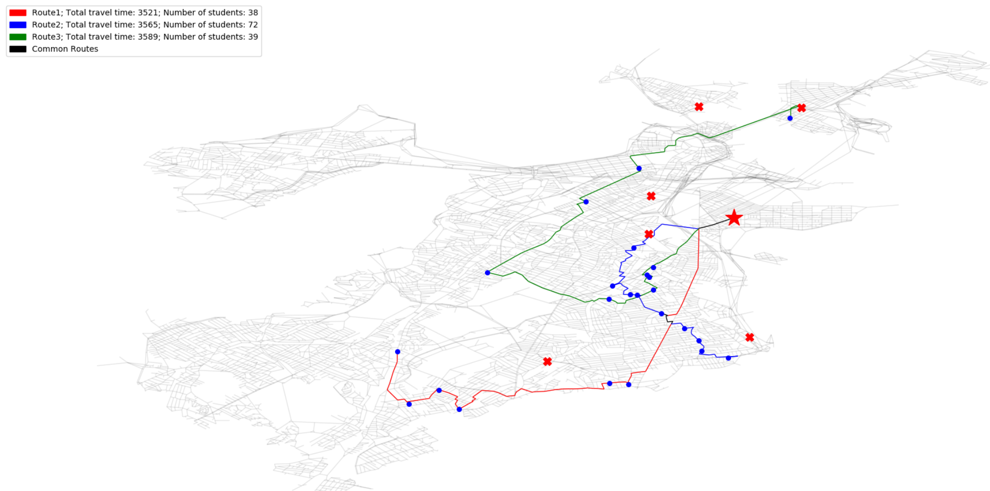


Figure B.4: Optimal school bus schedules for Frank Malzone

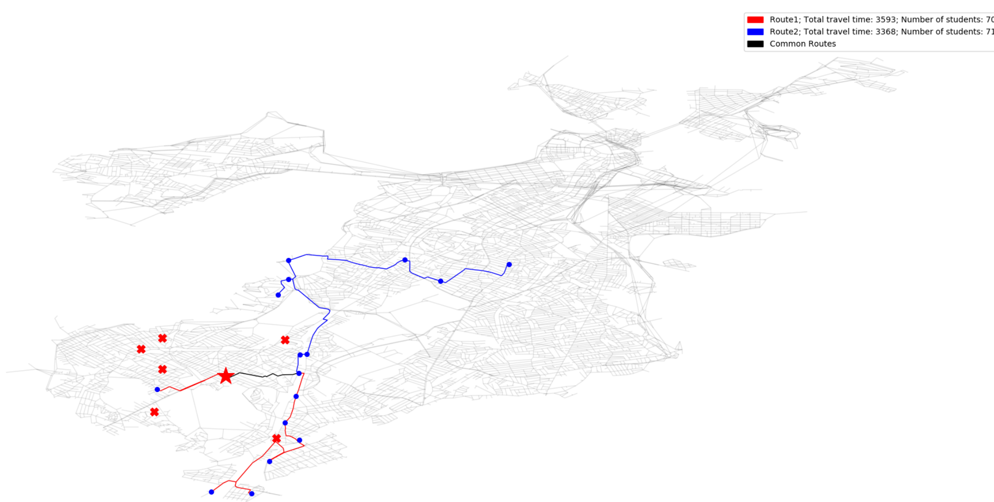


Figure B.5: Optimal school bus schedules for Dick Williams

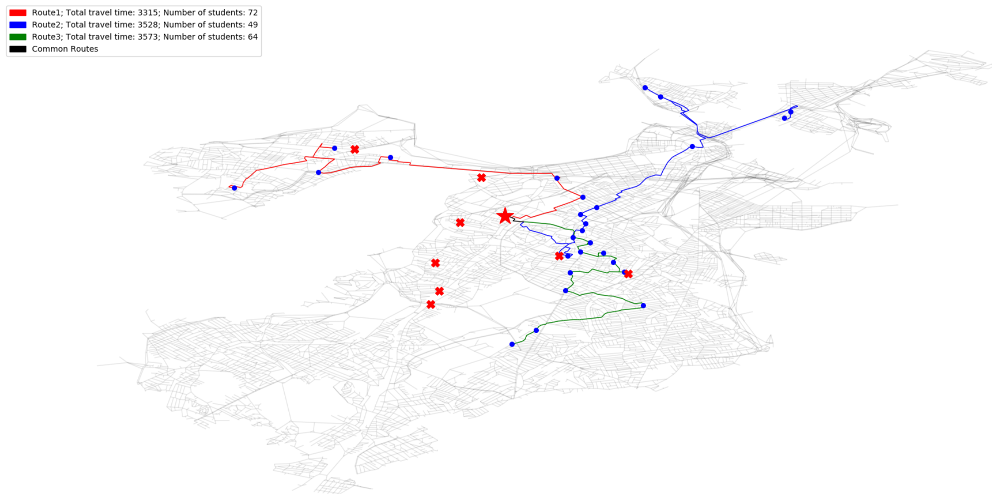


Figure B.6: Optimal school bus schedules for Dick Bresciani

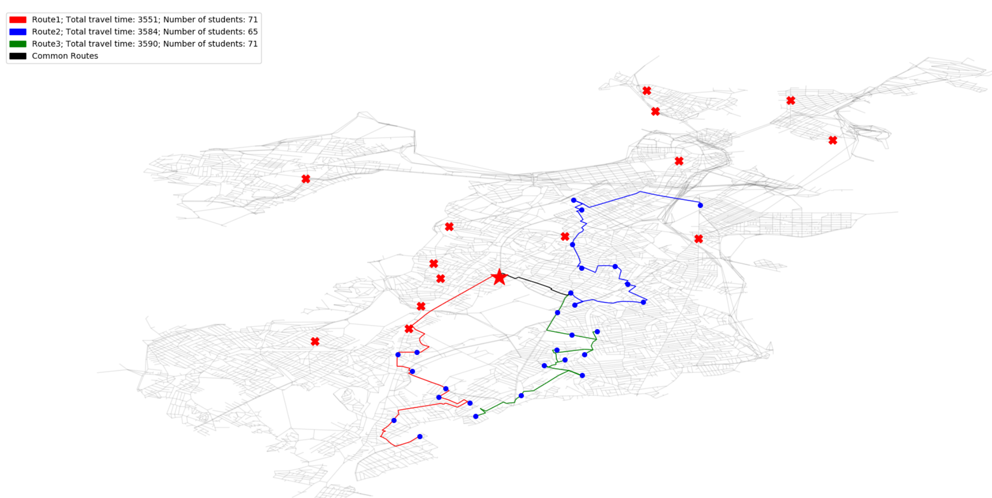


Figure B.7: Optimal school bus schedules for Dutch Leonard

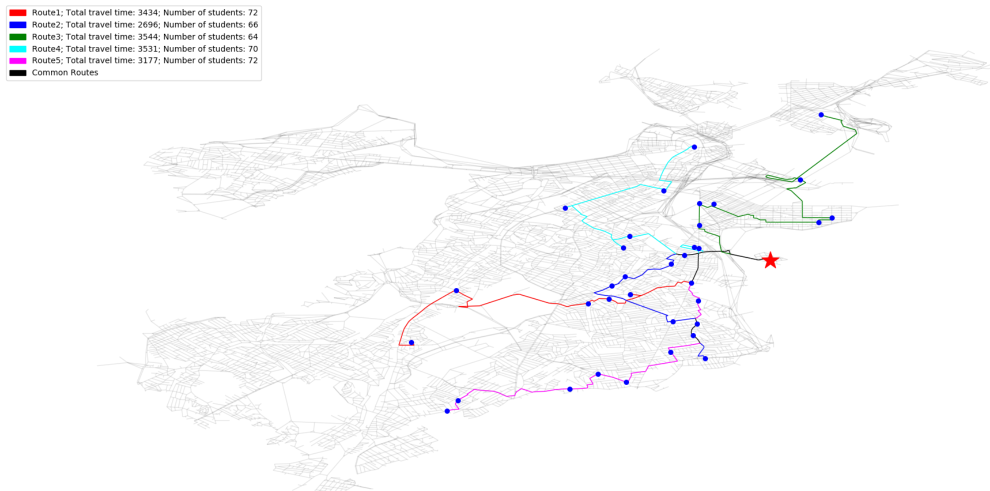


Figure B.8: Optimal school bus schedules for Christian Vazquez

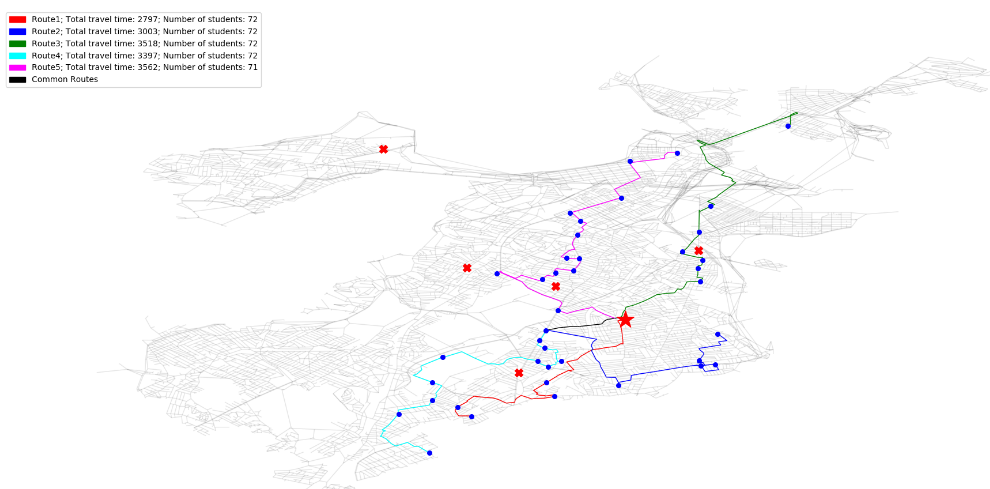


Figure B.9: Optimal school bus schedules for Dennis Eckerley

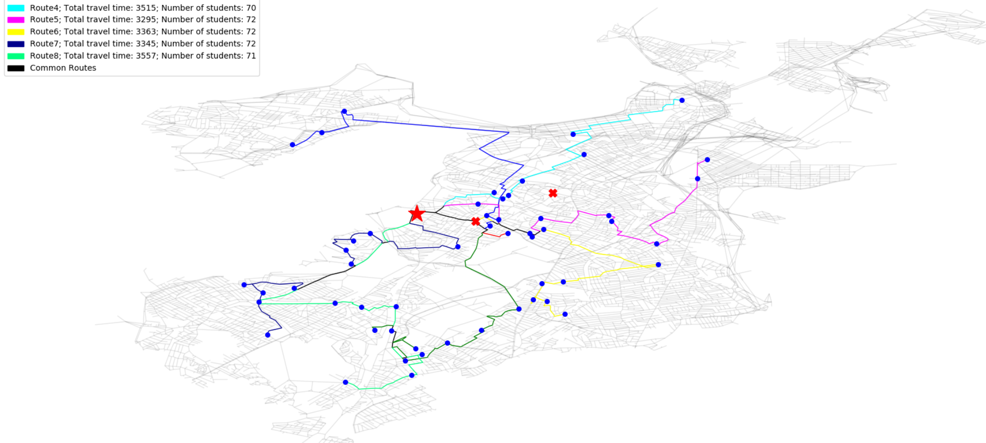
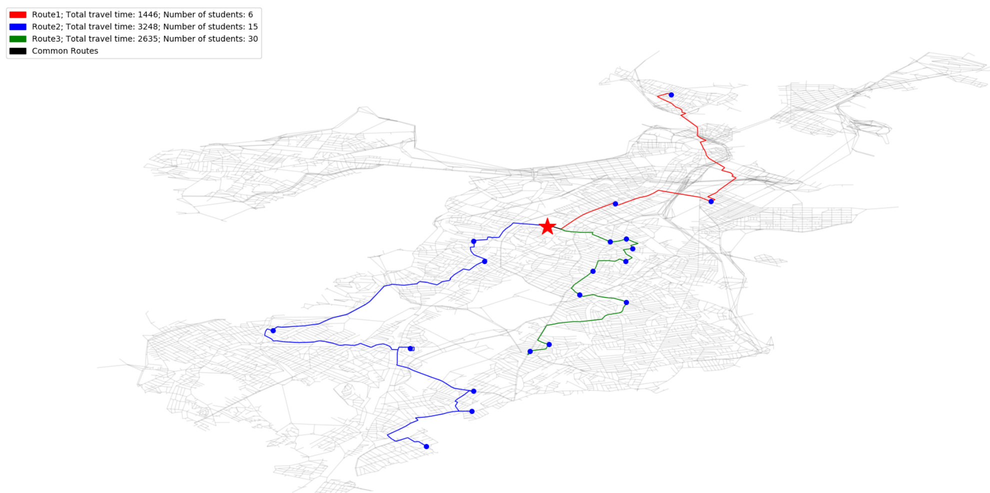


Figure B.10: Optimal school bus schedules for Rick Ferrell

APPENDIX C

BENCHMARK RESULTS

This section states the school bus schedules comparisons between the shareability network based decomposition approach and BiRD algorithm. For figures in the following, red stars represent school locations and blue dots denote student locations.



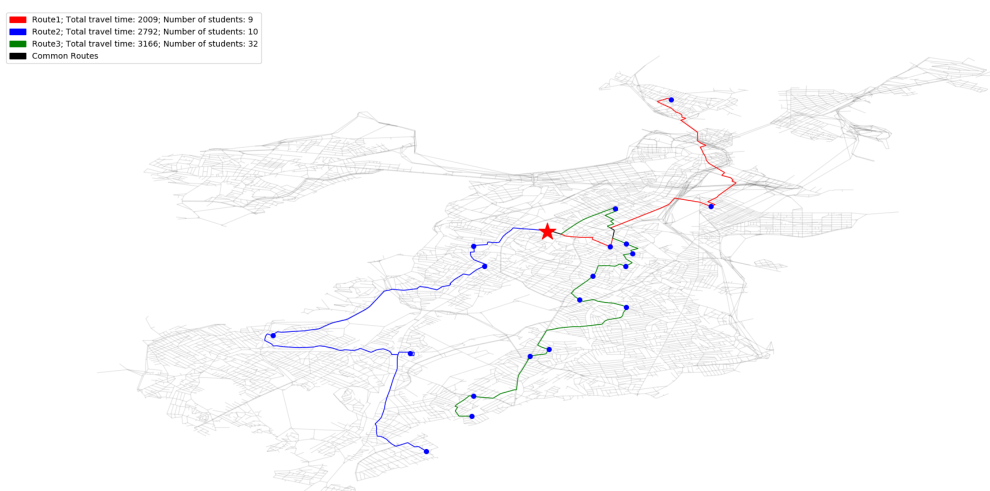
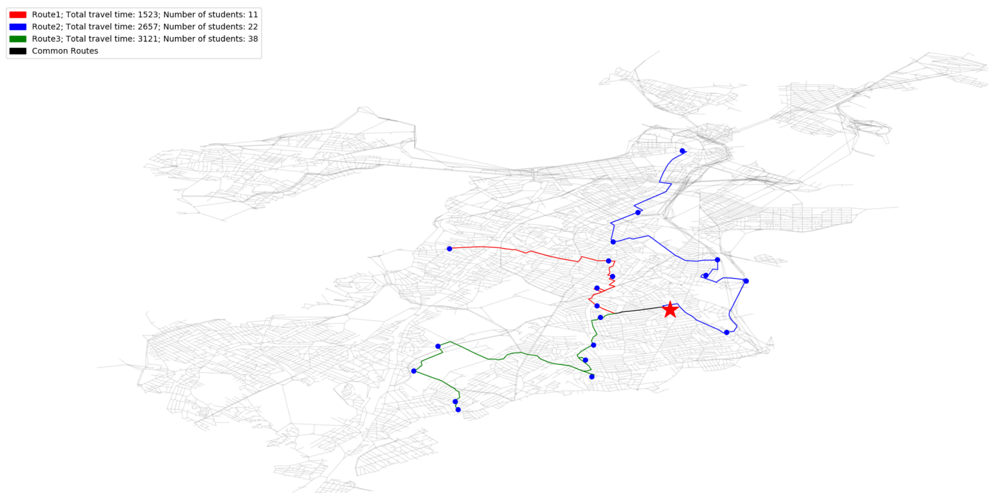


Figure C.1: Bus routes comparison for Tommy Harper (SND on the left)



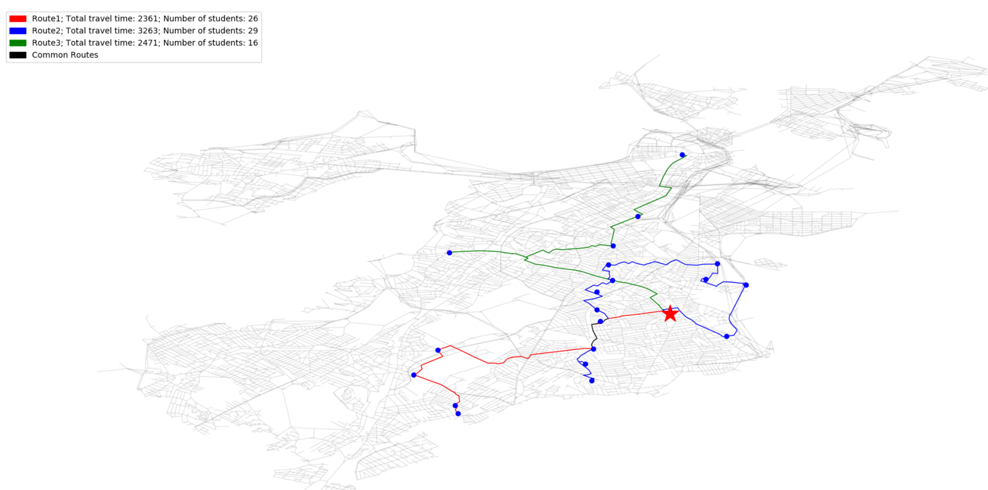
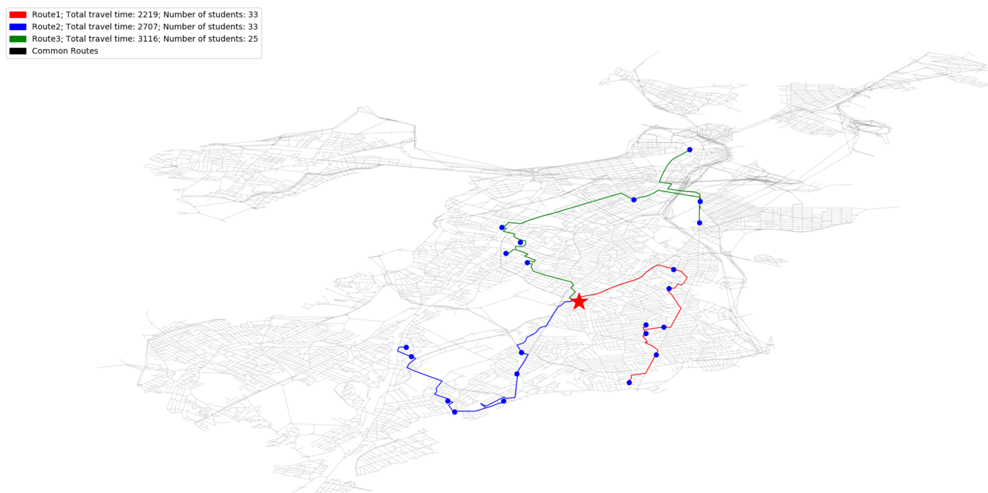


Figure C.2: Bus routes comparison for Craig Kimbrel (SND on the left)



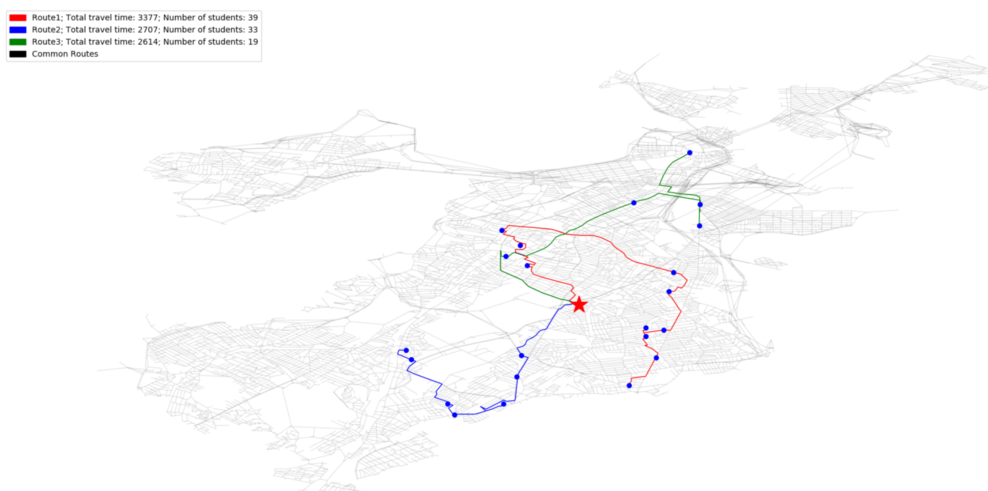
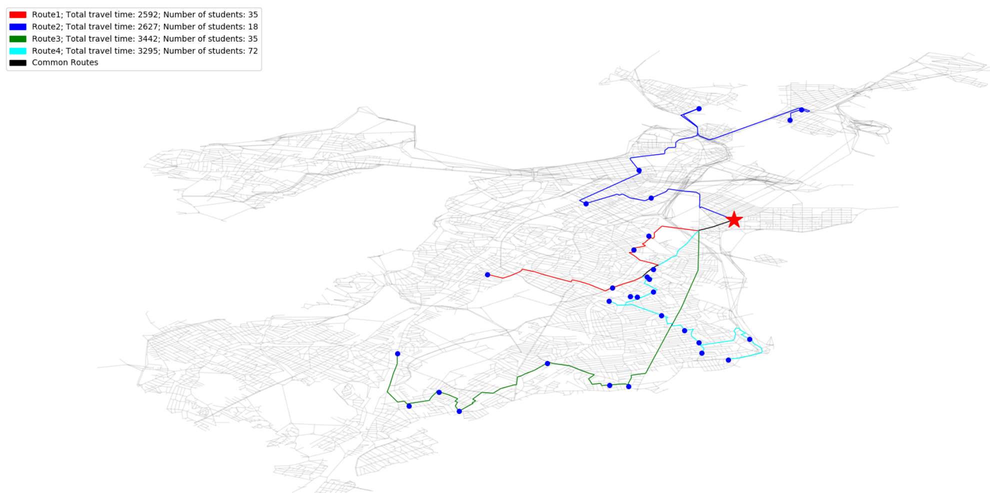


Figure C.3: Bus routes comparison for Deven Marrero (SND on the left)



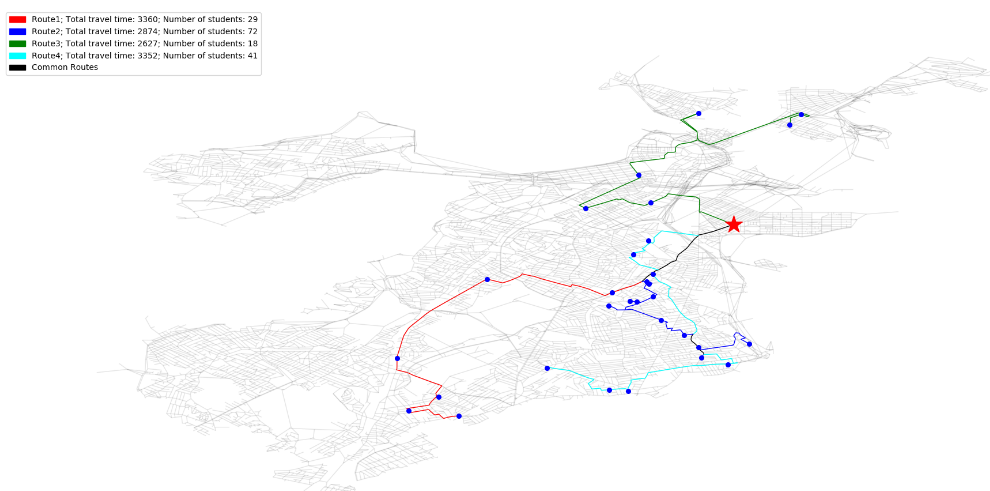
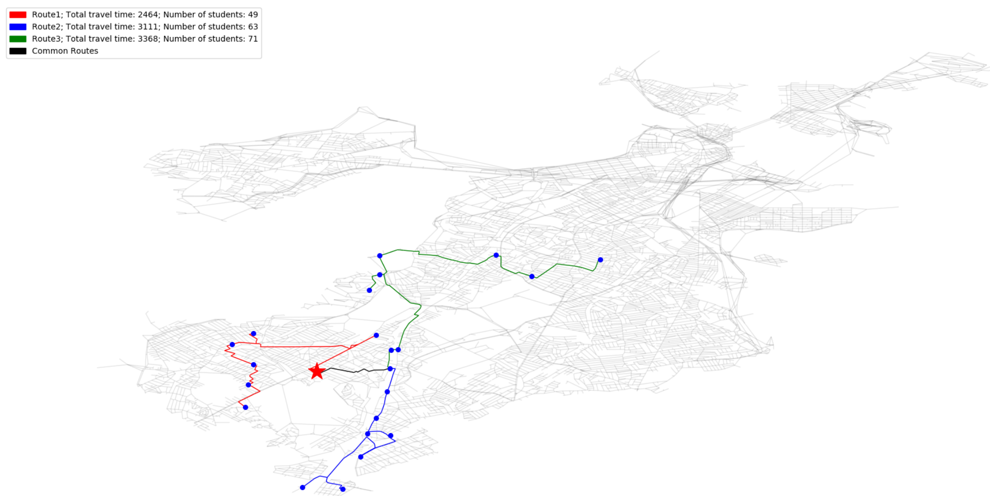


Figure C.4: Bus routes comparison for Frank Malzone (SND on the left)



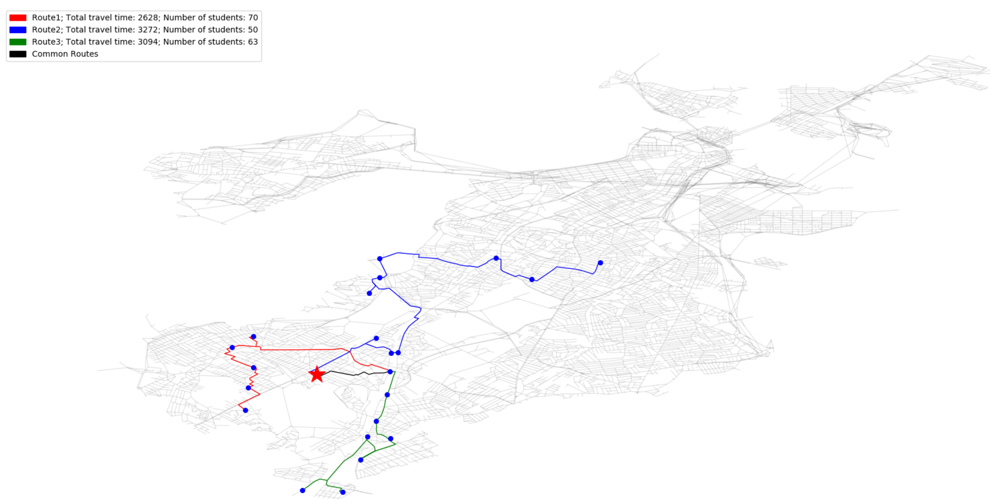
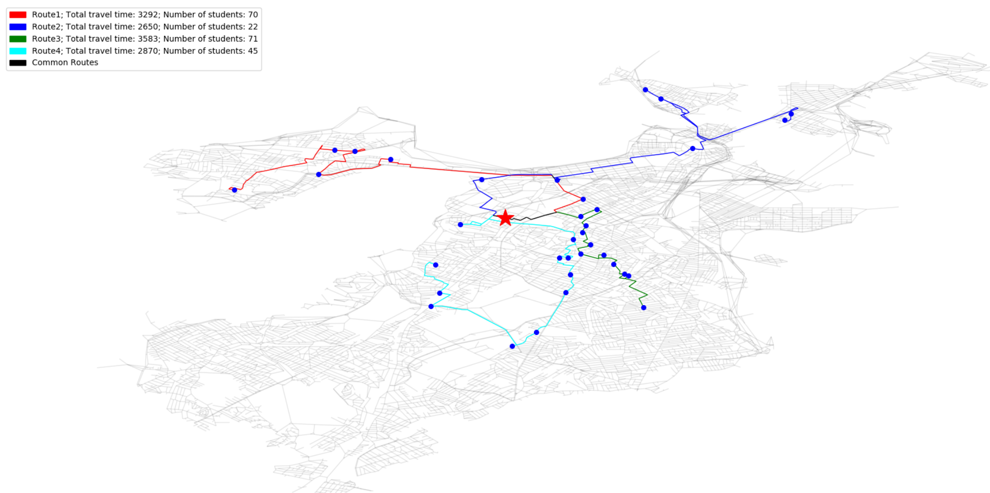


Figure C.5: Bus routes comparison for Dick Williams (SND on the left)



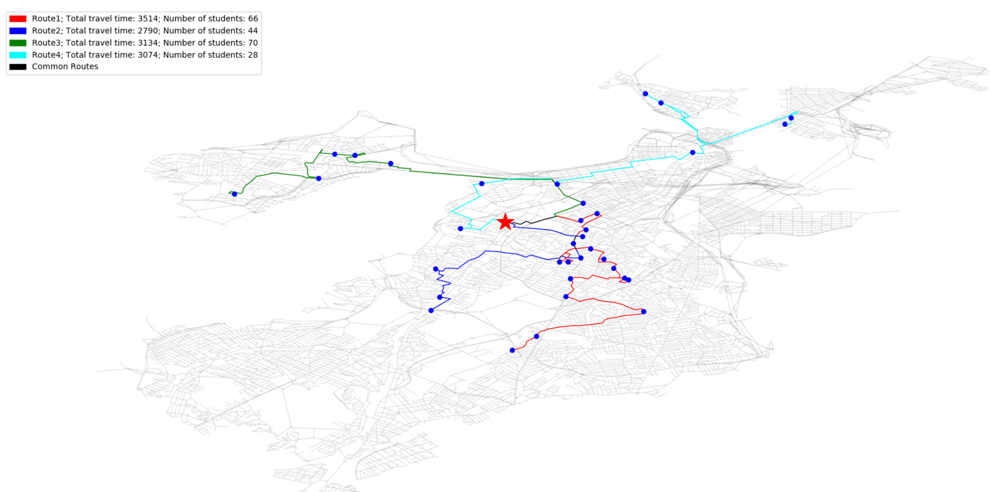


Figure C.6: Bus routes comparison for Dick Bresciani (SND on the left)

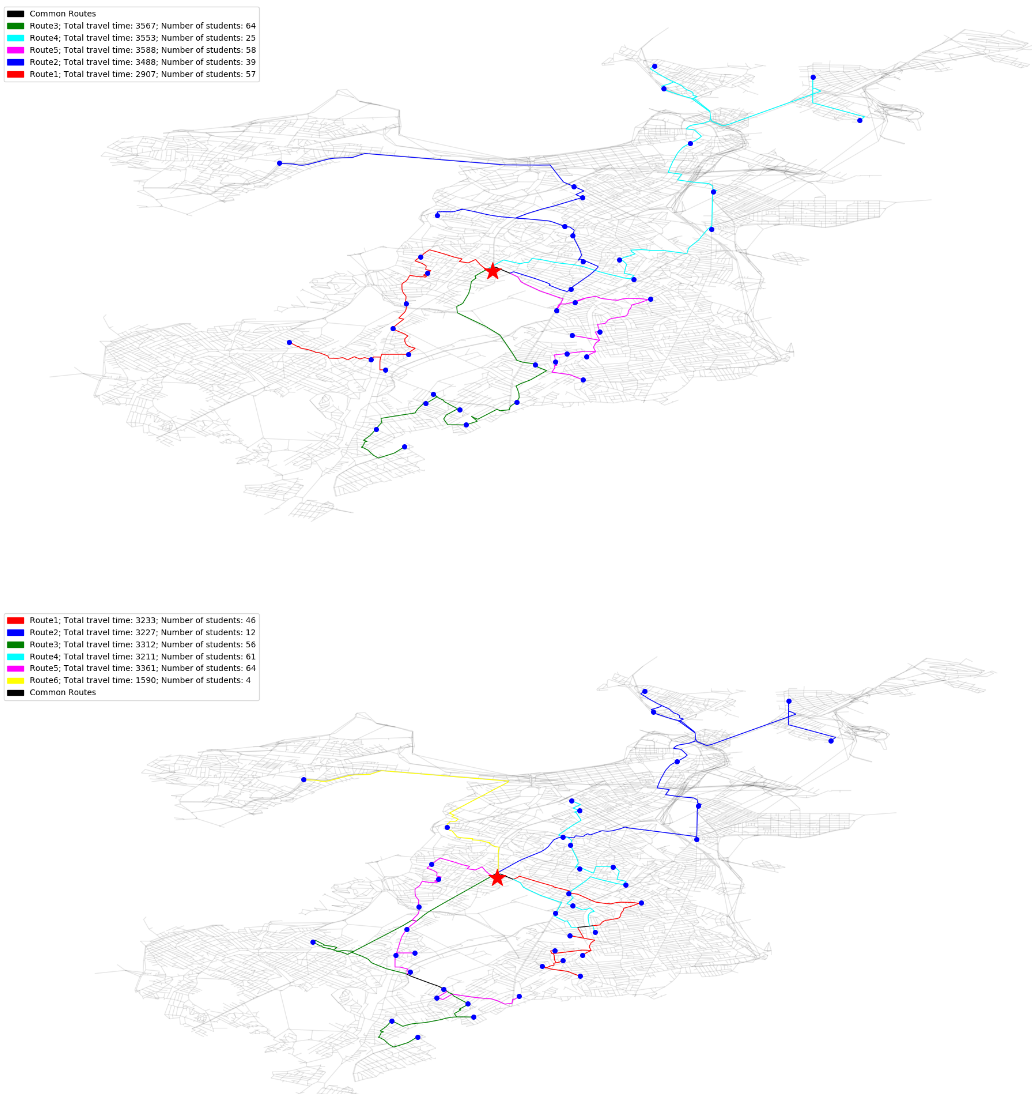
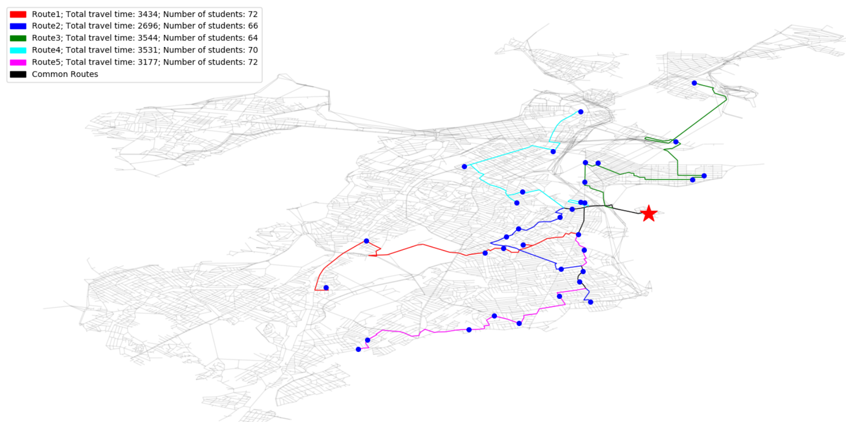


Figure C.7: Bus routes comparison for Dutch Leonard (SND on the left)



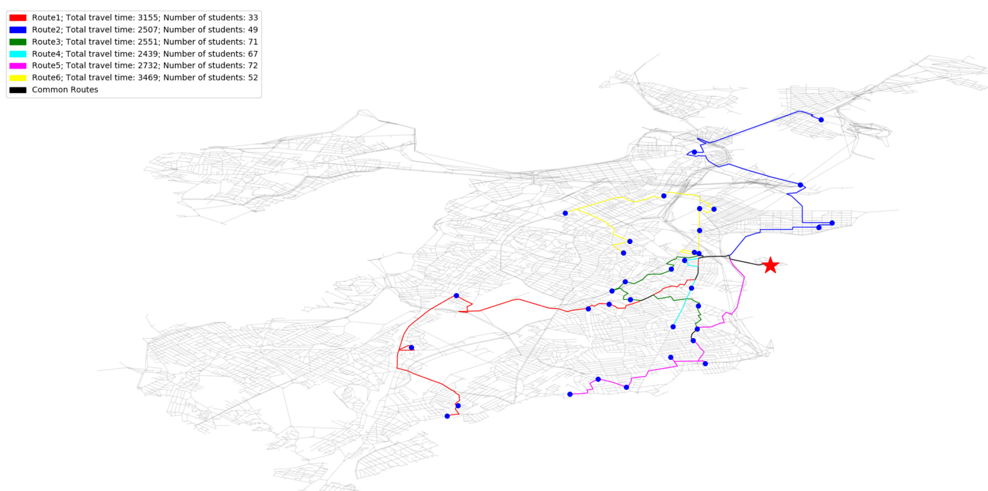
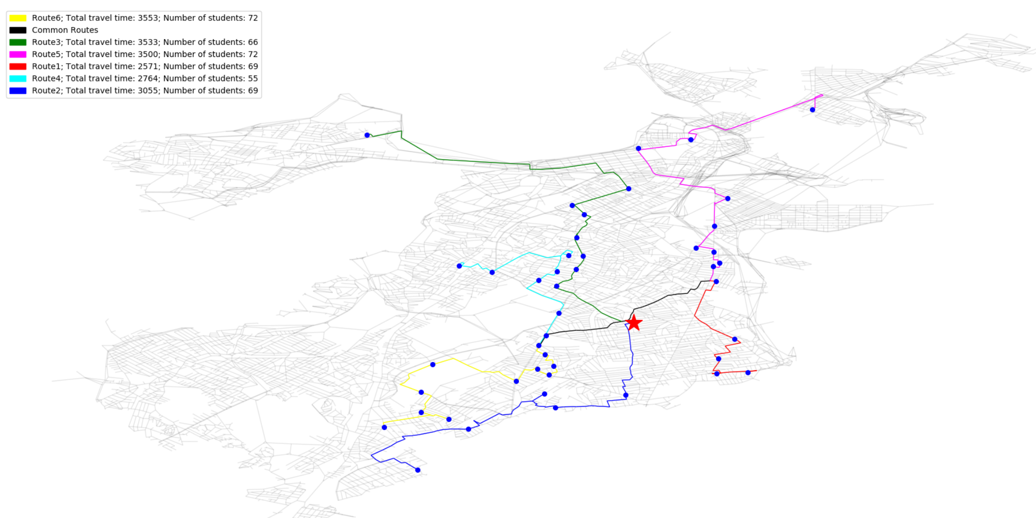


Figure C.8: Bus routes comparison for Christian Vazquez (SND on the left)



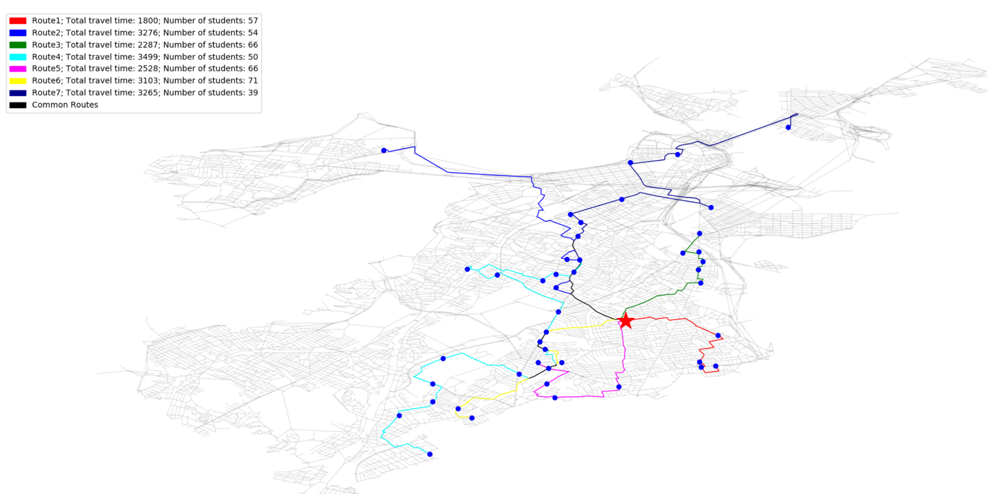
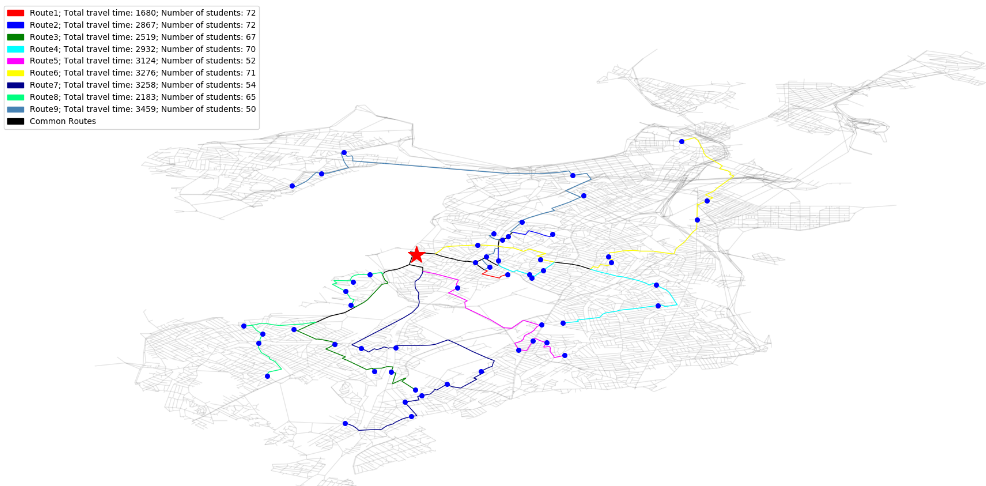


Figure C.9: Bus routes comparison for Dennis Eckerley (SND on the left)



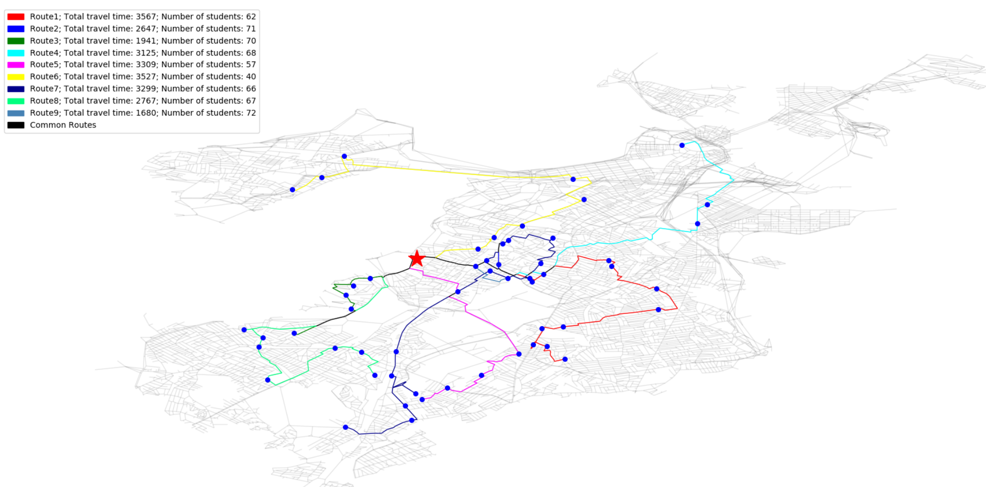


Figure C.10: Bus routes comparison for Rick Ferrell (SND on the left)

BIBLIOGRAPHY

[1] Alleghany county public schools. School bus per-mile operating cost analysis, 2016.

[2] Javier Alonso-Mora, Samitha Samaranayake, Alex Wallar, Emilio Frazzoli, and Daniela Rus. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, 114(3):462–467, 2017.

[3] American School Bus Council. Environmental benefits. `http://www.americanschoolbuscouncil.org/issues/environmental-benefits`, 2013.

[4] T. Bektaş and Seda Elmastaş. Solving school bus routing problems through integer programming. *Journal of the Operational Research Society*, 58(12):1599–1604, Dec 2007.

[5] Dimitris Bertsimas, Arthur Delarue, and Sebastien Martin. Optimizing schools' start time and bus routes. *Proceedings of the National Academy of Sciences*, 116(13):5943–5948, 2019.

[6] Geoff Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126 – 139, 2017.

[7] Jeffrey Braca, Julien Bramel, Bruce Posner, and David Simchi-levi. A computerized approach to the new york city school bus routing problem. *IIE Transactions*, 29(8):693–702, Aug 1997.

[8] Gurobi Optimization, Inc. Gurobi optimizer reference manual, 2016.

[9] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using networkx. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA, 2008.

[10] Rita M. Newton and Warren H. Thomas. Design of school bus routes by computer. *Socio-Economic Planning Sciences*, 3(1):75 – 85, 1969.

[11] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org . `https://www.openstreetmap.org`, 2017.

[12] Junhyuk Park and Byung-In Kim. The school bus routing problem: A review. *European Journal of Operational Research*, 202(2):311 – 319, 2010.

[13] Junhyuk Park, Hyunchul Tae, and Byung-In Kim. A post-improvement procedure for the mixed load school bus routing problem. *European Journal of Operational Research*, 217(1):204 – 213, 2012.

[14] Ridester Staff. The comprehensive guide: How much does uber cost? `https://www.ridester.com/uber-rates-cost`, 2018.

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

[16] Paolo Santi, Giovanni Resta, Michael Szell, Stanislav Sobolevsky, Steven H. Strogatz, and Carlo Ratti. Quantifying the benefits of vehicle pooling with shareability networks. *Proceedings of the National Academy of Sciences*, 111(37):13290–13294, 2014.

[17] Patrick Schittekat, Joris Kinable, Kenneth Srensen, Marc Sevaux, Frits Spieksma, and Johan Springael. A metaheuristic for the school bus routing problem with bus stop selection. *European Journal of Operational Research*, 229(2):518 – 528, 2013.

[18] Ali Shafahi, Zhongxiang Wang, and Ali Haghani. Solving the school bus routing problem by maximizing trip compatibility. *Transportation Research Record: Journal of the Transportation Research Board*, 2667:17–27, 2017.

[19] Paolo Toth and Daniele Vigo. *The vehicle routing problem*. Society for Industrial and Applied Mathematics, Philadelphia, 2002.

[20] U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. *Digest of Education Statistics: 2016*, 2017.