Mobility21 Final Research Report:
Joint Optimization of School Bus Routes and Last-mile Services

Peter Zhang

July 2021

Contents

1 Executive Summary

According to the Shared Transportation Guide of Allegheny County, per state law, elementary school students in Pennsylvania may need to walk up to a mile and a half to a school bus stop, and secondary school students may need to walk up to two miles. Such first- and last-segments of students' daily commute create a "first and last-mile" problem (or last-mile for short) for students' daily travel. The unavailability of first-mile and last-mile transportation service is one of the main deterrents to the use of school buses for certain families, and also a primary cause for safety and security concerns to parents, especially during bad weather, early morning, and evening time. Although the selection of bus lines and stops take into account the pick-up and drop-off locations for each student, the walking distance to the bus stop for some students will be still long under cost consideration given a limited fleet of school buses. In other words, the "last-mile" problem is always serious for some students.

With the rapid development and popularization of mobile and wireless communication technologies, ride-sourcing companies, such as Uber, Lyft, and Z-Trip, have been able to leverage internet-based platforms to operate transportation services with flexible routes and schedules. These companies can connect passengers and vehicles in real-time. We propose to explore the feasibility and actionable operational plans to use ride-sourcing vehicles as feeding last-mile services to the school bus. With integrated optimization models to consider both school bus routing and stops selection, as well as ride-sourcing feeding shared routes and schedules, we may provide more efficient and safer transportation technologies to students. To this end, we provide a market analysis to understand the landscape of ride-sourcing services specifically designed to transport students in the United States, via review of publicly available information such as company data, school district contract with ride-sourcing companies, and also private interviews with a leading company in the field. In addition, we also include reference to a technical paper that was created with the support of this award, focusing on the allocation of vehicle fleet across multiple service regions when demand is highly uncertain.

A second stream of inquiry is on the potential to reduce transportation cost by creating shared transportation routes between schools. While the cost-saving aspect is conceptually straightforward, it is unclear whether such collaborations would become feasible. From the transportation providers' point of view, merging transportation services may lead to the termination of contracts between some schools and providers. From the schools' point of view, it is unclear how to attribute the costs and benefits among a transportation coalition. Computing the optimal routes under different scenarios is already a computationally challenging task, and creating the best collaboration plan adds another layer of complexity. To this end, we provide a tractable method to check the feasibility of route sharing among any coalition of schools, assuming every school would be responsible for the costs proportional to the ratio of their own students on each (shared) route. We do so by leveraging classical results in transportation research, concepts from collaborative game theory, and computational methods from machine learning.

Our findings are as follows.

1. Ride-sourcing for students is an emerging market. Due to both cost and regulation reasons, it is uncommon for schools to outsource student transportation at a large scale to student ride-sourcing companies. But schools are actively seeking help from these companies to solve problems related to unexpected demand changes, and special needs transportation.

2. Demand uncertainty can add complexity to and change the optimal vehicle fleet allocation decisions for last-mile transportation. We design an analytical decision tool for vehicle fleet allocation under uncertainty across multiple service regions.

3. We call a shared route for a coalition of schools *feasible* if the route (and cost sharing plan) makes everyone better off than they are without sharing. Shared transportation presents a significant cost-saving opportunity, and we design an analytical decision tool that can quickly screen the feasibility of shared routes between any coalition of schools.

4. Route sharing also creates many issues that cannot be solved by computational methods,

and it is less clear how such issues should be attributed to each school on that level. Because shared services involve families, drivers, and administrators that are each responsible for different aspects of the problem.

We organize this report as follows.

- In Section 3, we provide a review of current practices and literature, to show how students are transported in the United States.

- In Section 4, we list the academic and policy outcomes of effort supported by this award.

- In Section 5, we give details to a route sharing and cost sharing decision-support tool.

2   Project Description

School bus routing across multiple districts requires the coordination of many resources. In this project, we plan to understand the challenges in current school bus routing in Allegheny County, Pennsylvania, and improve the mobility of students via decision analytical models and deployment of information technology tools. Our main focus is to find efficient, safe, and implementable multi-modal transportation methods to help children move to and from schools. Last-mile and first-mile transportation is an existing problem in this context, due to the 1-2 mile walking distance between home and bus stops for some students, which poses safety and accessibility concerns. Our goal is to design an analytical model to support the joint optimization of bus routes and last-mile and first-mile mobility services for students, and provide a deployment technology that can help connect students, parents, buses, and ridesharing services to achieve reliable transportation. In addition, we also provide an analytical tool, to efficiently check the feasibility of route sharing between schools and school districts.

According to the Shared Transportation Guide of Allegheny County, per state law, elementary school students in Pennsylvania may need to walk up to a mile and a half to a school bus stop, and secondary school students may need to walk up to two miles. Such first- and last-segments of students' daily commute create a "first and last-mile" problem (or last-mile for short) for students' daily travel. The unavailability of first-mile and last-mile transportation service is one of the main deterrents to the use of school buses for certain families, and also a primary cause for safety and security concerns to parents, especially during bad weather, early morning, and evening time. Although the selection of bus lines and stops take into account the pick-up and drop-off locations for each student, the walking distance to the bus stop for some students will be still long under cost consideration given a limited fleet of school buses. In other words, the "last-mile" problem is always serious for some students.

With the rapid development and popularization of mobile and wireless communication technologies, ride-sourcing companies, such as Uber, Lyft, and Z-Trip, have been able to leverage internet-based platforms to operate transportation services with flexible routes and schedules. These companies can connect passengers and vehicles in real-time. We propose to explore the feasibility and actionable operational plans to use ride-sourcing vehicles as feeding last-mile services to the school bus. Specifically, a vehicle from Z-Trip with a pre-arranged route is able to bring several students from their home to the bus stop in one shared trip. Therefore, with integrated optimization models to consider both school bus routing and stops selection, as well as ride-sourcing feeding shared routes and schedules, we may provide more efficient and safer transportation technologies to students.

3 Literature and Practice: How Students Are Currently Transported in the U.S.

According to the 2017 National Household Travel Survey[1], 54.2% of students travel to/from school by personal vehicle, 33.2% by school bus, 1.4% by walk/bike, and 2.2% by transit and other means. That is about 480,000 school buses transporting 33% of students in the United States[2]. This is by far the transportation fleet with the most vehicles, more than three times the size of the second largest fleet of public transit, which has 140,000 vehicles.

According to the National Center for Education Statistics, the total spending for transportation has doubled from $12.6 million to $25.3 million from 1980 to 2016. Yet the share of students who receive such transportation service has declined. Overall, this translates to the average transportation cost per student increasing from $567 a year to $982 in 2016[3].

3.1 Transportation Modes

We give more background to three typical student transportation modes.

District Provided Services is the most common. Districts are in charge of all elements of school transportation.

The **advantages** of district provided services include

- Direct and complete control decision making;
- Ownership over buses;
- Possible investments for reducing transportation costs: invest in tech to improve efficiency.

These elements contribute to cost reduction and stability.

The **disadvantages** of district provided services include

- Reduced funding for education;
- Increased responsibility for students supports;
- Complex routing logistics;
- Lack of capital for buses.

These elements contribute to escalating costs and increased complexity of routing.

Contracted Services is the second most common mode – contracting with privately owned transportation provider for yellow bus service.

The **advantage** of contracted services is the potential for school districts to negotiate favorable contract terms for potential saving if the supply side of the market is competitive.

The **disadvantages** of contracted services include

- Possible extra costs not revealed in contracts;
- No ownership in fleets thus contractors can cancel any time;

---

[1] Children's Travel to School https://nhts.ornl.gov/assets/FHWA_NHTS_%20Brief_Traveltoschool_032519.pdf
[2] NCES Student Fleet https://nces.ed.gov/programs/digest/d18/tables/dt18_236.90.asp?current=yes
[3] NCES Transportation Fast Facts https://nces.ed.gov/fastfacts/display.asp?id=67

- No sufficient consideration over routing optimization.

Public Transit Services are more common in urban areas where public transit systems are robust and well developed.

The **advantages** of public transit for student transportation include

- Less transportation spending and demand for drivers;

- Increased attendance rates (*e.g.*, Baltimore)

The **disadvantages** of public transit for student transportation include

- Disctricts need to provide discounted passes for low-income families;

- Minimum age requirements for students to use such service;

- Additional costs for transporting special need students;

- Varying reliability of public transit options.

These three aforementioned modes are not mutually exclusive. For example, the Cincinnati Public School system uses a **hybrid mode**: school bus transportation is used for grades K-6, public transit is used for grades 7-12, and contractors are used on an ad hoc basis for additional needs exceeding capabilities.

3.2 Challenges

School districts and families need to take students safety and special needs students into full consideration. In addition, the escalating costs, complex education system, and complex routing logistics lead to reduction in service quality and less frequent system upgrades. We summarize challenges faced by stakeholders in this system.

Parents face challenges transporting their children. According to a survey by the student last-mile transportation company HopSkipDrive Parent Survey[4], 38% of parents spend more than 5 hours per week driving their children to and from schools; 67% say driving their children disrupts their work on a regular basis.

Unequal access to schools due to transportation limitations. Lower performing schools tend to locate in neighborhoods where low income students live. These students need to spend more time and money on transportation to schools for better education and school choice[5].

Driver Shortage. Districts and contractors struggle to hire bus drivers, due to low school bus driver wages and job responsibilities. According to the School Bus Fleet Survey and the U.S. Bureau of Labor Statistics[6]:

- School Bus Drivers

  - Average Wage: $16.05 / hr
  - Requirements: Commercial Drivers License (CDL); Drug and Alcohol Testing; Additional training before transporting Children(10-40 hours)

- Transit and Intercity Bus Drivers

---

[4]HopSkipDrive Survey https://www.hopskipdrive.com/blog/back-to-school
[5]Urban Institute https://www.urban.org/features/segregated-neighborhoods-segregated-schools
[6]https://fleetimages.bobitstudios.com/upload/_migratedeecms/files/stats/SBF1118-survey-webpress-lores.pdf

Table 1: Percentage of respondents experiencing school bus driver shortage.

Shortage Level	Districts	Contractors
Desperate	7%	5%
Sever	24%	20%
Moderate	36%	40%
Mild	24%	29%
None	9%	5%

- Average Wage: \$21.47 / hr
- Requirements: Commercial Drivers License (CDL)

- Heavy and Tractor-trailer Truck Drivers

  - Average Wage: >\$21.91 / hr
  - Requirements: Commercial Drivers License (CDL)

As a result, most districts and contract experience some level of school bus driver shortage (Table 1).

Additional challenges include the lack of ridership data and environmental concerns.

3.3 Emerging Market of Last-Mile Ride-Sourcing

We provide our findings about three main players in the emerging market of ride-sourcing for students. The three companies are HopSkipDrive, Zum, and Bubbl.

HopSkipDrive was launched in 2013 by Joanna McFarland.

- **Location**: They currently operate in Los Angeles, San Francisco, San Diego, Sacramento, Colorado, and Washington, D.C.

- **Target market**: Children between the ages six years old who are legally restricted to use other transportation network companies. (Homeless and foster kids).

- **Business model**: ToC: families with direct contacts; ToB: School and county contracts. (70% revenue channel)

- **Funding**: In 2016 raised 10.2 million in series A. In 2017, raised 7.4 Million. In 2020, raised 22 million from VCs.

- **Driver**: 90% females with five year or above child care experiences. Contractors.

- **Safety**: Duo Authentication Process, real time tracking, abnormality detectors.

- **Pricing**: no membership fee, charge per ride. Single family: \$1.25/estimated mile + \$0.50/actual min + \$3 booking fee. \$18 minimum fare. \$11 minimum fare base for two families, \$9 for three, and \$7 for four.

Zum

- **Location**: San Francisco, LA. San Diego, Miami, Phoenix, Dallas, Chicago and the D.C. area.

- **Target market**: reserve rides in advance for children of working families.

- **Channel**: Appointment the day before, not on demand.

- **Safety**: One stop tech platform – route optimization, vehicle and quality tracking, and real-time vehicle dashboards. Drivers go through Federal Bureau of Investigation and Department of Justice-level background checks and fingerprints. Drivers fully vetted with three or more year of experiences.

- **Business model**: ToB: 150 school districts, private schools (same funding source for school bus to pay for Zum). ToC: families appointments.

- **Funding**: Raised 44 million in series C.

- **Pricing**: $10 for carpool rides (per child for a one-way trip) and $19.50 for a single (non-carpool) ride. Monthly package also offered.

Bubbl was launched in 2016 by the former Dallas Deputy Chief of Police, Craig Miller. All drivers are retired or off-duty first-responders.

- **Location**: Dallas, Texas, Plano/Castle Hills, Texas, Frisco, Texas, Round Rock/Austin, Texas, Emerald Coast, Florida, Greenwich, Connecticut, McLean, Virginia.

- **Target market**: children over 8, seniors, passengers with special needs.

- **Drivers**: retired or off-duty first-responders. Fully vetted and trained.

- **Channel**: Mobile App, phone calls and website appointment.

- **Business model**: ToC.

4   Resulting Publications, Policy Reports, and Data

Publications and Reports With the support of this award, we developed several analytical models for optimizing transportation efficiency. They can be seen as general-purpose tools that focus on:

1. Vehicle fleet allocation across multiple service regions for last-mile transportation (*Fleet Sizing and Allocation for On-demand Last-Mile Transportation Systems*, draft available at http://www.optimization-online.org/DB_HTML/2020/10/8080.html).

   - **Summary** The last-mile problem refers to the provision of travel service from the nearest public transportation node to home or other destination. Last-Mile Transportation Systems (LMTS), which have recently emerged, provide on-demand shared transportation. In this paper, we investigate the fleet sizing and allocation problem for the on-demand LMTS. Specifically, we consider the perspective of a last-mile service provider who wants to determine the number of servicing vehicles to allocate to multiple last-mile service regions in a particular city. In each service region, passengers demanding last-mile services arrive in batches, and allocated vehicles deliver passengers to their final destinations. The passenger demand (i.e., the size of each batch of passengers) is random and hard to predict in advance, especially with limited data during the planning process. The quality of fleet-allocation decisions is a function of vehicle fixed cost plus a weighted sum of passenger's waiting time before boarding a vehicle and in-vehicle riding time. We propose and analyze two models—a stochastic programming model and a distributionally robust optimization model—to solve the problem, assuming known and unknown distribution of the demand, respectively. We conduct extensive numerical experiments to evaluate the models and discuss insights and implications into the optimal fleet sizing and allocation for the on-demand LMTS under demand uncertainty.

2. Data collection and analysis pipeline for mapping household commute patterns and income levels (*Commuting Distances and Times under Different Transportation Modes for U.S. Households*. Draft and code available at https://github.com/peteryz/employment-od).

   - **Summary** Efficient and fair transportation planning creates opportunities and equity for jobs, health care, and education. Therefore, data consolidation for transportation systems provides basis for evidence based policies. In this work, we construct a dataset that documents home-to-job commuting time and distance information for the 100 most populated U.S. urban areas. Our dataset builds on the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Dataset U.S. Census Bureau (2018), which provides origin (home) and destination (job) location information for households. For these origin-destination (OD) pairs, we derive commuting time and distance information under different travel modes, each a combination of walking, public transit, and ridesourcing. We construct data under different modes so policymakers and researchers have information about alternatives and can perform what-if analysis. Towards the end of this data sheet, we document a sample use case to illustrate this goal.

3. Collaborative game theory model on the feasibility of route sharing and cost sharing among multiple schools (see Section 5 of this report).

The first two tools have been developed into academic papers currently under review. We refer the readers to the links above. For the third analytical tool, we describe the technical problem, model, and method developed in the next section.

   In addition, with the support of this award, we also developed:

4. A simulation model to understand how transportation patterns affect epidemic spread. The resulting paper, entitled *Travel Cadence and Epidemic Spread*, has been accepted for publication in the Proceedings of the Winter Simulation Conference (2021). The draft is available at https://arxiv.org/abs/2107.00203, and the code is available at https://github.com/lpstreitmatter/Travel-Cadence-Epidemic-Spread.

- **Summary** In this paper, we study how interactions between populations impact epidemic spread. We extend the classical SEIR model to include both integration-based disease transmission simulation and population flow. Our model differs from existing ones by having a more detailed representation of travel patterns, without losing tractability. This allows us to study the epidemic consequence of inter-regional travel with high fidelity. In particular, we define *travel cadence* as a two-dimensional measure of inter-regional travel, and show that both dimensions modulate epidemic spread. This technical insight leads to policy recommendations, pointing to a family of simple policy trajectories that can effectively curb epidemic spread while maintaining a basic level of mobility.

5. An analytical tool to model and prescribe shutdown and re-opening strategies. We produced a policy memo to the President of the Republic of Indonesia. Our recommendation was adopted to support the decision to gradually and cautiously re-open on July 25, 2021, after a one-month shutdown related to the surge of COVID-19 Delta variant. Analysis supporting the memo (data redacted) is available at https://github.com/peteryz/dynamic-SIR.

Data The data accompanying the report is *Commuting Distances and Times under Different Transportation Modes for U.S. Households* is available at https://github.com/peteryz/employment-od

5 School Bus Route and Cost Sharing: A Collaborative Game Theory Perspective

In the last three decades, there has been a significant increase in school transportation costs in U.S. The main reasons for this growth are the changes in student school choices, strict federal and state mandates, and high fuel prices. In addition, transportation fundings are being cut. To accommodate transportation needs, schools use complex routing policies that eventually result in long bus riding times for students and worse student experience. One possible solution to improve school bus transportation is collaboration between schools and school districts. In this project, we study the tactical design of school bus collaboration. We derive closed-form approximations as metrics to evaluate the performance of school bus collaboration, such as the total bus cost and students travel time under collaboration. We evaluate the conditions under which the benefits from collaboration among schools is worthwhile. We also analyze the impact of collaboration on the social welfare of students and find the conditions under which collaboration is the most beneficial. Overall, this analysis is in the spirit of a cooperative game, where we consider multiple levels of players (i.e., schools, school districts) with different objectives, and aim to design a solution that aligns with the objectives of different players at each level.

5.1 Background

In the United States, over 480,000 school buses are currently used to provide transportation to millions of students to and from schools everyday (Burgoyne-Allen et al., 2019). According to the Bellwether report (Burgoyne-Allen et al., 2019), the average per student transportation costs have increased by 73% since the 1980's to approximately $2000 per year. The main reasons for increased transportation costs are the growing vehicle and gas costs and more complex routing. Most of the school districts have school buses which are quite old and often malfunction, resulting in high maintenance costs. School districts can overcome this issue by replacing the old school buses with newer ones. However, reduction in federal funding to schools has made the situation worse. Moreover, there are different state and federal policy mandates which puts additional stress on the existing school bus routing system, leading to complex routing and thereby increasing the overall transportation cost for the school districts. Another change in the school choices have also negatively impacted the school transportation system. In recent decades, different school choices such as open enrollment, charter schools, private schools etc., have emerged, giving families multiple choices to choose from. As a consequence of having multiple choices, students end up traveling adjoining district's schools, resulting in more complex routing.

At the same time, significant overlaps in student distributions from different schools pose an opportunity for schools to collaborate on student transportation. The goal of this paper is to provide a general framework to quantify the benefits of school bus sharing, for both schools (in terms of total money saved) and students (in terms of total welfare). We also apply our framework to analyze a specific case with data from Allegheny County, Pittsburgh, Pennsylvania. In Allegheny County, the state law mandates that each school district provide transportation services to students residing in the district and going to public and charter schools within the district as well as charter/private schools that are within 10-mile radius from the district's furthest border. This 10 mile radius on an average adds 2-5 extra districts to be served in Allegheny County. However, the impact of the 10 mile radius can be worse depending on the area of a given school district. For instance, the Salisbury Township School district in Pennsylvania Lehigh Valley, is 11.3 square miles and has two elementary schools but it provides school transportation to 46 schools (AlliesForChildren, 2017). The school district spends approximately $2.5k annually per student. Figure 1 below shows 7 different school districts that lie within 10-mile radius from the selected charter school in Allegheny County. Thus, the selected yellow school districts need to provide bus transportation to students going to Environmental Charter School at Frick Park. The work done in this paper is in partnership with Allies for Children, Mobility21 and Buhl Foundation, who helped us get data from Allegheny County school districts.

While the aforementioned problems seems to be more associated with school district, there are
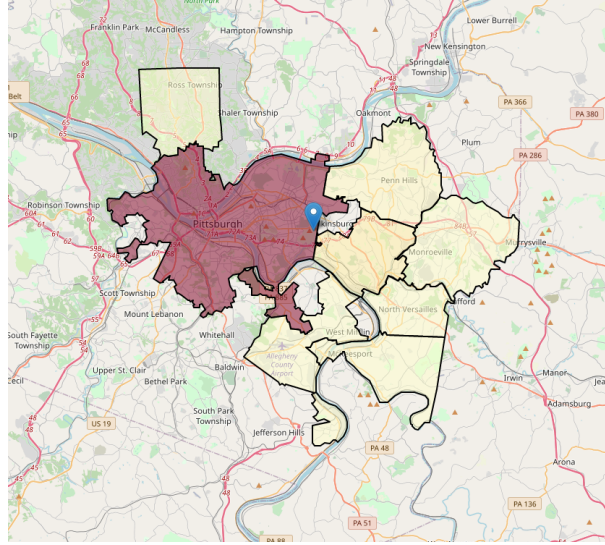
Figure 1: Enrolled Students Resident School districts data going to Environmental Charter School at Frick Park from Allegheny County. The blue marker represents the selected charter school. Red region is the school district in which the charter school is located. The neighboring yellow regions are enrolled students resident school districts.

few hidden side-effects of these problems on students. The complex routing and policy constraints indirectly increases the average bus riding time per student which directly impacts the health of the students negatively. In addition, several studies have been conducted in literature (Evans, 2014, Galliger et al., 2009, Hendrix et al., 2019, Vincent et al., 2014, Walters et al., 2020) that shows the complex social interactions such as bullying, and other forms of misconduct occurring on school buses. Moreover, increased bus riding times indirectly influences families/students school choices, thus rendering the option of having multiple school choices irrelevant.

Based on previous discussion, one can summarize the key challenges involved in the school bus transportation services as the following:

1. Schools: Rising transportation costs (due to high fuel prices, rising contracting costs (Cassell, 2000, Thompson, 2011), multiple school choices leading to complex routing, state and federal mandates, rising maintenance costs, funding cuts).

2. Students/Families: Long bus riding times, unequal access to different schools due to inefficient transportation services and complex social interactions on schools eventually causing worse student experience.

There are different case studies DetroitGoalLine (DetroitGoalLine), OSSC (OSSC) conducted in multiple cities in US that present possible practical solutions to tackle the above challenges. The broad categories of plausible solutions are the following:

1. Public Transit: The District of Columbia allows students to ride public transit for free, irrespective of the type of school the students attend. This initiative has caused more than 60% of the students to ride public buses to school.

2. Centralized Drop-off points: DetroitGoalLine (DetroitGoalLine) initiative provides transportation and after school programming services to students going to 14 partnered schools. Students can board at the nearest GoalLine bus stop to get to and from the schools in the mornings, afternoons and evening and then return to the stop nearest to their home.

3. Regional Transportation Coordinating System: OSSC (OSSC) is an association of 20 school districts that share resources via web-based scheduling application, allowing schools to collaborate on bus routes going to same location.

4. Sharing Transportation Services (zTrip/YellowCab): According to a news article in Pittsburgh Post-Gazette Behrman (2018), zTrip and yellowcabs are collaborating with 9 local schools districts and few private schools to provide 70-80 rides per day for students.

Most of the work done in School Bus Routing Problems (SBRPs) literature focuses on optimizing the overall transportation costs for school districts. Few of the previous works tries to incorporate some constraints on maximum bus riding time, maximum walking distance to school bus stop, etc., in order to improve the overall student experience. In practice, these constraints do turn out to be beneficial for sufficiently large fraction of the students while the remaining students are left worse off than before. The goal of this project is to redesign the school bus transportation system in order to minimize the overall operating costs for schools and school districts as well as improve the student welfare. In order to achieve our goal, we try to find answers to the following sub-questions:

1. Optimize efficiency which can be measured as total route length for schools.

2. Ensure fairness among students by maximizing the total welfare of students, via a social welfare objective function.

One way to solve the above problems would be to see if there is any benefit from collaboration among schools on certain school bus routes. In order to analyze the collaboration benefits, a pilot study was conducted with five participating school districts from Allegheny County. According to the study, an estimate of 161 school bus routes are used to transport students going to charter and other non-public schools across all five districts. The study showed that there was a potential overall savings of up to 20 routes from collaboration among the school districts on the 161 school bus routes. In other words, only 141 routes are needed to transport all students going to charter and non-public schools from the five participating school districts via collaboration. The promising results obtained from this study motivates us to further investigate the benefits from collaboration for SBRPs and also understand its impact on student welfare costs.

With above questions in mind, we approach the problem using the traditional SBRP technique, where our objective is twofold: analyze the benefit from collaboration in the school bus operating costs and student welfare. The model is subject to the general routing constraints such as school bell time, maximum walking distance, maximum bus riding time, etc. along with our proposed fairness consideration either in form of new constraints or a new objective function. Such problems with above considerations and constraints are called OR games and its theory constitutes for main approach behind our model. The potential contribution of the proposed approach to SBRP literature is the explicit fairness consideration among participants (school districts, students). On the other hand, analyzing multiple levels of fairness and incentives coordination for schools (reduced cost and fair division of cost) and students (reduced riding time and fair improvement in riding experience) is a contribution to the fairness literature. The distinction between our work and the past research in the field is that we look at the tactical design of school bus routing problems where the analysis relies on continuous approximations of TSP.

The rest of the paper is presented in the following manner. In Section 2, literature review on SBRP, Collaborative Vehicle Routing Problem (CoVRP) and TSP/VRP approximation papers is presented. In Section 3, we approximate the collaboration benefit between two schools for different cases of school bus capacity. We also discuss the benefit of collaboration for multiple schools and present a method to find the best coalition within the schools so that the overall route length is minimized. Further, in Section 4 we present the simulation work that supports our proposed analysis in the previous section. In Section 5, we briefly discuss on data from Allegheny County school districts and further analyze student welfare via social welfare function in Section 6. In Section 7, we do further analysis on the proposed approximations in Section 3 and finally present the experimental findings in Section 8.

5.2 Literature Review

This section reviews the past work done in traditional SBRP approach followed by Collaborative Vehicle Routing Problem and finally ends with the theory of TSP/VRP continuous approximations which are essential for our model.

5.2.1 SBRP Review

According to Desrosiers et al. (1981), the SBRP consists of four smaller sub-problems: *bus stop selection, bus route generation, school bell time adjustment and the route scheduling.* The *bus stop selection* problem find the locations of bus stops and assign students to them. In the next step, bus routes for individual schools are generated. The *school bell time adjustment* and the *route scheduling* problems are solved for multi-school scenario, where the routes generated for single schools are merged together to create bus schedules. These four problems are highly correlated but still are solved separately due to complexity and size of the problem. The different papers solving the above four problems can be found in the survey paper by Park and Kim (2010). They cite a wide range of papers categorized based on the problem characteristics (single or multiple schools, morning vs afternoon problem, allowance of mixed loading (Park et al., 2012), homogeneous vs heterogeneous fleet, constraints, different objectives (optimizing efficiency or ensuring equity), solution methods, etc.). Some of the existing approaches in the literature to solve school bus routing problems use metaheuristics (Schittekat et al., 2013), local search method (Spada et al., 2005), simulated annealing (Chen et al., 2015) and some special case routing heuristics (Braca et al., 1997).

The recent paper by Bertsimas et al. (2019) is the first paper that addresses the School Time Selection Problem (STSP) where the school bell time problem is studied in conjunction with the school bus routing problem. They develop a new school bus routing algorithm called biobjective routing decomposition (BiRD) which performed better than the existing approaches in literature on benchmark datasets and also proposed a mathematical formulation that can incorporate any number of objectives using BiRD. One of the potential pushback from the families regarding the proposed solutions by BiRD algorithm could be regarding fairness and student welfare concerns. Since there is no explicit fairness consideration in the formulation (in the objective function or the constraints), the results of BiRD algorithm can make a worse off student's experience (bus riding time or waiting time) even worse. In our proposed approach, we are making an effort to ensure that such a solution does not exist.

5.2.2 CoVRP Review

CoVRP is responsible for optimizing the overall operational costs by increasing the efficiency via cooperation. In the context of SBRP, one can study the potential gains from collaboration among different schools (maybe across different school districts) that can help design better bus routing policies. Another important aspect of collaborative vehicle routing is to propose a fair cost allocation strategy so that none of the schools want to break off their current coalition/allocation. The survey paper by Gansterer and Hartl (2018) extensively discusses the literature on CoVRP. They review three major streams of CoVRP literature: centralized collaborative routing (collaboration under complete information), decentralized planning without auctions (collaboration without central authority and complete information) and auction-based decentralized routing (collaboration with central authority and complete information). Another survey paper by Guajardo and Rönnqvist (2016) discusses the profit-sharing in collaborative routing literature based on cooperative game theory. Most of the work in this domain uses one of the three methods:

1. *Shapley value*: (introduced by Shapley (1953)) Vanovermeire and Sörensen (2014).

2. *Nucleolus method*: (introduced by Schmeidler (1969)) Dragan (1981), Guajardo and Jörnsten (2015), Maschler et al. (1979), Tamir (1989).

3. *Proportional methods*: Özener et al. (2013).

In our work, we are going to explicitly define fairness in game theoretic solution concepts for VRPs/SBRPs by determining the conditions under which the core (a feasible imputation) of our cooperative game is non-empty (Göthe-Lundgren et al., 1996, Maschler et al., 1979, Tamir, 1989) for given sets of different objectives.

5.2.3 TSP/VRP Approximation Review

A TSP aims at finding the shortest closed path through a given set of points. This problem has been extensively studied in the past but despite having so much work done in this area, there was no explicit/workable formula that could provide the value of the optimal tour until 1959. Beardwood et al. (1959) found a simple asymptotic formula for the length of the shortest tour when the number of customers n is large. The authors proved that the length of the shortest tour connecting n customers (independently chosen points over a set) in a bounded plane region of area a is "almost always" asymptotically proportional to $\sqrt{na}$ when n (≥ 200) is large. In addition, they gave numerical bounds on the constant of proportionality (depends on the dimension of the space and independent of the shape of the region) for various values of dimension. The authors extend this result to bounded Lebesgue measurable sets in k-dimensional Euclidean space.

Another approach to finding the optimal tour length for the symmetric TSP using 1-trees (variant of spanning tree) was proposed by Held and Karp (1970). The authors define a family of lower bounds on the optimal tour cost and obtain the maximum of the lower bounds when a certain linear program has optimal integer solution. The lower bound obtained is also known as Held-Karp (HK) lower bound which is basically the solution to the LP relaxation version of the original TSP with integer formulation. Further, the authors suggest different approaches (column-generation and ascent method) to find the H-K bound in Held and Karp (1970, 1971). Multiple papers such as Hansen and Krarup (1974), Johnson et al. (1996), Smith and Thompson (1977), Volgenant and Jonker (1982) have done computational experiments that showed, on an average, the H-K lower bound is very close (gap of less than 0.8% for randomly generated instances and less than 2% for real-world datasets) to the length of the optimal tour.

Similar to TSP, the paper by Haimovich and Rinnooy Kan (1985) finds asymptotically optimal bounds on the total distance travelled for a capacitated vehicle routing problem (each cab has capacity q), serving a set of n customers on the Euclidean plane. The authors also perform asymptotic analysis of the optimal solution value for different functions of cab capacity ($q = $ constant and $o(\sqrt{n})$), different distributions of points and present the results in the paper. Later, a paper by Bompadre et al. (2006) improves the bound presented by Haimovich and Rinnooy Kan (1985) and use it to develop best-to-date approximation algorithms.

5.3 Collaborative Game Problem Statement

We want to analyze the setting under which collaboration among different schools on common routes to transport students is beneficial. In addition, we would like to have that none of the students under the proposed busing system are worse off than before while ensuring that schools who collaborate are benefited from doing so. Moreover, we want the schools to be incentivized to collaborate with other schools on common routes to gain benefits (if any) from collaboration as well as ensure that such routes are stable (do not leave the coalition/collaboration to form smaller sub-coalitions) enough to not cause schools to leave the system. In order to solve the above described problem, we employ the continuous approximation of TSP/VRPs to quantify the benefit of collaboration and understand the collaboration properties using cooperative game theory.

Consider a game $G = (N, v)$ where N represents the number of players (for example schools) involved, $v(S)$ is the payoff obtained by forming sub-coalition $S \subseteq N$. Let core be the set of individually rational allocations with payoff vector x (also known as an imputation) such that

$\sum_{i \in N} x_i = v(N), \sum_{i \in S} x_i \geq v(S) \ \forall S \subseteq N$. If one can find conditions under which the core is non-empty, then the grand coalition is stable and no player would want to deviate and form different sub-coalition. The school busing system can be thought of as a cooperative game, where we have different players (can be schools, school districts or students), and we want to find a set of feasible routing plans/allocations and the set of conditions such that the core is non-empty (in other words, conditions under which collaboration is beneficial).

Formulating the school bus routing problem as a cooperative game and understanding the different properties of this game might be a bit tricky. There are some questions that need to be answered first before we can go ahead with the proposed approach:

1. Is it even tractable to calculate or approximate $v(S)$ for all $S \subseteq N$?

   - In general, one cannot approximate submodular $v(S)$ better than $O(n^{1/3})$ with polynomial number of function calls, but yes if the game is simple (Balcan and Harvey, 2011).
   - Is it possible to approximate TSP/VRP types $v(S)$ efficiently? Also, can we do the same for SBRP?

2. What is the core for the collaborative routing games and can we efficiently (or approximately) describe it?

3. Since for our problem we have multiple types of players with different objectives, we end with different cores for different games and the question now becomes to identify a feasible solution in the intersection of cores for all the games. Can we efficiently do that?

In the proposed analysis, we also try to incorporate fairness among students under collaboration using a general social-welfare objective function (Bertsimas et al., 2011, Hooker and Williams, 2012, Zukerman et al., 2005) which has fairness embedded in its definition. The idea is to understand how the student welfare is affected by collaboration and is it better/worse-off than the no-collaboration setting.

We start the analysis by finding answers to questions concerning the collaborative routing games and approximating payoff values for TSP/SBRPs.

5.4   Approximating the Benefit of Collaboration

There are two problems that we want to primarily focus on in this section, particularly

1. Can we approximate $v(S)$ for TSP/VRP for a given coalition S of schools? If so, what would be workable formula for it?

2. Once we have the above information, under what conditions does the core of the game is non-empty?

In order to find solution to the first question, we use the results obtained in the paper by Beardwood et al. (1959) and Haimovich and Rinnooy Kan (1985). The key objective/motivation behind our work is that the schools, if collaborate on common routes, can achieve higher levels of efficiency (measured either in terms of travel time, total transportation cost, etc.). One way to measure efficiency could be by calculating the total tour length serving all the students. Since we are interested in an approximate solution, we use the results from Beardwood et al. (1959) paper, which proves that the length of the optimal TSP tour for a set of n points in a bounded plane of area a is proportional to $\sqrt{na}$. Thus, the cooperative game involving different schools as players has the objective of minimizing the total travelled distance while serving students and following general constraints. In order to come up with a workable solution, we perform the following analysis for two schools, where we analyze the payoff values for different types of coalition and determine under what conditions does the grand coalition (complete collaboration) is better.

5.4.1 Collaboration Between Two Schools

Consider two schools (for now in the same school district), school A (s_A) and school B (s_B) whose geographic locations are deterministic. Let the number of students going to school A be denoted by n_A and similarly n_B denotes the number of students going to school B. Suppose the geographic locations of each of the n_A students is independently sampled from a distribution with bounded support and p.d.f f_A. Similarly we have a distribution with bounded support and p.d.f f_B for sampling n_B students independently. Assume that $n_A + n_B (= n, let)$ points are chosen in a bounded area a in the Euclidean plane. Both the schools, s_A and s_B are located within area a. One thing to note here is that we want to find a reasonable formula that can gives us the approximate length of the optimal tour, riding time for students, etc., serving n students, some of the students going to a different school, as a function of depot, school bus capacity (assume homogeneous fleet with capacity c), p.d.f of the geographic distribution of student's location and some other parameters. Since finding a universal formula might be difficult, we break the problem into different cases and find approximate solutions for them.

Assumption 5.1. Students that go to schools A and B are located in the same contiguous, bounded region with area a.

Assumption 5.2. Schools are in the same region (including the boundary).

We later on relax the assumptions and discuss the extension of our results. Note that under these two assumptions, we do not restrict the shape of the area, and in particular it can be non-convex.

School buses with infinite capacity. We consider two cases, collaboration vs no-collaboration between the two schools, and derive conditions under which collaboration is beneficial. The key result is that due to the concavity of total travel length as a function of the number of students, collaboration is indeed beneficial if schools are not too far apart. Derivation details are shown below. The result is similar to the previous section, but the benefit is attenuated by a factor proportional to the number of buses required to transport students from the smaller school.

1. **No-Collaboration**: Assume we have two depots, located at s_A and s_B. Since there is no collaboration between schools, two buses, each serving n_A and n_B students start at their respective school/depot, pickup and drop-off the students at the desired school. The total approximate length of the optimal tour denoted by L_{NC} would be:

$$\text{L}_{\text{NC}} = TSP(n_A) + TSP(n_B) \tag{1}$$
$$\implies \text{L}_{\text{NC}} = \beta\sqrt{n_A a} + \beta\sqrt{n_B a} \tag{2}$$

   where $TSP(x)$ is a function that outputs the approximate length of optimal tour containing x customers using the Beardwood et al. (1959) TSP approximation. β_A, β_B are the proportionality constants that depend only on the dimension and independent of the shape of the bounded region.

2. **Collaboration**: Assume we have 1 depot here and w.l.o.g, let's consider the depot to be at s_A. Since there is collaboration among schools and we have school buses of infinite capacity, the school bus would leave the depot, pickup all the n students, bring them back to s_A and then take the n_B students to s_B. Thus, the total approximate length of the optimal collaborative tour denoted by L_C would be:

$$\text{L}_{\text{C}} = TSP(n) + \text{distance}(s_A, s_B) \tag{3}$$
$$\implies \text{L}_{\text{C}} = \beta\sqrt{na} + \text{distance}(s_A, s_B) \tag{4}$$

   where $\beta_{A \cup B}$ is the proportionality constant when the TSP tour comprises of n students, serving together $n_A + n_B$ students. Note that the distance(s_A, s_B) is the euclidean distance between the two schools.

Once we have found the approximate total length of the tours under collaboration and no-collaboration, we further investigated the conditions under which collaboration would be beneficial, i.e.,

$$L_C \leq L_{NC} \tag{5}$$

$$\implies \beta\sqrt{na} + \text{distance}(s_A, s_B) \leq \beta\sqrt{n_A a} + \beta\sqrt{n_B a} \tag{6}$$

Note: There are no explicit formulas for constants $\beta_{A \cup B}, \beta_A$ and β_B but there are bounds defined on their values, i.e.,

$$(2 + k^{-1})![(k/2)]!6(1/k)/(2 * \sqrt{\pi}(k^{0.5} + k^{-0.5})) \leq \beta_k \leq 6^{-0.5}12^{1/2*k} \tag{7}$$

$$\implies 0.625 \leq \beta_2 \leq 0.922 \tag{8}$$

where k is the dimension of the problem (in our case, $k=2$). There are other papers such as Few (1955), Steinerberger (2015) which show results on improved lower and upper bounds. In order to deduce some intuition from inequality (6), for now, let us assume that both n_A and n_B points are sampled independently from a uniform distribution with same mean and variance in $[0, 1]^2$. Then by BHH theorem (Beardwood et al., 1959), there exists a universal constant β with probability 1 and $a = 1$. Substituting these values in inequality (6) we get,

$$distance(s_A, s_B) \leq \beta(\sqrt{n_A} + \sqrt{n_B} - \sqrt{n_A + n_B}) \tag{9}$$

where the r.h.s of (9) is always ≥ 0. The inequality (9) in other words implies that there can be a potential gain from collaboration as long the distance between the two schools is less than or equal to the the gain from cooperation. Now, let us consider the following scenarios to understand few implications of inequality (9):

- $n_A \gg n_B$: In this case, the r.h.s of (9) approximately tends to 0 and thus for (9) to hold true, l.h.s must be equal to 0. This makes sense in practice, i.e., when either of the school has very few students compared to the other, collaboration would be beneficial only if the schools are nearby. If not, schools are better off serving their students independently. The same thing holds true when $n_A \ll n_B$.

- n_A, n_B both are very large: Since r.h.s of (9) increases as n_A, n_B increases, one can deduce that, for a fixed distance(s_A, s_B) and for some particular values of n_A, n_B, if we further increase the number of students, then collaboration among schools would be beneficial. In other words, as the number of students going to both schools increases in the bounded planar region, the probability of benefit from collaboration increases (as $\sqrt{n_A + n_B} \leq \sqrt{n_A} + \sqrt{n_b}$).

- distance(s_A, s_B) is very large: Intuitively one would think that collaboration would be infeasible if the schools are far away. We observed similar result from inequality (9) when $distance(s_A, s_B)$ tends to infinity and thus l.h.s dominates r.h.s.

- distance(s_A, s_B) is very small: In this case, one might expect collaboration to be more beneficial and we observed similar result when distance(s_A, s_B) tends to 0. Thus, for some particular values of β's, total length of the tour is smaller under collaboration.

Note: The uniformity assumption on the locations of students is not a necessary condition. As long as the points are sampled from a distribution with bounded support, BHH theorem holds true, i.e., if the students are independently distributed with respect to some probability on a measurable set γ and if f is the absolutely continuous part of the measure, then with probability 1

$$\lim_{n \to \infty} L(X_1, ..., X_n)/\sqrt{n} = \beta \int_\gamma f^{0.5}(v)dv \tag{10}$$

where $L(X_1, ..., X_n)$ is the optimal length of the TSP tour containing n customers. According to Beardwood et al. (1959), Steinerberger (2015), the necessary condition for the BHH theorem to hold for random distribution of points with unbounded support is that all the absolute $k/(k-1)$

th moments of the probability distribution should exist (k is the dimension of the Euclidean space) when $\int_\gamma f^{0.5}(v)dv$ is finite.

School buses with finite capacity. Assume that we have 2 schools as above but instead of buses with infinite capacity, we have enough buses of finite capacity c such that each bus does one round to pickup students, returns to the depot and stay there forever. We will again consider two cases: collaboration vs no-collaboration among the schools and analyze the results as performed in previous section.

1. **No-Collaboration**: Assume we have two depots, located at s_A and s_B. Further, partition the region into m_A and m_B sub-regions such that each sub-region has not more than c students. Partitioning the region in this manner would allow each depot to send m_i ($i \in \{A, B\}$) buses to pickup all the students and bring them back to the depot. For simplicity, assume that both $n_A(\geq c), n_B(\geq c)$ are multiples of c. Since there is no collaboration between schools, buses start off at their respective depots, serving n_A and n_B students and drop-off them at the desired school. The total approximate length of the optimal tour denoted by L_{NC} would be:

$$L_{NC} = m_A TSP(n_A/m_A) + m_B TSP(n_B/m_B) \tag{11}$$
$$\implies L_{NC} = \beta m_A \sqrt{n_A/m_A * v/m_A} + \beta m_B \sqrt{n_B/m_B * v/m_B} \tag{12}$$
$$\implies L_{NC} = \beta\sqrt{n_A v} + \beta\sqrt{n_B v} \tag{13}$$

where $m_A = n_A/c, m_B = n_B/c$. Note that the area v also changes as we partition the region into smaller sub-regions for the analysis. For bus routing the students of school A, the region is partitioned into m_A sub-regions, each sub-region consisting of c number of students. Thus, school A requires m_A number of school buses to pickup the students and bring them to s_A, creating m_A TSP tours serving c students per tour. Here the bounded region is divided into m_A sub-regions and so the v for each TSP tour is technically v/m_A instead of v. Similar calculations are performed for school B.

2. **Collaboration**: Assume we have 1 depot here and for simplicity, let the depot be either the s_A or s_B depending on $min\{n_A, n_B\}$. If $n_A \leq n_B$, s_B is the selected depot; otherwise, s_A is the depot. Since there is collaboration among schools and we have school buses of finite capacity, we partition the bounded region containing entire n students into $m_{A\cup B}$ sub-regions, each region having c students. $m_{A\cup B}$ number of school buses leave the depot, pickup all the n students, brings them back to the selected depot and then $min\{m_A, m_B\}$ buses are used to transport either n_A or n_B students (depending on the chosen depot) to s_A or s_B respectively. Thus, the total approximate length of the optimal collaborative tour denoted by L_C would be:

$$L_C = m_{A\cup B} TSP(n/m_{A\cup B}) + m_B distance(s_A, s_B) \tag{14}$$
$$\implies L_C = \beta m_{A\cup B}\sqrt{n/m_{A\cup B}v/m_{A\cup B}} + m_B distance(s_A, s_B) \tag{15}$$
$$\implies L_C = \beta\sqrt{nv} + (min\{n_A, n_B\}/c)distance(s_A, s_B) \tag{16}$$

where $m_{A\cup B} = n/c$, $m_A = n_A/c$ and $m_B = n_B/c$.

After finding approximations for the length of the optimal tour under collaboration and no-collaboration, we further investigate the conditions under which collaboration would be beneficial, i.e.,

$$L_C \leq L_{NC} \tag{17}$$
$$\implies \beta\sqrt{nv} + (min\{n_A, n_B\}/c)distance(s_A, s_B) \leq \beta\sqrt{n_A v} + \beta\sqrt{n_B v} \tag{18}$$

Further, in order to infer some results from inequality (18), let us assume that the n points are sampled independently from a uniform distribution with same mean and variance in $[0,1]^2$. Thus

by BHH theorem (Beardwood et al., 1959), there exists a universal constant β with probability 1 and $a = 1$. Substituting these values in (18), we get:

$$(min\{n_A, n_B\}/c)\text{distance}(s_A, s_B) \leq \beta(\sqrt{n_A} + \sqrt{n_B} - \sqrt{n_A + n_B}) \tag{19}$$

Inequality (19) is more stronger than inequality (9) found in the case where buses are of infinite capacity (as $n_A \geq c$ and $n_B \geq c$). We can do similar analysis as the ones done in section 4.1.1 and will obtain similar results here. One thing to note is that partitioning the region into smaller sub-regions and then dispatching buses to serve $\leq c$ students in every sub-region is one way to find the approximate length of the optimal tour. In other words, the obtained approximate lengths are upper bounds on the actual optimal tour under collaboration and no-collaboration. Using different routing policies will result in different approximate lengths, thereby different inequalities and hence the future analysis might differ.

To summarize, the above analysis serves two purposes in our work: we have a clean workable formula for length of the optimal tour under collaboration/no-collaboration as well as further analysis showed that these formulas are reasonable and coherent when tested for different scenarios. One can extend the above analysis to the case when we have finite fleet of school buses with fixed capacity $c > 0$ and would find similar results as above with different approximation for the total length of the tour. The next step is to use these approximation formulas and find properties of the cooperative game such as whether the characteristic function is sub-modular/super-modular, conditions under which the core is non-empty, or use other measures of efficiency such as total riding time for students, number of buses needed, etc.

School buses with finite capacity and tiered routes. In the previous subsections, we assumed an infinite supply of school buses such that no bus has to do multiple rounds in order to serve the students. In this subsection, we discuss the case when we have a fixed number of school buses, let say N, where each school bus has a fixed capacity c and some buses will be used more than once. This is a more complex optimization problem than before since we have smaller number of buses and tiered routes (each bus has a set of tiered routes to serve in order). We have a similar setup of two schools as before and we consider two cases: collaboration vs no-collaboration among the schools and analyze the results.

In order to find the approximate optimal tour length, we partition the area into smaller sub-regions as done in the previous case for school buses with finite capacity and will eventually end up with similar results. It does not matter if we have two buses running on two routes or one bus doing two rounds on those two tiered routes. The overall route length remains the same. Thus, there is no impact on the total route length as compared to section 4.1.2. However, there may or may not be an increase in the total driver hours depending on the number of available drivers, constraints on driving time, etc. For students, there will be an impact on their pickup and drop-off time which thereby affects the bell times for schools (or vice-versa) but the bus riding time for each student remains unaffected when compared with the solutions from section 4.1.2. We defer the detailed analysis to computational experiments later.

5.4.2 Multiple Schools

For multiple schools, it is not immediately clear that the grand coalition is the most beneficial. In fact, consider three schools, schools 1 and 2 are located close to each other and the third one is far apart. It may be the case that the optimal collaboration is to partition the three schools into $\{1, 2\}$ and $\{3\}$. Thus we can pose the general question: how should we partition the schools so we can minimize total route length of picking up all students and dropping them off at the schools?

We can formulate this optimization problem as a matching problem. In particular, we can upperbound the optimal cost by solving a weighted matching problem on a graph, where the nodes are schools, and edge weights are the pairwise collaboration benefits. Such problems can be solved via a generalization of Edmond's Blossom algorithm (Edmonds, 1965, Lovász and Plummer, 1986).

5.5 Simulation

Our approach employs continuous approximations in order to preserve simplicity of the analysis and in the previous section we have discussed ways in which our proposed approach could be used to understand when collaboration would be beneficial. However, in practice, one would want to see how far is the approximate solution from the one obtained via existing OR approaches to solve SBRPs in literature. To demonstrate the above comparison, we would want to do a computational case study to analyze the accuracy of the proposed approach and its potential practical use. We can also compare the student welfare costs in both the approaches and see if our model is better at capturing fairness compared to the existing models in literature.

Another interesting experiment we want to conduct is to understand the combined impact of overlap between the geographical locations of students going to both the schools and the distance between the two schools on the absolute value of the difference between L_C and L_{NC}. We can randomly generate coordinates for the students (assuming different distributions) and come up with different metrics to measure overlap between the geographical distribution of n_A, n_B students locations. Further, we can add more variations (1 depot or 2 depots, different routing policies, etc.) to the above experiments and conduct multiple experiments in order to ensure robustness of the results. The results obtained would give us intuition on how the two factors affect collaboration among schools. Also, conducting these experiments would help us validate the proposed approximations on the optimal tour lengths and its practicality.

Our experimental study has the following setup. We randomly choose two school locations and then sample 200 students locations for each school such that all the students locations are within a circle of radius r from the school's location. After we have the list of students locations for both the schools, we need to solve a Bus_Stop_Selection problem in order to find the appropriate bus stops and the corresponding demand to be served at each of the stop. The Bus_Stop_Selection problem is an Integer Program solved for each school, where the objective is to minimize the number of stops and the total student walking distance (similar to (Bertsimas et al., 2019)). The next step is to find the tour length under collaboration and no-collaboration setting among schools and we use the TABU_SEARCH metaheuristic algorithm for it. TABU_SEARCH method employs a local search heuristic to iteratively move from one plausible solution to an improved solution within a certain neighborhood of the previous solution unless some stopping criteria has been met. In our experimental analysis, we first implement TABU_SEARCH to find the approximate route length under no-collaboration for both the schools. Further, the solution obtained for the no-collaboration is later used as an input to initialize the TABU_SEARCH method used to solve the collaboration scenario. The proposed TABU_SEARCH method does not guarantee optimal results but definitely provides reasonable approximate solutions.

5.6 Data description

To conduct the above simulation study, we use synthetic as well as actual data from Allegheny County school districts. The actual data comprises of information about:

- Contractor(s) that each school district utilizes

- Amount paid to each contractor

- Districts utilizing Port Authority Transportation

- Amount paid to Port Authority and individuals

- Number of students transported (broken down by type)

- Vehicles used by the school district

- Different types of schools (Public, Non-Public, Charter) along with their addresses in Allegheny County

- Students enrollment in each of these schools for 2017-2018 and 2018-2019 academic years along with enrolled students resident districts

- Current school bus routes, operational costs and annual miles traveled

5.7 Future Work

The next step in the project is to understand the properties of the cooperative game with the above defined approximate characteristic function and determine conditions under which the core is non-empty. Once we are able to achieve that, we would want to further use the above approximations and determine the core for the cooperative game involving students, where the objective is to improve the student welfare cost. In order to do this, we will use the proposed continuous approximations to determine utilities for every student (utility can be a function of waiting time, riding time, etc.). Once we have these utilities defined for every student, we would want to allocate route choices such that the overall social welfare objective (either α fairness or the Hooker-Williams social welfare function, etc.) is maximized, ensuring fair improvement over the existing bus routes in service. Finally, we would want to propose bus routes which are in sync with the outputs obtained as a result of optimizing the student welfare costs as well as the overall operational costs for the schools/school districts, if it exists (otherwise approximate).

References

AlliesForChildren (2017). At the Crossroads: School Transportation in Allegheny County. Technical report, Allies for Children.

Balcan, M.-F. and N. J. Harvey (2011). Learning submodular functions. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pp. 793–802.

Beardwood, J., J. H. Halton, and J. M. Hammersley (1959). The shortest path through many points. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 55, pp. 299–327. Cambridge University Press.

Behrman, E. (2018). *Pittsburgh's zTrip targets new market: students.* https://www.ztrip.com/2018/06/05/pittsburghs-ztrip-targets-new-market-students/.

Bertsimas, D., A. Delarue, and S. Martin (2019). Optimizing schools' start time and bus routes. *Proceedings of the National Academy of Sciences 116*(13), 5943–5948.

Bertsimas, D., V. F. Farias, and N. Trichakis (2011). The price of fairness. *Operations research 59*(1), 17–31.

Bompadre, A., M. Dror, and J. B. Orlin (2006). Improved bounds for vehicle routing solutions. *Discrete Optimization 3*(4), 299–316.

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

Burgoyne-Allen, P., K. Boone, J. Squire, and J. O. Schiess (2019, 07). The Challenges and Opportunities in School Transportation Today. Technical report, Bellwether Education Partners.

Cassell, M. (2000). Taking them for a ride: an assessment of the privatization of school transportation in ohio's public school districts. *Published by the Association of Federal, State, County and Municipal Employees*.

Chen, X., Y. Kong, L. Dang, Y. Hou, and X. Ye (2015). Exact and metaheuristic approaches for a bi-objective school bus scheduling problem. *PloS one 10*(7), e0132600.

Desrosiers, J., J. Ferland, J. Rousseau, G. Lapalme, and L. Chapleau (1981). An overview of a school busing system. nk jaiswal, ed., scientific management of transport systems.

DetroitGoalLine. *GOAL Line Detroit.* Available at https://www.goaldetroit.org/.

Dragan, I. (1981). A procedure for finding the nucleolus of a cooperativen person game. *Zeitschrift für Operations Research 25*(5), 119–131.

Edmonds, J. (1965). Maximum matching and a polyhedron with 0, 1-vertices. *Journal of research of the National Bureau of Standards B 69*(125-130), 55–56.

Evans, N. (2014). Exploring the school bus as an environment for bullying: A phenomenological study.

Few, L. (1955). The shortest path and the shortest road through n points. *Mathematika 2*(2), 141–144.

Galliger, C. C., M. S. Tisak, and J. Tisak (2009). When the wheels on the bus go round: Social interactions on the school bus. *Social Psychology of Education 12*(1), 43–62.

Gansterer, M. and R. F. Hartl (2018). Collaborative vehicle routing: a survey. *European Journal of Operational Research 268*(1), 1–12.

Göthe-Lundgren, M., K. Jörnsten, and P. Värbrand (1996). On the nucleolus of the basic vehicle routing game. *Mathematical programming 72*(1), 83–100.

Guajardo, M. and K. Jörnsten (2015). Common mistakes in computing the nucleolus. *European Journal of Operational Research 241*(3), 931–935.

Guajardo, M. and M. Rönnqvist (2016). A review on cost allocation methods in collaborative transportation. *International transactions in operational research 23*(3), 371–392.

Haimovich, M. and A. H. Rinnooy Kan (1985). Bounds and heuristics for capacitated routing problems. *Mathematics of operations Research 10*(4), 527–542.

Hansen, K. H. and J. Krarup (1974). Improvements of the held—karp algorithm for the symmetric traveling-salesman problem. *Mathematical Programming 7*(1), 87–96.

Held, M. and R. M. Karp (1970). The traveling-salesman problem and minimum spanning trees. *Operations Research 18*(6), 1138–1162.

Held, M. and R. M. Karp (1971). The traveling-salesman problem and minimum spanning trees: Part ii. *Mathematical programming 1*(1), 6–25.

Hendrix, J. A., E. K. Kennedy, and J. V. Trudeau (2019). The rolling hotspot? perceptions of behavioral problems on school buses among a nationally representative sample of transportation officials. *Journal of school violence 18*(3), 455–467.

Hooker, J. N. and H. P. Williams (2012). Combining equity and utilitarianism in a mathematical programming model. *Management Science 58*(9), 1682–1693.

Johnson, D. S., L. A. McGeoch, and E. E. Rothberg (1996). Asymptotic experimental analysis for the held-karp traveling salesman bound. In *Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, Volume 341, pp. 350. ACM Press San Francisco.

Lovász, L. and M. Plummer (1986). Matching theory, volume 29 of ann. *Discrete Math. North-Holland Publishing Co., Amsterdam*.

Maschler, M., B. Peleg, and L. S. Shapley (1979). Geometric properties of the kernel, nucleolus, and related solution concepts. *Mathematics of operations research 4*(4), 303–338.

OSSC. *Ohio Shared Services Collaborative*. http://www.ohiossc.org/.

Özener, O. Ö., Ö. Ergun, and M. Savelsbergh (2013). Allocating cost of service to customers in inventory routing. *Operations Research 61*(1), 112–125.

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

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

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

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on applied mathematics 17*(6), 1163–1170.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games 2*(28), 307–317.

Smith, T. H. and G. L. Thompson (1977). A lifo implicit enumeration search algorithm for the symmetric traveling salesman problem using held and karp's 1-tree relaxation. In *Annals of Discrete Mathematics*, Volume 1, pp. 479–493. Elsevier.

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

Steinerberger, S. (2015). New bounds for the traveling salesman constant. *Advances in Applied Probability 47*(1), 27–36.

Tamir, A. (1989). On the core of a traveling salesman cost allocation game. *Operations Research Letters 8*(1), 31–34.

Thompson, O. (2011). The estimated cost impact of privatizing student transportation in minnesota school districts. *Public Choice 146*(3-4), 319–339.

U.S. Census Bureau (2018). United States Census Bureau: Longitudinal Employer-Household Dynamics Datasets. https://lehd.ces.census.gov/data/. [Accessed 06-01-2021].

Vanovermeire, C. and K. Sörensen (2014). Integration of the cost allocation in the optimization of collaborative bundling. *Transportation Research Part E: Logistics and Transportation Review 72*, 125–143.

Vincent, J. M., C. Makarewicz, R. Miller, J. Ehrman, and D. L. McKoy (2014). Beyond the yellow bus: Promising practices for maximizing access to opportunity through innovations in student transportation. *Center for Cities & Schools*.

Volgenant, T. and R. Jonker (1982). A branch and bound algorithm for the symmetric traveling salesman problem based on the 1-tree relaxation. *European Journal of Operational Research 9*(1), 83–89.

Walters, G. D., J. Kremser, and L. Runell (2020). The fear of being bullied on the school bus: Perceptions, correlations, and sex differences. *International Journal of Bullying Prevention*, 1–12.

Zukerman, M., L. Tan, H. Wang, and I. Ouveysi (2005). Efficiency-fairness tradeoff in telecommunications networks. *IEEE Communications Letters 9*(7), 643–645.