FÁTIMA MACHADO DE SOUZA LIMA

A MIXED LOAD RURAL SCHOOL BUS ROUTING PROBLEM WITH HETEROGENEOUS FLEET: A STUDY FOR THE BRAZILIAN PROBLEM

Belo Horizonte
02 de julho de 2015

FÁTIMA MACHADO DE SOUZA LIMA

A MIXED LOAD RURAL SCHOOL BUS ROUTING PROBLEM WITH HETEROGENEOUS FLEET: A STUDY FOR THE BRAZILIAN PROBLEM

Tese apresentada ao Curso de Pós-Graduação em Engenharia de Produção da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Doutor em Engenharia de Produção.

ORIENTADOR: SAMUEL VIEIRA CONCEIÇÃO
CO-ORIENTADOR: RICARDO SARAIVA DE CAMARGO

Belo Horizonte
02 de julho de 2015

FÁTIMA MACHADO DE SOUZA LIMA

A MIXED LOAD RURAL SCHOOL BUS ROUTING PROBLEM WITH HETEROGENEOUS FLEET: A STUDY FOR THE BRAZILIAN PROBLEM

Thesis presented to the Graduate Program in Engenharia de Produção of the Universidade Federal de Minas Gerais in partial fulfillment of the requirements for the degree of Doctor in Engenharia de Produção.

ADVISOR: SAMUEL VIEIRA CONCEIÇÃO
CO-ADVISOR: RICARDO SARAIVA DE CAMARGO

Belo Horizonte
July 2, 2015

UNIVERSIDADE FEDERAL DE MINAS GERAIS

FOLHA DE APROVAÇÃO

A mixed load rural school bus routing problem with heterogeneous fleet: A study for the Brazilian problem

FÁTIMA MACHADO DE SOUZA LIMA

Tese defendida e aprovada pela banca examinadora constituída por:

Ph. D. Samuel Vieira Conceição – Orientador
Universidade Federal de Minas Gerais

Ph. D. Ricardo Saraiva de Camargo – Co-orientador
Universidade Federal de Minas Gerais

Ph. D. Gilberto de Miranda Júnior
Universidade Federal de Minas Gerais

Ph. D. Rodney Rezende Saldanha
Universidade Federal de Minas Gerais

Ph. D. Alexandre Xavier Martins
Universidade Federal de Ouro Preto

Ph. D. Henrique Pacca Loureiro Luna
Universidade Federal de Alagoas

Belo Horizonte, 02 de julho de 2015

Abstract

The underdevelopment of Brazilian rural families is largely explained by their historical process of formation and by their poor access to a functional education and transportation systems. In the last decade, the federal government has been encouraging the nucleation of rural schools to offer better structured schools to the rural students. Multi-grade rural schools, often located closer to the rural families but with students of different grades being taught by the same teacher at the same class, are being shutdown and transfered to bigger, better installed facilities located near to the counties' downtown area. The success of such endeavor relies on offering a transportation system for the rural students. Hence the Brazilian federal government has been making a great effort to support local administrators to provide better transport to rural students. One of such efforts gave rise to a central decision support system which solves the mixed load capacitated rural school bus routing problem with heterogeneous fleet. The mixed load feature allows students from different schools to ride the same bus during at the same time. This is an important but neglected problem in vehicle routing literature. In this thesis, four based meta-heuristic algorithms are devised and embedded into the support system. The computation performance of the proposed algorithms was assessed on solving four different datasets, including a real case from Brazil. The proposed methods were also compared with one known method from the literature. The attained cost savings and reduction of the number of buses required to serve the rural students showed the suitability of the mixed load approach over the single load one for the Brazilian rural context. Furthermore four based meta-heuristic based multi-objective algorithms to solve the multi-objective capacitated mixed load rural bus routing problem with heterogeneous fleet were also devised. The three involved objectives were the routing costs, the average weighted riding distances and the routes balance. The proposed multi-objective methods were compared with one from literature adapted for the problem and evaluated by assessing the metrics of cardinality, coverage and hyper-volume, followed by a statistical analyses. The work also introduces a new approach to help decision makers to selected a suitable solution from a Pareto set. All of the four devised multi-objective heuristics outperformed the literature procedure.

Keywords: Capacitated rural school bus routing problem, mixed loading, heterogeneous fleet, Meta-heuristic methods, multi-objective, decision support systems.

Aos meus pais, Lorival e Regina,
meu tesouro.

Acknowledgments

A quem mais agradecer senão a Deus? Por Ele tudo foi e sempre será possível.

Meu Deus, meu tudo...

Agradeço ao meu orientador Prof. Samuel Vieira Conceição pela oportunidade, por sua orientação, conselhos e preocupação que iam além do meu trabalho.

Ao meu co-orientador Prof. Ricardo Saraiva de Camargo, por sua orientação, paciência, por ter sempre me motivado e sempre me lembrado de que sou capaz.

Ao Prof. Gilberto de Miranda Jr, que de forma indireta teve papel fundamental nesse trabalho.

Aos membros da banca por disponibilizarem seu tempo e por suas valiosas contribuições.

Agradeço em especial ao Davi Doro e André Maravilha por sempre estarem dispostos a me ajudar e a contribuir com esse trabalho.

Agradeço a meus pais, Lorival e Regina, exemplos de santidade, fé, humildade, sensatez e bondade. Mesmo longe sempre se fizeram presentes e sempre me sustentaram nas dificuldades desse doutorado. Por meio do exemplo de vocês venho descobrindo o que é essencial nessa vida!

Aos meus irmãos, Júnior, Marcelo e Sandra que igualmente sempre se preocuparam e me ajudaram em todos os momentos. Saibam que aprendo muito com vocês.

À Viliv e ao Fábio, vocês moram no meu coração.

Vítor e Davi, amo vocês!

Àquele cuja espera mudou a direção do meu caminho, revelando-me uma estrada que eu não conhecia, mais leve, mais bonita, mais alegre e cheia de luz! A seus pais e sua familia por todo amor e carinho, que é recíproco!

Agradeço às minhas amigas que sempre me deram apoio nas horas mais difíceis! Aos colegas do departamento, pelas discussões e pelo suporte emocional! Aos funcionários de DEP por estarem sempre disponíveis.

Por fim, gostaria de agradecer ao Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsa de estudo.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

With 27 states, Brazil has a remarkable feature which is its diversity, revealed in a rich variety of cultures wide spread on an extensive territory and an area bigger than 8.500 millions km^2 divided in 5.560 counties. The cultural differences and contrast among counties can be perceived in its population density due to its colonial occupation and to the trend of having economic activities concentrated in great urban areas.

This thesis comprises one of the biggest states in Brazil: Minas Gerais (figure 1.1). Its area represents 7% of Brazilian territory, equal to 2.55 times the size of Great Britain and 0.92 times the size of France, the state ranks as the second most populous and the fourth largest by area in the country. It has over 20 millions inhabitants of which 85.3% live in urban centers and 14% live in rural areas (IBGE, 2014).

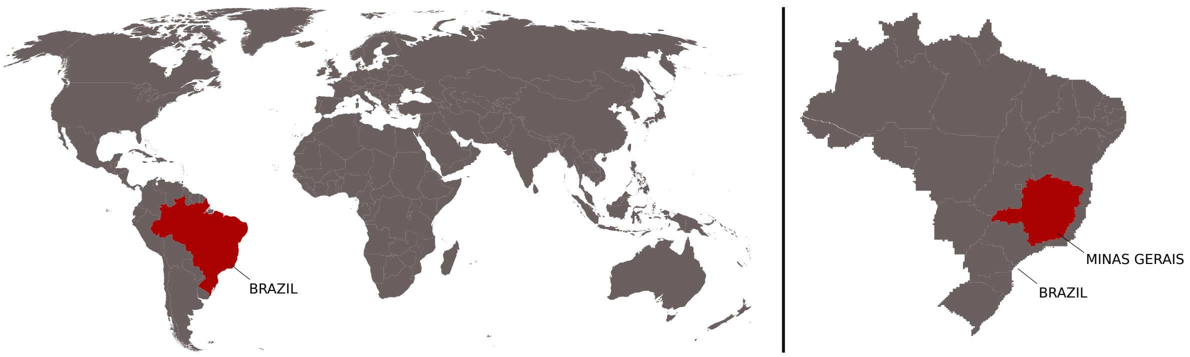


Figure 1.1: Brazil and Minas Gerais state.

For being less densely populated, rural areas has been forgotten by social and political policies leading to less opportunities of growth and development. However, besides having a smaller number of inhabitants the rural community have the same citizens rights of been benefited with basic infrastructure and services such as health, public transport and education (Carvalho et al., 2010).

To reverse the actual scenario is necessary to facilitate the access to education once the development of a municipality or county is directly connected with the education of their dwellers. With good education is possible to ensure good jobs opportunities, a better life quality and a promising future.

Unfortunately, the access to school of many rural students has been prevented due to the nucleation process. The schools used to be located in small cities and villages and have students of many grades in the same class. In order to organize the educational system in rural areas and increase the technical and economical support to the children and teachers, those units were shutdown and the teaching were concentrated in central schools.

The process, exemplified in figure 1.2, represents the schools before being closed, where students could attend classes close to their residences and all the children were taught at the same classroom (a). After the nucleation process, the students start to be settled in different classes according to their grades, in a central and farthest schools, figure (b).

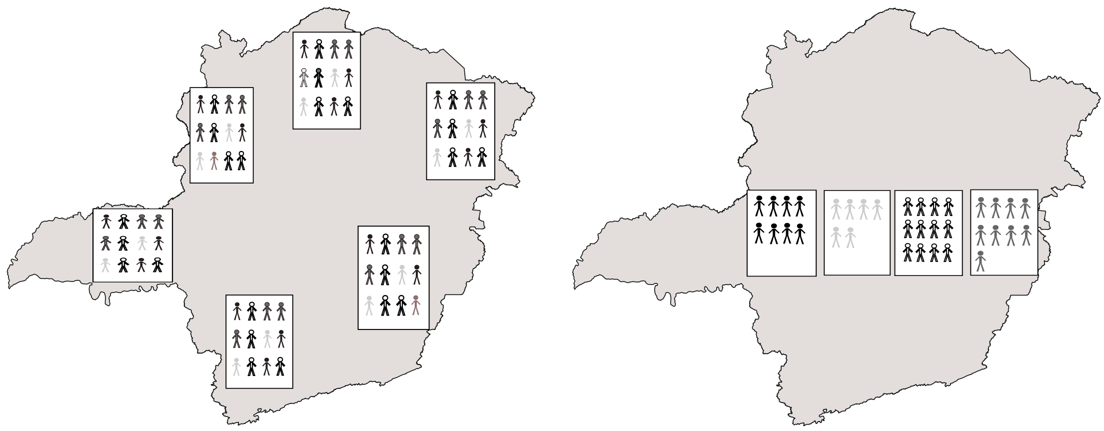


Figure 1.2: (a) Multi-graded classes before nucleation process, (b) Single graded classes after nucleation process.

So, the feasibility of the nucleation process and the development of rural areas are highly dependent of the students and rural community transportation. Then, to make up the lack of schools, a free public transportation is guaranteed by law for students who assist classes in urban areas, but it is one of the biggest barriers they face to attend classes in distant schools.

Although been guaranteed by law, the provided rural transportation has a very low quality, which is consequence of a set of factors. The routes, defined annually according to the children's residences, ought to have a suitable service level such as maximum travel time and maximum number of students per vehicle (Sanches and Ferreira, 2003).

However, due to the lack of a good system and qualified technicians, those limits are often disregarded and the students have to endure long travel times along unpaved roads (figure 1.3) and overcrowded vehicles without maintenance or any convenience (figure 1.4). The students dispersion force them to make long walks to reach the school or the bus stop, reducing their performance at classes.

The schools starting times also increases the difficulty of defining routes. In countries such as United States (Park et al., 2012) or Germany (Fügenschuh, 2009) the schools have its classes starting at different times so it is possible to have reuse of buses with short distance and time travels. But in Brazil, besides having the classes starting at the same time, the distances to be traveled are bigger, so the fleet must have higher number of buses to attend all of the students, increasing the final costs, distance and time travels. This scenario suggest the consideration of a heterogeneous fleet to attend more than one school at a time, i.e., transporting students from multiple schools.

Figure 1.3: Roads on raining season.

Figure 1.4: Overcrowded buses in Minas Gerais state.

In most Brazil rural counties the transportation system is deficient and the final expenses for it are high because there is a lack of qualified workers to manage the transportation network which includes: establish routes, select students, designate and assign buses to the routes, consider cost, safety, time window, travel time and vehicles capacity. Instead, the work is performed by the "feeling" of the responsible for the transportation. Thus, the task of routing is barely done, increasing the expenses, the travel time, number of required buses and its maintenance, the drivers work hours and reducing the students performance and their assiduity to classes.

Therefore, research about rural transportation is helpful for social and spatial integration. According to Thangiah et al. (2013), the significance of the school bus routing problem is directly attributed to its social impact, besides the economic objectives. So, the purpose of this thesis is to contribute to increase the service level of the rural transportation related to travel time, safety, routes planning and final costs in order to increase the life quality of rural population. Also, present mathematical models, single and multi-objective heuristics to represent and solve the rural student transportation problem.

1.2 Purpose of the thesis

1.2.1 General

The main objective of this thesis is to develop approaches to solve the capacitated rural school bus transportation problem in Brazil by respecting the envolved constraints so that good solutions in a reasonable time with lower costs and adequate service level.

1.2.2 Specifics

Specifically it is intended to:

- study the literature about school vehicle routing problem and focusing in the rural transportation,
- propose mathematical formulations for the problem,
- implement heuristic algorithms single objective case of the problem,
- analyze the impact of using single and mixed load in the final costs,
- propose a multi-objective version to solve the problem,
- propose enhancement features for a meta-heristc iterades local search framework.

1.3 Thesis Organization

At chapter 2 a literature review is presented with specific topics of rural transportation and the Brazilian problem addressing mainly the papers which deal with school routing problem, the chapters in sequence present a deeper review about the related issue.

Chapter 3 presents the related works and a heuristic methods to deal with a single objective problem. Five meta-heuristics are presented to solve the problem and tested with four data sets. Chapter 4 shows a review and a solution approach for multi-objective rural school bus vehicle routing problem, where more then one function is managed at the same time. A new approach for helping decision makers to find good solutions among Pareto set is also presented in this chapter. Finally, chapter 5 brings the findings, conclusions and future research for the work.

Chapter 2

Literature Review

2.1 The School Bus Routing Problem

2.1.1 Introduction

This section brings a literature review about the School Bus Routing Problem (SBRP). The concerned issue is related to the process of routes construction for the rural school transportation which has its own particularities and requirements.

Schittekat et al. (2013) divide the problem in three subproblems: find a set of bus stop to visit, determine the bus stop to which each student should be moved to and determine routes to visit the defined bus stops minimizing the total distance traveled by buses.

Desrosiers et al. (1981) divide the problem in five subproblems, allowing it to be solved in steps, which are: data preparation, bus stop selection, routes generation, time windows adjustment and route scheduling (buses assignment). The literature review is based on Desrosiers et al. (1981) division because, besides being more complete, encompasses larger amount of papers enabling a better state of the art about the problem.

The first step, data preparation, specify the data network and is compounded by the students, schools and garage geographic locations and an origin-destination matrix. In bus stop selection step, the bus stops for boarding and drop-off the children are defined according to the easiness access to it. Depending on the classes starting time it is required to adjust the routes with the time windows and then, is necessary to assign the buses to the routes. The subproblems are detailed below.

Data preparation: The first stage is to prepare the data to the following subproblems. The roads are specified in addition to the students and schools geographic locations, kind of fleet and distance matrix which must include the lowest times or the distance between two pair of nodes. It can be calculated through geographic information systems (GIS) or shortest path algorithms (Gallo and Pallottino, 1988).

The students data include their residence location, the destination school and if the student is handicapped or not. Schools information include its location, the maximum student riding time and the classes starting and ending times to define the buses arrival. In a few works, this information is given by schools, but when is not available it can be settle by the author. The fleet information comprehend their original point (garage) and capacities.

Bus stop selection: The bus stop selection phase defines the bus stops (pick up and drop-off points) and assign the students to it. In rural school bus routing problem authors consider that the bus stops are the students residences (Park and Kim, 2010), while in a urban problem the children are allowed to walk to the stops. This consideration is made because the distances that rural students need to travel are already too long to and back from school, if they still need to walk to get the bus, the travel would be even longer and weariness what can reduce their performance in school or increase school dropout.

However, in some cases, the buses access might be unfeasible due to the roads condition, so it is acceptable to settle a common board point to the students, which is also an alternative when the students residences are close to each other. So, as in the urban cases, the children walk from their houses to the bus stop.

Only few authors apply heuristics to calculate the points (Desrosiers et al. (1981), Bowerman et al. (1995), Schittekat et al. (2006), Schittekat et al. (2013)), but usually, this step is disregarded because it is considered that the bus stops are given.

Routes generation: The heuristics used to construct the routes are classified in "cluster-first, route-second" or "route-first, cluster-second" approach (Bodin and Berman, 1979b). "Cluster-first" is based on the method proposed by Min et al. (1998) and has two phases. First build clusters of students then, for each cluster, a TSP is solved to build the routes respecting the problems constraints (Dulac et al. (1980), Chapleau et al. (1985) and Bowerman et al. (1995)).

"Route-first" approach, also known as route partitioning, is based on the method proposed by Beasley (1983) for the Vehicle Routing Problem (VRP). As the former this is also compounded by two phases, initially the vehicle capacity is relaxed and a "giant tour" is build (TSP tour), then the TSP tour is split into feasible trips applying different operators (Newton and Thomas (1969) and Bodin and Berman (1979b)).

In both cases, after define the initial solution, improvement heuristics are applied to enhance the solution. Newton and Thomas (1969), Dulac et al. (1980), Chapleau et al. (1985) and Desrosiers et al. (1986) apply 2-*opt* method, while Bennett and Gazis (1972) and Bodin and Berman (1979b) apply 3-*opt*. The author define whether use best or first improvement and the search stops in a local optimum when none improvement can be made.

Time window adjustment: Many works treat the starting and ending times as constraints of the problem, however, some authors consider that those times are different for each school, allowing them to adjust the buses travel according with the starting and ending time classes.

Considering that times are different, the buses can be schedule to more then one trip, reducing the travel costs. This flexibility does not exist in Brazil the vehicle can only be used in one route and the fact of having long routes can increase the final costs. Fügenschuh (2009) deal with the problem of scheduling school starting times allowing student transshipment from a route to another, but the transshipment issue is discussed latter.

Route scheduling: This stage aim to define the routes sequence. Some authors specify the start and ending time of each route to create a sequence of routes for the same bus.

Newton and Thomas (1974) admit different period times for the classes, so they develop a model for multiple schools to construct routes for a scholar district using less buses and saving in final costs.

Bodin and Berman (1979b) does the same assumption so his problem could be solved by time period. The final route is combined with a posterior step or with the starting time of the next class, allowing use buses in two or more routes. This approach can not be applied if the starting times overlap, because the time required for one path may exceed the start time of the next school class leading to children delay.

Braca et al. (1997) adopt a different method from literature to solve the problem. While the most works solve the problem for each school separately, they deal with it considering the set of schools of New York at the same time. Li and Fu (2002) apply the shortest path algorithm to generate the initial route and an improvement algorithm to enhance it.

Spada et al. (2005) consider multiple schools and presented a heuristic method to solve the problem. The schools are arranged according to its starting times and the routes are constructed considering a greedy method. If possible the routes are moved in together and the result improvement is made using Simulated Annealing or Tabu Search.

Most of the works about school transportation deal with those subproblems separately and in sequence, what does not mean they are independents. Rather, they are highly correlated, however, need to be solved in stages because of their complexity (NP-hard) (Park and Kim, 2010). Although being complex they can be applied as a combination of two or more subproblems and can be considered as a variant of a known optimization problem, what makes hard to define them some times.

2.1.2 Classification

The target of this section is review the technical features of the School Bus Vehicle Routing Problem. Among a countless number of approaches for the SBRP, only few of them were chosen to be reviewed, the choice was made based on the most important works found in the literature and those which are more relevant to this research.

Number of schools: The SBRP can be solved considering one school or multiple schools at the same time. Real problems usually deals with a network compounded by several schools, however, for the simplicity of deal with one school at a time, this approach is more common in the literature (Gavish and Shlifer (1978), Bowerman et al. (1995), Corberán et al. (2002), Li and Fu (2002), Andersson and Lindroth (2005), Schittekat et al. (2006), Pacheco and Martí (2006), Martinez and Viegas (2011), Ledesma and Gonzalez (2012) and Euchi and Mraihi (2012)).

To solve multiple schools problems and generate the routes the author can consider the students residences or the schools (Spada et al., 2005). For students residences approach, the nodes are inserted in the routes in order to attain the best arrangement of visits (Braca et al., 1997). The set of nodes include the schools which is also inserted where the result for cost is smaller. This method allows students from different schools to be transported by the same bus what is called in literature by mixed load and is applied by Verderber (1974), Chen et al. (1990), Braca et al. (1997), Thangiah et al. (2008), Spada et al. (2005), Andersson and Lindroth (2005) and Park et al. (2012). When considering the schools, the set of routes is generated for each school at a time, the buses are assigned to them and the school bell time adjustment is made according to their time windows and constraints.

Urban and Rural problem: The solution method to be applied depends on the context of the problem. Urban and rural areas are manage differently because they have distinct characteristics mentioned by several authors.

Bodin and Berman (1979b), Chapleau et al. (1985), Bowerman et al. (1995) and Simchi-Levi et al. (2005) state that in urban areas the bus capacity is reached before the limit travel time due to the high volume of students in those areas. This statement can not be applied to rural areas, because the population density in rural areas is smaller and even if the maximum riding time is exceeded is unusual to reach the bus capacity.

In Brazilian rural areas, however, the bus capacity is frequently reached and there are cases that students have no option but travel standing. The lack of buses to attend the demand has increased because of the nucleation process that rural areas have being passing through (Sanches and Ferreira (2006), Sales (2013)) and will be explained latter in this chapter.

Single and Mixed loads: This characteristic covers the allowance of students from different schools travel in the same bus or not. The single load plan does not allow the transportation of students from different schools at the same bus. This hypotheses is very restrictive and can lead to an excessive use of buses when the students are further afield.

When mixed load is allowed the routes flexibility and the use of buses are increased and the final costs tend to reduce (Braca et al., 1997). Bodin and Berman (1979b) claim that because of the low population density and the possibility of using fewer buses, the mixed load system is more suitable in rural than urban areas. Despite being addressed by several authors (Verderber (1974), Chen et al. (1990), Spasovic et al. (2001), Thangiah et al. (2008), Park et al. (2012) and Ledesma and Gonzalez (2012)), only Braca et al. (1997) has proposed an algorithm to solve it. His approach is based initially in a insertion rule, verifying two consecutive nodes and if a bus stop can be inserted between them.

Chen and Kallsen (1988) developed a system for rural schools considering routing and bus scheduling. The routing phase define routes for each bus and school, while the scheduling consider the operational time of the buses. The vehicles can be used in multiple travels and mixed load is allowed, the solution has to attend the routes balance, satisfy the constraints of bus capacity, students travel time and time window and has to minimize the number of vehicles and total travel time.

Fleet: When adopting homogeneous fleet, the author assumes that the fleet has the same characteristics (Corberán et al. (2002), Pacheco and Martí (2006), Martinez and Viegas (2011), Euchi and Mraihi (2012), Kim et al. (2012)).

Newton and Thomas (1974) consider a homogeneous fleet capacity, however, the maximum load of each bus is defined according to the schools policies about the number of students sitting or standing. Bowerman et al. (1995) do the same assumption, but they consider that each children has different weight, so each one represent a different type of load. For the authors, a student in the first grade occupies just 2/3 of a regular student, therefore, two buses with same capacity can transport different amount of students.

Problems that admit heterogeneous fleet assume that vehicles have different characteristics such as capacity, fixed and variable cost and riding times (Li and Fu (2002), Thangiah et al. (2008), Spada et al. (2005), Andersson and Lindroth (2005), Fügenschuh (2009) e Ledesma and Gonzalez (2012)). The heterogeneous fleet problem is similar to the VRP whit heterogeneous fleet, thus, is considered its variant.

Transshipment: This assumption allows that students transshipment from one route to another. Usually, a small vehicle performs the first stage of the travel, taking the students who live in remote places or whose residences are located in roads of difficult access . They are driven to a transfer point where a bigger vehicle gather all the students from that area and does the last part to school.

There is not many works about transshipment for school bus routing problem. In the performed literature review only three papers about the theme were found: Baldacci et al. (2004), Andersson and Lindroth (2005) and Fügenschuh (2009).

Objectives: The objectives to be adopted depend on the problem and the priorities of the decision maker. The objective function in mathematical models ensure that limited resources are being used in the best way in order to attend the demand, reducing costs, increasing profits or service level.

Usually, the minimization of the number of buses (Li and Fu (2002) and Pacheco and Martí (2006)) and the total travel time or distance (Li and Fu (2002) and Schittekat et al. (2006)) are the most common objectives adopted, separately or together (Corberán et al. (2002), Li and Fu (2002), Thangiah et al. (2008), Schittekat et al. (2006), Fügenschuh (2009), Park et al. (2012) e Ledesma and Gonzalez (2012)).

Despite not being the most explored objective, some articles address methodologies to analyze the quality of the service level, after all, school transportation is a public service and it might be verified. Savas (1978) and Bowerman et al. (1995) discuss about three factors to evaluate the service level: efficiency, effectiveness and equity.

Efficiency is defined as the ratio of the service level to the cost of the resources required to provide such service. For a fixed service level, the efficiency can be determined by its cost.

Effectiveness can be measured by customer satisfaction, how good the demand is attended. An effective school bus system should be available to all eligible student with a great service level. The effectiveness of a school network service can be determined measuring the total riding time Spada et al. (2005) or the total walking distance from his residence to the pick up point (Chapleau et al. (1985) and Bowerman et al. (1995)).

Equity analyze the fairness or impartiality of providing the service in question. An efficient solution can offer lower costs and travel times, however it can be unacceptable due to the unequal service level, such as uncomfortable trips. The measure of service level has been omitted in school transportation so as in others public services as education and health. However its importance has been recognized with the increasing amount of papers which consider this parameter as objective functions or constraints (Chapleau et al. (1985), Li and Fu (2002), Bowerman et al. (1995) and Jozefowiez et al. (2009)).

Constraints: So as objective functions, the constraints can vary with the problem specificities. There is a large number of constraints in literature once most papers deal with specific problems leading to an wide variety of them. Braca et al. (1997) and Spada et al. (2005) show some examples:

- vehicle capacity - refers to the maximum number of students that can be transported in a bus at the same time (Baldacci et al. (2004),Schittekat et al. (2006), Baldacci et al. (2007), Euchi and Mraihi (2012)),
- maximum riding or travel time - refers to the maximum time that the student can stay in a bus to reach the school (Verderber (1974), Chen et al. (1990), Martinez and Viegas (2011), Park et al. (2012)),
- maximum walking distance - refers to the limit distance the student can walk to the bus stop (Bowerman et al. (1995), Braca et al. (1997)),
- time window - refers to the time range within the buses has to drop of the student at the school (Braca et al. (1997), Spada et al. (2005), Andersson and Lindroth (2005), Kim et al. (2012)),
- constraint about the minimum number of students necessary to create a route (Braca et al. (1997)),
- boarding time - is related to the earlier time that a bus can pick up the first student (Braca et al. (1997), Martinez and Viegas (2011)).

Occasionally, those constraints can be found as objective functions. Bennett and Gazis (1972) and Li and Fu (2002), for instance, consider the maximum travel time as objective in order to minimize the total riding time spent by the students. Bowerman et al. (1995) adopt as objective function the total walk distance, while Desrosiers et al. (1981), Bodin et al. (1983) and Fügenschuh (2009) assume as decision variable the school time window, they claim that adjusting the school time window can reduce the number of buses needed.

2.1.3 Mathematical formulations and Solutions methodology

Bowerman et al. (1995) prove that the two subproblems, bus stops selection and routes generation, combined or not, are NP-hard problems. In bus stops selection each student has to be assigned to a bus stop and each bus stop has its capacity. Using this constraints, the subproblem can be converted in a generalized assignment problem which is also NP-hard (Fisher et al., 1986). The problem of routes generation with capacity constraint and maximum travel time corresponds to the capacity and distance constrained Open VRP which is also known as a NP-hard problem. Because of the computational complexity of those problem, most authors use to solve them with heuristic approaches rather than exact.

Usually, mathematical models are developed as mixed integer programming (MIP) or as nonlinear mixed integer programming (NLMIP). However, most of them have not been used directly to solve the problem but only as part of it.

Gavish and Shlifer (1978) consider for a single school problem a column generation and present a NLMIP for the problem, generate upper bounds and solve a sequence of assignment problem, the optimal solution is define using branch-and-bound procedure. Bowerman et al. (1995) consider simultaneously the bus stop selection and routes generation subproblems. He present a NLMIP model but the mathematical formulation is not used to solve the problem.

Li and Fu (2002) develop a multi-objective NLMIP to generate routes, while Ripplinger (2005) develop a MIP model for the SBRP, however, the mathematical formulation also is not used to solve none of the both problems. The mathematical formulation of Kara and Bektas (2006) is assumed for a single school as a travel salesman problem with multiples depots and single destination. Bektas and Elmastas (2007) work in a formulation for single school and use it to solve the problem.

Schittekat et al. (2006) assume a homogeneous fleet and that school and garage are located at the same location, i.e., the starting and ending points of the routs are the same, however, the model is very simple and do not consider practical constraints as maximum travel time. A MIP model is developed to solve a problem with 10 stops and 50 students.

Ledesma and Gonzalez (2012) address the Multiple Vehicle Traveling Purchaser Problem, a variation of the school routing problem, that considers certain constraints on each bus route, such as bounds on the distances traveled by the students, bounds on the number of visited bus stops, and bounds on the minimum number of students that a vehicle has to pick up. They propose a branch-and-price algorithm to solve a three index variable formulation. Park et al. (2012) propose a mixed load improvement algorithm and some benchmark problems. They measure its effects on the number of required vehicles.

Schittekat et al. (2013) develop a mixed integer problem model where they have to decide the set of stops to visit, the routes along the stops and assign the students. They apply an exact algorithm to assign the students and a meta-heuristic GRASP+VND to define the stops and routes.

The most common heuristic methods used to solve the SBRP are reviewed in section 2.1.2 in bus stop selection and routes generation section. The application of meta-heuristics such as Simulated Annealing (SA), Deterministic Annealing (DA), Tabu Search (TS), Genetic Algorithms (GA), Ant Colony Optimization (ACO) and Neural Networks (NN), have been proved to be a good approach to solve combinatorial optimization and more specifically VRP problems (Gendreau et al. (2002), Langevin and Riopel (2005), Bin et al. (2009)), however, few papers apply them to solve school bus routing problems.

Thangiah and Nygard (1992) apply Genetic Algorithms (GA) to minimize the vehicle fleet and the travel distance using the GENROUTE system to generate the routes of two scholar districts, while Corberán et al. (2002) apply a Scatter Search to improve the initial solution generated by two heuristics based on clustering mechanisms.

Spada et al. (2005) use Simulated Annealing and Tabu Search to improve the initial solution generated by heuristic insertion. Ripplinger (2005) use the clustering algorithm to generate an initial feasible solution and Tabu Search is used to improve it. Pacheco and Martí (2006) construct a set of feasible initial solution using heuristics from Corberán et al. (2002) and Fisher et al. (1986) besides an insertion mechanisms, the obtained solutions are improved using Tabu Search. Thangiah et al. (2013) propose a formulation to the problem and use Genetic Algorithm to find the initial solution and improve it with intra and inter routes movements.

2.2 The rural bus vehicle routing problem and similar problems

Many real problems involve the issue of vehicle routing, thus the increasing of scientific studies about the theme has led to an indescribable use of terminologies to define different classes of it. Therefore, the aim of this section is not to do a review about the School Bus Vehicle Routing similar problems but elucidate the differences between them.

The definition of the Vehicle Routing Problem (VRP) (Dantzig and Ramser, 1959) states that m vehicle located at the depot have to attend n customers aiming to minimize the overall transportation cost. The solution is a set of routes starting and ending at the depot, satisfying the constraints of capacity and that all the customers must be visited only once.

Another intensively studied problem in Combinatorial Optimization is the Traveling Salesman Problem (TSP), which consists in finding among a set of nodes the shortest route that visits each one exactly once and returns to the original point. There is no capacity constraint or previous known demand (Gendreau et al., 2002).

The Capacitated Vehicle Routing Problem has a deterministic demand, previously known, which has to be attended. Its definition consider the Vehicle Routing Problem (VRP), however the m vehicles have identical capacity C which can not be exceed. The delivery has to be accomplished at a minimum total cost and does not have any other constraint such as maximum travel time, time window or maximum riding time.

2.2.1 Pick up and Delivery

For pick up and delivery problems different approaches are observed, all of them, however, focus on the efficient use of a vehicles fleet that must meet the customers demands. Because of the similarities of the pick up and delivery problems, Savelsbergh and Sol (1995) established the General Pick up and Delivery Problem which associates several features found between these problems, whose objective is define a set of routes that meet the demands of pick up and/or delivery minimizing the transportation costs.

Parragh et al. (2008) divide the General Pick up and Delivery Problem into two classes. In the first class, denoted as Vehicle Routing Problems with Back-hauls (VRPB), the transportation is made between customers and depot, i.e., the goods are transported from the depot to line-haul customers and from back-haul customers to the depot. For this class the author consider four subtypes which are: Vehicle Routing Problem with Divisible Delivery and Pickup (VRPDDP), the customers which demand delivery and pickup service can be visited twice, the Vehicle Routing Problem with Simultaneous Delivery and Pickup (VRPSDP), where the customers demanding both services have to be visited exactly once, the Vehicle Routing Problem with Clustered Back-hauls (VRPCB) where all line-hauls are executed before back-hauls, and the Vehicle Routing Problem with Mixed line-hauls and Back-hauls (VRPMB), allows any sequence of line-hauls and back-hauls permitted.

The second class, deals with those problems where goods are transported between pick up and delivery nodes, which are divided into three types of problems: the Pickup and Delivery Vehicle Routing Problem (PDVRP), where pick up and delivery nodes are unpaired, identical good is considered and each unit picked up can be used to fulfill the demand of any other delivery customer, the classical Pickup and Delivery Problem (PDP) and the Dial-A-Ride Problem (DARP). Both types consider requests associated with an origin and a destination, resulting in paired pickup and delivery nodes. The PDP deals with the transportation of goods while the DARP deals with passenger transportation.

2.2.2 Ring-star

The Capacitated m-Ring-star problem (CmRSP) is a variation of the Capacitated Vehicle Routing Problem with single depot, where the customer can be at the route, connected to another customer or to a node denoted Steiner node, defined as a transition points (Hoshino and Souza, 2012).

The problem consists of designing a set of m cycles (rings), corresponding to the number of buses, with capacity Q to attend all customers and minimize the final costs, which is computed by the costs related to edges in the route and the arcs connected to the points in the ring (Baldacci et al., 2007). Each solution includes a central depot, a certain number of customers, and maybe the transition points that can be used to save routing costs.

2.2.3 Multi-echelon Vehicle Routing Problem

In multi-echelon Vehicle Routing Problems, the delivery from the depot to the customers is managed by rerouting and consolidating the freight through different intermediate satellites. The general objective is to ensure an efficient and low-cost operation of the system, while the freight is delivered on time and the total cost of the traffic on the overall transportation network is minimized. Usually, capacity constraints on the vehicles and the satellites are considered (Baldacci et al., 2007).

More precisely, in the Multi-echelon VRPs the network can be decomposed into $k \geq 2$ levels. The 1st level connects the depots to the first level satellites, the intermediate levels interconnecting the satellites and the last level the freight is delivered from the satellites to the customers. Each transportation level has its own fleet to manage the delivery and the vehicles assigned to a level can not be reassigned to another one. The most common version of multi-echelon VRPs applied is the Two-Echelon Vehicle Routing Problem, where just two levels are considered (Perboli et al., 2011).

2.3 The rural bus vehicle routing problem in Brazil

Transport student to and from school is one of the biggest challenges for the Brazilian educational authorities. In urban centers the issue is not that serious, the students receive tickets for free or with discounts to get the shuttle transportation to go to and back school.

The deal is, indeed, in rural areas where education units have been shut down for economics and social reasons and the students have been moved to school in central areas. This transportation is constitutionally guaranteed but in most rural counties they are deficient and the expenses are too high in order of the low population density and bad roads conditions. It is one of the most big barriers that students face to attend classes.

Many of those counties do not have qualified workers to define routes, select students, designate and assign buses to the routes and still consider cost, safety, time window, travel time and vehicles capacity. Thus, the task of routing is barely done in those areas increasing the expenses, travel time, buses use and its maintenance, drivers work hours and reducing the students performance and their assiduity to classes.

Besides, the rural school bus transport has some negative aspects such as poor roads condition, inappropriate and old vehicles, overcrowded buses and a lack of proper planning of routes evidenced by excessive long paths. Long distances is a critical factor in Brazilian rural areas because, eventually, children need to walk to the bus stops. The maximum allowed walking distance found in literature was around 2 or 3 kilometers (GEIPOT (1995), Vasconcellos (1997)) and 45 minutes was the maximum travel time (Arantes, 1986). However, in Brazil, some routes has over 140 kilometers, of which 30% of those has more than 50 Km, 32% are 60 and 90 minutes long, 13% has routes over 2 hours and a few of them last more than 4 hours one way (CEFTRU (2007), CEFTRU (2009)).

More than 90% of the roads used by the rural scholar buses are unpaved or does not receive maintenance very frequently, so the uncomfortable travels let the children stressed and tired reducing their performance or their assiduity to school (Carvalho et al., 2010). Depending on the season the buses have to cross muddy roads requiring special kind of fleet. When the roads are impassible and the buses are not able to cross it, the students remain unable to go to school until the season ends.

The students dispersion is another particularity of the Brazilian problem. This factor reduces the number of children per stop and increases the travel time because the buses have to stop in a higher number of nodes and travel longer distances to pick up the infants. Thus it demand higher number of vehicles to attend all the students. The fixed starting and ending times and the short time window between the morning and afternoon classes of Brazilian schools avoid the reuse of buses requiring, then, a bigger fleet for this scenario.

An approach which could deal with the students dispersion and small time windows would be he transshipment points strategically located and where the students would be transfered from a vehicle to another. The small vehicle would pick up the children and bring them to the transfer points where a bigger bus would make the remaining way to school.

People are convinced that urban transportation requires a special attention because its importance in facilitating the transport of people, and reducing traffic and hazard gases emission. However the efforts for rural community has not to be smaller, the needy of investments is so big as it is in urban transportation because this service can avoid the spacial segregation, rural flight and support their inclusion in urban areas and maintain the cultural and social life of rural community.

However, only few studies has being conducted in Brazil about the access and transportation of students to school in rural areas. The pioneer studies were conducted by an already closed Enterprise of Transportation Planning (GEIPOT) in 1995 (Carvalho et al., 2010) followed by Vasconcellos (1997), Pegoretti and Sanches (2004), Carvalho et al. (2010), Carvalho and Yamashita (2011) and Mandujano et al. (2012).

Therefore, studies in rural transportation are helpful in social and spatial integration. For that purpose this thesis aim to present mathematical models, single and multi-objective heuristics to represent and solve the rural student transportation problem.

2.4 Research Opportunities

Many efforts have been made to improve the service level, routing costs and travel times in rural school transportation system, but this service has a lot to be improved yet. Thus, several issues still remain open for future research.

Because of its computational complexity, few papers deal with exact methods (Letchford et al. (2006) and Bektas and Elmastas (2007)) what makes it a promising field of research. The incorporation of school location decisions combined with maximum riding distance constraints is one of the propositions which can be made for rural areas.

Another especial issue for those areas is the possibility of creating transition points. The transshipment allows that smaller vehicles reach areas that bigger ones can not, driving the students from home to these points where buses with higher capacity would do the second part until the school. The needy of studies about transshipment is stated by Park and Kim (2010) and Park et al. (2012). For extensive areas where students and schools are dispersed, as the Brazilian case, the schools could work as transshipment points reducing the number of required buses.

Chapter 3

Heuristic algorithms for the Brazilian context

Chapter information

The content of this chapter was subjected to the Expert Systems with Applications journal on january 24th of 2015 and was slightly modified in order to suit the thesis standard.

3.1 Introduction

Brazil has over 50 million students enrolled in a complex public educational system that encompasses more than 198 thousand state and municipal public schools (elementary, middle and high schools). 14% of these pupils are located on rural areas which are usually served by multigrade rural public schools (24% of the total number of schools), i.e. schools in which groups of students of different grades are placed in a single classroom and are typically taught by only one teacher (INEP, 2013; Carvalho et al., 2010).

On one hand multigrade schools have the advantages of having: flexible schedules, proximity to the community where they are placed, and the development of unique programs to meet student's individual needs in order to offer opportunity for them to become independent learners. On the other, multigrade schools generally have the disadvantages of having: inadequate facilities, poor trained teachers, scarcity of varied levels and types of materials, limited or no access to different types of more advanced curriculum activities, and the absence of sport infra-structure (Vincent, 1999).

Concerned with these downsides, the Brazilian federal government has been making a great effort over the last twelve years to provide a better education to the rural population by encouraging and financing state and municipal authorities to close their multigrade rural schools and transfer those affected students to better structured, centrally located, school facilities with single-grade classes. To achieve that several government programs have been devised to assist this endeavor. One in particular referred to as the *Way to School program* is responsible for improving the mobility of the students by providing new better adapted to the often severe operational conditions buses to the municipal administrations. The city is then responsible to carry back and forth the students from their homes to their respective public schools (state and municipal).

However, due to the lack of qualified technicians, to the great social and cultural diversity found in many regions of Brazil, and to its extensive land, it has been a great challenge for public authorities to manage rural school transportation services (Carvalho et al., 2010). As the provided resources are scarce, it has not been possible to dedicate buses to a single school as typically found in the school bus routing literature (Park and Kim, 2010). Public managers are having to deal with a more complex problem in which buses are required to serve to both state and municipal schools at the same time. Students from different schools have to ride on the same bus at the same time in order to go to their respective schools, i.e. buses carry mixed loads of students prior to drop them off at different destinations before returning to the garage. To complicate matters even more, managers have often to plan the routing for a heterogeneous fleet.

Figure 3.1 illustrates an example of such problem: Thirteen students (stick figures) scattered in six bus stops must be carried to their respective schools represented by triangles numbered 7-9. The students are associated to each school by a color scheme, e.g. filled stick students go to the filled triangle, unfilled stick students go to the unfilled triangle and so on. Two bus routes leaving the garage (square node), picking up the students and delivering them to their respective schools prior to returning to the garage are pictured as filled and dotted arrows. Note that (*i*) students of different schools are carried by the same bus, (*ii*) the picking operation is carried out prior of the delivery, (*iii*) each student bus stop is visited exactly once; (*iv*) each school node can be visited by more than one bus, though each bus can visit a school only once; (*v*) the schools visited by a bus depend only on the students being carried by it, i.e. a bus is not required to visit all of the schools. These features give raise to a problem known as the rural school bus routing problem with heterogeneous fleet and with mixed loading (Park and Kim, 2010).

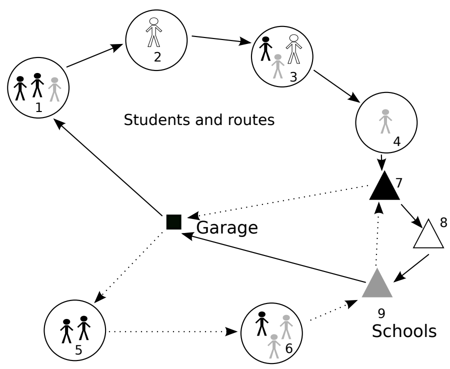


Figure 3.1: Example of a school bus routing problem with mixed loading.

The mixed loading assumption is first considered by Bodin and Berman (1979b) who point out that it can occur frequently in rural areas, however they do not present any methodology to deal with it. Chen and Kallsen (1988) also remark that having buses to transport students of just one school at time (single load) can result in the deployment of an excessive number of vehicles, specially when dealing with remote pupils who live in low density areas. They propose an expert system to aid the routing and scheduling of buses for a rural school system in which routes are manually generated.

Actually Braca et al. (1997) are the first authors to directly address the mixed load bus routing problem by proposing an insertion procedure. The devised method constructs each route by randomly selecting a bus stop and inserting it and its respective associated school into the route at the best cost estimation possible, but making sure that the time window and capacity constraints are satisfied. The objective is to minimize the number of used buses only, and not the routing costs. The authors also state, without reporting quantitative data, that an increase of flexibility and cost savings can be achieved whenever buses carry mixed loadings.

Spada et al. (2005) devise a decision-aiding methodology for a school bus routing and scheduling system in which the number and the types of buses are given a priori. The authors explicitly optimize the level of service provided by the bus operator while allowing mixed loading. The level of service is represented by two objectives: the student's *time loss* and the *maximum time loss*. The time loss of a student is calculated as the sum of the delays and the waiting time. The delays are the difference between the actual journey time of each student and the shortest possible time between their bus stop and their respective school. While the waiting time is the time spent by the pupils waiting for class to begin at their school. Spada et al. construct initial routes by sorting the bus stops associated to each school in decreasing order of their distance to their respective school. Then the bus stops are inserted in a greedy fashion to form the routes, while respecting the considered constraints (time and capacity). A local search which exchanges non-bus stops from different routes is then applied to improve the initial solution. This local search is embedded into two heuristic strategies (simulated annealing and tabu search) to see which one perform better (simulated annealing).

Park et al. (2012) improve the method of Braca et al. (1997) by devising a post improvement procedure similar to the neighborhood of Spada et al. (2005). Starting from a solution with a dedicate fleet per school obtained from a sweep based algorithm (Gillett and Miller, 1974a), the procedure reallocates one bus stop at a time in a greedy way until routes can be merged or deleted. Time windows and capacity constraints are respected during the bus stop reassignments. They also adapt a formulation for the pickup and delivery problem with time windows and heterogeneous fleet. They split each bus stop with students of different schools into new stops, one for each school, while preserving the location of the original bus stop (e.g. in Figure 3.1 node one would be divided into two new nodes, one for each school). Then, for each student bus stop, a virtual node is created having the location of the school associated to the student bus stop so that precedence constraints can be imposed in order to ensure that these virtual nodes are only visited after their respective student bus stops. Buses' capacities, time windows for each school, and the maximum riding time for the students are enforced by proper constraints. The objective of the formulation is to find a set of feasible routes with a minimum number of vehicles while allowing mixed loads. The proposed formulation allows for a bus to visit the same school several times if the time constraints permit.

In most mixed loading problem variants addressed in the literature (Bodin and Berman, 1979b; Chen and Kallsen, 1988; Braca et al., 1997; Spada et al., 2005; Park et al., 2012), it is assumed that schools have different starting times spanned in a large time interval. For instance, Park et al. (2012) consider on their computational tests that the starting time of each school can occur on an interval from 7 to 11AM with a 30 minute time window; while Fügenschuh (2009) uses an interval from 5 to 9AM. This allows for buses to be reused on different routes at the same day, though school time window constraints need to be imposed. By exploiting these large time spans, some authors (Desrosiers et al., 1981,9; Fügenschuh, 2009) have an indirect way to minimize the number of vehicles. They propose methods to adjust the school bell times, i.e. to determine new starting and ending times for schools in order to maximize the number of routes that can be done by the same bus. That way, the number of buses employed can be reduced.

Unfortunately, the aforementioned assumptions are not valid for all contexts, specially in Brazil where all (municipal or state) schools start at the same time because of labor regulations. Moreover due to Brazil's road conditions and to the dispersion of the population on its rural areas, the imposition of maximum riding times for the students is not always possible. Carvalho et al. (2010) present a good description of the situation of the Brazilian rural bus school transportation services. According to them, 92% of the rural roads are unpaved and ill maintained, being one of the main reasons for the large riding times, e.g. 33% of the rides last from 60 to 90 minutes, but there are travels with time lengths greater than 2 hours (15% of them). Further, 28% of the routes are longer than 37 miles, with many routes being as long as 125 miles.

Moreover, as all Brazilian's public schools start at the same time, it is not on the best interest of the welfare of the students to arrive early at the schools so that buses can be reused to do other routes. So Brazil has adopted the policy that all buses must pick the students up first and then deliver them to their respective schools with no student pickups between schools. One may wonder about how this policy may shorten the possibility of cost savings when student pickups between school deliveries are disregarded. In fact, this is indeed true. Nevertheless cost savings can still be achieved if fleet and routing costs are both taken into account in a mixed load routing plan. Though Braca et al. (1997) and Park et al. (2012) state that the reduction of one bus results in a larger cost savings than reducing the routing costs; for the Brazilian context, this turned out to be the opposite. In fact, for low density populated regions with student bus stops scattered far apart from each other and far from the school locations, the routing costs play an important role in the overall costs, as can be seen in a real case presented in Section 3.4.5.

In the present article, five different meta-heuristics are devised to solve the rural school bus routing problem with heterogeneous fleet and with mixed loads for the Brazilian context: The first one is an adaptation of the Record-to-Record Travel algorithm for solving the heterogeneous fleet vehicle routing problem proposed by Li et al. (2007); the other two are based on an Iterated Local Search (ILS) meta-heuristic (Lourenço et al., 2010); while the last two use a Variable Neighborhood Search (VNS) (Hansen and Mladenović, 2001; Hansen et al., 2010) strategy. All meta-heuristics employ five neighborhood structures which are organized or in a fixed or in a random Variable Neighborhood Descent (VND) local search (Mladenović and Hansen, 1997). To assess the performance of the devised algorithms, three different types of data sets, totaling 150 test problems ranging from 250 to 2000 bus stops and transporting from 3,204 to 27,594 students, are used on the computational experiments. Besides showing that the ILS algorithm with a random VND out-performs the other methods, the results also show that the policy of having mixed loads on the buses, while minimizing fleet and routing costs, allows for greater saving costs when compared with the procedure of Park et al. (2012) for minimizing the number of vehicles or with solutions having buses with single loads, and a dedicated fleet per school. Furthermore, a real instance extracted from a county located in the state of Minas Gerais, Brazil, is presented to illustrate the suitability of the aforementioned policy for the Brazilian context.

The remainder of this chapter is organized as follows. Section 3.2 presents the notation and the proposed formulation for the problem. In Section 3.3, the developed meta-heuristics are described, while Section 3.4 reports the achieved results on the benchmark instances. Concluding remarks are drawn in the last Section 3.5.

3.2 Notation, definitions and formulation

The rural school bus routing problem with heterogeneous fleet and with mixed loads can be defined on an undirected graph $G = (N, E)$ in which the set of nodes $N = \{0\} \cup P \cup H$ is the union of the garage node with the sets of nodes $P = \{1, \dots, n_p\}$ and $H = \{n_p + 1, \dots, n_p + n_s\}$ representing where the students (pupils) and schools are located, respectively; and E is the set of edges. For each node $i \in P$, there is a set of schools requisitions $K_i \subseteq H$ in which $d_{ik} > 0$ students from node $i \in P$ must be driven to school $k \in K_i$. All of the $q_i = \sum_{k \in K_i} d_{ik}$ students of node $i \in P$ must be picked up by the same bus at the same time and driven to the schools of set K_i. Let B be a set of heterogeneous buses stationed at garage node 0 and with capacities Q_b and fixed costs a_b, for all $b \in B$. Without loss of generality we assume that $Q_1 \leq Q_2 \leq \dots \leq Q_{|B|}$ and $a_1 \leq a_2 \leq \dots \leq a_{|B|}$. Let also every edge $(i, j) \in E$ have a nonnegative cost $c_{ij}^b = \tau_b \ell_{ij}$ associated to it depending on which bus $b \in B$ is using it, where τ_b is the cost per unit of traveled distance for bus $b \in B$ and ℓ_{ij} is the distance between nodes i and j, $(i, j) \in E$.

For a subset F of E, let $G_b(F)$ denote the subgraph $G_b(N(F), F)$ of bus $b \in B$ induced by F, in which $N(F)$ is the set of nodes incident to at least one edge of F. A *route* of a bus $b \in B$ is then defined as a nonempty subset $R_b \subset E$ of edges for which the induced subgraph $G_b(R_b)$ is a simple cycle having node 0 and such that the total demand of nodes in $N(F) \setminus (\{0\} \cup H)$ does not exceed the capacity of bus b. Such cycle represents the route of bus b leaving the garage, picking up the students at nodes in $N(F) \setminus (\{0\} \cup H)$ first, delivering them to the school nodes in $N(F) \setminus (\{0\} \cup P)$ and returning to the garage. The cost of such a route is given by the sum of the costs of the edges forming the cycle and by the fixed cost of bus b. The problem objective is then to design the least cost routes such that all of the student nodes are visited exactly once and delivery to their respective school nodes.

One way of formulating the aforementioned problem is by extending the two index formulation introduced by Laporte et al. (1985) for the capacitated vehicle routing problem and combining it with a variable size bin packing model (Dyckhoff, 1990). Let $y_i^b \in \{0,1\}$ be a integer variable equal to 1 if node $i \in N \setminus \{0\}$ is assigned to bus $b \in B$; 0, otherwise. Let $z^b \in \{0,1\}$ be a integer variable equal to 1 if bus $b \in B$ is set to do a route; 0, otherwise. Let $x_{ij}^b \in \{0,1\}$ be an integer variable equal to 1 if edge $(i,j) \in E$ is used on the route of bus $b \in B$; 0, otherwise; The formulation for the mixed load capacitated rural bus routing problem can be formulated as:

$$\min \phi(z,y,x) = \sum_{b \in B} \left[a_b z^b + \sum_{(i,j) \in E} c_{ij}^b x_{ij}^b \right] \tag{3.1}$$

$$\text{s.t.:} \sum_{b \in B} y_i^b = 1 \qquad \forall i \in P \tag{3.2}$$

$$\sum_{i \in P} q_i y_i^b \leq Q_b z^b \qquad \forall b \in B \tag{3.3}$$

$$y_i^b \leq y_k^b \qquad \forall i \in P, k \in K_i, b \in B \tag{3.4}$$

$$\sum_{(i,j) \in E} x_{ij}^b + \sum_{(j,i) \in E} x_{ji}^b = 2y_j^b \qquad \forall j \in N \setminus \{0\}, b \in B \tag{3.5}$$

$$\sum_{(0,j) \in E} x_{ij}^b = 2z^b \qquad \forall b \in B \tag{3.6}$$

$$\sum_{(i,j) \in E : i \in P, j \in H} x_{ij}^b = z^b \qquad \forall b \in B \tag{3.7}$$

$$\sum_{(i,j) \in E : i,j \in S} x_{ij}^b \leq \sum_{i \in S \setminus \{k\}} y_i^b \qquad \forall b \in B, k \in S, S \subseteq N \setminus \{0\} \tag{3.8}$$

$$x_{ij}^b \in \{0,1\} \qquad \forall (i,j) \in E, b \in B \tag{3.9}$$

$$y_i^b \in \{0,1\} \qquad \forall i \in N \setminus \{0\}, b \in B \tag{3.10}$$

$$z^b \in \{0,1\} \qquad \forall b \in B \tag{3.11}$$

The objective function (3.1) minimizes the total cost of setting the buses and forming the routes. Constraints (3.2) assure that each student node is assigned to a bus, while constraints (3.3) ensure that the capacity of each bus is not exceeded. Constraints (3.4) guarantee that whenever a student node is assigned to a bus, the schools associated to this node will also be visited by the same bus. Constraints (3.5) and (3.6) are the degree constraint for each customer visited by a bus and for the garage node whenever a bus is set. Constraints (3.7) allow for only one edge connecting student to school nodes to be used whenever a bus is set. This guarantees that the picking operation is carried out prior to the delivery of the students to their schools. Constraints (3.8) are the well-known sub-tour elimination constraints (SECs) which ensure that the nodes assigned to a bus are connected. Finally, (3.9)-(3.11) are domain constraints.

3.3 Solution approaches

Five meta-heuristics are devised to solve the mixed load rural bus routing problem with heterogeneous fleet. The proposed algorithms are thought to have the least number of parameters to be tunned and to be easy of implementation and use. Further, all designed methods employ an adapted Clark and Wright savings algorithm (Clark and Wright, 1964) for generating a starting solution. Before presenting the devised methods, a few remarks about the employed solution representation are in order.

3.3.1 Solution Representation

Each solution consists of a set of routes which are stored as a doubly-linked list structure proposed by Li et al. (2007). Each route has two parts: one for the student bus stops, and another for the school nodes. This division is important because each student node is visited only once in a solution, while each school can be visited multiple times. Further, this structure stores the routes as the predecessor and the successor nodes of each bus stop of a route, greatly reducing the computational time when performing the local search movements. Moreover, a sorted list of the closest bus stops to each node is assembled. This list speeds up the search process during the local search. Each list has the nodes that are within 60% of the largest distance between the bus stops of the instance being addressed. The aforementioned representation is used in the developed algorithms which require a starting feasible solution.

3.3.2 Starting Solution

To generate a starting feasible solution s, a modified Clark and Wright savings procedure, depicted in Algorithm 1, is developed. Initially, each node $i \in P$ is assigned to an individual route r_i with the smallest vehicle capacity available and capable of accommodating the number of students q_i at node i (line 1). Route r_i is constructed such that it visits all of the schools associated to node i, i.e. schools in the set K_i. An estimation of the savings (γ_{ij}) of merging each pair of routes r_i and r_j is calculated by pondering a parameter λ (line 2), where $\phi(\cdot)$ is the cost function. Parameter λ implicitly controls the number of route pairs which will be candidates for merging. For small (large) values, λ increases (decreases) the number of pair candidates to be merged. It is then possible to produce many different initial solutions by varying a single parameter (λ). The symbol $\|$ (line 2) represents the best possible merge of the routes respecting the bus capacity and schools to be visited. A student bus stop of r_j is inserted into a route r_i by observing the cheapest insertion possible and the bus capacity. If necessary and possible, a larger bus can be used, though having its cost accounted for. School nodes are also inserted, if required, in the cheapest way possible. Recall that the student bus stops are visited first, and school nodes later. After computing the savings γ, pairs of routes are merged in a descending order (lines 3-4) of the savings.

To merge two routes, say r_i and r_j, each r_j student bus stop is attempted to be inserted into another route r_i by observing its list of closest nodes which are already in route r_i, and then choosing the cheapest insertion cost. If necessary and possible, a larger bus can be set. Otherwise, the node is not inserted. If all of r_j nodes are inserted into r_i, route r_j is deleted. Otherwise route r_j is left with the nodes not inserted. A smaller bus for route r_j can also be set in place of the previous one if necessary and possible. In case school nodes are supposed to be inserted into route r_i, these insertions are carried out in the cheapest way. The algorithm continues with the next pair of routes until no more routes can be merged. Whenever a route can not be further merged, a $2Opt$ optimization procedure (please refer to subsection 3.3.3) is performed to improve the final cost of the resulting route (line 5). During the merge process, it is always observed that the student bus stops are visited first, before the school nodes. At the end, the route is appended to solution S (line 6). This adaptation of the Clark and Wright savings procedure is embedded in all of the devised meta-heuristics which employ, following their respective strategies of solution, neighborhood structures to improve the initial solution.

Algorithm 1 Generating an initial feasible solution

Require: λ
Ensure: s
1: Assign one bus to each node $i \in P$ to obtain route R_i
2: $\forall i, j \in P : i \neq j$ **do** $\gamma_{ij} \leftarrow \phi(R_i) + \phi(R_j) - \lambda\, \phi(R_i \parallel R_j)$
3: Sort γ_{ij} in a descending order
4: Merge routes in the savings order respecting the buses' capacities, to form a new route R_b
5: Improve each route R_b by performing a $2Opt(R_b)$ procedure
6: $s \leftarrow \bigcup_{b \in B} R_b$

3.3.3 Neighborhood Structures

Four neighborhood structures were adapted from the capacitated vehicle routing problem literature (Laporte and Semet (2002), Gendreau et al. (2002)) to improve the initial solution s:

- One Point Move - $\mathcal{N}^1(s)$: this neighborhood structure reallocates a student bus stop in a different position on the solution. It can be applied within and between routes (Li et al., 2007). When performed within a route, this neighborhood may not have a significant impact on the solution or its cost. However when carried on between routes, both affected routes may be significantly altered afterwards, because a group of schools associated to the node being moved may now be required to be visited in the inserted route. Likewise this same group of schools may not need to be visited anymore on the previous route. In Figure 3.2, the left original solution consists of two routes formed by dashed and straight arcs. The routes start and end at the garage (black square), though returning arcs to the garage are not here represented for sake of presentation. The routes visit the student nodes (circles) first and then the school nodes (triangles). Each student node is associated to a school by a color scheme. For instance, gray nodes 1, 2 and 5 go to the gray triangle. In the example, node 5 is moved from the dashed to the straight arc route yielding the solution in Figure 3.2(a). Note that now the dashed route does not visit the gray triangle school anymore.

- Two Point Move - $\mathcal{N}^2(s)$: this neighborhood structure swaps one student node from a route for another one of a different route. Once again, both affected routes may be very different in the end given the group of schools associated to the nodes being moved. In Figure 3.2, student nodes 3 and 5 are swapped in the left original solution obtaining the solution of Figure 3.2(b). Note that the gray triangle school is removed from the dashed arc route, as it is the black triangle for the straight arc route.

- Cross-Exchange - $\mathcal{N}^3(s)$: this neighborhood structure removes one arc from two different routes and reconnects the involved nodes by cross-linking the heads and tails. For instance, if two arcs (i, j) and (i', j') from two different routes are selected, the cross-exchange neighborhood would reconnect the nodes as (i', j) and (i, j') forming two different routes. In Figure 3.2, arcs $(3, grayschool)$ and $(6, 4)$ are selected from the left original solution to be removed and the involved nodes are reconnected as $(6, whiteschool)$ and $(3, 4)$ forming the solution of Figure 3.2(c). Note also that schools are removed and/or inserted on the involved routes.

- 2Opt - $\mathcal{N}^4(s)$: this neighborhood structure removes two non-consecutive arcs of a route and reconnects the involved nodes by linking the heads and the tails together, e.g. given arcs (i, j) and (i', j'), the nodes would be reconnected as (i, i') and (j, j'). An example is provided in Figure 3.2 on the right. Arcs $(5, 6)$ and $(4, 7)$ of the right original solution are removed from the route and the nodes are reconnected as $(5, 4)$ and $(6, 7)$ as illustrated in Figure 3.2(d).

Besides the task of selecting among the nodes or arcs which will produce a better solution than the current one, the first three neighborhood structures require also an extra effort of checking if schools will be inserted into or removed from the affected routes whenever a movement is evaluated. These insertions and removals are very time consuming, and represent a further computational burden. To speed up the process, each neighborhood structure is implemented as a combination of two procedures: one for evaluating the move and another to execute it. It's important to notice that most of the evaluated movements are not executed. So the first function does only the absolutely necessary tasks to assess the feasibility and the improvement on the solution. Once an evaluated move is allowed, the function to execute it does not need to recalculate the savings and how the nodes and arcs will be affected. Further, bus capacities are increased or decreased as needed and possible, but the proper corresponding costs are always accounted for. Finally, though there are other more elaborated neighborhood structures that can be adapted from the capacitated vehicle routing problem literature (Laporte and Semet, 2002; Gendreau et al., 2002), these required adaptations represent a computational challenge that needs to be overcome due to the further effort of the school removals and insertions.

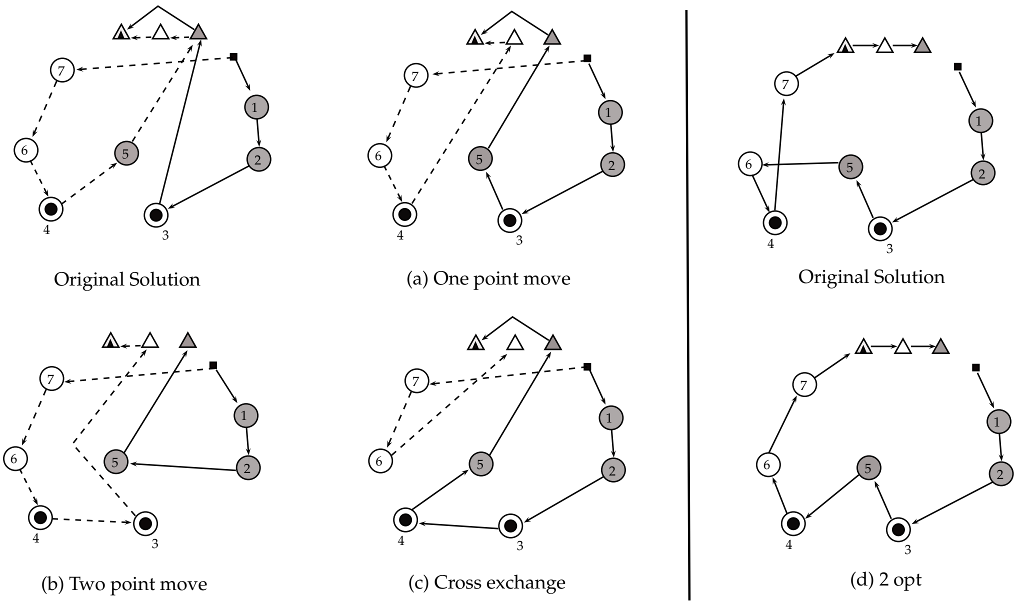


Figure 3.2: Result examples for the local search operators.

3.3.4 Variable Neighborhood Descend Local Search Methods

The four neighborhood structures are organized in a Variable Neighborhood Descend (VND) local search procedure (Mladenović and Hansen, 1997). VND is a meta-heuristic which consists of a permutation $\mathbb{P} = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_{|\mathscr{N}|}\}$ of a set $\mathscr{N} = \{\mathscr{N}_1, \ldots, \mathscr{N}_{|\mathscr{N}|}\}$ of neighborhood structures. These local searches are applied to an initial solution s, one at a time in the order defined by the permutation $\mathbb{P}$, with the objective of finding better solutions. Starting at $\mathfrak{p}_1$, the VND algorithm goes from a local search $\mathfrak{p}_i$ to the next $\mathfrak{p}_{i+1}$, whenever $\mathfrak{p}_i$ fails to find an improving solution. Every time a better solution than s is found, s is updated and the procedure recommences the search at $\mathfrak{p}_1$. The algorithm stops when no improving solution is found by $\mathfrak{p}_{|\mathscr{N}|}$. Algorithm 2 illustrates the aforementioned steps.

As a complete exploration of a neighborhood $\mathfrak{p}_i$ on s may be too time consuming, an alternative adopted policy is to accept the first improved move by a local search rather than the overall best (Mladenović and Hansen, 1997). Furthermore, permutation $\mathbb{P}$ is usually sorted *a priori* by the increasing complexity of computation in finding improved solutions. However a random arrangement of $\mathbb{P}$ is also possible originating what is known as the Random Variable Neighborhood Descend (RVND) algorithm.

3.3.5 Implemented meta-heuristics

Five different meta-heuristics are implemented. The first to be presented is an adaptation of the Record-to-Record Travel (RRT) method proposed by Li et al. (2007) for the heterogeneous fleet vehicle routing problem; followed by the other four algorithms which are the result of embedding the VND and RVND approaches as the local search method into an Iterated Local

Algorithm 2 Variable Neighborhood Descend Algorithm

Require: $\mathbb{P} = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_{|\mathcal{N}|}\}, s$

$k \leftarrow 1$
while $k \leq |\mathbb{P}|$ **do**
    $s' \leftarrow \mathfrak{p}_k(s)$
    if $\phi(s') < \phi(s)$ **then**
        $s \leftarrow s'$
        $k \leftarrow 1$
    else
        $k \leftarrow k + 1$
    end if
end while

Search (ILS) (Lourenço et al., 2010) and a Variable Neighborhood Search (VNS) (Hansen and Mladenović, 2001; Hansen et al., 2010) strategies.

3.3.5.1 A Record-to-Record Travel algorithm

The implemented RRT algorithm (Li et al., 2007), illustrated in Algorithm 3, resembles a deterministic simulated annealing methodology which iterates between two phases: an Uphill (lines 5-8), which can be considered as a perturbation stage; and a Downhill (lines 11-20), which can be seen as an intensification, refinement step. In both phases, the neighborhood structures are used as a local search method to improve a given solution. Though instead of functioning as in a VND algorithm, each neighborhood structure is applied only once in the order prescribed by the permutation $\mathbb{P}$ given rise to the function *OnePass* (lines 6 and 12). Further, each phase adopts a different acceptance criterion for updating the current best solution. While during the Uphill phase, solutions worst than the best current solution, but within a given threshold (δ), are accepted as candidate solutions to be further improved; on the Downhill phase only solutions better than the overall best are accepted. The method iterates for a maximum of *MaxIter* iterations without improvements; while the Uphill phase lasts for *MaxPert* iterations. The simplicity behind the method and its fast running times for generating good solutions for instances of the classical vehicle routing problem with hundreds of nodes are the main motivations for the RRT procedure to be here adapted for the mixed load rural school bus routing problem with heterogeneous fleet.

Algorithm 3 Record-to-Record Travel Algorithm

Require: $\mathbb{P}, s, MaxIter, MaxPert$

1: $i \leftarrow 0$
2: **while** $i < MaxIter$ **do**
3:     $p \leftarrow 0$

{Uphill}
while $p < MaxPert$ **do**
    $s \leftarrow OnePass(s, \mathbb{P}, \delta)$
    $p \leftarrow p + 1$
end while
$stop \leftarrow FALSE$
{Downhill}
while $stop = FALSE$ **do**
    $s' \leftarrow OnePass(s, \mathbb{P})$
    if $\phi(s') < \phi(s)$ **then**
        $s \leftarrow s'$
        $i \leftarrow 0$
    else
        $i \leftarrow i + 1$
        $stop \leftarrow TRUE$
    end if
end while
end while

3.3.5.2 An Iterated Local Search algorithm

The implemented Iterated Local Search (ILS) (Lourenço et al., 2010), represented in Algorithm 4, starts with a solution s and performs, at each iteration, a perturbation followed by a local search procedure. The perturbation function (line 3) randomly selects any neighborhood structure of $\mathbb{P}$, but the $\mathcal{N}^4(s)$ (2Opt), and executes a randomly move. As suggested by Li et al. (2007), the 2Opt was set to be the last neighborhood structure to be executed since it is more interesting to change the nodes of the routes before optimizing the order of visitation. The local search function (line 4) can be based on the VND or on the RVND of Section 3.3.4, resulting then into two different meta-heuristic algorithms. After the excution of the local search function, the current best solution s is updated if the attained solution s' is better than it (lines 5-10). The algorithm iterates for a maximum of $MaxIter$ iterations without improving the current best solution.

Algorithm 4 Iterated Local Search Algorithm

Require: $\mathbb{P}$, s, $MaxIter$
1: $i \leftarrow 0$
2: **while** $i < MaxIter$ **do**
3:     $s' \leftarrow Perturbation(s, \mathbb{P})$
4:     $s' \leftarrow LocalSearch(s', \mathbb{P})$
5:     **if** $\phi(s') < \phi(s)$ **then**
6:         $s \leftarrow s'$
7:         $i \leftarrow 0$
8:     **else**
9:         $i \leftarrow i + 1$
10:     **end if**
11: **end while**

3.3.5.3 A Variable Neighborhood Search algorithm

The Variable Neighborhood Search (VNS), shown in Algorithm 5, deploys a different strategy than the ILS for reducing the risk of being trapped in non-promising search areas. Instead of randomly selecting a neighborhood structure from permutation $\mathbb{P}$ to perform the perturbation of a solution s as in the ILS method (line 3 of Algorithm 4), the VNS procedure applies one neighborhood structure for the perturbation and the local search functions at a time (lines 3 and 4 of Algorithm 5) following the order of permutation $\mathbb{P}$. Further, while in the ILS algorithm the local search function (line 4 of Algorithm 4) is actually a VND method, in the VNS the local search function consists in the use of only one neighborhood at a time. Whenever a neighborhood structure is able to find an improving solution (lines 5-10 of Algorithm 5), the VNS algorithm restarts the procedure by applying the neighborhood structure of the first position of permutation $\mathbb{P}$. The method stops when no further improvement on solution s can be achieved. When using the RVND version, the local search function randomly selects a neighborhood structure which will be used.

Algorithm 5 Variable Neighborhood Search Algorithm

Require: $\mathbb{P}, s$
1: $k \leftarrow 1$
2: **while** $k \leq |\mathbb{P}|$ **do**
3: $\quad s' \leftarrow Perturbation(s, \mathfrak{p}_k)$
4: $\quad s' \leftarrow LocalSearch(s', \mathfrak{p}_k)$
5: $\quad$ **if** $\phi(s') < \phi(s)$ **then**
6: $\quad\quad s \leftarrow s'$
7: $\quad\quad k \leftarrow 1$
8: $\quad$ **else**
9: $\quad\quad k \leftarrow k + 1$
10: $\quad$ **end if**
11: **end while**

3.3.6 A Mixed Load Improvement Algorithm (Park et al., 2012)

In the rural school bus routing problem literature, the works of Park et al. (2012) and Braca et al. (1997) are the only ones which address the mixed load rural bus routing problem with heterogeneous fleet. Actually, the mixed load improvement (MLI) procedure of Park et al. can be considered as an enhancement to the algorithm devised by Braca et al.. It starts with a set of single load routes constructed by using a variant of the sweep algorithm (Gillett and Miller, 1974a) followed by an assignment problem of routes to buses. Then the method attempts to move all student bus stops from a route to other routes through a relocation operator. If it succeeds, the empty route is deleted from the solution. Otherwise, it restores the stops to the original route and selects a different one. The algorithm then iterates until no further routes can be eliminated. Student bus stops are inserted into the other routes at the best possible saving costs, though no further optimization is performed on the routes. One of the key steps of the relocation operator is the feasibility checking of the routes. After moving a bus stop from one route to another, the algorithm has to check whether the bus capacity and the maximum riding time are exceeded or not, and if a new school will be required to be visited or not. When a route is set to stop at a new school, the school node is usually appended to the end of the trip. All of these operations of performing the node relocations without the guarantee of the suppression of the route at the end, together with the route feasibility testings, require a great deal of computational effort, being critical to the method. It is important to remark that Park et al. (2012)'s MLI procedure minimizes the number of buses only, not performing route optimization.

3.4 Computational results

All six algorithms (RRT,ILS-VND,ILS-RVND,VNS-VND,VNS-RVND,MLI) were coded in C++, compiled with GCC 4.8.1., and tested on an Intel Xeon 2.53 GHz with 24GB RAM running Linux Mint 16. For the experiments, three bus types with capacities of 20, 30 and 40 seats were made available to transport the students. The daily fixed depreciation costs (fixed costs) and the routing costs were set to \$100, \$150, and \$200, and \$1.00, \$1.20 and \$1.40 per unit of traveled distance, respectively, for each bus type. The daily depreciation costs were estimated by assuming a bus lifespan of 10 years, while the distances between nodes were considered to be Euclidean.

Four data sets were used in the experiments: *(a)* PARK - 24 instances of Park et al. (2012) were used in the calibration phase. These instances are classified into two types: *(i)* one with randomly distributed schools and bus stops, and *(ii)* other with schools and bus stops gathered in several clusters. The number of students ranges from $3,204$ to $27,594$ scattered in bus stops within the set $\{250, 500, 1000, 2000\}$. The number of schools varies according to the set $\{6, 12, 25, 50, 100\}$. All bus stops are dispersed in an area of $960mi^2$. *(b)* MOD - from each PARK instance 5 new ones were created by modifying node demands. For each node i, the number of associated schools $|K_i|$ was uniformly generated from 1 up to 3 schools. The schools were then uniformly selected among the available schools to form set K_i. For each school $k \in K_i$, a new demand d_{ik} was uniformly created within the range of 1 to 3 students. A total of 120 new instances were built and used on the comparison tests. *(c)* RAND - although instances from Park et al. (2012) are considered to be rural, they do not represent the Brazilian reality. So, besides the PARK and the MOD instances, five new ones were also created to characterize Brazilian counties of different sizes. A total of 2000 students were randomly placed on bus stops scattered over an area with sizes ranging in the set $\{3, 6, 12.5, 18.5, 31\}$ (mi^2). From 10% up to 30% of the bus stops as well as all of the 10 available schools were located inside an imaginary downtown area with a radius of 1.85 miles. The remaining nodes were located on the outskirts of this radius. The demands were randomly set following the same logic of the MOD instances, that is, from 1 up to 3 students were allowed for each bus stop. *(d)* REAL - a real case for the city of Governador Valadares, located in the state of Minas Gerais, Brazil (please refer to Section 3.4.5), was studied.

As four (ILS-VND,ILS-RVND,VNS-VND,VNS-RVND) of the five proposed methods have a random nature, each one of the instances was executed 30 times with a different seed ($\{1,...,30\}$) being supplied to the generator of random numbers for each run. Altogether more than 4,500 runs were executed to assess the computational performance of the devised meta-heuristics. Differences in the solution costs and in the average weighted riding distance of the students were also recorded to compare the mixed load approach with single load solutions. Further, the overall best proposed method was compared with the MLI procedure presented by Park et al. (2012).

3.4.1 Calibration phase

To calibrate the proposed meta-heuristics, the PARK dataset (Park et al., 2012) was used as the control instances to set the main parameters or configurations. The idea is to identify which parameter value and/or settings perform better for the methods. Park et al. devise their instances to allow bus stops with demands up to 66 students. In the present experiments, the used bus capacities were smaller (20, 30 and 40 available seats), so bus stops with demands greater than 40 pupils were split into new nodes according to respective their schools, but respecting the least feasible capacity available.

a) **Parameter λ of Algorithm 1:** values in the set $\{0.4, 0.6, ..., 2.0\}$ were used for λ as suggested by Li et al. (2007). Of the 24 instances, $\lambda = 0.4$ returned 19 best solutions, i.e., the best solution was found in 79.17% of the times, against 5 (20.83%) for the value $\lambda = 0.6$. For higher values for λ, the obtained solutions were worse 5.43% in average than the best ones found. This is somehow expected since λ controls the number of pairs of routes to be candidates to be merged. Larger values of λ imply that bus stops which are far apart may now be selected to be grouped, which per se may greatly increase the routing distances. Based on these results, λ was set to 0.4.

b) **Neighborhood structure search order for the VND:** Numbering the neighborhood structures as 1 for OnePointMove, 2 for TwoPointMove, 3 for CrossExchange, and 4 for 2Opt, and leaving the 2Opt search structure as the last neighborhood to be executed, six possible orders for the VND were investigated to assess which one performs better when an initial solution is provided. The 2Opt was set to be the last one because as pointed out by (Li et al., 2007) it is more interesting to change the nodes of the routes prior to optimize the sequence of visitation. Though statistically equivalent, the order 2314 performed slightly better than the other orders, presenting an improvement of 0.64% on average over the supplied initial solutions.

c) **RRT Parameter *MaxIter* of Algorithm 3:** The values 5, 10, 15 were assessed as the maximum number of iterations (*MaxIter*) for the RRT algorithm, while adopting the order 2314 on the permutation set $\mathbb{P}$ for the *OnePass* function. Value 15 was able to return the best solutions for all instances of the PARK dataset when compared with the values 5 and 10. These values were only able to find the best solutions in 87.5% of the cases. For sake of fairness in the comparison, the parameter $MaxIter$ of the ILS meta-heuristic (Algorithm 4) was set also to 15.

The number of perturbation iterations $MaxPert$ was set to 5 as suggested by Li et al. (2007). Recall that during the *Uphill* phase, solutions up to 10% worse than the current solution are allowed as new current solutions. If the number of perturbation iterations is greatly increased, a random search behavior may be perceived, yielding probably a poor behavior of the algorithm.

3.4.2 Heuristics performance on the MOD dataset

The MOD dataset has 120 instances. Each one was solved 30 times for each method (ILS-VND,ILS-RVND,VNS-VND,VNS-RVND), but the RRT, with a different seed being supplied to the random generator of numbers for each run. As the RRT procedure is a deterministic algorithm, each instance was solved only once by it. Given the large number of executions and the obtained records, a more interesting and compact presentation of the results was here favored. Ribeiro et al. (2002) propose a concise and elegant methodology to report computational experiments of different meta-heuristics with a random nature. Prior to introduce the used metrics to assess the performance of the methods, some notation is required. Let $\mathscr{I} = \{1, \ldots, 120\}$, $\mathscr{R} = \{1, \ldots, 30\}$, and $\mathscr{M} = \{$ILS-VND,ILS-RVND, VNS-VND,VNS-RVND, RRT$\}$ be the sets of instances, runs, and evaluated methods, respectively. Let also $\phi(s_{irm})$ and $\phi(s_i^*)$ be the costs of the solution of run $r \in \mathscr{R}$ for instance $i \in \mathscr{I}$ by method $m \in \mathscr{M}$, and of the overall best solution for instance $i \in \mathscr{I}$ ($\phi(s_i^*) = \min_{r \in \mathscr{R}, m \in \mathscr{M}} \{\phi(s_{irm})\}$), respectively.

Ribeiro et al. (2002) show that is possible to assess how good a heuristic is when compared with others whenever optimal solutions are not available for the instances. They suggest the use of the overall best solution (s_i^*) for each instance as reference during the computation of the following metrics:

- **#Best:** it is the number of instances in which the regarded method finds the overall best solution after all the runs for each instance. More precisely,

$$\textbf{\#Best}_m = |\{i \in \mathscr{I} : \phi(s_i^*) = \min_{r \in \mathscr{R}} \phi(s_{irm})\}| \qquad \forall m \in \mathscr{M}.$$

  A large value indicates a good performance by the studied heuristic.
- **%Best:** it represents the same of #Best, but in a percentage format.
- **%DevMed:** it estimates the precision of the regarded method. It is the average percentage of the deviations of the achieved solutions for each run from the overall best solution of each instance, or:

$$\%\textbf{DevMed}_m = 100 \frac{\sum_{i \in \mathscr{I}} \sum_{r \in \mathscr{R}} \left(\frac{\phi(s_{irm}) - \phi(s_i^*)}{\phi(s_i^*)} \right)}{|\mathscr{I}|\ |\mathscr{R}|} \qquad \forall m \in \mathscr{M}$$

A small value indicates a more precise method.

- **%DevMin:** it assesses the accuracy of the referred algorithm. It is the average percentage of the deviations of the best solution obtained by the procedure from the overall best solution for each instance, or:

$$\textbf{\%DevMin}_m = 100 \frac{\sum_{i \in \mathscr{I}} \left(\frac{\min_{r \in \mathscr{R}} \phi(s_{irm}) - \phi(s_i^*)}{\phi(s_i^*)} \right)}{|\mathscr{I}|} \qquad \forall m \in \mathscr{M}$$

A small value indicates a more accurate algorithm.

- **Score:** it indicates how the other methods performed better than the referred algorithm over the instances. For each instance, it is computed how many algorithms obtained a better solution than the observed method. These values are then summed to represent the final score of the procedure, or:

$$\textbf{Score}_m = \sum_{i \in \mathscr{I}} |\{ j \in \mathscr{M} : j \neq m \wedge \min_{r \in \mathscr{R}} \phi(s_{irj}) < \min_{r \in \mathscr{R}} \phi(s_{irm}) \}| \qquad \forall m \in \mathscr{M}$$

The lower the score value the better the performance of the algorithm. For example, for a given instance, if four methods obtained a better solution then the current algorithm, its score would be equal to four.

- $\%\boldsymbol{\Delta}(\boldsymbol{s^0})$: it indicates the average percentage on how much the initial solution (s^0) was improved by the referred method or

$$\%\boldsymbol{\Delta}(\boldsymbol{s^0})_m = 100 \frac{\sum_{i \in \mathscr{I}} \sum_{r \in \mathscr{R}} \left(\frac{\phi(s^0_{irm}) - \phi(s_{irm})}{\phi(s^0_{irm})} \right)}{|\mathscr{I}| \, |\mathscr{R}|} \qquad \forall m \in \mathscr{M}$$

- $\mathbf{min}\,\%\boldsymbol{\Delta}(\boldsymbol{s^0})$: it indicates the minimum percentage on how much the initial solution (s^0) was improved by the referred method or

$$\mathbf{min}\,\%\boldsymbol{\Delta}(\boldsymbol{s^0})_m = 100 \min_{i \in \mathscr{I}, r \in \mathscr{R}} \left(\frac{\phi(s^0_{irm}) - \phi(s_{irm})}{\phi(s^0_{irm})} \right) \qquad \forall m \in \mathscr{M}$$

- $\mathbf{max}\,\%\boldsymbol{\Delta}(\boldsymbol{s^0})$: likewise, it indicates the maximum percentage on how much the initial solution (s^0) was improved by the referred method or

$$\mathbf{max}\,\%\boldsymbol{\Delta}(\boldsymbol{s^0})_m = 100 \min_{i \in \mathscr{I}, r \in \mathscr{R}} \left(\frac{\phi(s^0_{irm}) - \phi(s_{irm})}{\phi(s^0_{irm})} \right) \qquad \forall m \in \mathscr{M}$$

Table 3.1 shows the attained results for the RAND instances. Though the four first methods presented similar results, the ILS-RVND had the overall best values for the considered metrics; while the RRT algorithm got worst. ILS-RVND was able to provide the greatest improvements on the given initial solutions. The performance of the ILS-RVND can be explained by observing Table 3.2 which presents the averages of the averages of the best solutions obtained by each method and of the averages of the running times, group by the instance indexes of the PARK data set (Park et al., 2012). The numbers of student bus stops and schools are also displayed for each group. The ILS-RVND expended more time in the search in half of the groups, taking on average 15.10% more time on average than the other methods.

Table 3.1: Summarized results for the MOD instances.

	ILS-VND	ILS-RVND	VNS-VND	VNS-RVND	RRT
$\%\Delta(s^0)$	3.02	3.07	3.02	2.97	2.24
min $\%\Delta(s^0)$	1.20	1.20	1.03	1.20	0.88
max $\%\Delta(s^0)$	6.96	8,24	6.90	6.24	4.76
%DevMin	0.26	0.21	0.27	0.31	1.07
%DevMed	0.86	0.85	0.89	0.89	1.07
%Best	28.33	33.33	18.33	15.83	4.17
#Best	34	40	22	19	5
Score	190	143	196	224	444

Table 3.2: Average solution costs and running times for the MOD instances.

Group #	$\|P\|$	$\|H\|$	ILS-VND $\bar{\phi}(s)$	ILS-VND $\bar{T}$ [secs.]	ILS-RVND $\bar{\phi}(s)$	ILS-RVND $\bar{T}$ [secs.]	VNS-VND $\bar{\phi}(s)$	VNS-VND $\bar{T}$ [secs.]	VNS-RVND $\bar{\phi}(s)$	VNS-RVND $\bar{T}$ [secs.]	RRT $\bar{\phi}(s)$	RRT T [secs.]
1	250	6	9740.12	0.86	9742.97	0.83	9748.43	0.80	9740.02	0.83	9730.55	0.70
2	250	12	12231.41	0.77	12233.02	0.80	12231.86	0.73	12239.92	0.70	12214.07	0.71
3	500	12	24556.08	7.55	24563.39	7.46	24573.99	7.66	24568.42	7.69	24694.69	4.07
4	500	25	31324.25	6.62	31311.68	6.81	31331.86	6.39	31343.41	6.13	31429.03	4.12
5	1000	25	59480.14	57.78	59447.84	59.35	59466.17	58.82	59486.20	55.70	59605.21	25.05
6	1000	50	65529.26	60.72	65516.16	61.15	65521.43	58.55	65522.01	58.67	65474.12	32.25
7	2000	50	118913.06	364.10	118917.64	359.70	118944.60	342.36	118935.49	353.81	119080.07	122.20
8	2000	100	137804.13	481.62	137842.42	440.13	137813.47	496.34	137755.46	485.05	137937.14	207.88
9	250	6	9581.32	0.94	9584.42	0.92	9582.34	0.90	9582.62	0.90	9614.87	0.73
10	250	12	14001.53	0.97	14015.44	0.93	14019.46	0.84	14021.71	0.89	14074.93	0.75
11	500	12	27119.63	6.13	27112.65	6.05	27132.14	5.94	27127.14	5.85	27175.41	3.73
12	500	25	33084.98	5.81	33076.05	6.09	33091.71	5.41	33089.59	5.66	33181.48	3.55
13	1000	25	64099.23	49.40	64079.12	50.44	64091.54	50.06	64106.90	49.82	64287.94	21.54
14	100	50	68091.49	46.66	68065.75	47.93	68115.86	43.33	68129.75	43.15	68114.90	23.80
15	2000	50	127711.59	464.96	127736.95	475.29	127801.76	432.96	127771.51	436.62	128042.89	154.70
16	2000	100	136756.13	480.19	136859.17	413.62	136885.88	410.32	136858.99	427.54	136942.61	189.30
17	250	6	16139.34	0.91	16137.11	0.90	16151.68	0.82	16141.08	0.85	16226.56	0.63
18	250	12	19460.62	1.08	19447.12	1.07	19457.36	1.01	19478.22	0.98	19561.92	0.76
19	500	12	34140.12	5.81	34139.02	5.85	34133.50	5.85	34151.52	5.62	34242.59	2.97
20	500	25	37421.48	6.86	37435.20	6.85	37461.24	6.36	37453.64	6.76	37455.58	3.87
21	1000	25	68487.04	67.80	68457.63	70.34	68498.86	65.85	68495.21	67.22	68570.18	30.06
22	1000	50	75930.22	69.70	75914.02	69.49	75946.30	65.98	75938.44	67.75	75848.57	31.52
23	2000	50	141522.49	384.21	141363.28	450.40	141477.10	415.37	141446.48	413.63	142000.92	123.42
24	2000	100	146317.18	444.21	146231.92	486.00	146269.62	466.71	146276.14	479.51	146402.66	191.97
	Average		61643.45	125.65	61634.58	126.18	61656.17	122.89	61652.49	124.22	61746.20	49.18

3.4.3 ILS-RVND-ML vs ILS-RVND-SL vs MLI

The MLI procedure of Park et al. (2012) does not take into account the routing costs, having the only objective of minimizing the number of used buses. Park et al. argue that the routing costs can be disregarded because they are much smaller than the fixed bus costs. Unfortunately this is not true for all the situations. Given the lifespan of a bus and as the picking and delivery of students are routinely done over the school year, on the long run, the routing costs may start to weight on public spendings. Further, when they are not taken into account, some collateral distortions may occur. For instance, the average weighted riding distance of the students may be much larger than when these costs are embedded into the solution process. Recall that the average weighted riding distance of the students is the sum of all of the traveled distance of each student from his bus stop to his school divided by the total number of pupils. These effects are more widely perceived when schools do not have the flexibility of having different starting times, as it assumed in the study of Park et al. (2012). Due to labor policies and union treats, schools in Brazil have to start at the same time. To show the influence of these effects, the PARK dataset was solved by the ILS-RVND with mixed load (ML) of students, and by a modified version of the ILS-RVND allowed to handle only single load (SL). The attained results were compared with the ones obtained by the MLI procedure configured to disregard time windows. The same homogeneous fleet of Park et al. with capacity of 66 seats was used during the experiments. The fixed and the routing costs per unit of traveled distance were set to \$200 and \$1.00, respectively. Each instance of the PARK dataset was solved 30 times by the ILS-RVND-ML and ILS-RVND-SL with a different seed for each run.

Table 3.3 presents the total costs, the routing and the fixed costs, the number of buses, and the average weighted riding distance of the students ($\bar{\omega}_\ell$) for the obtained results. The MLI procedure provided solutions with the smaller number of buses for every instance. However, by disregarding the routing costs during the search, it achieved solutions with the largest total costs when these costs are considered afterwards. For all methods, it is interesting to note that whenever the number of bus stops are large, (e.g. instances $6-9$, $15-16$ and $23-24$) the routing costs are larger than the fixed costs, contradicting the assumptions of Park et al. (2012). The attained $\bar{\omega}_\ell$ for the ILS-RVND-SL and ILS-RVND-ML algorithms were much smaller than the MLI procedure. This metric indicates that more students are riding for longer distances in the MLI's solutions than in the other two methods. Furthermore, the approach with a SL of students had a tendency to provide the best $\bar{\omega}_\ell$, but, depending on how the bus stops were scattered over the studied region, the ML version provided smaller values (e.g. instances 2, 3, 4, 7, and 11). On average, the $\bar{\omega}_\ell$ values of the SL solutions were only %6.33 smaller than the ML ones, though the ML solutions had total costs and the number of buses that were on average %6,53 and %7.72 smaller. Another interesting fact worth of remarking is that the MLI's solutions were worse than the ones got by ILS-RVND-SL. In only three instances, the MLI algorithm generated a better solution (see rows 2, 6 and 9). This indicates that for the cases in which schools start at the same time, there is no advantage of adopting the MLI procedure over a strategy of having a dedicate fleet per school.

Table 3.3: ILS-RVND-SL vs. ILS-RVND-ML vs. RC.

			ILS-RVND-SL					ILS-RVND-ML					MLI				
Park inst.	$\|P\|$	$\|H\|$	TC[1] ($)	RC[2] ($)	FC[3] ($)	#B[4]	$\bar{\omega}_\ell$ (mi.)	TC ($)	RC ($)	FC ($)	#B	$\bar{\omega}_\ell$ (mi.)	TC ($)	RC ($)	FC ($)	#B	$\bar{\omega}_\ell$ (mi.)
1	250	6	20424.43	7024.43	13400	67	43.60	20067.13	7067.13	13000	65	45.47	20934.86	8734.86	12200	61	70.26
2	250	12	21974.96	10574.96	11400	57	52.18	19760.28	9560.28	10200	51	48.24	21613.38	11413.38	10200	51	72.39
3	500	12	42368.23	19368.23	23000	115	53.65	39503.67	17703.67	21800	109	50.40	42913.56	21913.56	21000	105	85.25
4	500	25	52866.31	27066.31	25800	129	55.26	48408.95	24408.95	24000	120	51.57	57228.13	34028.13	23200	116	116.95
5	1000	25	109422.58	52622.58	56800	284	46.12	103059.18	50059.18	53000	265	47.62	118645.05	66245.05	52400	262	93.54
6	1000	50	130939.52	65139.52	65800	329	44.98	120480.16	60680.16	59800	299	46.87	130876.78	72876.78	58000	290	76.59
7	2000	50	181398.64	89398.64	92000	460	52.34	169529.73	83329.73	86200	431	48.23	218427.35	134427.35	84000	420	134.37
8	2000	100	204615.44	105215.44	99400	497	53.50	181895.45	94695.45	87200	436	56.21	222124.30	136324.30	85800	429	117.53
9	250	6	22530.63	7930.63	14600	73	32.84	21998.29	7598.29	14400	72	34.01	22369.92	9169.92	13200	66	63.88
10	250	12	27424.85	12224.85	15200	76	49.68	25546.57	11546.57	14000	70	51.44	27706.06	14306.06	13400	67	81.82
11	500	12	37681.59	17681.59	20000	100	54.38	35557.67	16557.67	19000	95	52.79	42062.04	23662.04	18400	92	110.71
12	500	25	44437.74	23037.74	21400	107	54.39	40820.82	21820.82	19000	95	60.64	56226.92	37426.92	18800	94	171.53
13	1000	25	98227.06	50627.06	47600	238	49.24	93064.34	48464.34	44600	223	49.95	109750.06	66350.06	43400	217	110.02
14	1000	50	124985.90	66585.90	58400	292	47.42	114404.48	61204.48	53200	266	47.19	129038.29	78038.29	51000	255	98.74
15	2000	50	187061.08	94661.08	92400	462	49.21	176985.20	90985.20	86000	430	52.51	219547.53	135947.53	83600	418	128.54
16	2000	100	175446.78	88846.78	86600	433	59.43	155747.12	80947.12	74800	374	62.89	200766.59	127166.59	73600	368	147.51
17	250	6	22212.15	10812.15	11400	57	39.58	22127.96	11127.96	11000	55	48.69	24167.34	13567.34	10600	53	82.44
18	250	12	27645.70	14645.70	13000	65	35.46	25954.49	14154.49	11800	59	42.56	29335.98	17735.98	11600	58	88.15
19	500	12	44840.37	21840.37	23000	115	41.77	43276.21	21476.21	21800	109	49.47	56951.70	36151.70	20800	104	142.48
20	500	25	48723.59	24723.59	24000	120	47.52	46044.94	23444.94	22600	113	56.32	59195.54	38395.54	20800	104	146.85
21	1000	25	95364.69	47764.69	47600	238	36.16	89591.24	45991.24	43600	218	45.06	114951.72	72751.72	42200	211	128.32
22	1000	50	91896.77	47496.77	44400	222	45.32	83724.96	44724.96	39000	195	57.26	108247.07	70647.07	37600	188	140.28
23	2000	50	186682.92	95882.92	90800	454	38.74	178230.70	93630.70	84600	423	45.75	230430.42	148230.42	82200	411	133.81
24	2000	100	228412.04	115212.04	113200	566	37.83	207849.95	107249.95	100600	503	45.45	242955.51	144155.51	98800	494	94.52

1: Total cost. 2: Routing Costs. 3: Fixed Costs. 4: Number of buses.

3.4.4 Comparisons for the RAND instances

Most of Brazil's counties had their rural schools transfered to their urban limits. So to mimic this situation the RAND dataset was constructed. Each instance was then solved 30 times by the ILS-RVND-SL and ILS-RVND-ML algorithms to assess which strategy (SL or ML) was more suitable to the Brazilian context. Once again, a different seed was used for each run. The obtained results are presented in Table 3.4 where, besides the involved costs, some other metrics are also shown.

The ILS-RVND-ML algorithm was able to generate solutions with much lower costs than the ILS-RVND-SL. On average they were 8.77% cheaper; though requiring much more computer running time to be obtained. On average, 3.67 more time was required to solve an instance than its counterpart. The school insertion and checking after each neighborhood structure moves of the ILS-RVND-ML method is indeed very time consuming. The average number of students per bus (avg. stdt./bus column) was also greater for the ML's solutions than the SL ones. On average, the SL solutions had 7.26% less students per vehicle than the ML. This indicates that the ML approach allows for a better use of the installed bus infra-structure. The average weighted riding distance of the students ($\bar{\omega}_\ell$) as well as the average route length were much lower in the ML case, on average 33.75% and 28.07%, respectively. This implies that students tend to ride less time on the ML solutions than on the SL ones. Furthermore, not only the obtained solutions had a smaller number of buses (on average 7.25% smaller), but the ML approach was able to induce solutions with a more homogeneous fleet, which would probably ease the maintenance and the management of the transportation system. For this type of instances, the ML approach was able to provide very interesting solutions.

Table 3.4: ILS-RVND-SL vs. ILS-RVND-ML for the RAND instances.

Inst.	Time (secs.)	TC[1] ($)	RC[2] (mi.)	FC[3] ($)	Avg. stdt /bus	$\bar{\omega}_\ell$ (mi.)	Avg. RL[4] (mi.)	Total capacity (seats)	# buses/type 20	30 (seats)	40	#B[5]
					ILS-RVND-SL							
R1	42.33	11224.30	874.30	10350	35.71	17.00	15.60	2070	5	7	44	56
R2	41.44	11574.40	1274.40	10300	37.04	27.20	23.60	2060	3	4	47	54
R3	56.38	12503.30	2103.30	10400	35.71	42.10	37.60	2080	6	4	46	56
R4	62.96	13060.40	2860.40	10200	36.36	59.10	52.00	2040	6	4	45	55
R5	38.72	14923.40	4723.40	10200	37.04	95.50	87.50	2040	3	6	45	54
					ILS-RVND-ML							
R1	113.09	10786.60	736.60	10050	39.22	14.50	14.40	2010	1	1	49	51
R2	167.32	10958.80	908.80	10050	39.22	18.90	17.80	2010	1	1	49	51
R3	232.92	11494.30	1444.30	10050	39.22	29.70	28.30	2010	1	1	49	51
R4	154.73	11647.50	1597.50	10050	39.22	30.80	31.30	2010	1	1	49	51
R5	195.55	12568.80	2518.80	10050	39.22	51.40	49.40	2010	1	1	49	51

1: Total cost. 2: Routing Costs. 3: Fixed Costs. 4: Average route length. 5: Number of buses.

3.4.5 Real case: the city of Governador Valadares, Minas Gerais, Brazil

Located in South America, Brazil is the largest country of the continent. It has 26 states, being Minas Gerais the fourth largest state by area. Minas Gerais has an area equal to 2.55 (0.92) times the size of Great Britain (France). The state ranks as the fourth largest by area and the second most populous in the country. It has a total of 853 municipalities, being Governador

Valadares the ninth most populous of these cities. Governador Valadares, depicted in Figure 3.3, has a population size of 275,000 inhabitants and an area of 906.61 mi^2 (2,348.1 km^2) of which 9.42 mi^2 (24.4 km^2) (IBGE, 2014) correspond to the urban center of the county, represented by the dark gray shape in Figure 3.3. The geographic coordinates of 541 rural residences, scattered in the rural area of the county, were collected. These bus stops correspond to a demand of 944 students which were attended by 25 different schools located near or at the urban center in the year of 2014. In Figure 3.3, the red dots represent the bus stops, while the empty triangles and circle are the schools and the county's bus garage, respectively.

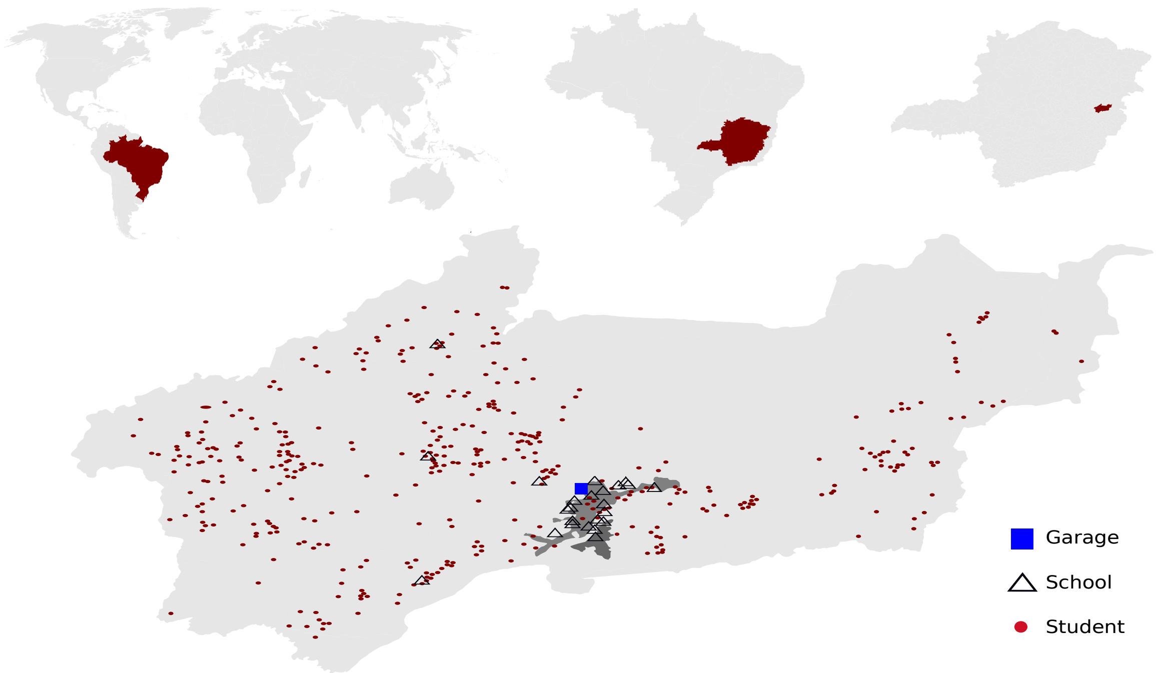


Figure 3.3: The county of Governador Valadares, Minas Gerais, Brazil.

For the test problem of Governador Valadares, it was investigated which policy was more suitable: the single or mixed load approaches. The instance was then solved by the ILS-RVND-SL and ILS-RVND-ML algorithms as well as by the MLI procedure. Once again, 30 runs were carried out with different seeds. Table 3.5 shows the obtained results. Both the ILS-RVND-ML and MLI methods provided cheaper solutions than the single load scheme. To have a single fleet per school requires much more buses of different types, being ridden by fewer students (row avg stdt./bus), i.e. buses did not have their full capacities used. The only advantages of the SL approach were the average weighted riding distance of the students (row $\bar{\mu}_\ell$) and average route length (row avg. route length) which were much smaller than the other two methods. Note that though the MLI procedure attained a solution with fewer buses than the ILS-RVND-ML algorithm, the fixed costs were the same. Further, the routing cost of the MLI's solution was much worse (69.68%) than the one got by the ILS-RVND-ML algorithm. Observe also that students rode for much more time and for longer distances in the MLI's solution.

The results for Governador Valadares were similar to those of the RAND dataset. The adoption of a mixed load transportation has many advantages when compared with a single load approach. It is less expensive, it requires less buses, and it allows for the management of a more homogeneous fleet. However the large values for the average route length and for the average weighted riding distance may suggest that constraints for limiting the maximum riding distance should be enforced. This may be indeed true for most cases, but for the instance of Governador Valadares this would be worthless since there is a considerable group of bus stops (34.38%) which are far apart (distances greater than 50 miles) from the downtown area. In other words, it would be more effective and better for the students if the Brazilian government would review its efforts of moving the rural schools to the urban centers.

Table 3.5: Results for Governador Valadares instance

	ILS-RVND-SL	ILS-RVND-ML	MLI
Time (secs.)	11.67	40.69	2.03
Total cost	8716.02	6629.93	7939.86
Fixed cost	6150.00	4750.00	4750.00
Routing cost	2566.02	1879.93	3189.86
# buses	40	25	24
avg. stdt./bus	23.60	37.76	39.33
$\bar{\mu}_\ell$ (mi)	47.28	67.03	151.52
avg. route length (mi)	64.15	75.20	132.91
Total capacity	1230	950	950
# buses (20 seats)	15	2	0
# buses (30 seats)	7	1	1
# buses (40 seats)	18	22	23

3.5 Final Remarks

Five meta-heuristics were proposed for the mixed load capacitated rural school bus routing problem with heterogeneous fleet. This is an important problem that has been neglected by the routing literature. The adoption of mixed loads has shown to be very effective and cost savings, specially for the Brazilian context. Of the five devised methods, the ILS-RVND-ML has provided the best results given the analyzed metrics. The attained solutions by the ILS-RVND-ML have proved to be better than the ones got by the MLI procedure of Park et al. (2012). The statement of Park et al. that routing costs can be disregarded turn out to be not so true for larger instances. For these cases, the routing costs play an important role which can not be disregarded. Some insights have also been provided about the Brazilian policy of transferring the rural schools to the urban areas. Given Brazil's dimensions, this policy should be revised or carried out more judiciously, because the riding distances can be prohibitively long even if limiting constraints are imposed, in which case infeasible solutions would be probably rendered.

Chapter 4

A Multi-objective capacitated rural school bus routing problem with heterogeneous fleet and mixed loads

4.1 Introduction

Brazil has around 50 million elementary students enrolled in its public educational system. Of these students, 13% are located on rural areas (INEP, 2013) that are served by schools with multi-grade classes, i.e. pupils from different grades are arranged in the same classroom and taught by the same teacher. Tough debatable, this type of classrooms is often considered to be unsuitable for todays advanced curricula, because the classrooms have usually fewer resources than necessary. Further it is often difficult to attract qualified, skilled teachers to these schools due to their usually remote location from larger towns (Vincent, 1999).

To offer a more suitable environment with better school infrastructures, a richer curriculum with more courses and activities, and single-grade classrooms, the federal government has been doing a great effort to nucleate some rural schools by placing them closer to the counties downtown area. This endeavor is allowing students to have access to better, larger facilities and better teaching. But, at the same time, it has transfered the burden of the transportation of the rural students to the local municipality, which has now to provide a transportation system. Students are required to be picked up at their homes, brought to their respective schools, and delivered back to their residence later on the day.

To easy the process and facilitate for public servants to plan the local bus routes, the Brazilian government is not only financially supporting the acquisition of new buses, but it is also encouraging the use and development of decision support systems. The idea is to guarantee a suitable transportation service level for the students, while observing the drivers working hours and labor union polices, and keeping the involved costs at their lowest possible. This gives rise to a multi-objective optimization problem with three distinct and conflicting objectives that must be optimized all together.

The rich literature of rural school bus routing problems (Park and Kim, 2010) has some interesting works– e.g. Thangiah and Nygard (1992), Corberán et al. (2002), Schittekat et al. (2006), and Pacheco et al. (2013)– that deal with a single objective only. Though of great value and importance, these works fall short to support decision makers who often have to handle more than one objective. In real cases, the problem scope often goes beyond one goal.

Fewer authors have addressed a multi-objective approach. Bowerman et al. (1995) devise a multi-objective framework to solve an urban school bus routing problem with four different objectives: total bus route length, student walking distance, load balancing, length balancing. Assuming a homogeneous fixed size fleet and known weights of importance for each objective, they propose to first group the students into clusters in each district, and then plan the bus stops and routes by a set covering algorithm followed by a traveling salesman procedure. They apply their method to an instance extracted from the Wellington County, Ontario, with 138 students, 3 buses with 50 available seats.

Recently, Pacheco et al. (2013) propose a bi-objective procedure for school buses in rural areas which seeks the minimization of a normalized objective function consisted of the longest route and the total distance traveled. Like Bowerman et al. (1995), the weights are also assumed to be known a priori. They devise a solution algorithm based on a tabu search with a multi-objective adaptive memory programming and compare it with an implementation of a non-dominated sorting genetic algorithm. Assuming also a homogeneous fixed size fleet, the framework is used to solved 16 instances corresponding to middle schools in the Province of Burgos, Spain. The largest test set has 57 bus stops, 429 students, and 15 buses. The results show that their approach dominates the general framework based on genetic algorithm.

Both aforementioned works adopted the weighted sum form of multi-objective optimization. Though largely employed because its simplicity, the gathering of the multiple objectives into a single weighted objective requires the previous knowledge of importance of each objective and, as consequence, the values for their weights. The selection of these values can be difficult and trick because one has to discern between setting weights to compensate for differences in magnitudes of the objectives and setting weights to indicate the importance of an objective over the others (Marler and Arora, 2010). One alternative is to find multiple trade-off solutions with a wide range of values for the objectives and to let the decision makers chose the most suitable solution for their needs.

Corberán et al. (2002) use an evolutionary method to solve a multi-objective rural school bus routing problem with heterogeneous fleet, single load and destination for which the fleet is dedicated. seeking minimizing the number of buses and the time that a given student spends in route.

However, given the large dimensions of Brazil and how rural students are sparsely scattered over it, the deployment of a homogeneous fleet as well as the idea of having a dedicate fleet per school may not the most adequate for the Brazilian context (Carvalho et al., 2010). A homogeneous fleet with large vehicles will most likely have many routes with vacant seats. Filling these vacant seats by increasing the length of the journeys so that more students can be picked up may not be in the best interests of the pupils and the drivers. Long routes on poorly maintained roads are usually tiresome. On the other hand, adopting small or mid-sized buses in a homogeneous fleet may result in numerous vehicles with many drivers. Hence the use of heterogeneous fleets allows balancing the length of the routes, and the sizes of the vehicles and fleets. One way to further improve the capacity utilization of the buses is by using mixed loads during the rides. Instead of having a dedicated fleet per school (i.e. vehicles carry a single type of load), students of different ages and schools are transported at the same time on the same vehicle (i.e. mixed load).

Both assumptions (heterogeneous fleet and mixed loads) have been addressed for single objective optimization problems only. Pacheco and Martí (2006) adapt four known different heuristics for the same problem. Bodin and Berman (1979a) are the first authors to discuss about mixed loads. They point out that mixed loads occur frequently in rural areas, though they do not devise any procedure to handle mixed load buses. Chen et al. (1988) state that an excessive number of buses can occur when single loads are adopted, specially when dealing with remote rural students located in poorly populated areas. Braca et al. (1997) address the mixed load bus routing problem by proposing an insertion procedure. The method constructs each route by randomly selecting a bus stop and inserting it and its respective associated school into the route at the best cost estimation possible, but enforcing the capacity constraints. The single objective minimizes the number of used buses. Park et al. (2012) improve the method of Braca et al. (1997) by devising a post improvement procedure. Starting from a solution with a dedicate fleet per school obtained from a sweep based algorithm (Gillett and Miller, 1974b), the procedure reallocates one bus stop at a time in a greedy way until routes can be merged or deleted.

Li and Fu (2002) claim to have addressed a multi-objective optimization algorithm for the urban bus school routing problem with heterogeneous fleet in which four different objectives are minimized: the total number of buses; the students waiting time at the bus stops; the total riding time; and the load unbalance of the routes. However they neither use a objective function consisted of a weighted sum of objective functions during; nor they supply a pool of non-dominated solutions so that the decision maker can select one from them at the end. They devise one optimization heuristic for each adopted objective. The procedures are then executed sequentially sorted by the relative relevance of each objective. Each procedure receives as input the solution obtained by the immediate predecessor procedure. At the end of Li and Fu framework, only one solution is returned. Neither the decision maker has opportunity to prescribe weights for the objective functions, nor he is able to select from a pool of non-dominated solutions the one that is the most interesting for him.

Motivated by the social relevance of the application, and the scarce literature on the theme, four multi-objective heuristics are here devised for the rural school bus routing problem with heterogeneous fleet and mixed loads. The problem seeks the minimization of three conflicting objectives: the total routing and fixed costs; the total weighted traveling time of the students; and the unbalance riding time of the routes. The devised methods return non-dominated solutions in a set named Pareto frontier from which the decision maker can select the most suitable one for his problem.

The first heuristic is based on the multi-objective iterated local search (MOILS) framework proposed by Assis et al. (2013). The second method is an improved version of the MOILS. The last three heuristics have different strategies embedded into the second version to improve the quality of the achieved frontier of non-dominated solutions. The third one has a procedure which is based on a simple idea which resembles a path-relinking procedure (PR) (Glover et al., 2000); while the fourth has an actual PR procedure. The last version combines the two PR schemes. Statistical analysis on three known metrics (cardinality, coverage, hyper-volume) used in multi-objective optimization followed by graphical analysis are carried out to assess the performance of the proposed heuristics. As the number of candidate solutions in a Pareto frontier can be large, finding a satisfactory one among them becomes a burden to the decision maker. Therefore, besides the devised heuristics, a simple new visual approach to support decision makers on selecting a good solution is also proposed.

The next sections are divided as follows. §4.2 describes the problem and its mathematical formulation. §4.3 explains the implementation details of the procedures used in the solution approaches devised to solve the problem §4.4. §4.5 presents the new approach for helping decision makers to find good solutions. §4.6 shows the attained results. Finally in Section 4.7 conclusions and future researches are presented.

4.2 Notation, definitions and formulation

The problem consists in transporting rural students geographically dispersed with a heterogeneous fleet back and forth from their homes to their respective schools every day. The problem seeks the minimization of the unbalance riding time of the routes, of the total weighted students riding time, and of the total routing and fixed costs. As there are more than one school involved, mixed loads are allowed.

The problem uses the following definitions: Let $G = (N, A)$ be an directed graph with $N = \{0\} \cup P \cup H$, where 0 represents the garage node; $P = \{1, \ldots, n_p\}$ be the set of student nodes, where n_p is the number of student nodes; $H = \{n_p + 1, \ldots, n_p + n_s\}$ be the set of schools, where n_s is the number of schools; and $B = \{1, \ldots, n_b\}$ be the set of available buses, where n_b is the number of buses. Let also $A = \{(i, j) : i, j \in N, i \neq j\}$ and $E = \{(i, j) \in A : i < j\}$ be the arc and edge sets. Let $d(i) = s$ be a function which given a student node $i \in P$ returns the school $s \in S$ associated to it. Here it is assumed that each student node has only one school associated to it. If a student node has more than one school associated to it, this node is then split into other student nodes, one per school. Each node $i \in P$ has number of students q_i to be picked up and delivered to their respective schools. There are also an associated riding time $t_{i,j}$ to arc $(i, j) \in A$, and a cost $c^b_{ij} \geq 0$ for a bus $b \in B$ to ride edge $(i, j) \in E$. The capacity and the fixed costs of the buses are given by Q^b and a_b, respectively.

The following decision variables are used to model the problem. $z^b \in \{0, 1\}$ is equal to one if bus $b \in B$ is used; zero otherwise. $y^b_i \in \{0, 1\}$ is equal to one if the node $i \in P \cup H$ is visited by bus $b \in B$; zero otherwise. $x_{ij}^b \in \{0,1\}$ is equal to one if the edge $(i,j) \in E$ is crossed by bus $b \in B$; zero otherwise. $f_{ij}^{lb} \geq 0$ is the percentage of the the demand q_l of node $l \in P$ which rides on bus $b \in B$ over arc $(i,j) \in A$. $\tau_b \geq 0$ is used to measure the travel time of bus $b \in B$. Function $\Gamma(\tau)$ computes the difference of the longest by the shortest riding time of the routes. Given these parameters and variables, the mixed load, multiple destination rural school bus with heterogeneous fleet and three objectives can be formulated as follows:

$$\min F(x,y,z,f,\tau) = (\varphi^1, \varphi^2, \varphi^3) \tag{4.1}$$

$$\varphi^1(z,y,x) = \sum_{b\in B}\left(a_b z^b + \sum_{(i,j)\in E} c_{ij}^b x_{ij}^b\right) \tag{4.2}$$

$$\varphi^2(z,y,x,f) = \sum_{b\in B}\sum_{l\in P}\sum_{(i,j)\in A} q_l t_{ij} f_{ij}^{lb} \tag{4.3}$$

$$\varphi^3(z,y,x,\tau) = \Gamma(\tau) \tag{4.4}$$

$$\text{s.t.:} \sum_{b\in B} y_i^b = 1 \qquad \forall i \in P \tag{4.5}$$

$$\sum_{i\in P} q_i y_i^b \leq Q_b z^b \qquad \forall b \in B \tag{4.6}$$

$$y_i^b \leq y_{d(i)}^b \qquad \forall i \in P, b \in B \tag{4.7}$$

$$\sum_{(i,j)\in E} x_{ij}^b + \sum_{(j,i)\in E} x_{ji}^b = 2y_j^b \qquad \forall j \in N \setminus \{0\}, b \in B \tag{4.8}$$

$$\sum_{(0,j)\in E} x_{ij}^b = 2z_b \qquad \forall b \in B \tag{4.9}$$

$$\sum_{\substack{(i,j)\in E:\\ i\in P, j\in H}} x_{ij}^b = z^b \qquad \forall b \in B \tag{4.10}$$

$$\sum_{\substack{(i,j)\in E:\\ i,j\in S}} x_{ij}^b \leq \sum_{i\in S\setminus\{k\}} y_i^b \qquad \forall b \in B, k \in S, S \subseteq N \setminus \{0\} \tag{4.11}$$

$$\sum_{(i,j)\in E} t_{ij} x_{ij}^b = \tau_b \qquad \forall b \in B \tag{4.12}$$

$$\sum_{(l,j)\in A} f_{lj}^{lb} = y_l^b \qquad \forall l \in P, b \in B \tag{4.13}$$

$$\sum_{\substack{(i,j)\in A:\\ i\neq d(l)}} f_{ij}^{lb} = \sum_{\substack{(j,i)\in A:\\ i\neq l}} f_{ji}^{lb} \qquad \forall l \in P, j \in N \setminus \{0, l, d(l)\}, b \in B \tag{4.14}$$

$$\sum_{(i,d(l))\in A} f_{id(l)}^{lb} = y_{d(l)}^b \qquad \forall l \in P, b \in B \tag{4.15}$$

$$f_{ij}^{lb} + f_{ji}^{lb} \leq x_{ij}^b \qquad \forall l \in P, (i,j) \in E, b \in B \tag{4.16}$$

$$x_{ij}^b \in \{0,1\} \qquad \forall (i,j) \in E, b \in B \tag{4.17}$$

$$y_i^b \in \{0,1\} \qquad \forall i \in N \setminus \{0\}, b \in B \tag{4.18}$$

$$z^b \in \{0,1\} \qquad \forall b \in B \tag{4.19}$$

$$f_{ij}^{lb} \geq 0 \qquad \forall l \in P, (i,j) \in A, b \in B \tag{4.20}$$

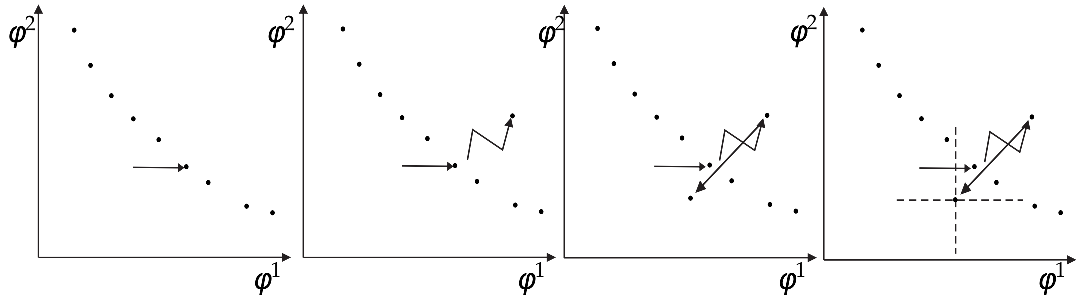
$$\tau_b \geq 0 \qquad \forall b \in B \tag{4.21}$$

The objective function (4.1) is consisted of three objectives which minimize: the fixed and routing costs (4.2), the total weighted traveling time (4.3) and the unbalance of the riding time of the routes (4.4). Constraints (4.5)-(4.7) ensure that each student node has to be visited by some vehicle; that the bus capacity can not be exceeded by the students riding on it; and that if a student node $i \in P$ is visited by a bus $b \in B$, then the associated school $d(i) \in H$ to this node has to be also visited by the same bus, respectively. Constraints (4.8)-(4.11) are responsible for forming the bus routes. Constraints (4.8) guarantee the degree of a bus stop if this node is visited by a bus $b \in B$. Constraints (4.9) assures that if a bus $b \in B$ is activated then it has to leave and return to the garage node. Constraints (4.10) prevent intermediate deliveries prior of picking up more students. In other words, the students have to be picked up first and then delivered to their respective schools. Constraints (4.11) are sub-tour elimination constraints. Constraints (4.12) compute the bus riding time of each bus. Constraints (4.13)-(4.16) are flow balancing constraints responsible for computing the path riding time from each student node to their respective schools. Constraints (4.17)-(4.21) show the domain of the variables. Formulation (4.1)-(4.21) is a multi-objective mixed integer mathematical program of difficult solution.

4.3 Implementation details

Four heuristics based on the MOILS framework introduced by Assis et al. (2013) are devised to solve the multi-objective mixed load rural bus routing problem with heterogeneous fleet. The algorithms are thought to have the least number of parameters to be tunned and to be easy of implementation and use. Before presenting the proposed heuristics, some important concepts and implementation details have to be introduced. The MOILS general idea is depicted in Figure 4.1 for two objectives. A solution with the largest crowding distance (Deb et al., 2002) is selected from a non-dominated frontier set (Figure 4.1(a)). The selected solution is randomly perturbed to obtain a new solution that is most likely to be dominated by the others in the non-dominated frontier set (Figure 4.1(b)). Then a local search is applied on this new solution (Figure 4.1(c)) to get an improved solution. A dominance checking (Figure 4.1(d)) is performed to verify if the attained solution is dominated or dominates any solution in the non-dominated solution set.

The proper functioning of the MOILS is depicted in Algorithm 6. The method starts with the set of non-dominated solutions containing only extreme solutions (line 3). How these extreme solutions are got is explained in §4.3.2. For a fixed number of iterations $\mathbb{I}_{\max}$, a solution with the largest crowding distance (see §4.3.3) is selected from the frontier set $\mathbb{S}$ of non-dominated solutions (line 6). The crowding distance allows less explored regions to have higher selection priority during the search. Given a selected solution, the method iterates between a perturbation phase (line 9) followed by a local search step (line 10) as long as it is possible to insert



(a) selection (b) perturbation (c) local search (d) frontier insertion

Figure 4.1: MOILS general ideal.

non-dominated solutions in the frontier $\mathbb{S}$ (line 11) or the maximum number of iterations $\mathbb{C}_{\max}$ without updating the frontier is reached. The attained solution after the local search is only inserted into the frontier set if Pareto dominance checking is successful (§4.3.3). Assis et al. (2013) suggest to perform the local search and then do the dominance checking as well as the use of several neighborhood structures for each considered objective. These suggestions are here altered and improved given rise to four different heuristics based on the MOILS framework of Algorithm 6.

Algorithm 6 MOILS framework

1: $\mathbb{H}_{\max}$: maximum number of iterations
2: $\mathbb{C}_{\max}$: maximum number of iterations without frontier insertion
3: $\mathbb{S} \leftarrow InitialExtremeSolutions()$
4: $h \leftarrow 0$
5: **while** $h < \mathbb{H}_{\max}$ **do**
6:     $s \leftarrow CrowdingDistanceSelection(\mathbb{S})$
7:     $c \leftarrow 0$
8:     **while** $c < \mathbb{C}_{\max}$ **do**
9:         $s' \leftarrow Perturbation(s)$
10:         $s'' \leftarrow LocalSearch(s')$
11:         **if** $DominanceCheckingInsertion(\mathbb{S}, s'') = True$ **then**
12:             $s \leftarrow s''$
13:             $c \leftarrow 0$
14:         **else**
15:             $c \leftarrow c + 1''$
16:         **end if**
17:     **end while**
18:     $h \leftarrow h + 1$
19: **end while**

4.3.1 Solution representation

Each solution consists of a set of routes with a double-linked structure as proposed by Li et al. (2007). This structure reduces computational time because it stores the routes as the predecessor and the successor nodes of each bus stop of a route. Moreover a sorted, fixed-length neighbor list is assembled for each node with the bus stops that are within 60% of the largest distance among the bus stops of the instance being addressed. This neighbor list is restricted to have at most 20% of the number of total nodes.

4.3.2 Extreme solutions: initial solutions

The first solutions to be inserted in the frontier of non-dominated solutions are the extreme solutions, i.e. solutions posited as the best ones for each objective. They represent the extremest points of the Pareto set for each objective. When an extreme solution for a particular objective is calculated, the other two objective are set to zero. To generate each extreme solution, different strategies are adopted:

Total routing and fixed costs: The extreme solution for the total cost is obtained after running a heuristic devised at chapter 3 for the mixed load capacitated rural school bus routing problem with heterogeneous fleet problem. The heuristic is based on a Iterated Local Search (ILS) meta-heuristic which combines a Clark and Wright savings procedure (Clark and Wright, 1964) to generate the initial solutions and a random variable neighborhood descent (Hansen and Mladenović, 2001) search with four different neighborhood structures. The best overall solution is inserted into the frontier of non-dominated solutions.

Total weighted traveling time: To generate routes with the least weighted traveling time, one bus is assigned to each student node. This bus has the smallest capacity possible such that it can still serve the student node demand.

Routes riding time unbalance: To generate the extreme solution for the routes length unbalance objective, one bus is firstly allocated for each student node. Each bus is selected with the smallest possible capacity such that it can still serve the student node demand. Pair of routes are then merged, followed by a 2Opt procedure to avoid artificial balancing. If the merge improves the overall balance then this merge is executed and the balance is updated. If the merge exceeds the bus capacity, a larger available bus is used. Otherwise the merge is not carried on. The procedure keeps verifying the merge of pair of routes until no balance betterment is possible. The routes riding time unbalance is measured by subtracting the largest riding time by the shortest one.

4.3.3 Crowding distance and dominance checking

The crowding distance ω_s of a solution s of a set $\mathbb{S}$ of non-dominated solutions provides an estimation of the density of the solutions surrounding solution s (Deb, 2001). It is computed as the summation of each normalized distance ϖ_s^j for each objective $j = 1..n_o$, where n_o is the number of objectives. ϖ_s^j is calculated as the distance difference of the two adjacent solutions of a solution s normalized by the largest distance of the referred objective j being addressed. Distance is interpreted as the difference in value of the objective function of two different solutions. Algorithm 7 describes how the crowding distance is computed for each solution $s \in \mathbb{S}$. Lines 7 and 8 computes the normalized distances on the borders; while line 10 computes the normalized value from adjacent solutions to a solution s. Line 14 shows the summation of the crowding distance for solution s. A larger crowding distance indicates a solution which has fewer solutions around it, being a natural candidate to have its neighborhood searched for other solutions. This way the diversity of the frontier set of non-dominated solutions is most likely to be increased.

Algorithm 7 Crowding Distance

1: n_o : number of objectives
2: $\mathbb{S}$: set of non-dominated solutions
3: $n_s \leftarrow |\mathbb{S}|$
4: **for** $j = 1$ to n_o **do**
5:     sort S in non-descending order by φ^j
6:     $\max_\varphi \leftarrow \varphi^j_{n_s} - \varphi^j_1$
7:     $\varpi^j_1 \leftarrow 2(\varphi^j_2 - \varphi^j_1) / \max_\varphi$
8:     $\varpi^j_{n_s} \leftarrow 2(\varphi^j_{n_s} - \varphi^j_{(n_s-1)}) / \max_\varphi$
9:     **for** $s = 2$ to $(n_s - 1)$ **do**
10:         $\varpi^j_s \leftarrow (\varphi^j_{(s+1)} - \varphi^j_{(s-1)}) / \max_\varphi$
11:     **end for**
12: **end for**
13: **for** $s = 1$ to n_s **do**
14:     $\omega_s \leftarrow \sum_{j=1}^{n_o} \varpi^j_s$
15: **end for**

When a new solution is obtained, it has to be verified if it is dominated or not by the other solutions in the frontier set of non-dominated solutions. Hence during the search for non-dominated solutions, a dominance checking is performed every time a candidate solution is found. This dominance checking is based on the Pareto dominance of Definition 1.

Definição 1. *Pareto Dominance: A solution $s_1(\varphi^1_1, \ldots, \varphi^{n_o}_1)$ dominates a solution $s_2(\varphi^1_2, \ldots, \varphi^{n_o}_2)$, denoted by $s_1 \prec s_2$, if and only if $\varphi^j_1 \leq \varphi^j_2, \forall j \in \{1, \ldots, n_o\}$, and $\exists r \in \{1, \ldots, n_o\}$ such that $\varphi^r_1 < \varphi^r_2$ (Deb, 2001).*

4.3.4 Neighborhood Structures

Six neighborhood structures inspired on the works of Laporte and Semet (2002) and Gendreau et al. (2002) are used to find non-dominated solutions to populate the frontier of solutions. The structures are divided in two types: inter-routes (Figure 4.2) and intra-routes (Figure 4.3).

Inter-routes structures search for improving solutions doing movements between pair of routes. Four different structures are proposed: **(a) The one-point move** relocates a student node from a route to a different route. The Figure 4.2(a) shows that the student node 5 is moved from a dashed line route to the bold line route. The order of the visited schools needs to be rearranged since the bold line route is not required to visit the gray school anymore. **(b) The two-point move** exchanges one student node of a route by another of a different routes. In the example on Figure 4.2(b), the student nodes 5 and 3 are exchanged from their original routes. The order of schools visit has also to be rearranged. **(c) cross-exchange move** removes arcs (i, j) of (i', j') two different nodes and replace them by reconnecting as (i, j') and (i', j). Figure 4.2(c) shows that arcs $(6, 4)$ and $(3, school)$ are reconnect as two new arcs $(6, school)$ and $(3, 4)$. **(d) merge routes move** two random routes are chosen and merged respecting the largest bus capacity. The last node of a route is connected with the first one of another route. If the movement improves the objective function than it is executed; otherwise, another merge is tried. In Figure 4.2(d) dashed and bold line routes are merged through the connection of nodes 3 with node 7.

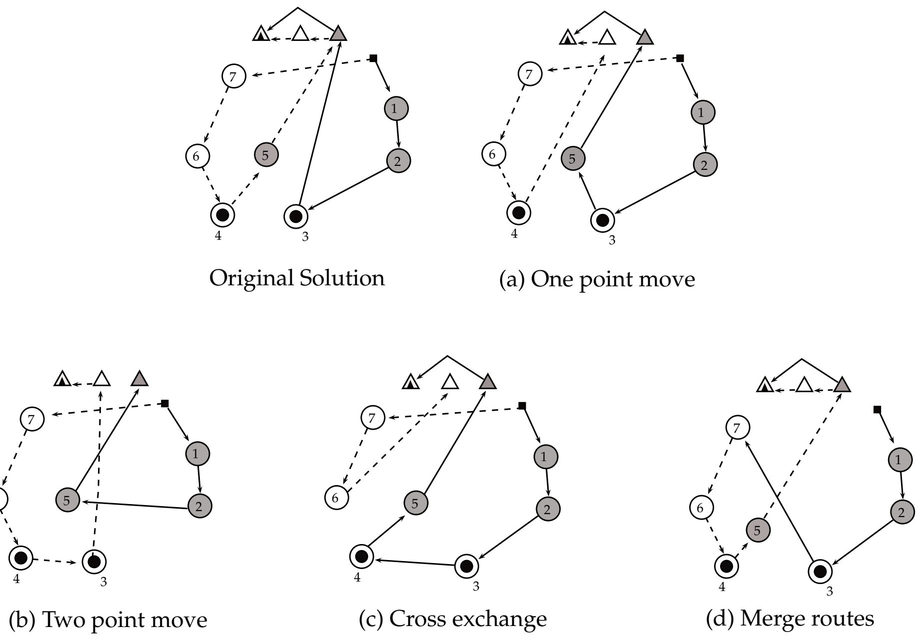


Figure 4.2: Inter-routes local search operators.

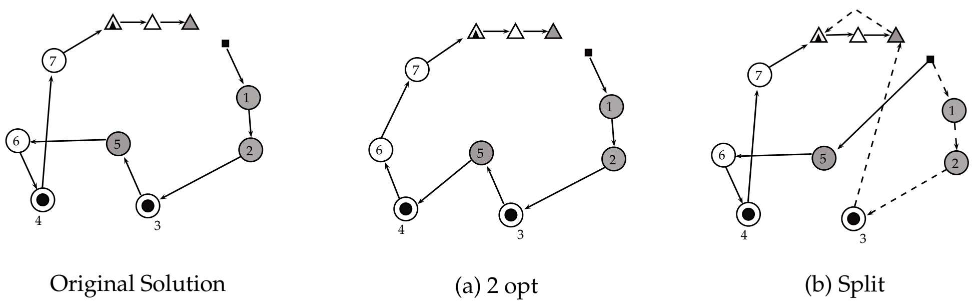


Figure 4.3: Intra-routes local search operators.

Intra-routes apply a movement in a route. Only two structures are devised: **(a) 2opt move** removes two non-consecutive arcs,e.g. (i, j) and (i', j'), of a route and reconnects them by linking the heads and the tails together, or (i, i') and (j, j'). In Figure 4.3(a), the original solution is modified by removing arcs $(5, 6)$ and $(4, 7)$, replacing them for new ones $(5, 4)$ and $(6, 7)$. **(b)**

split move randomly selects a route and divides it into two new routes on the arc next to the middle of the route. In Figure 4.3(b) the original solution has its route split into two new ones by removing the arc $(3,5)$, the arc closest to the middle of the route.

Is important to remark that the split move is applied for all of the objectives but the total cost one, because this movement does not improve the total cost function. One has to pay for an additional bus. The first improving movement strategy is used for all the neighborhood structures. The movements are applied in the pickup section, and for each considered movement some modifications in the delivery section may be necessary. For that purpose it is observed that there are only three cases where the delivery section of the route will need to be changed: (i) if schools are no longer needed to be visited in a route; (ii) if new schools are required to be visited in a route; and (iii) if the last visited student node is changed. Checking for these three cases is very time consuming. These neighborhood structures are organized in a random VND.

Given a current or initial solution and a list of available neighborhood structures, in a random VND, a neighborhood structure is randomly selected among the list of available structures. Then a local search is performed in this neighborhood with the hope of improving the current solution. If successful, the current solution is updated. The list of available movement structures is also augmented with any removed neighborhoods. In case of not getting a better solution, the current movement structure is removed from the list and the process restarted. These steps go on until the list of available neighborhood structures is not empty.

4.3.5 *Walk* procedure

Generally speaking, the Path Relinking (PR) (Glover et al., 2000) strategy consists in generating new solutions from a given solution, but having as target a solution randomly selected from an elite set. The elite set usually has a restricted number of the best overall solutions that are structurally different, i.e. that have some degree of dissimilarity. The idea is to "walk" from a guided solution over a path of solutions towards a guiding solution in the elite set by performing small structural changes in each step. At each step, a local search is carried out with the hope of obtaining better solutions. In the present case, the dissimilarity of a pair of solutions is measured by the summation of the total number of different arcs used in the routes.

For the devised heuristics, two different schemes based on the PR method are implemented. One is an actual PR phase embedded into MOILS based heuristics. This PR phase maintains an elite set for each objective which is constantly updated if necessary. The other scheme resembles a PR method, but instead of keeping an elite set for each objective, the adjacent solutions of the solution selected by the crowding distance procedure is used as guiding and guided solutions. This scheme greatly increases the diversity of the frontier set of non-dominated solutions.

Both PR based schemes rely on a *Walk* procedure. This procedure executes the best possible movement among splitting or merging routes, and inter and intra node moves. These movement structures are done in separated phases. In the beginning, the *Walk* procedure effort is toward having the guided solution with the same number of routes of the guiding solution. This is accomplished by applying splitting or merging routes routines. After that, routes in the guided solution are relabeled so that each route has a correspondent route in the guiding solution. Inter route node movements are then performed so that at the end routes of the guided solution will have the same nodes of the guiding solution. A node is inserted in a route after observing the best improvement possible. Finally, intra route node movements are done so that the routes will have the same order of visitation. After each phase, a random VND is realized with the hope of obtaining better solutions and improving the quality of the frontier set of non-dominated solutions. As the PR based schemes are computationally time consuming, a pre-specified number of iterations is actually set to limit this effort. So at the end the implemented PR based procedures may not have an exact match, but a guided solution that resembles the guiding solution. At each non-dominated solution found during the *Walk* procedure, the frontier of non-dominated solutions is updated accordingly.

4.4 Proposed heuristics based on the MOILS

The four devised heuristics are based on the MOILS framework of Assis et al. (2013) illustrated in Algorithm 6, but with minor modifications and enhanced procedures to improve the overall quality of the attained solutions. Assis et al. suggest the adoption of many distinct specialized neighborhoods for each addressed objective. They propose the use of all of these movement structures during the local search phase which can possibly result in a zigzag phenomenon. Depending on how the neighborhood structures are organized or selected, a neighborhood may improve one objective, while the next one to be used in sequence will most likely go in a different direction, i.e. it will improve a conflicting different objective, and will most probably worsen the previous gain obtained by the former movement structure. Further, as the achieved solution is only verified to be inserted into the frontier of non-dominated solutions at the end of the local search phase, a failure in this test results in a waste of computational effort and time. To circumvent the implementation of many different neighborhood structures, the zigzag phenomenon, and the possibility of failure after the local search phase, a modified MOILS algorithm is proposed.

4.4.1 A modified multi-objective iterated local search (MOILS-M)

To increase the success rate of generating good non-dominated solutions, the neighborhood structures of §4.3.4 were slightly modified. The policy of adopting the first improving move was changed to the first improving move which results in a non-dominated solution, i.e. the frontier set is passed as an input to the local search phase so that the dominance checking (§4.3.3) can be done along the search inside the neighborhood structure. Further, instead of having different specialized neighborhood structures for each objective, each neighborhood is set to work with any objective. To accomplish that, the search direction (e.g. the objectives costs, weighted riding time, routes unbalancing) has to be inform in the beginning of the local search. For instance, the one point move in Figure 4.2(a) modifies a route in its cost, in the total weighted traveling time, and the unbalance of the routes. Hence it can be used to obtain good solutions with respect to any of the objectives, given that the appropriate direction is adopted. So when trying to improve the unbalance of the routes, the search direction of the moves carried out by the one point move has to be in this direction, and not, for example, toward decreasing the costs. To prevent the zigzag phenomenon, the local search is performed separately for each objective one at a time.

The aforementioned gives rise to the modified MOILS presented in Algorithm 8. This pseudo-algorithm has the same steps of Algorithm 6, but with the difference depicted by lines 11-14, and variables $\mathbb{I}$ and Υ. The local search is done in lines 11-14. Note that it receives as input not only the perturbed solution but the frontier set of non-dominated solutions, the current objective being addressed, and variable $\mathbb{I}$ and parameter $\mathbb{I}_{max}$. As the added enhancements to the MOILS greatly increases the cardinality of the frontier (see §4.6.3 for the results), parameter $\mathbb{I}$ counts the number of insertions made for the current objective. An insertion limit $\mathbb{I}_{max}$ is established to prevent too many insertions in the neighborhood of a single solution. If any solution is added to the frontier set $\mathbb{S}$, then variable Υ receives the value true. Otherwise Υ is set to false. If the maximum number of insertions are made, then Υ is also set to false. Whenever the number of unsuccessful insertions reach $\mathbb{C}_{max}$, a new solution is selected by the crowding distance selection procedure of line 7. The algorithm iterates for $\mathbb{H}_{max}$ iterations when it returns the frontier set to the decision maker for analysis.

Algorithm 8 Modified MOILS (MOILS-M)

1: $\mathbb{H}_{max}$: maximum number of iterations
2: $\mathbb{C}_{max}$: maximum number of iterations without frontier insertion
3: $\mathbb{I}_{max}$: maximum number of insertions in the frontier
4: $\mathbb{S} \leftarrow InitialExtremeSolutions()$
5: $h \leftarrow 0$
6: **while** $h < \mathbb{H}_{max}$ **do**
7:     $s \leftarrow CrowdingDistanceSelection(\mathbb{S})$
8:     $c \leftarrow 0$
9:     **while** $c < \mathbb{C}_{max}$ **do**
10:         $s' \leftarrow Perturbation(s)$
11:         **for** $j = 1$ to n_o **do**
12:             $\mathbb{I} \leftarrow 0$
13:             $\Upsilon \leftarrow LocalSearch(s', \varphi^j, \mathbb{S}, \mathbb{I}, \mathbb{I}_{max})$
14:         **end for**
15:         **if** $\Upsilon = True$ **then**
16:             $c \leftarrow 0$
17:         **else**
18:             $c \leftarrow c + 1$
19:         **end if**
20:     **end while**
21:     $h \leftarrow h + 1$
22: **end while**

4.4.2 Multi-objective iterated local search with a standard path relinking (MOILS-PR)

In this variant of the MOILS-M (Algorithm 8), a standard path relinking (PR) procedure (Martí et al., 2011) is embedded into the method. The PR integrates intensification and diversification in the search process, exploring paths that connect previously found solutions (guided solutions) with high quality solutions (guiding solutions) stored in an elite set (Ξ). The idea is similar to the MOILS-PRA (Algorithm 10). However in the standard PR, the guiding solution is chosen from an elite set (see lines 13 of Algorithm 9) instead of the adjacent solutions of the solution with the largest crowding distance. At every $\mathbb{W}$ iterations, an elite set Ξ is assembled for every objective (line 12). Then an elite solution s_e is randomly selected from the elite set Ξ (line 13). Then the method proceeds in the same way as Algorithm 10. The elite set Ξ is assembled with the best overall solutions of the frontier $\mathbb{S}$ with respect with every objective, but with at least 10% of dissimilarity.

Algorithm 9 MOILS with standard path relinking procedure (MOILS-PR)

1: $\mathbb{H}_{\max}$: maximum number of iterations
2: $\mathbb{C}_{\max}$: maximum number of iterations without frontier insertion
3: $\mathbb{I}_{\max}$: maximum number of insertions in the frontier
4: $\mathbb{S} \leftarrow InitialExtremeSolutions()$
5: $h \leftarrow 0$
6: $w \leftarrow 0$
7: **while** $h < \mathbb{H}_{\max}$ **do**
8: $\quad s \leftarrow CrowdingDistanceSelection(\mathbb{S})$
9: $\quad$ **if** $w = \mathbb{W}$ **then**
10: $\quad\quad w \leftarrow 0$
11: $\quad\quad$ **for** $j = 1$ to n_o **do**
12: $\quad\quad\quad \Xi \leftarrow EliteSet(\mathbb{S}, \varphi^j)$
13: $\quad\quad\quad s_e \leftarrow RandomSelection(\Xi)$
14: $\quad\quad\quad \mathbb{D} \leftarrow DissimilarityDegree(s, s_e)$
15: $\quad\quad\quad Walk(s, s_e, \varphi^j, \mathbb{S}, \mathbb{I}_{\max}, \mathbb{D})$
16: $\quad\quad\quad Walk(s_e, s, \varphi^j, \mathbb{S}, \mathbb{I}_{\max}, \mathbb{D})$
17: $\quad\quad$ **end for**
18: $\quad$ **else**
19: $\quad\quad w \leftarrow w + 1$
20: $\quad$ **end if**
21: $\quad c \leftarrow 0$
22: $\quad$ **while** $c < \mathbb{C}_{\max}$ **do**
23: $\quad\quad s' \leftarrow Perturbation(s)$
24: $\quad\quad$ **for** $j = 1$ to n_o **do**
25: $\quad\quad\quad \mathbb{I} \leftarrow 0$
26: $\quad\quad\quad \Upsilon \leftarrow LocalSearch(s', \varphi^j, \mathbb{S}, \mathbb{I}, \mathbb{I}_{\max})$
27: $\quad\quad$ **end for**

        if $\Upsilon = \mathit{True}$ **then**
            $c \leftarrow 0$
        else
            $c \leftarrow c + 1$
        end if
    end while
    $h \leftarrow h + 1$
end while

4.4.3 A multi-objective iterated local search with a path relinking procedure for the crowding distance adjacency (MOILS-PRA)

An enhanced feature is added to Algorithm 8 given rise to Algorithm 10. A path relinking based procedure or *Walk* procedure (§4.3.5) is executed every $\mathbb{W}$ iterations. The idea is to increase the diversity of the frontier $\mathbb{S}$ on its most sparse regions. Further, by doing a random VND along the *Walk* procedure, not only diversification but intensification as well is also sought to further improve the quality of the frontier set. In Algorithm 10, the immediate predecessor (s_p) and successor (s_p) solutions (i.e. adjacent solutions) of the solution (s) with the current largest crowding distance, see lines 10-20, are selected to be used as the guiding and guided solutions. The degree of dissimilarity $\mathbb{D}$ (line 10) between both solutions is used as a stopping criterion for the *Walk* procedure. If in $\mathbb{D}$ moves the solutions do not match, the procedure is stopped. The walk along the "path" of solutions from s_p to s_s is done in both directions for every objective. Recall from §4.3.5 that on every change of movement phase of the *Walk* procedure, a random VND is executed. The number of allowed insertions into the frontier of non-dominated solutions $\mathbb{S}$ during the random VND calls is given by $\mathbb{I}_{\max}$. The remainder of the algorithm is the same of Algorithm 8.

Algorithm 10 MOILS with a path relinking procedure for the criterion of the crowding distance adjacency (MOILS-PRA)

1: $\mathbb{H}_{\max}$: maximum number of iterations
2: $\mathbb{C}_{\max}$: maximum number of iterations without frontier insertion
3: $\mathbb{I}_{\max}$: maximum number of insertions in the frontier
4: $\mathbb{S} \leftarrow \mathit{InitialExtremeSolutions}()$
5: $h \leftarrow 0$
6: $w \leftarrow 0$
7: **while** $h < \mathbb{H}_{\max}$ **do**
8:     $s \leftarrow \mathit{CrowdingDistanceSelection}(\mathbb{S})$
9:     **if** $w = \mathbb{W}$ **then**
10:         $w \leftarrow 0$
11:         $s_p \leftarrow \mathit{pred}(s)$
12:         $s_s \leftarrow \mathit{succ}(s)$
13:         $\mathbb{D} \leftarrow \mathit{DissimilarityDegree}(s_p, s_s)$

        for $j = 1$ to n_o **do**
            $Walk(s_p, s_s, \varphi^j, \mathbb{S}, \mathbb{I}_{\max}, \mathbb{D})$
            $Walk(s_s, s_p, \varphi^j, \mathbb{S}, \mathbb{I}_{\max}, \mathbb{D})$
        end for
    else
        $w \leftarrow w + 1$
    end if
    $c \leftarrow 0$
    while $c < \mathbb{C}_{\max}$ **do**
        $s' \leftarrow Perturbation(s)$
        for $j = 1$ to n_o **do**
            $\Upsilon \leftarrow LocalSearch(s', \varphi^j, \mathbb{S}, \mathbb{I}_{\max})$
        end for
        if $\Upsilon = True$ **then**
            $c \leftarrow 0$
        else
            $c \leftarrow c + 1$
        end if
    end while
    $h \leftarrow h + 1$
end while

4.4.4 Multi-objective iterated local search combining MOILS-PRA with MOILS-PR (MOILS-PRA-PR)

The last devised algorithm is a combination of both Algorithms 10 and 9. For sake of presentation, it is not shown here. Basically, the algorithm executes both path relinking strategies at every $\mathbb{W}$ iterations.

4.5 Visualization tool for selecting solutions

At the end of a multi-objective optimization, a decision maker is often faced with a great number of different non-dominated solutions from which he has to chose the most suitable for his needs. This process is not an easy task given the many possibilities available. The complexity further increases when the studied problem has a combinatorial nature. In these cases, graphical representations of the objectives do not show any clear tendency on which solution to pick up. This is specially true when the number of addressed objectives grow as in the present work.

Though there is a rich literature (Gettinger et al., 2013; Freitas et al., 2014) about developing elaborated tools for selecting the most suitable solution; a simple, easy to use method has a great appealing motivation. The general idea of the tool here introduced relies on a hierarchization of the frontier solutions after the computation of a weight for each solution. The weights are calculated by computing the area formed by polygon originated from the normalization of the objectives plotted in polar graph. This area is then elevated to a scalar which represents the total unbalance length (pairwise difference of the sides' length) of the sides of this polygon. In other words, the weights are a power law function of the area of the polygon by the total unbalance of the sides.

The rationality behind such scheme is simple, but intuitive. Usually, the objectives have different scale magnitudes what makes difficult to compare them. To understand the suitability of a solution against others the objective function values can be normalized, i.e. the values on different scales can be adjusted into a common one. After normalized, a given solution can be plotted in a polar graph to obtain a polygon. The closer this polygon is to the origin (zero value for the axis) the more suitable is the solution, since all objectives are at their minimum value. Here it is assumed that all objectives are of a minimization type; otherwise one has just to invert the axis orientation. Instead of browse every single polar plot for the most suitable (note that the number of solutions is more likely in the thousands), one can rely on a hierarchization of the area of the polygon formed by the plot elevated to the power of the unbalance of the sides. And then, only the first on this hierarchy will be analyzed. Using only the area in the hierarchization process is not enough to assess the quality of the solution since. For instance, in the present work which there is three objectives, it is possible to have two objectives closer to the zero and the remaining objective at its worst possible value. This will most likely generate an area with a small value, putting this solution at the top of the hierarchy list. To prevent these cases the area value is elevated to the total pairwise unbalance length of the polygon sides. This way the larger the absolute difference of the sides, the larger will be the weight.

More specifically, given the frontier set $\mathbb{S}$ of non-dominated solutions obtained after a multi-objective optimization. Let $n_s = |\mathbb{S}|$ be the number of solutions in $\mathbb{S}$, n_o be the number objectives or $(\varphi^1, \ldots, \varphi^{n_o})$. The normalization is done by dividing the values of each solution objective φ^j by the difference between the maximum by the minimum for each objective or $\Delta^j = \max_{s\in\mathbb{S}} s(\varphi^j) - \min_{s\in\mathbb{S}} s(\varphi^j)$, for all $j = 1, \ldots, n_o$ ors:

$$s^N(\varphi^1, \ldots, \varphi^{n_o}) = (\frac{\varphi^1}{\Delta^1}, ..., \frac{\varphi^{n_o}}{\Delta^{n_o}}) \qquad \forall s \in \mathbb{S} \tag{4.22}$$

With the solutions normalized, each objective can be assessed by visualizing them in a polar graphic. For a problem with three objectives, the polar graph shape is a triangle. Hence the weight w_s of a solution $s \in \mathbb{S}$ is set to be:

$$w_s = (A_s^{\delta_s})$$

where A_s is the area of the polygon in the polar graph and δ_s is the given by

$$\delta = 1 + \sum_{i\in\mathscr{S}} \sum_{j\in\mathscr{S}:i\neq j} |\ell_i - \ell_j|$$

where $\mathscr{S}$ is the set of sides of the polygon of the polar graph, and ℓ_i is the length of the side i of the polygon. After calculating the weights w_s for each solution $s \in S$, they solutions can be sorted in an increasing order. The first solutions can then be chosen to be plotted for the decision maker screening.

4.6 Computational experiments

This section describes the computational experiments performed to test the efficiency of the aforementioned heuristics. All algorithms were coded in C++, compiled with GCC 4.8.1., and tested on an Intel Xeon 2.53 GHz with 24GB RAM running Linux Mint 16. For the experiments, three bus types with capacities of 20, 30 and 40 seats were made available to transport the students. The fixed costs (daily fixed depreciation costs) and the routing costs for each bus type were set to $100, $150, and $200, and $1.00, $1.20 and $1.40 per unit of traveled distance, respectively. The daily depreciation costs (fixed costs) were estimated by assuming a bus lifespan of 10 years and the distances between nodes were considered to be Euclidean.

To test the devised heuristics, 15 random instances were generated to depict Brazilian counties with different sizes. The bus stops were scattered within the set $\{50, 75, 100, 150\}$ in an area of $155mi^2$(12,5mi x 12,5mi). The number of schools were varied according to the set $\{5, 10, 20\}$. A total of 12 instances were generated for assessing the devised heuristics. For the calibration phase, three instances with the number of stops and schools set to $(50, 5)$, $(75, 5)$ and $(100, 10)$ were also created. A total of 15 random instances were devised. Moreover 20% of the bus stops as well as all of 100% of the schools were located inside an imaginary downtown area with a radius of 1.24 miles. The remaining nodes were located on the outside of the radius. For each node, a school and its respective demand were uniformly selected. The demands q_l, $l \in P$, were generated within the range of 1 to 3 students.

In multi-objective optimization, one is interested in finding the set of frontier non-dominated solutions, which keeps the best compromise solutions among all the objectives. In most cases, it is not possible to compute the Pareto optimal set, thus the problems try to computing a set of non-dominated solutions which is as close as possible to the Pareto optimal set, which is a set of non-dominated solutions which is never been dominated by another explored solutions (Batista et al., 2014).

As the obtained subset of non-dominated solutions typically contains worse solutions than the Pareto optimal frontiers, there is a need to analyze the quality of these frontiers. One can use a metric for diversity or convergence to measure the Pareto set quality but once they can be conflicting, is advised to use more than one (Deb, 2001). For that purpose three metrics are considered:

1. **Cardinality (Car):** Number of solutions in the Pareto frontier. Its purpose is to verify the efficiency of the developed algorithm in finding non-dominated points. It is assumed that the decision maker prefer rather more options than fewer efficient solutions (Martí et al., 2011).

2. **Coverage of many sets(CS):** This metric was proposed by Batista et al. (2014) and is a generalization of the Coverage of Two Sets (Zitzler and Thiele, 1999). This metric indicates the percentage of a set $\mathbb{U}$ that is dominated by a set $\mathbb{S}$. It quantifies the domination

of a frontier over the union of the remaining ones. The function which gives this value is stated as:

$$CS(\mathbb{S}_i, \mathbb{U}_i) = \frac{|\{s_1 \in \mathbb{U}_i : \exists s_2 \in \mathbb{S}_i \wedge s_2 \preceq s_1\}|}{|\mathbb{U}_i|}$$

where $\mathbb{S}_i$ represents the Pareto frontier of algorithm i, for all $i = 1, \dots, n_a$, and n_a is the number of available algorithms, and $\mathbb{U}_i$ is the union of all Pareto frontier of all algorithms except algorithm i. This super set is defined as:

$$\mathbb{U}_i = \bigcup_{\substack{j=1 \\ j \neq i}}^{n_a} A_j \qquad \forall i = 1, \dots, n_a$$

The expression $s_2 \preceq s_1$ means that the solution s_2 is no worse than s_1 in all objectives and the solution s_2 is strictly better than s_1 in at least one objective. This concept defines weakly dominance. The function CS maps the pair ($\mathbb{S}_i$,$\mathbb{U}_i$) within the interval $[0, 1]$. The value $CS(\mathbb{S}_i, \mathbb{U}_i)$ equal to 1 means that all points in $\mathbb{U}_i$ are dominated by or equal to the points of $\mathbb{S}_i$. The opposite result $CS(\mathbb{S}_i, \mathbb{U}_i)$ to 0 represents that the frontier $\mathbb{S}_i$ of the algorithm ß is dominated by the others.

3. **Hyper-volume:** Introduced by Zitzler and Thiele (1999), this indicator measures the volume of the region dominated by the Pareto front with respect to a reference point which can be defined by a vector of the worst values for the objective function. In this work the results of hyper-volume metrics were normalized by dividing the difference between the current value and the minimum value of the objective function by the difference of the maximum and minimum value of the objective function. The reference point was set to $\mathscr{H}_0(1, 1, 1)$.

   Mathematically, for each solution $s \in \mathbb{S}$, a hyper-cube $\mathscr{V}_s$ is created accordingly to the reference point $\mathscr{H}_0$. The final result is the sum of all obtained hyper-cubes. Assuming minimization objectives, a higher value for the hyper-volume evinces a higher spread among solutions of the Pareto front as well as a higher convergence. A more detailed explanation about how to compute hyper-volume please refer to Beume et al. (2009) and Bradstreet (2011).

4.6.1 Calibration phase

1. **Maximum number of iterations ($\mathbb{H}_{\max}$):** To define the parameter $\mathbb{H}_{\max}$, required by all algorithms, 15 runs were executed with the values {50, 100, 150} for each one of the three calibration instances. The results showed that $\mathbb{H}_{\max}$ = 100 presented better results than $\mathbb{H}_{\max}$ = 50 for cardinality, hyper-volume and coverage in most methods. $\mathbb{H}_{\max}$ = 100 also had better performance than $\mathbb{H}_{\max}$ = 150 except for cardinality. For this metric, it was expected because the higher number of iterations of $\mathbb{H}_{\max}$ = 150 allows more solutions to

get into the Pareto set increasing its result. Thus, based on the results the $\mathbb{H}_{max}$ parameter was settled as 100.

2. **Maximum number of iterations without insertion ($\mathbb{C}_{max}$):** The parameter $\mathbb{C}_{max}$ (maximum number of iterations without any insertion in the front, algorithm 6) was tested at the same aforementioned calibration instances with the values {1, 2} and 5 execution for each algorithm and each instance. The best values for cardinality, coverage and hyper-volume were obtained when $\mathbb{C}_{max} = 1$ mainly because as many solutions are inserted in the frontier $\mathbb{S}$, the counter c takes a long time to reach the value $\mathbb{C}_{max} = 2$. In this case, many solutions are also inserted, however, they are close to each other, which does not improve the hyper-volume or the spread, besides preventing farther neighborhoods to be visited. Thus, this metric was settled in $\mathbb{C}_{max} = 1$.

3. **Maximum insertion ($\mathbb{I}_{max}$):** The value of $\mathbb{I}_{max}$ is set to the number of stops of the instance. As the instance size increases, the number of values to be inserted also increases.

4. **Iterations between *Walk* procedure calls ($\mathbb{W}$):** The *Walk* procedure is very time consuming procedure, thus it can not be applied at each iteration. However the number of iterations should not be set as a fixed number, instead it should vary with the instance size. Three different values were tried $\lceil(n/2)\rceil$, $\lceil(n/4)\rceil$ and $\lceil(n/8)\rceil$. The value $\lceil(n/8)\rceil$ presented the best overall results.

4.6.2 Statistical Analysis

To evaluate the performance of the methods, a statistical analysis was made to identify its differences and magnitude, if exists. For each metric a pairwise comparison was applied independently through T test, summing ten null hypothesis of absence of difference between algorithms.

Once many tests were performed a correction on alpha level was necessary to prevent **Type I** error, reject null hypothesis when it is true, caused by the inflation of the alpha level. This problem usually happens when many tests are executed and the more test are performed on a set of data, the more likely is to reject the null hypothesis as a consequence of the logic of hypothesis test which is: a null hypothesis is rejected if a rare event is witnessed. But the larger the number of tests, the easier is to find a rare event and therefore to think that there is an effect when there is none. Thus, making the alpha level smaller less error will be accepted, despite of also making harder to detect real effects (Salkind (2006)).

The solution for the alpha inflation was the Bonferroni correction (Hochberg, 1988) which consists on divide the final alpha significance level by the number of tests. As the final significance level desired is 5% and 10 hypothesis has to be tested, the significance level of 0.005 was applied for each T test. The statistical analysis were applied in the average of the results of 8 replication and 5 heuristics applied in the 12 instances in aleatory mode.

Table 4.1: Average (standard deviation) of cardinality.

	Cardinality					
Instance	MOILS-PRA-PR	MOILS-PR	MOILS-PRA	MOILS-M	MOILS	Smallest
A50	12.74 (1.08)	11.70 (1.70)	11.53 (1.25)	8.29 (1.59)	1.50 (0.32)	67
B50	11.59 (1.34)	11.79 (0.87)	10.31 (1.47)	8.19 (0.88)	1.33 (0.28)	79
C50	12.76 (1.56)	11.68 (1.22)	11.85 (1.71)	9.66 (1.72)	1.63 (0.47)	67
A75	20.45 (1.59)	17.13 (2.67)	17.51 (3.12)	13.72 (2.05)	1.45 (0.37)	73
B75	22.66 (3.03)	23.01 (3.73)	18.81 (2.34)	16.77 (2.38)	1.74 (0.37)	62
C75	22.05 (1.48)	20.57 (3.01)	18.29 (1.83)	15.77 (1.02)	1.25 (0.23)	59
A100	17.62 (2.46)	18.39 (3.29)	17.35 (2.21)	13.81 (2.12)	1.20 (0.16)	89
B100	16.98 (2.15)	17.36 (2.75)	16.42 (3.50)	12.08 (1.94)	1.37 (0.26)	101
C100	16.93 (2.62)	14.70 (2.54)	15.25 (1.19)	12.03 (1.54)	1.21 (0.19)	105
A150	54.52 (9.30)	48.38 (11.92)	41.38 (6.31)	34.74 (3.83)	1.51 (0.38)	55
B150	29.08 (5.08)	28.04 (5.58)	24.68 (3.39)	20.72 (2.96)	1.44 (0.30)	95
C150	31.97 (3.62)	31.94 (4.55)	27.52 (4.86)	23.96 (3.01)	1.52 (0.49)	81
Average	22.45 (2.94)	21.22 (3.65)	19.24 (2.77)	15.81 (2.09)	1.43 (0.32)	77.75

Table 4.2: Average (standard deviation) of coverage.

	Coverage				
Instance	MOILS-PRA-PR	MOILS-PR	MOILS-PRA	MOILS-M	MOILS
A50	0.60 (0.18)	0.46 (0.14)	0.40 (0.15)	0.19 (0.05)	0.06 (0.05)
B50	0.61 (0.13)	0.44 (0.14)	0.39 (0.20)	0.27 (0.15)	0.03 (0.01)
C50	0.57 (0.15)	0.42 (0.17)	0.41 (0.13)	0.26 (0.08)	0.09 (0.04)
A75	0.51 (0.10)	0.37 (0.17)	0.52 (0.20)	0.24 (0.12)	0.04 (0.03)
B75	0.55 (0.16)	0.40 (0.14)	0.45 (0.16)	0.24 (0.05)	0.06 (0.03)
C75	0.53 (0.07)	0.37 (0.09)	0.39 (0.14)	0.33 (0.04)	0.04 (0.03)
A100	0.40 (0.11)	0.37 (0.14)	0.50 (0.12)	0.29 (0.16)	0.06 (0.03)
B100	0.47 (0.14)	0.31 (0.12)	0.48 (0.17)	0.28 (0.11)	0.10 (0.06)
C100	0.42 (0.22)	0.34 (0.18)	0.51 (0.09)	0.32 (0.11)	0.05 (0.03)
A150	0.46 (0.16)	0.33 (0.14)	0.38 (0.08)	0.32 (0.14)	0.03 (0.02)
B150	0.39 (0.16)	0.29 (0.15)	0.45 (0.11)	0.38 (0.08)	0.06 (0.04)
C150	0.41 (0.13)	0.28 (0.08)	0.46 (0.13)	0.36 (0.10)	0.07 (0.05)
Average	0.51 (0.14)	0.38 (0.14)	0.44 (0.14)	0.27 (0.10)	0.06 (0.03)

4.6.3 Results

The results obtained from the computational tests are depicted at Tables 4.1-4.4 which report the average for the executions of each instance by each algorithm and the standard deviation, in parenthesis, for the metrics of cardinality, coverage and hyper-volume. The statistical analysis are summarized in Table 4.5.

For each instance, the results of cardinality (see Table 4.1) and time (refer to Table 4.4) were divided by the smallest value obtained among the 8 executions of the five methods to make easier the comparison between them. Observing the statistical analyze in Table 4.5, one can conclude that all proposed methods overcame the literature framework MOILS in all metrics. For the cardinality metric, the algorithm MOILS-M was outperformed by the variants

Table 4.3: Average (standard deviation) of hyper-volume.

	Hyper-volume				
Instance	MOILS-PRA-PR	MOILS-PR	MOILS-PRA	MOILS-M	MOILS
A50	0.93 (0.01)	0.92 (0.01)	0.92 (0.01)	0.90 (0.01)	0.87 (0.04)
B50	0.92 (0.01)	0.91 (0.01)	0.91 (0.01)	0.90 (0.01)	0.85 (0.03)
C50	0.92 (0.01)	0.92 (0.01)	0.92 (0.01)	0.92 (0.01)	0.89 (0.01)
A75	0.94 (0.01)	0.93 (0.01)	0.94 (0.01)	0.93 (0.01)	0.88 (0.02)
B75	0.93 (0.02)	0.93 (0.02)	0.92 (0.02)	0.92 (0.02)	0.86 (0.04)
C75	0.91 (0.01)	0.91 (0.01)	0.91 (0.01)	0.91 (0.01)	0.84 (0.03)
A100	0.90 (0.01)	0.89 (0.02)	0.90 (0.02)	0.90 (0.01)	0.85 (0.03)
B100	0.91 (0.01)	0.91 (0.01)	0.91 (0.02)	0.90 (0.02)	0.87 (0.03)
C100	0.90 (0.01)	0.90 (0.02)	0.91 (0.01)	0.91 (0.02)	0.86 (0.03)
A150	0.90 (0.02)	0.89 (0.02)	0.90 (0.01)	0.90 (0.02)	0.82 (0.04)
B150	0.91 (0.02)	0.90 (0.01)	0.91 (0.01)	0.91 (0.01)	0.86 (0.02)
C150	0.91 (0.01)	0.90 (0.01)	0.92 (0.01)	0.92 (0.01)	0.87 (0.03)
Average	0.92 (0.01)	0.91 (0.01)	0.92 (0.01)	0.91 (0.01)	0.86 (0.03)

Table 4.4: Average (standard deviation) of time.

	Time					
Instance	MOILS-PRA-PR	MOILS-PR	MOILS-PRA	MOILS-M	MOILS	Smallest
A50	20.19 (3.34)	17.87 (2.27)	7.95 (3.80)	3.38 (1.41)	1.46 (0.43)	19.12
B50	11.44 (2.35)	11.56 (2.04)	5.74 (1.88)	3.00 (1.24)	1.56 (0.71)	33.35
C50	15.04 (2.04)	12.44 (2.02)	7.49 (1.33)	4.05 (1.39)	1.89 (0.98)	23.09
A75	14.06 (2.93)	13.02 (3.01)	8.22 (1.44)	4.23 (1.29)	1.37 (0.48)	60.16
B75	18.61 (4.10)	15.81 (4.26)	6.23 (1.84)	4.20 (1.10)	1.80 (0.71)	75.97
C75	10.33 (1.38)	8.63 (2.27)	4.80 (1.19)	4.15 (1.53)	1.18 (0.21)	83.55
A100	11.61 (3.05)	13.17 (4.22)	7.69 (2.59)	3.75 (0.83)	1.80 (0.67)	143.05
B100	16.66 (6.89)	12.88 (4.06)	6.17 (2.62)	3.14 (1.32)	2.19 (0.81)	203.61
C100	13.81 (4.70)	10.76 (3.14)	6.66 (1.46)	3.33 (1.08)	2.02 (1.69)	161,10
A150	9.55 (1.64)	9.06 (1.31)	4.56 (1.49)	3.76 (1.56)	1.65 (0.63)	706.02
B150	8.60 (1.81)	7.00 (1.35)	4.39 (2.33)	2.78 (0.88)	1.68 (0.55)	1033.51
C150	9.86 (3.39)	9.56 (1.68)	4.41 (0.89)	3.10 (0.45)	1.56 (0.45)	743.29
Average	13.31 (3.13)	11.81 (2.64)	6.19 (1.90)	3.57 (1.17)	1.68 (0.69)	273.82

MOILS-PRA-PR, MOILS-PRA and MOILS-PR. Among these three methods, the MOILS-PRA was outperformed by the versions MOILS-PRA-PR and MOILS-PR, which was able to generate 7.12 and 5.61 times more solutions than MOILS-PRA, respectively. However, these last heuristics MOILS-PRA, MOILS-PRA-PR and MOILS-PR are statistically equal. Therefore is possible to conclude that the MOILS-PR heuristic embedded in the MOILS-M allows the proposed method to return more non-dominated solutions. Between MOILS-PRA-PR and MOILS-PR, the MOILS-PR is more interesting once it requires smaller computational time to find a better frontier, please refer to Table 4.4.

To examine the results for coverage (see Table 4.2) one can follow the same reasoning above. The MOILS-M was outperformed by the variants MOILS-PRA-PR, MOILS-PR and MOILS-PRA. The frontier of these later methods were cover 20%, 8% and 16% more solutions, re-

Table 4.5: Estimated difference in average performance between the row and column algorithms for the performance metrics.

	Cardinality				Coverage				Hyper-volume			
MOILS	PR	PRA	M	MOILS	PR	PRA	M	MOILS	PR	PRA	M	MOILS
PRA-PR	n.s.[1]	7.12	9.74	11.63	0.13	n.s.	0.20	0.44	0.004	n.s.	n.s.	0.054
PR		5.61	8.24	10.13		-0.08	0.08	0.31		-0.004	n.s.	0.049
PRA			2.62	4.50			0.16	0.39			0.005	0.054
A				1.89				0.23				0.049

1: not statistically significant.
OBS: Only results significant at 95% of confidence level adjusted for multiple hypothesis testing using Bonferroni correction (Hochberg, 1988) are shown. Positive values indicate higher average value for the algorithm in the row.

spectively, than of the version MOILS-M. MOILS-PR was outperformed by MOILS-PRA and MOILS-PRA-PR, which are not different statistically at 5% of significance. Thus, the MOILS-PRA has better performance for the coverage metric. Once MOILS-PR does not get to improve this metric, and that the MOILS-PRA took half of the time to return a good frontier, the MOILS-PRA is considered the best method for this metric.

The last metric to be evaluated is the hyper-volume (see Table 4.3). The statistical analyze showed that the MOILS-PRA-PR and MOILS-PR have no significant difference from MOILS-M. MOILS-PRA had a better performance than MOILS-M and MOILS-PR, with 0.5% and 0.4% bigger hyper-volume, respectively, and no significant difference from MOILS-PRA-PR, which in turn outperformed the MOILS-PR with 0.4% larger hyper-volume. The MOILS-PRA heuristic showed to be once again more interesting. The values obtained for this metrics confirm the above conclusion about the advantage got by MOILS-PRA to the coverage metric. With those results one can also state that the MOILS-PRA has also a better performance for the hyper-volume, which estimates both convergence and spread of the solutions.

The MOILS-PRA is considered the overall best proposed meta-heuristic. Even better than MOILS-PRA-PR because if one consider the three multi-objective metrics, which are coverage, spread and time, the former presented better performance for all of them. Figures 4.4 and 4.5 depict the obtained frontiers by MOILS and MOILS-PRA-PR for the instance A150. The graphics show the relation between each pair of objectives demonstrated by the six different view points of the frontier. The picture confirm the better performance of MOILS-PRA-PR over MOILS. Observing them one can state that MOILS-PRA-PR was able to return a more populated frontier than MOILS with higher diversity of solutions, which in turn are also closer to the axis showing a better convergence of the devised method.

The better performance of MOILS-PRA over MOILS-PR can be explained by the fact that MOILS-PRA explores new empty spaces guided by the crowding distance procedure while MOILS-PR might be exploring a space of good solutions, but already populated by others, not improving thus the hyper-volume or the coverage. Moreover, once the solution of the elite set for the MOILS-PR has to have 10% of different edges and the space of the extreme solution can be over crowded the solutions of the elite set might be not that good. Because if it space is over crowded, a solution with 10% of different edges can be far from the extreme one, so there is a likelihood of the search has been exploring poor paths.

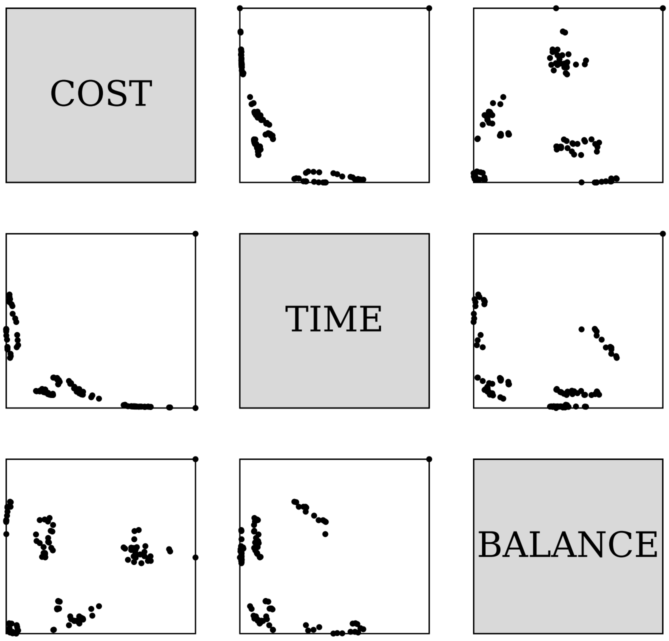


Figure 4.4: MOILS non-dominated solution frontier for instance A150.

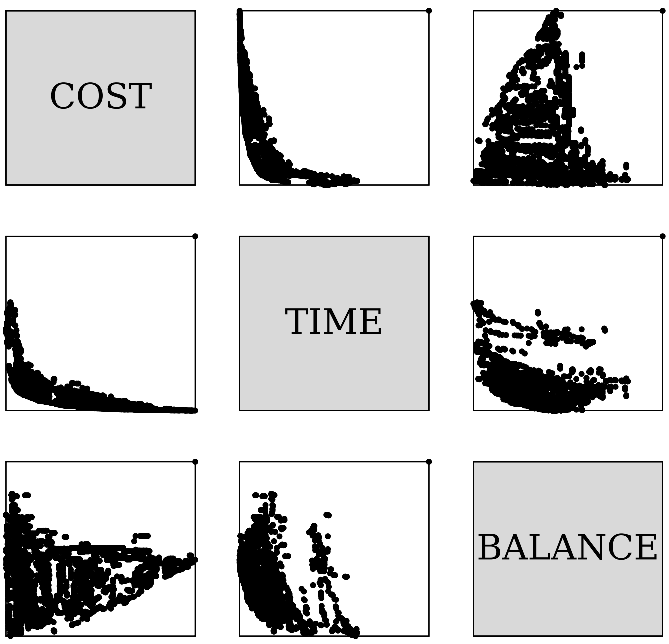


Figure 4.5: MOILS-PRA-PR non-dominated solution frontier for instance A150.

After generating the non-dominated frontiers, the decision maker has to find among all of the results that solution which best suits him. An example of the introduced approach is depicted in Figure 4.6 which shows the polar graphics for the three overall best solution for instance A150. Note that the first and second graphs are similar with a higher value for time and smaller values for cost, while the third has smaller value for time but higher for cost and the balance has values with a slightly difference in all of them. Thus the final conclusion about the result depends on the scenario and which objective will the decision maker prioritize.

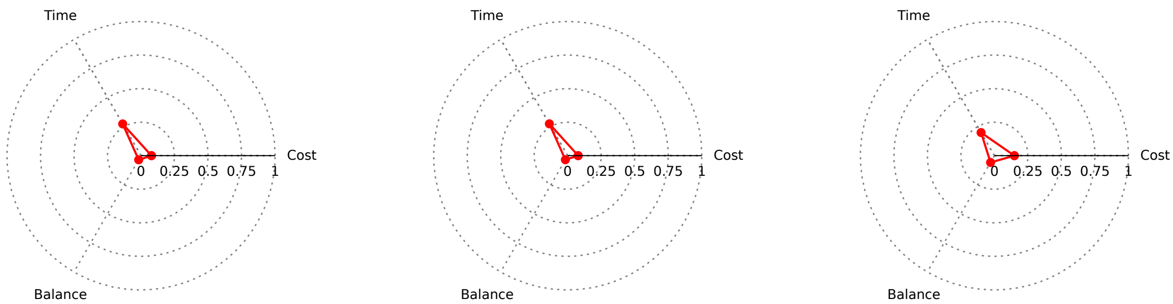


Figure 4.6: Polar graphs for the three best solutions of instance A150 for the MOILS-PRA method.

4.7 Final Remarks

A multi-objective approach for the capacitated rural school bus routing problem with heterogeneous fleet and mixed loads was proposed. The problem considers not only costs, but the average weighted riding distances as well as distance balance among drivers. This is an important problem that has been neglected by the government and literature. It is very common to have short and long routes in the set of paths, given rise to labor complains from the drivers, increasing the hazards for children, due to weariness of drivers and a source of nuisance to public management. However routing costs and travel time play an important role and can not be disregarded. So, a multi-objective problem comes to help decision makers on solving this tough issue. Four meta-heuristics (MOILS-M, MOILS-PR, MOILS-PRA, MOILS-PRA-PR) were proposed and compared among them on a test set of 15 instances randomly generated. A new approach to chose a good solution after generating the non-dominated solution frontier is also introduced. All of the proposed heuristics outperformed the framework proposed by the literature (Assis et al., 2013). The heuristic MOILS-PRA presented the overall best results in the metrics of coverage, spread and time, followed closely by MOILS-PRA-PR. Several issues still remain open for future research. The incorporation of school location decisions combined with maximum riding distance constraints is one of the propositions which can be made for rural areas. Another especial issue for those areas is the possibility of creating transition points. The transshipment allows that smaller vehicles reach areas that larger ones can not, driving the students from home to these points where buses with higher capacity would do the second part of the route until the school.

Chapter 5

Conclusion

This thesis aimed to propose a mathematical formulation and solution methodologies still not found in literature for the Brazilian rural school bus routing problem considering heterogeneous fleet and mixed load. The devised methods seek for a minimum cost routes generation, respecting the specific problem constraints, in order to automate the routes planning, decrease students travel time enabling them to have a higher performance at school. Besides the literature contribution, it also intend to provide better service level for rural population allowing the families of these regions having access to better education systems, life quality and equal opportunities.

To reach the intended goals the thesis presented heuristic and multi-objective methodologies to deal with the referred problem. For the heuristic approach five meta-heuristics were devised: The first one is an adaptation of the Record-to-Record Travel algorithm for solving the heterogeneous fleet vehicle routing problem proposed by Li et al. (2007). The others had as it main structure the Iterated Local Search (ILS) meta-heuristic and the Variable Neighborhood Search (VNS) strategy. Both had the Variable Neighborhood Descent (VND) and the Random Variable Neighborhood Descent (RVND) local search procedures embedded on it separately, giving rise to four meta-heuristics.

The work also introduce four multi-objective meta-heuristics for the . One adapted from literature, named modified multi-objective iterated local search (MOILS-M). The other three have the first one as its main structure with the following procedures: a path relinking procedure for the crowding distance adjacency (MOILS-PRA), the standard path relinking (MOILS-PR) and both of them embedded in MOILS-M. The methods were compared with a heuristc from literature, adapted for this problem, defined as Multi-objective Iterated Local Search (MOILS, Assis et al. (2013)) through statistical analysis and three metrics (cardinality, coverage of many sets and hyper-volume). The work also proposes a new approach to support decision makers on selecting a good solution.

The proposed methodologies has returned good solution for the Brazilian problem, thus attained the expected objectives. The mixed load approach appraised by the heuristic methods showed to be more suitable over the single one, because it attained lower cost savings and number of buses required to serve the students. Among the devised meta-heuristics, the the

ILS-RVND-ML has provided the best results, even than the literature procedure, given the analyzed metrics. An important statement to be done is that for large instances the routing costs matters at the final costs, not allowing it to be disregarded.

For the multi-objective approach, all of the proposed heuristics outperformed the literature one. The MOILS-PRA presented better results in most of the metrics (coverage and hypervolume) followed by MOILS-PRA-PR. The devised new method to help decision makers find a good solution had good performance as expected.

The main limitations of this work were the difficulty to adapt the formulations and find a good procedure that would suits the Brazilian needs. The referred issues might seem easy to deal, but when one have a problem such as the Brazilian, where there are many different scenarios with many different situations, the problem gets bigger and more complicated. So, the procedures devised at this work may not be able to solve the routing problem for all the Brazilian regions.

Thus, besides different scenarios to be investigated, another issues also remain open for future research such as the exact approach to solve smaller instances, exact approach associated with heuristc one. Implementation of transshipment points allowing smaller buses to reach farther students and driving them to the transition station. Also, the decision of school locations with students maximum riding distance as constraint would be a suitable issue for rural areas.

References

Andersson, M. and Lindroth, P. (2005). *Route optimization applied to school transports – A method combining column generation with greedy heuristics*. Department of Mathematics - Chalmers University of Technology and Göteborg University, Göteborg, Sweden.

Arantes, C. O. (1986). Planejamento de rede escolar: questões teóricas e metodológicas. Technical report, Ministério da Educação, Centro de Desenvolvimento e Apoio Técnico à Educação, Brasília.

Assis, L. P., Maravilha, A. L., Vivas, A., Campelo, F., and Ramírez, J. A. (2013). Multiobjective vehicle routing problem with fixed delivery and optional collections. *Optimization Letters*, 7(7):1419–1431.

Baldacci, R., Dell Amico, M., and Gonzalez, J. S. (2007). The capacitated m-ring-star problem. *Operations Research*, 55(6):1147–1162.

Baldacci, R., Hadjiconstantinou, E., and Mingozzi, A. (2004). An Exact Algorithm for the Capacitated Vehicle Routing Problem Based on a Two-Commodity Network Flow Formulation. *Operations Research*, 52(5):723–738.

Batista, L., Campelo, F., Guimarães, F., and Ramírez, J. (2014). The cone epsilon-dominance: an approach for evolutionary multiobjective optimization. Evol. Comput. (submitted).

Beasley, J. E. (1983). Route first - cluster second methods for vehicle routing. *Omega*, 11(4):403–408.

Bektas, T. and Elmastas, S. (2007). Solving school bus routing problems through integer programming. *Journal of Operational Research Society*, 12(58):1599–1604.

Bennett, B. T. and Gazis, D. C. (1972). School bus routing by computer. *Transportation Research*, 6:317–325.

Beume, N., Fonseca, C. M., López-Ibáñez, M., Paquete, L., and Vahrenhold, J. (2009). On the complexity of computing the hypervolume indicator. *Evolutionary Computation, IEEE Transactions on*, 13(5):1075–1082.

Bin, Y., Zhong-Zhen, Y., and Baozhen, Y. (2009). An improved ant colony optimization for vehicle routing problem. *European Journal of Operational Research*, 196:171–176.

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

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

Bodin, L. D., Golden, B., Assad, A., and Ball, M. O. (1983). Routing and scheduling of vehicles and crews: the state of the art. *Computers and Operations Research*, 10:63–211.

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

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

Bradstreet, L. (2011). *The hypervolume indicator for multi-objective optimisation: Calculation and use*. University of Western Australia.

Carvalho, W. L., Cruz, R. O. M., Câmara, M. T., and Aragão, J. J. G. (2010). Rural school transportation in emerging countries: The Brazilian case. *Research in Transportation Economics*, 29:401–409.

Carvalho, W. L. and Yamashita, Y. (2011). Rural school transport in Brazil: an indispensable service for the education policy. 12th International Conference on Competition and Ownershipin Land Passenger Transport.

CEFTRU (2007). *A methodology for characterizing rural school transportation, volume I*. Center for Interdisciplinary Studies in Transportation e Ceftru/UnB.

CEFTRU (2009). *Pilot survey regarding the cost of rural school transportation in Ceará*. Center for Interdisciplinary Studies in Transportation e Ceftru/UnB.

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

Chen, D., Kallsen, H. A., Chen, H., and Tseng, V. (1990). A bus routing system for rural school districts. *Computers and Industrial Engineering*, 19:322–325.

Chen, D. S. and Kallsen, H. A. (1988). School bus routing and scheduling: as expert system approach. *Computers and Industrial Engineering*, 15:179–183.

Chen, D.-S., Kallsen, H. A., and Snider, R. C. (1988). School bus routing and scheduling: An expert system approach. *Computers & Industrial Engineering*, 15(1-4):179 – 183.

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

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

Dantzig, G. B. and Ramser, J. (1959). The truck dispatching problem. *Management Science*, 6:80–91.

Deb, K. (2001). *Multi-objective Optimization Using Evolutionary Algorithms*. Wiley.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197.

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

Desrosiers, J., Ferland, J., Rousseau, J.-M., Lapalme, G., and Chapleau, L. (1986). *TRANSCOL - a Multiperiod School Bus Routing and Scheduling System*, pages 47–71. North-Holland.

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

Dyckhoff, H. (1990). A typology of cutting and packing problems. *European Journal of Operational Research*, 44(2):145–159.

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

Fisher, M. L., Jaikumar, R., and Van Wassenhove, L. N. (1986). A multiplier adjustment method for the generalized assignment problem. *Management Science*, 40:868–890.

Freitas, A. R., Fleming, P. J., and Guimarães, F. G. (2014). Aggregation trees for visualization and dimension reduction in many-objective optimization. *Information Sciences*.

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

Gallo, G. and Pallottino, S. (1988). Shortest path algorithms. *Annals of Operations Research*, 13(1):1–79.

Gavish, B. and Shlifer, E. (1978). An approach for solving a class of transportation scheduling problems. *European Journal of Operational Research*, 3:122–134.

GEIPOT, E. (1995). Assessing rural school transportation. *Brasília: Brazilian Company for Transportation Planning*. (in portuguese).

Gendreau, M., Laporte, G., and Potvin, J.-Y. (2002). *6. Metaheuristics for the Capacitated VRP*, chapter 6, pages 129–154. Toth, P. and Vigo, D.

Gettinger, J., Kiesling, E., Stummer, C., and Vetschera, R. (2013). A comparison of representations for discrete multi-criteria decision problems. *Decision support systems*, 54(2):976–985.

Gillett, B. E. and Miller, L. R. (1974a). A heuristic algorithm for the vehicle-dispatch problem. *Operations research*, 22(2):340–349.

Gillett, B. E. and Miller, L. R. (1974b). A heuristic algorithm for the vehicle-dispatch problem. *Operations research*, 22(2):340–349.

Glover, F., Laguna, M., and Martí, R. (2000). Fundamentals of scatter search and path relinking. *Control and Cybernetics*, 39:653–684.

Hansen, P. and Mladenović, N. (2001). Variable neighborhood search: Principles and applications. *European Journal of Operational Research*, 130(3):449 – 467.

Hansen, P., Mladenović, N., and Pérez, J. A. M. (2010). Variable neighbourhood search: methods and applications. *Annals of Operations Research*, 175(1):367–407.

Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4):800–802.

Hoshino, E. A. and Souza, C. C. (2012). A branch-and-cut-and-price approach for the capacitated m-ring–star problem. *Discrete Applied Mathematics*, 160:2728–2741.

IBGE (2014). Instituto Brasileiro de Geografia e Estatística - Cidades. `http:www.cidades.ibge.gov.br`. (Accessed Oct-01-2014).

INEP (2013). School census - 2013. http://portal.inep.gov.br/basica-censo. In Portuguese.

Jozefowiez, N., Semet, F., and Talbi, E.-G. (2009). An evolutionary algorithm for the vehicle routing problem with route balancing. *European Journal of Operational Research*, 195(3):761–769.

Kara, I. and Bektas, T. (2006). Integer linear programming formulations of multiple salesman problems and its variation. *European Journal of Operational Research*, 174.

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

Langevin, A. and Riopel, D., editors (2005). *New heuristics for the vehicle routing problem*. Langevin, A., Riopel, D. (Eds.), Logistics Systems: Design and Optimization. Springer.

Laporte, G., Nobert, Y., and Desrochers, M. (1985). Optimal routing under capacity and distance restrictions. *Operations Research*, 33(5):1050–1073.

Laporte, G. and Semet, F. (2002). *5. Classical Heuristics for the Capacitated VRP*, chapter 5, pages 109–128. Toth, P. and Vigo, D.

Ledesma, J. R. and Gonzalez, J. J. S. (2012). Solving a school bus routing using the multiple vehicle traveling purchaser problem: A branch-and-cut approach. *Computers and Operations Research*, 39:391–404.

Letchford, A. N., Lysgaard, J., and Eglese, R. W. (2006). A Branch-and-Cut Algorithm for the Capacitated Open Vehicle Routing Problem. *Emerald Management Reviews*.

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

Li, L. and Fu, Z. (2002). The school bus routing problem: a case study. *Journal of the Operational Research Society*, 53:552–558.

Lourenço, H. R., Martin, O. C., and Stützle, T. (2010). Iterated local search: Framework and applications. In Gendreau, M. and Potvin, J.-Y., editors, *Handbook of Metaheuristics*, volume 146 of *International Series in Operations Research & Management Science*, pages 363–397. Springer US.

Mandujano, P., Giesen, R., and Ferrer, J. (2012). Model for Optimization of Locations of Schools and Student Transportation in Rural Areas. *Transportation Research Record: Journal of the Transportation Research Board*, 2283:74–80.

Marler, R. T. and Arora, J. S. (2010). The weighted sum method for multi-objective optimization: new insights. *Structural and multidisciplinary optimization*, 41(6):853–862.

Martí, R., Campos, V., Resende, M. G., and Duarte, A. (2011). Multi-objective grasp with path-relinking. *AT&T Labs Research Technical Report*.

Martinez, L. M. and Viegas, J. M. (2011). Design and Deployment of an Innovative School Bus Service in Lisbon. *Procedia Social and Behavioral Sciences*, pages 120–130.

Min, H., Jayaraman, V., and Srivastava, R. (1998). Combined location-routing problems: A synthesis and future research directions. *European Journal of Operational Research*, 108:1–15.

Mladenović, N. and Hansen, P. (1997). Variable neighborhood search. *Computers & Operations Research*, 24(11):1097 – 1100.

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

Newton, R. M. and Thomas, W. H. (1974). Bus routing in a multi-school system. *Computers and Operations Research*, 1:213–222.

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

Pacheco, J. and Martí, R. (2006). Tabu search for a multi-objective routing problem. *Journal of the Operational Research Society*, 57:29–37.

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

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

Parragh, S. N., Doerner, K. F., and Hartl, R. F. (2008). A survey on pickup and delivery models. *Journal für Betriebswirtschaft*.

Pegoretti, M. S. and Sanches, S. P. (2004). Análise da acessibilidade dos alunos da zona rural que freqüentam escolas urbanas considerando as variáveis distância de caminhada e tempo de viagem. In *Latin American Real Estate Society*. IV Seminário Internacional da LARES Latin American Real Estate Society, IV Seminário Internacional da LARES Latin American Real Estate Society.

Perboli, G., Tadei, R., and Vigo, D. (2011). The two-echelon capacitated vehicle routing problem: models and math-based heuristics. *Transportation Science*, 45(3):364–380.

Ribeiro, C. C., E., U., and Werneck, R. F. (2002). A hybrid grasp with perturbations for the steiner problem in graphs. *INFORMS Journal on Computing*, 14:228–46.

Ripplinger, D. (2005). Rural school vehicle routing problem. *Transportation Research Record*, pages 105–110.

Sales, S. S. (2013). Análise de um processo consolidado: histórico do percurso da nucleação de escolas rurais brasileiras. In *XXVII Simpósio Nacional de História, ANPUH 2013, July 30*.

Salkind, N. J. (2006). *Encyclopedia of measurement and statistics*. Sage Publications.

Sanches, S. and Ferreira, M. (2003). Avaliação do padrão de acessibilidade de um sistema de transporte de alunos da zona rural. In *Anais do XVII Congresso de Pesquisa e Ensino em Transportes*, pages 931–942.

Sanches, S. and Ferreira, M. (2006). Avaliação multicriterial de um sistema de transporte de alunos da zona rural. In *Anais do XX Congresso de Pesquisa e Ensino em Transportes*.

Savas, E. (1978). On equity in providing public services. *Management Science*, 24:800–808.

Savelsbergh, M. and Sol, M. (1995). The general pickup and delivey problem. *Transportation Science*, 29 (1)(17-29).

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

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

Simchi-Levi, D., Chen, X., and Bramel, J. (2005). *A case study: school bus routing*, chapter The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management. In: The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management. Springer, New York.

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

Spasovic, L., Chien, S., and Feeley, C. K. (2001). A Methodology for Evaluating of School Bus Routing - A Case Study of Riverdale, New Jersey. In *Transportation Research Board*. Transportation Research Board - 80th Annual Meeting.

Thangiah, S. R., Bryan, W., Anthony, P., and William, M. (2008). School bus routing in rural school districts. In *Computer-aided Systems in Public Transport*, pages 209–232. Springer.

Thangiah, S. R., Forsythe, J., Vinayagamoorthy, R., Wajda, T., and Bair, P. (2013). Reducing student transportation cost through efficient routing of school buses. In *18th International Conference of Hong Kong Society for Transportation Studies (HKSTS*.

Thangiah, S. R. and Nygard, K. E. (1992). School bus routing using genetic algorithms. In *Aerospace Sensing*, pages 387–398. International Society for Optics and Photonics.

Vasconcellos, E. A. (1997). Transporte rural: o resgate de um tema esquecido. *Revista dos Transportes Públicos*, 75:31–48.

Verderber, W. J. (1974). Automated pupil transportation. *Computers and Operations Research*, 1(2):235–245.

Vincent, S., editor (1999). *The multigrade classroom: A resource handbook for small, rural schools*, volume 1. Northwest Regional Education Laboratory. Book 1: Review of the research on multigrade instruction.

Zitzler, E. and Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. *Evolutionary Computation, IEEE Transactions on*, 3(4):257–271.