Relazione sui risultati dell'attività di ricerca svolta presso l'Istituzione straniera

Luglio 2010

Le attività del progetto sono state svolte dall'Ing. Luigi Moccia in collaborazione con ricercatori del Centre Interuniversitaire de Recherche sur les Reseaux d'Entreprise, la Logistique et le Transport di Montréal. Le attività svolte sono state le seguenti:

1. Rassegna sullo stato dell'arte sul tema del location-routing multi-obiettivo per inquadrare una applicazione al servizio autobus casa-lavoro di un centro di ricerca italiano.

2. Definizione di un modello di ottimizzazione per l'applicazione.

3. Individuazione di un algoritmo metaeuristico per il modello di cui al punto precedente.

4. preliminare validazione dell'algoritmo proposto con esperimenti computazionali.

Si presenta nella Sezione 1 la descrizione dell'applicazione, nella Sezione 2 la rassegna sullo stato dell'arte, mentre il modello di ottimizzazione è proposto nella Sezione 3. Questo lavoro è stato svolto in collaborazione con il Prof. Jean-Francois Cordeau (CIRRELT), il Prof. Gilbert Laporte (CIRRELT), e l'Ing. Alessandro Perugia (mobility manager dell'ENEA).

Firma del Proponente (Prof. Domenico Talia)

Firma del Fruitore (Ing. Luigi Moccia)

1 Introduction

We discuss the design of a bus service plan for the home-to-job transportation of employees of a large research center in Italy. The ENEA Casaccia Research Center is located in the outskirts of Rome, see Figure 1, and employs around 1500 full term contract workers. Currently, the bus service plan consists of 22 routes with 300 stops located in the Rome metropolitan area. Each bus ends its morning route at the center at 8h00, and follows the reverse job-to-home route, leaving the center at 16h00. ENEA does not own the buses and outsources the bus service. The bus stops and the bus routes are the specifications of the bus service plan, and transportation companies make their proposals according to this plan. The number and candidate locations of bus stops are long-term decisions and are considered as inputs in this study. The *Home-to-Job Transportation Problem* (HJTP) focuses on the route design: it determines where to locate bus stops among equivalent locations and determines the bus routes. The two objectives are the minimization of total costs and the maximization of passenger perceived quality of service while taking equity aspects into account. A key feature of the proposed model is the imposition of time windows on the earliest arrival time of a bus at a bus stop. As will be shown, these time windows enforce equity among all users in terms of service quality. Another specific feature is that a bus stop can be equivalently located at several vertices in the graph representing the road network. This captures two real-life issues. First, it correctly models the routing in a dense urban network where turn penalties and prohibitions can apply. Second, this feature enhances flexibility in locating bus stops.

The remainder of this paper is organized as follows. Section 2 discusses the relevant literature on this problem, while Section 3 describes an optimization model for the HJTP.

2 Literature review and contributions

We first review in Section 2.1 the literature pertaining to management and passenger viewpoints in our application. Buses must be routed on an urban street network in which turn penalties and prohibitions must be considered. The literature relevant to this problem is discussed in Section 2.2. We then consider the literature related to an application similar to ours in Section 2.3. Finally, Section 2.4 covers the research related to the proposed solution method. Each section ends with pointers to the contributions of this paper.

2.1 Literature on management and passenger viewpoints

The HJTP consists in designing a bus service plan that must be acceptable for both the management and the workers of the ENEA research center. ENEA is a governmental agency and the bus service plan is negotiated between the management and the unions representing the employees (researchers, technical and administrative staff). Hence, the bus service is not a public service open to the population at large.

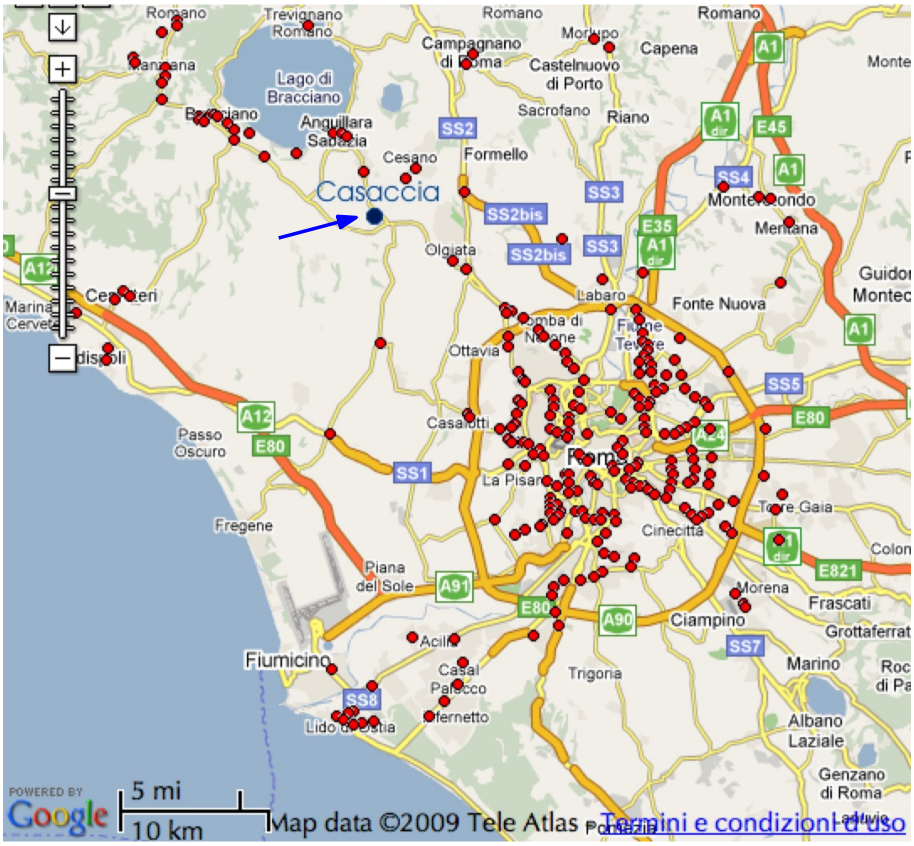

Figure 1: Map of the Rome metropolitan area with bus stops represented as dots and the destination indicated by the arrow. The spatial extension of the problem can be appreciated by the scale reported in the lower left corner.

However, the decision process is similar to those of public services: the bus plan serves many users, all of them having the same rights (as employees), and the plan is subject to collective agreements. In view of this, the framework of Savas (1978) for evaluating performances of public services is particularly relevant. More specifically, Savas (1978) discussed the importance of assessing three criteria that are potentially conflicting: efficiency, effectiveness, and equity. *Efficiency* is easily defined as a ratio between the amount of service provided and its cost. *Effectiveness* measures how well the need for the service is satisfied. *Equity* refers to the fairness or impartiality of the service. Efficiency is usually related to the service provider viewpoint, whereas effectiveness and equity are users' concerns. It is common to consider effectiveness as the only users' criterion, thus dealing with a bi-objective model. For example, Ghoseiri et al. (2004) have introduced a mathematical model for the passenger train-scheduling problem which optimizes efficiency and effectiveness. Tzeng and Shiau (1988) describe a bi-objective optimization model for the bus system of Taipei, Taiwan. The service operator and passengers viewpoints were here modeled by disutility functions to be minimized. The explicit modeling of equity is less common; one of the most notable exception is the work of Mandell (1991) which presented bi-objective optimization models to assess the trade-offs between effectiveness and equity resulting from different allocations of resources to delivery sites. The illustrative application was the allocation of new books among the branches of a public library system.

The field of location theory has extensively dealt with equity issues in public facility location. For reviews see Marsh and Schilling (1994), and Eiselt and Laporte (1995). Siting facilities in a populated area usually generates inequalities when the effects of the facility (positive or negative) vary with the distance from it. This interplay between facility location and effects determined by people's residences is similar to that of our application where the riding time induced by a route (an HJTP decision variable) is perceived as "fair" or "acceptable" depending on the direct home-to-job travel time. The reviews of Marsh and Schilling (1994), and Eiselt and Laporte (1995) assess 20 indexes used in the location theory literature to measure equity. Marsh and Schilling (1994) have introduced a classification scheme for equity indexes and showed that they can be obtained by using different reference distributions, scales, and metrics. Moreover, these authors have examined seven criteria used in the literature to choose among equity indexes. A similar conclusion is derived by both Marsh and Schilling (1994), and Eiselt and Laporte (1995) is that there is little consensus on how to model equity in facility location.

One of the main sources of debate is the use of a min-max objective function as an index for equity. Since the seminal work of Hakimi (1964) this type of function has been one of the most popular way of modeling equity in the literature (usually referred to as p-center models). The philosophical background is rooted in the work of Rawls (1971) who related justice or fairness with the objective of making the least well-off element of a population as well-off as possible. In addition to the critique of Harsanyi (1975) to Rawls's theory, the min-max objective as equity index has been challenged because it does not satisfy the *principle of transfers* (see e.g. Mandell (1991)). This principle is also called Pigou-Dalton condition according to the two economists that introduced it for assessing equity in income distributions. It states that a transfer of wealth from a subgroup to any relative worse-off subgroup should result in improvement in the equity index. A min-max objective does not satisfy this principle because the improvements are registered only when they affect the worse-off subgroup. In the economic literature a widespread equity measure that satisfies the principle of transfers is the Gini index. Following an interpretation suggested by Sen (1973), Mandell (1991) showed that the Gini index can be viewed as measure of average "perceived net envy level" of an income distribution. Marsh and Schilling (1994) moderately questioned the appropriateness of the principle of transfers in location theory. In fact, an underlining assumption of equity measures in income distribution is that there could exist an "ideal" state of equity where there are not wealth differences, i.e. equity as equality. However, in facility location it is not usually possible or desirable to relocate a geographically distributed population to achieve equality in terms of distance from the facility.

The research on *Dial-a-Ride Services* (DARSs) also deals with similar issues. DARSs are part of the more general transportation-on-demand services (Cordeau et al., 2007) where routes and schedules for vehicles are determined to serve the on-demand requests from several pickup points to several destinations. Usually, dial-a-ride services are provided to people with reduced mobility, hence the emphasis on *Quality of Service* (QoS) which incorporates the notions of effectiveness and equity, but is not limited to them. Paquette et al. (2009) provide a general discussion of QoS issues in dial-a-ride services. This study indicates that the temporal aspects are the predominant ones when customers assess this type of transportation service. In the DARS literature the minimization of the excess riding time over the direct time for all the passengers is a rather common objective, see e.g. Diana and Dessouky (2004). In Section 3.2 we will follow these studies in our definition of effectiveness for the HJTP. However, the reviewed literature is not conclusive regarding the modeling of equity. Moreover, we observe that there is a further dimension to be taken into account, namely decision making is often *incremental* in the sense that planning is usually not made from scratch, but with respect to an existing situation. Fairness should then also be defined in terms of the requested change with respect to an existing service plan and considering the independent residence decisions made by the users. The proposed modeling of equity presented in Section 3.2 attempts to tackle these issues in our application.

2.2 Literature on turn penalties and prohibitions

A road network can be modeled as a directed graph where a vertex represents an intersection or a demand point, an arc denotes a road segment, and the arc orientation is the allowed direction of movement. In urban street networks, there are also delays at intersections (usually when turning left in right driving roads) and turn prohibitions (e.g. U-turns). These networks can still be modeled as graphs, but turn restrictions induce penalties when using two consecutive arcs corresponding to a turn delay, or induce constraints preventing the use of two consecutive arcs when there is a turn prohibition. Routing models require shortest path computations between demand points. The literature on the *Shortest Path Problem with Turn Penalties and Prohibitions* (SPP-TPP), see e.g. Pallottino and Scutellà (1998), includes two main classes of methods. A first class transforms the original graph into an auxiliary graph in which the arcs of the original graph become vertices and the arcs of the auxiliary graph correspond to the allowed turns, eventually penalized. Thus standard shortest path algorithms can be used on the auxiliary graph. This approach was first suggested by Caldwell (1961). Añez et al. (1996) showed that a "link-based" representation of transportation networks (the auxiliary graph) can model not only turn charactheristics but also multiple operators and modes. A second line of research develops shortest path algorithms that explicitly consider turns. Kirby and Potts (1969) introduced a modified Bellman equation based on arcs instead than vertices. Easa (1985) presented an algorithm for the *Shortest Path Problem with Turn Prohibitions* (SPP-TP), i.e. turn penalties are not considered. Ziliaskopoulos and Mahmassani (1996) described a label correcting algorithm for the SPP-TPP. These authors stated that the label correcting algorithm proved to be more efficient than a label setting version in their computational experiments. More recently, Gutiérrez and Medaglia (2008) have introduced a label setting algorithm for the SPPTP. However these authors do not mention the more general algorithm of Ziliaskopoulos and Mahmassani (1996). In fact, shortest path algorithms with turn penalties and prohibitions are not always sufficiently acknowledged. For example, recently Fidler and Einhoff (2004) introduced in the computer science literature a graph transformation procedure (the turnnet concept) similar to the auxiliary graph of Caldwell (1961) without any mention to this paper. To give proper merit to the work of Fidler and Einhoff (2004) we mention that it generalizes to several objective function structures (additive, multiplicative and concave metrics) than previous research.

The arc routing literature has extensively discussed these turn characteristics, see Clossey et al. (2001) and references therein. Laporte (1997) proposed a graph transformation that allows solving arc routing problems with turn penalties and prohibitions as *Traveling Salesman Problems* (TSPs). Clossey et al. (2001) have introduced heuristic algorithms to directly handle turns in arc routing problems. These heuristics proved to be more effective than exact or heuristic algorithms applied to the equivalent TSPs.

The effect of turn prohibitions in a vertex routing problem was highlighted by Laporte et al. (1989) who studied the design of mailbox collection routes in urban areas. The mailboxes are usually located at street corners, and can typical be reached from several possible directions. Because of U-turn prohibitions, the distance between a pair of mailboxes depends on the direction they are reached. This problem can be modeled as a *Generalized Traveling Salesman Problem* (GTSP); see also Laporte et al. (1996). Our application deals with a case similar to that discussed in Laporte et al. (1989). We have to locate bus stops in a road network where the direction of movement has to be explicitly modeled. We observe that the issue of location-routing in networks with turn penalties and prohibitions has not received a wide attention in the literature. In view of this we devote Section 3.4 to it.

2.3 Literature on related applications

Our problem belongs to the area of location-routing (Laporte, 1988) and is similar to the *School Bus Routing Problem* (SBRP) where students must be transported from several pickup points to a unique location within a prescribed time. For a comprehensive review of the SBRP literature the interested reader is refered to the recent survey of Park and Kim (2010). In the following we present a short review highlighting mainly three aspects that are common with our application: the modeling of quality of service issues (effectiveness and equity), the modeling of bus stop location (if any), and the solution method (which has to be related to the size of the considered instances).

Bowerman et al. (1995) discuss several criteria for the SBRP following the classification of Savas (1978). They consider as guiding principle for evaluating effectiveness whether the level of service is acceptable to the public. The student eligibility for bus transport (which in the authors' case study depended on the distance from student's home to the school) was then used as an individualized effectiveness criterion. However, this aspect is exogenous to the optimization model which explicitly minimizes a cost measure (the number of bus routes) and several other terms for equity purposes. These are the "load-balance" and the "length-balance" of routes where the deviations with respect to the average route load and length are minimized. Furthermore, the authors have introduced as equity criterion the minimization of total student walking distance. However, in our opinion, this objective is related to the effectiveness of the service. The optimization model also considers equity by the following modeling of the routes. The designed routes allow a "first-on-first-off" policy, i.e. the students who are picked up earlier in the morning must also be those who are dropped off first in the afternoon. This policy is clearly limited to an urban setting in which routes are short, and it must not excessively penalize students who live close to the school in terms of extra-riding time. The solution method was based on several constructive heuristics applied to the two following phases. In a first phase a districting algorithm creates clusters of students. In this phase the two objectives that were expressively minimized are the number of routes and load-balancing. This phase also uses a cluster compactness measure as in Chapleau et al. (1985). In the second phase the bus stop locations and the bus routes are determined for each cluster. Computational results were provided for a case study at one school located in Wellington County, Ontario, and involving 183 students.

Braca et al. (1997) comprehensively described the computerized system used to solve the New York City *multi-School Bus Routing Problem* (m-SBRP). This is a more difficult problem occuring when a bus fleet serves more than one school, and schools have different starting and ending times. These differences can be exploited to reduce the bus fleet size in dense urban area where routes are usually short. The authors recognized the importance of optimal bus stop location, but no attempt was made to address this issue and bus stops were taken as inputs according to the past experience. The model minimizes the number of required buses and it includes QoS constraints such as maximal distance constraint (each student cannot be on the bus for more than 5 miles), school arrival time constraints (buses must arrive at a school between 5 and 25 minutes before school start), and pickup time constraint (the earliest pick-up must not be before 7h00). The solution method was based on a randomized heuristic derived from a heuristic for the *Capacitated Vehicle Routing Problem* (CVRP) introduced in Bramel and Simchi-Levi (1995).

Li and Fu (2002) have presented a constructive heuristic algorithm for a SBRP in Hong Kong. Their heuristic considers four objectives: the total number of buses required, the total travel time spent by the students, the total bus travel time, and the largest difference in loads and travel times between buses.

Corberán et al. (2002) have solved a SBRP in a sparse rural area where routes are usually longer than in urban areas. In view of this, the authors chose as QoS index the maximal time a student spends in the bus. The model minimizes this QoS index and the number of deployed buses. This bi-criteria optimization model was solved by means of a scatter search metaheuristic. The proposed methodology aimed at drawing the efficient frontier of the two objectives, thus highlighting trade-offs between different levels of service and costs. The authors showed that their algorithm could find better routes than the current bus plan in terms of QoS without increasing the number of buses. Pacheco and Martí (2006) have solved the same problem by using a tabu search algorithm with an intensification phase based on the path relinking methodology, which yielded better solutions than in Corberán et al. (2002).

Spada et al. (2005) describe a decision tool for a m-SBRP. This paper distinguishes itself from the SBRP literature by focusing exclusively on QoS issues. Students can face additional time losses than the extra-riding time. In fact, because a bus may perform several tours, some students can be dropped at their school earlier than the school starting time. This waiting time at the school increases the student time losses. The authors minimize two objective functions: the total time loss of all students, and the maximum time loss among all students. The problem was modeled as a non-linear integer program and solved heuristically. A solution was built by a constructive heuristic and improved by a metaheuristic. The authors have compared two metaheuristics: simulated annealing, and tabu search. Simulated annealing performed slightly better in their computational experiments.

Bektaş and Elmastas (2007) have studied a SBRP for an elementary school in Ankara. In this application there were 29 bus stops and up to 26 available vehicles. The aim was to minimize the total cost of the transportation service expressed as a sum of variable costs induced by traveled distances and fixed costs for activated vehicles. The authors modeled the problem as an *Open VRP* (see e.g. Aksen et al. (2007) and Li et al. (2007)) with capacity and route duration constraints. The route duration constraints ensured that a QoS index (the maximum time a student travels) was satisfied. The small size of the problem made it possible to solve the model to optimality by a commercial integer linear programming software. The optimal solution resulted in a 29% savings in total cost as compared to the previous routing scheme.

The reviewed literature considered equity issues by using min-max approaches either as objectives or as constraints. Moreover, the issue of bus stop location is often exogenous to the model and there is no mention to routing in road networks with turn restrictions. In Sections 3 we provide an integrated model for bus stop location and routing where equity issues are handled in a finer way than in classical min-max approaches and where turn characteristics are dealt with. In their review of the SBRP, Park and Kim (2010) observed that many of the solution methods of this literature stream are problem dependent and there is a need for more general methods based on some of the recent metaheuristics for vehicle routing. The usual motivation for problem dependent methods in SBRP is that there are structural differences between classical cost minimization routing models and passenger transportation. In Sections 3.2 and 3.3 we show that time windows can be used as a modeling tool to bridge this gap. Time windows are already used in freight transportation to handle customer satisfaction among other issues, see e.g. Moccia et al. (2010b) and references therein.

2.4 Literature on solution methods

The proposed optimization model for the HJTP is a bi-objective *Generalized Vehicle Routing Problem with Time Windows* (GVRPTW). An exact algorithm for the single-objective GVRP was introduced by Bektaş et al. (2009). To the best of our knowledge, there are no available exact algorithms for the GVRPTW. Moreover, in our application we deal with a large scale case study. Therefore, a heuristic algorithm is the favored solution method. Ibaraki (2010) discusses the importance of developing heuristic algorithms that can be adapted to several classes of problems. In our opinion, the tabu search heuristic of Moccia et al. (2010a) satisfies this criterion. It extends to the cluster routing case the *Unified Tabu Search* (UTS) heuristic of Cordeau et al. (2001, 2004) and of Cordeau and Laporte (2001). The UTS heuristic was successfully applied to the *Vehicle Routing Problem with Time Windows* (VRPTW) and several of its extensions such as the multi-depot VRPTW, the periodic VRPTW, and the site-dependent VRPTW. Moreover, UTS was adapted to an application very different from the classical VRPTW, the berth allocation in container terminal, both at the operational and at the tactical level, see Cordeau et al. (2005) and Giallombardo et al. (2010), respectively.

3 Optimization model

We first discuss how we model the efficiency and effectiveness criteria of Section 3.1 and the equity criterion of Section 3.2. The methodology introduced to model equity has multiple uses that we detail in Section 3.3. In fact, one of the guiding principles of the proposed model consists in favoring concepts that simultaneously tackle several issues at once. This is also the case of the modeling of turn characteristics, described in Section 3.4, which is also useful to the modeling of bus stop location (see Section 3.5). Finally, we present the full optimization model in Section 3.6.

3.1 Modeling efficiency and effectiveness

In our application the number of bus stops to be served and their potential locations are given, and thus the efficiency criterion can be modeled by minimizing the cost of the plan. We measure the effectiveness of a service plan by the total *extra-time* for all passengers, where the extra-time is the difference between the bus riding time and the time of a direct trip. The direct trip is the shortest path between the bus stop location and the destination. The extra-time is a common QoS measure used in transportation, as highlighted in our literature review (Sections 2.1 and 2.3). Observe that passengers would always prefer to be served from the last bus stop of a route because this choice results in no extra-time. However, this would mean deploying a number of buses equal to the number of bus stops, which would be equivalent to a taxi system. Such an option is prohibitive in our application setting. The efficiency objective favors the downsizing of the number of routes (i.e. using a smaller fleet) which results in longer routes and thus higher extra-times. The trade-off between efficiency and effectiveness can be assessed by analyzing the Pareto frontier of these two objectives. We observe that some improvements in effectiveness can be achieved by different routing decisions without any cost increase. For example, suppose that a bus route serves the bus stops labeled as A and B in Figure 2 and then reaches the destination C. The two possible routes are $R1$ and $R2$, where in $R1$ stop A is served before stop B, and vice versa for route $R2$. Assuming that the arc travel times and costs are all equal, both routes will have the same cost and duration. However, if the number of passengers served at stop A is larger than that at stop B, then route $R2$ will result in less total extra-time. Therefore, the explicit modeling of effectiveness is significant even if the efficiency objective is predominant in the decision making.

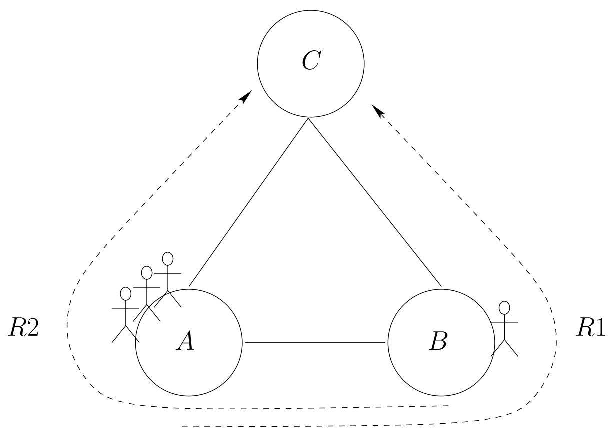

Figure 2: Example of different routing decisions improving effectiveness without worsening efficiency

3.2 Modeling equity

We now discuss the proposed modeling of equity issues. The main source of passenger criticism in the type of transportation service considered here is related to how the extra-time is distributed. The effectiveness objective introduced in the previous section strives for the minimization of the total extra-time. However, some passengers may be heavily penalized at the expense of others. In our application context, which is bus transportation in the large geographical area of Figure 1, the routes must be radial, and all converge to a common destination. This structural characteristic means that higher extra-times are to be expected at the locations that are more distant from the destination. Because of this, the fairness in extra-time distribution must be evaluated with respect to the minimal home-to-job distance.

We report in Figure 3 the distribution of the extra-times (y-axis) according to the direct travel times (x-axis) in the current service plan. We focus on the worst extra-time for a given value on the x-axis. We provide a lower envelope of these worst case extra-times, and in the following we call this lower envelope the *baseline capping function*. It can be observed that there is a parabolic increase of this function up to a certain critical point, and then a slight decrease. In Figure 3, the critical point occurs at the coordinates $(23, 32)$. Since the travel time is the sum of direct time and extra-time, the travel time cor-

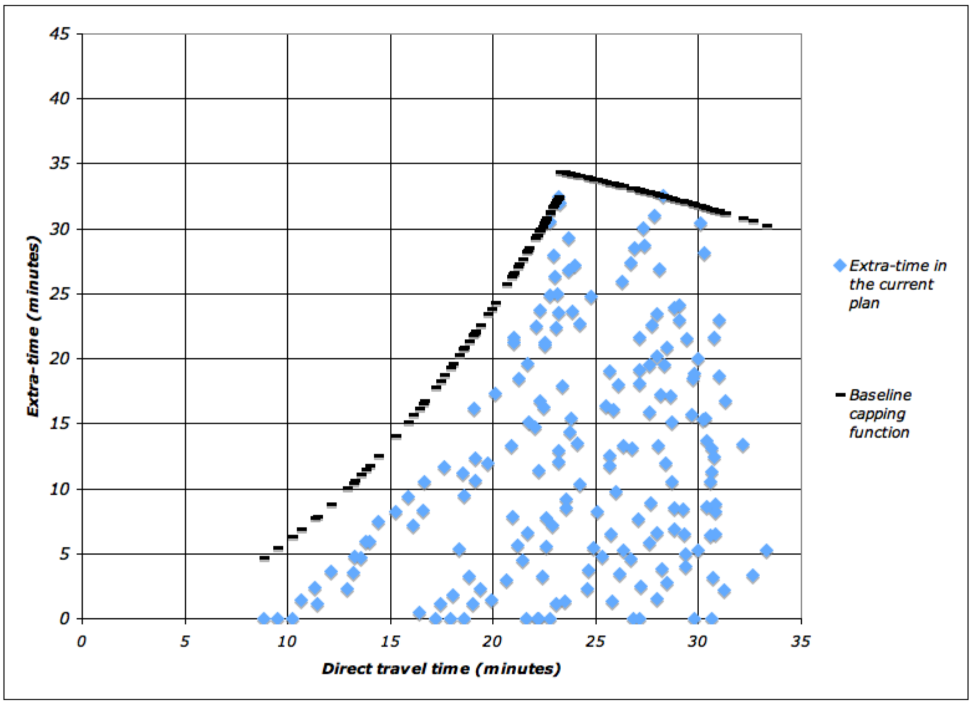

Figure 3: The extra-times in the current service plan

responding to this critical point is 55 minutes. The analysis of these data can be synthesized as follows. The worst-off passengers in the current plan have some sort of distributional equity among them. In fact, the worst-off passenger at a certain direct distance from the destination is better-off than the worst-off passengers residing further, and this holds up to a critical value of the direct distance (23 min.). Passengers more distant that this value incur an extra-time such that their total travel time does not exceed one hour. Our main idea is to use in our model this baseline capping function. We can enforce an upper bound on the extra-time as a function of the direct distance to destination. Since the arrival time at the destination is fixed, this results in a time window on the earliest arrival time at a bus stop (a more precise definition accounting for the bus stop location issues will be given in Section 3.6). The impact of these time windows constraints can be parametrically analyzed by varying the capping function by means of appropriate coefficients. Figure 4 illustrates the baseline capping function and four functions labelled as $0.8, 0.9, 1.1$, and 1.2, where 0.8 indicate a capping function scaled by 0.8, etc.

We now discuss properties of this equity model. Let assume first that we use the baseline capping function to define the time windows at the bus stops. By so doing, we can guarantee that every user in a new plan will not be worse-off than a peer already is in the current plan, where a peer is another user served from a bus stop located at the same direct distance from the destination. This is an easy to explain criterion and therefore increases the appropriateness of the measure in the sense of Marsh and Schilling (1994).

We can assess the impact of these constraints on the other objectives by using as a capping function a better one from the user viewpoint (e.g. the 0.8 and 0.9 in Figure 4) or a worse one. We observe that

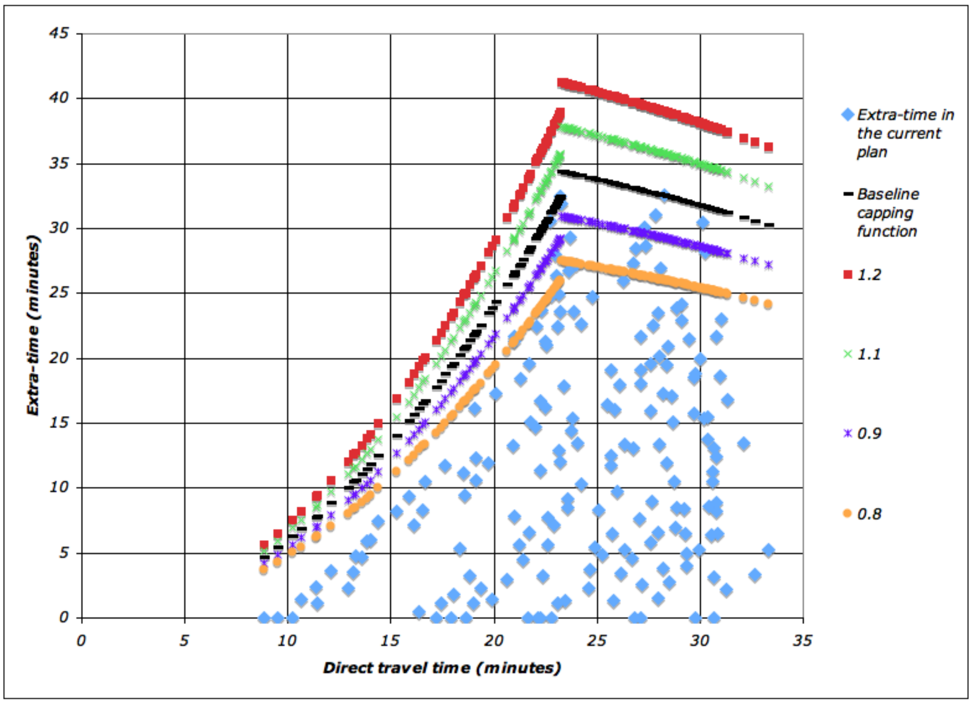

Figure 4: Examples of scaling the baseline capping function

parametrically changing these constraints is equivalent to assessing an objective in multi-objective analysis. We now show that the implicit objective of the time window constraints is the minimization of the scalar coefficient of the capping function. In the following of this section we use the framework and notation of Marsh and Schilling (1994) to classify the equity objective. We define how users are grouped (groups are indexed by i), the effect E_i of the model decisions on group i, the reference value D_i of the effect on group i, and a group-specific attribute A_i for scaling purpose. We group passengers according to the their loading bus stop. Thus a group of passengers i equivalently denotes their bus stop. The measured effect E_i is the extra-time for passengers served at bus stop i, and the reference value is zero extra-time for all users, i.e. $D_i = 0$, for all i. The group-specific attribute A_i is the value of the upper bound on the extra-time at the bus stop i given by the baseline capping function. The time windows impose constraints $E_i \leq \gamma A_i$, for all i, where γ is the scaling factor of the selected capping function. Thus, for each selected capping function we have $\gamma \geq E_i/A_i$. Therefore, parametrically changing γ implicitly assesses the objective $\max_i E_i/A_i$. This is equivalent to $[\sum_i |E_i/A_i - D_i|^{\infty}]^{1/\infty}$, thus highlighting the relation with the L_∞ metric. This type of objective can be classified according to the scheme of Marsh and Schilling (1994), but it is new. Because it uses the L_∞ metric it belongs to the class of min-max approaches. As any equity measure based on the L_∞ metric it does not satisfies the principle of transfers. The index registers an improvement if and only if a better-off group transfers a reduction in extra-times to *all* worse-off groups lying on the capping function (i.e. all the groups i such that $E_i = \gamma A_i$). It should be noted that this objective is a generalization of the classical min-max one where the baseline capping function would be a constant. By using a more complex baseline capping function we consider *rela-*

tive effects as should be the case in problems of this type where equity as equality is not appropriate. Moreover, the baseline capping function extracted from an existing scenario has the advantage of considering the performances that users were used to. It provides us with a frontier of what has historically been judged acceptable by the users. Drawing a similar function according to utility theory would have been another modeling option. However, the existence of historical data helps in communicating the objectives of the plan, whereas a newly computed function could be questioned by some groups.

3.3 Modeling additional features by time windows

We have discussed the usefulness of time windows to the modeling of equity. We now consider other of their uses. First, observe that time windows on the earliest arrival time at the more distant bus stops implicitly define an upper bound on the duration of the routes since the arrival time at destination if fixed. The route duration constraint is usually required in practice. Second, the upper bound on the extra-time on some bus stops can be set to specific values for special purposes. For example, suppose that a bus stop is located at a parking place where passengers arrive by car and wait for the bus for the last leg of their home-to-job trip, as in a park-and-ride setting. These passengers have already experienced some discomfort on the first part of their trip, and it would therefore make sense to give to this bus stop a higher priority (less worst case extra-time) than that resulting from the capping function. This is straightforward to achieve within the proposed modeling framework. A specific value for the arrival time at a bus stop could also be motivated by the need to synchronize the timetable with that of other mass transportation modes at certain locations (e.g. train stations). These are examples of the finer control on the characteristics of the plan provided by the time window concept.

3.4 Modeling turn characteristics

As mentioned in Section 2.2, our problem requires the computation of shortest paths between vertices representing demand points in a graph with turn penalties and prohibitions. We now discuss how this has an impact on the model. Kirby and Potts (1969) observed that a shortest path in such a network has an origin and destination pair of *arcs*, instead of vertices. Since arcs represent directed street segments a shortest path between two demand vertices is then associated to the directions both at the origin and at the destination. Whenever we compute a shortest *path* between a pair of vertices in a network with turn restrictions we implicitly assume that there are two artificial arcs with zero costs and no turn restrictions entering the origin vertex and leaving the destination vertex, respectively. The computed shortest path between these two artificial arcs is then the "ideal" path which considers turn restrictions *after* leaving the origin vertex. These ideal paths are not in general useful to determine a shortest *route* connecting several demand vertices. The sum of the lengths of the ideal paths between demand vertices is a lower bound of the route real length. We illustrate this by the example depicted in Figure 5. We indicate by capital letters (A, B, and C) three demand vertices, and by lower case letters (d, e, f, and g) four intersection vertices. Every arc has a unit length and U-turn prohibitions apply to every vertex, i.e. we cannot use a sequence of two arcs if the head and tail of the first arc correspond to the tail and head of the second arc (e.g. arc (e, B) cannot precede arc (B, e)). The three demand vertices can be visited by a route in two ways:

1. A precedes B, which then precedes C; for this route we have to consider the distances AB, BC, and CA (where we denote by AB the distance between the vertices A and B, etc.); this route is indicated as ABC;

2. A precedes C, which then precedes B; for this route we have to consider the distances AC, CB, and BA; this route is indicated as ACB.

The ideal shortest path lengths between the three demand vertices are reported in the upper-left corner of Figure 5. Computing the total length of these two routes by the ideal lengths results in a value of 7 for both of them. However, the ABC route has a length of 8 when taking the U-turn prohibitions into account. This is because the direction of movement at B resulting from the ideal path between A and B (which is (A, e, B)) is not compatible with that of the ideal path between B and C (which is (B, e, C)). In fact, the path (A, e, B, e, C) would require a U-turn at B. Instead, route ACB is the shortest with a total length of 7. To correctly model this we have to explicitly associate a direction of movement to a demand

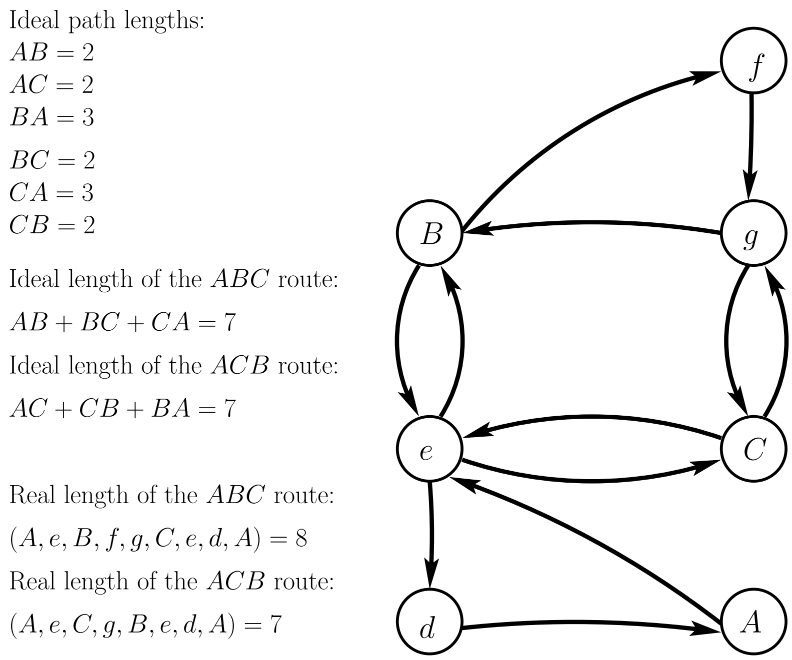

Figure 5: An example of routing with turn constraints. Demand vertices are labeled with capital letters, whereas intersection vertices have lowercase letters. Every arc has an unitary length and U-turn prohibition applies to every vertex. The ideal shortest path lengths between demand vertices are reported in the upper-left corner, followed by ideal route lengths, and real route lengths

vertex as in Laporte et al. (1989). We then need to represent a demand vertex as a *cluster* of vertices and requiring that exactly one vertex per cluster be visited by a route. The modeling framework becomes that of the *Generalized Traveling Salesman Problem* (GTSP) for one route, and of *Generalized Vehicle Routing Problem* (GVRP) for multiple routes. For other applications of the GVRP, see Baldacci et al. (2010).

In the above mentioned example there are two possible directions of movement at vertices B and C, whereas vertex A has only one. Figure 6 illustrates the routing example on a new graph which models the equivalent GTSP. A vertex of Figure 6 represents a direction of movement at a demand vertex of Figure 5, and is labeled by the corresponding arc of the graph of Figure 5. An arc of Figure 6 indicates the shortest path between the origin and destination arcs in the original graph. For example, the arc between (d, A) and (e, B) is the shortest path originating at A with direction of movement from d and ending at B with direction of movement from e. This path has a length of 2. We report in Figure 6 the subset of arc lengths useful for our example. As required, the shortest route for the sequence ACB has a length of 7, whereas the shortest route for the sequence ABC has a length of 8.

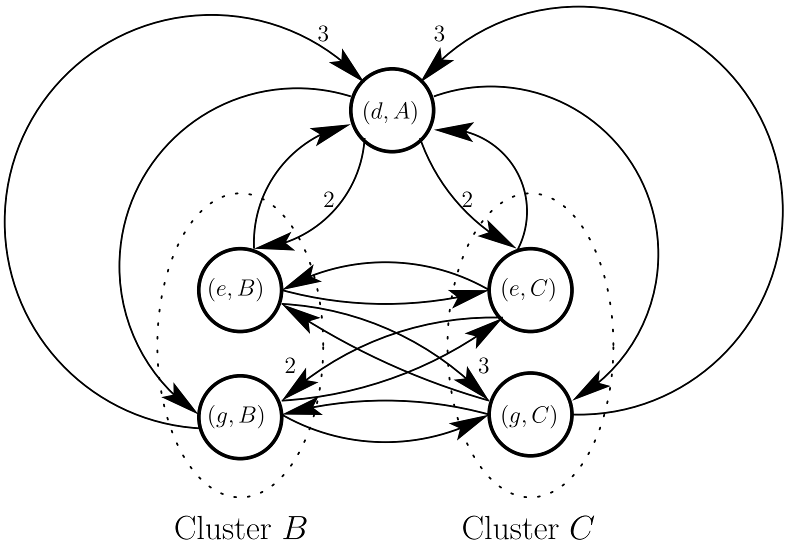

Figure 6: Graph for the cluster routing problem equivalent to the example of Figure 5

We observe that this problem could be modeled without resorting to clusters if the demand points were specifically associated to street sides (and hence to unique directions of movement). This condition can pertain to several applications such as the transportations of pupils who are not supposed to cross streets. However, we remark that in this case the turn characteristics could still require a cluster routing model. In fact, if the demand points are located at street corners, then there could be more than one direction of movement arriving at these street corners. Another condition that would make the above issue irrelevant would be a road network without U-turn prohibitions or in which U-turns have a zero or negligible cost. However, this condition is rare in urban road networks.

3.5 Modeling bus stop locations

In the previous section we have shown that the flexibility of equivalently locate a bus stop at one of the two sides of a two-ways street, or at one of the corners of an intersection, requires a *cluster routing* model. Cluster routing, as opposed to vertex routing, adds flexibility to the model because it becomes possible to consider several physical locations where a set of passengers can be equivalently collected. Thus, several aspects of the bus stop location problem can be integrated as routing decisions. However, we observe that a necessary condition for the appropriateness of the GVRP framework for bus stop location is that vertices in a cluster must be equivalent. This is the case in our application, but this is not necessary true in applications where the bus stop location must comply with some covering constraints of passenger demand.

3.6 Mathematical formulation

We define the HJTP on a directed graph $G = (V, A)$ as follows. We indicate by I the set of n vertices representing the candidate bus stop locations, and by B the set of bus stops that must be located. For each bus stop $b \in B$ there exists a set $I(b) \subset I$ of equivalent candidate bus stop locations. A set $I(b)$ is also called a cluster and the set I is the union of $|B|$ disjoint clusters. The HJTP requires that exactly one vertex belonging to a cluster be visited by a bus route. The vertices o, and d represent the bus depot vertex, and the workplace vertex, respectively. The set of digraph vertices is then $V = I \cup \{o, d\}$, and the set of arcs A is the union of three disjoint subsets: $A(I) = \{(i, j) \in A \mid i \in I, j \in I\}$, $A(o) = \{(o, j) \in A \mid j \in I \cup \{d\}\}$, and $A(d) = \{(i, d) \in A \mid i \in I \cup \{o\}\}$. For each vertex $i \in I$ the set $\delta(i)^+$ represents the vertices $j \in V$ such that $(i, j) \in A$. Similarly, $\delta(i)^-$ represents the vertices $j \in V$ such that $(j, i) \in A$. We are given the travel times along the arcs, expressed as $t_{ij}, (i, j) \in A$.

At each candidate bus stop $i \in I$ is associated a value p_i, number of expected passenger, and s_i is the time required to perform the stop and load the p_i passengers. These values are equal for all vertices of a cluster, $p_i = p_j$ if $i, j \in I(b)$. The buses arrive at the workplace at a given time $\bar{T}_d$. The arrival time at a bus stop b cannot be earlier than a given value a_b, defined as follows. Let $\bar{t}_b$ be equal to $\bar{T}_d - \min_{i \in I(b)}\{t_{id} + s_i\}$; then $\bar{t}_b$ represents the "ideal" bus arrival time for a passenger boarding at a bus stop $i \in I(b)$. In fact, because a bus ends its route at the workplace at the time $\bar{T}_d$, if it arrives at the bus stop $i \in I(b)$ at time $\bar{t}_b$, it achieves the highest quality of service (zero extra-time) for the passengers using this stop. The equity criterion introduced in Section 3.2 defines an upper bound on the extra-time according to a given capping function. Let $W(b)$ the value of the capping function for a bus stop b. We set $a_b = \bar{t}_b - W(b)$, which is the earliest bus arrival time such that the extra-time is considered as fair for the passengers. The other uses of the time window concept discussed in Section 3.3 would determine specific values for a_b.

We consider a fleet of m buses and this set of vehicles is indicated as K. The fleet can be heterogeneous, and therefore bus specific parameters must be introduced. Bus $k \in K$ can serve at most q^k passengers, has a fixed cost f^k, and variable cost $c^k_{ij}, (i,j) \in A$ when traversing the arc (i,j). With this notation the HTJP can be modeled as a generalized vehicle routing problem with time windows. The main difference is the existence of a second objective taking into account the quality of service offered to the passengers by means of extra-times. Let k be the bus serving the vertex $i \in I(b)$, and let T^k_i be the arrival time of the bus k at this vertex i. The extra-time at the vertex i is measured by the value w^k_i, computed as $\bar{t}_b - T^k_i$. Observe that w^k_i is always non-negative, and that $0 \leq w^k_i \leq \bar{t}_b - a_b = W(b)$, where $w^k_i = 0$ corresponds to the "ideal" quality of service, whereas $w^k_i = W(b)$ represents the largest allowed extra-time by the selected capping function. We then define the effectiveness objective function to be minimized as the sum of the extra-times weighted by the number of passengers p_i. As discussed in Section 3.1, the cost and extra-time are two conflicting objectives.

We now introduce the decision variables:

- $x^k_{ij} \in \{0, 1\}, \forall (i,j) \in A, k \in K$, where $x^k_{ij} = 1$ if bus k uses arc (i,j), and $x^k_{ij} = 0$ otherwise;
- $y^k \in \{0, 1\}, \forall k \in K$, where $y^k = 1$ if the bus k is activated, and $y^k = 0$ otherwise;
- $T^k_i \in \Re^+, \forall k \in K, i \in V$, indicates the arrival time of bus k at vertex i;
- $w^k_i \in \Re^+, \forall k \in K, i \in I$, is equal to $\bar{t}_b - T^k_i$ if i belongs to the route k, and it is equal to zero otherwise.

The mixed-integer linear programming formulation is as follows:

$$\text{minimize} \sum_{k \in K} (f^k y^k + \sum_{(i,j) \in A} c^k_{ij} x^k_{ij}) \tag{1}$$

$$\text{minimize} \sum_{k \in K} \sum_{i \in I} w^k_i p_i \tag{2}$$

subject to

$$\sum_{j \in \delta^+(o)} x_{oj}^k = 1 \qquad \forall k \in K, \tag{3}$$

$$\sum_{i \in \delta^-(d)} x_{id}^k = 1 \qquad \forall k \in K, \tag{4}$$

$$\sum_{j \in \delta^+(i)} x_{ij}^k - \sum_{j \in \delta^-(i)} x_{ji}^k = 0 \qquad \forall k \in K, i \in I, \tag{5}$$

$$\sum_{k \in K} \sum_{i \in I(b)} \sum_{j \in \delta^+(i)} x_{ij}^k = 1 \qquad \forall b \in B, \tag{6}$$

$$T_i^k + t_{ij} + s_i - T_j^k \le (1 - x_{ij}^k) M_i \qquad \forall k \in K, (i,j) \in A \tag{7}$$

$$a_b \le T_i^k \qquad \forall k \in K, \ i \in I, \tag{8}$$

$$T_d^k = \bar{T}_d \qquad \forall k \in K, \tag{9}$$

$$\bar{t}_b - T_i^k \le w_i^k \qquad \forall k \in K, \ i \in I, \tag{10}$$

$$\sum_{i \in I} \sum_{j \in \delta^+(i)} p_i x_{ij}^k \le q^k y^k \qquad \forall k \in K, \tag{11}$$

$$T_i^k \ge 0 \qquad \forall k \in K, i \in V, \tag{12}$$

$$w_i^k \ge 0 \qquad \forall k \in K, i \in V, \tag{13}$$

$$x_{ij}^k \in \{0, 1\} \qquad \forall (i,j) \in A, k \in K. \tag{14}$$

Here M_i takes the value $\bar{T}_d - \min_{b \in B}\{a_b\}$. The objective function (1) minimizes the total cost of the bus service and the objective function (2) minimizes the total extra-time. Constraints (3) and (4) define the degree of the origin and destination vertices, respectively. Flow conservation for the bus stop vertices is ensured by constraints (5). Constraints (6), together with (5), mean that each bus stop b is located at a vertex belonging to $I(b)$, and is served by exactly one bus route. The propagation of time variables T_i^k is enforced by constraints (7), while earliest arrival times at the bus stops, and fixed arrival time at the destination are enforced by constraints (8) and (9), respectively. Constraints (10) and non-negative coefficients in the objective function to be minimized define variables w_i^k. Finally, capacity constraints over the number of passenger per route is ensured by (11) which also defines the bus activation variables y^k.

References

Añez, J., De La Barra, T., and Pérez, B. (1996). Dual graph representation of transport networks. *Transportation Research Part B: Methodological*, 30(3):209–216.

Aksen, D., Ozyurt, Z., and Aras, N. (2007). Open vehicle routing problem with driver nodes and time deadlines. *Journal of the Operational Research Society*, 58(9):1223–1234.

Baldacci, R., Bartolini, E., and Laporte, G. (2010). Some applications of the generalized vehicle routing problem. *Journal of the Operational Research Society*, 61(7):1072 – 1077.

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

Bektaş, T., Erdoğan, G., and Ropke, S. (2009). Formulations and branch-and-cut algorithms for the generalized vehicle routing problem. Technical report, Technical University of Denmark.

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

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

Bramel, J. and Simchi-Levi, D. (1995). A location based heuristic for general routing problems. *Operations Research*, 43(4):649–660.

Caldwell, T. (1961). On finding minimum routes in a network with turn penalties. *Communications of the ACM*, 4(2):107–108.

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.

Clossey, J., Laporte, G., and Soriano, P. (2001). Solving arc routing problems with turn penalties. *Journal of the Operational Research Society*, 52(4):433–439.

Corberán, A., Fernández, E., Laguna, M., and Martí, R. (2002). Heuristic solutions to the problem of routing school buses with multiple objectives. *Journal of the Operational Research Society*, 53(4):427–435.

Cordeau, J.-F. and Laporte, G. (2001). A tabu search algorithm for the site dependent vehicle routing problem with time windows. *INFOR*, 39(3):292–298.

Cordeau, J.-F., Laporte, G., Legato, P., and Moccia, L. (2005). Models and tabu search heuristics for the berth-allocation problem. *Transportation Science*, 39(4):526–538.

Cordeau, J.-F., Laporte, G., and Mercier, A. (2001). A unified tabu search heuristic for vehicle routing problems with time windows. *Journal of the Operational Research Society*, 52(8):928–936.

Cordeau, J.-F., Laporte, G., and Mercier, A. (2004). Improved tabu search algorithm for the handling of route duration constraints in vehicle routing problems with time windows. *Journal of the Operational Research Society*, 55(5):542 – 546.

Cordeau, J.-F., Laporte, G., Potvin, J.-Y., and Savelsbergh, M. W. P. (2007). Transportation on demand. In Barnhart, C. and Laporte, G., editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, chapter 7, pages 429 – 466. Elsevier, Amsterdam.

Diana, M. and Dessouky, M. M. (2004). A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. *Transportation Research Part B: Methodological*, 38(6):539 – 557.

Easa, S. M. (1985). Shortest route algorithm with movement prohibitions. *Transportation Research Part B: Methodological*, 19(3):197–208.

Eiselt, H. A. and Laporte, G. (1995). Objectives in location problems. In Drezner, Z., editor, *Facility Location: A Survey of Applications and Methods*, Springer Series in Operations Research and Financial Engineering, chapter 8, pages 151 – 180. Springer-Verlag, New York.

Fidler, M. and Einhoff, G. (2004). Routing in turn-prohibition based feed-forward networks. In *Networking 2004*, volume 3042 of *Lecture Notes in Computer Science*, pages 1168–1179. Springer, Berlin.

Ghoseiri, K., Szidarovszky, F., and Asgharpour, M. J. (2004). A multi-objective train scheduling model and solution. *Transportation Research Part B: Methodological*, 38(10):927 – 952.

Giallombardo, G., Moccia, L., Salani, M., and Vacca, I. (2010). Modeling and solving the tactical berth allocation problem. *Transportation Research Part B: Methodological*, 44(2):232–245.

Gutiérrez, E. and Medaglia, A. (2008). Labeling algorithm for the shortest path problem with turn prohibitions with application to large-scale road networks. *Annals of Operations Research*, 157(1):169–182.

Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459.

Harsanyi, J. C. (1975). Review: Can the maximin principle serve as a basis for morality? A critique of John Rawls's theory. *The American Political Science Review*, 69(2):594–606.

Ibaraki, T. (2010). A personal perspective on problem solving by general purpose solvers. *International Transactions in Operational Research*, 17(3):303–315.

Kirby, R. F. and Potts, R. B. (1969). The minimum route problem for networks with turn penalties and prohibitions. *Transportation Research*, 3:397–408.

Laporte, G. (1988). Location-routing problems. In Golden, B. L. and Assad, A. A., editors, *Vehicle Routing: Methods and Studies*, volume 16 of *Studies in management science and systems*, pages 163–197. North-Holland, Amsterdam.

Laporte, G. (1997). Modeling and solving several classes of arc routing problems as traveling salesman problems. *Computers & Operations Research*, 24(11):1057–1061.

Laporte, G., Asef-Vaziri, A., and Sriskandarajah, C. (1996). Some applications of the generalized travelling salesman problem. *Journal of the Operational Research Society*, 47(12):1461–1467.

Laporte, G., Chapleau, S., Landry, P.-E., and Mercure, H. (1989). An algorithm for the design of mailbox collection routes in urban areas. *Transportation Research Part B: Methodological*, 23(4):271–280.

Li, F., Golden, B. L., and Wasil, E. A. (2007). The open vehicle routing problem: Algorithms, large-scale test problems, and computational results. *Computers & Operations Research*, 34(10):2918 – 2930.

Li, L. Y. O. and Fu, Z. (2002). The school bus routing problem: A case study. *Journal of the Operational Research Society*, 53(5):552–558.

Mandell, M. B. (1991). Modelling effectiveness-equity trade-offs in public service delivery systems. *Management Science*, 37(4):467–482.

Marsh, M. T. and Schilling, D. A. (1994). Equity measurement in facility location analysis: A review and framework. *European Journal of Operational Research*, 74(1):1–17.

Moccia, L., Cordeau, J.-F., and Laporte, G. (2010a). An incremental tabu search heuristic for the generalized vehicle routing problem with time windows. *Journal of the Operational Research Society*, Submitted.

Moccia, L., Cordeau, J.-F., Laporte, G., Ropke, S., and Valentini, M. P. (2010b). Modeling and solving a multimodal transportation problem with flexible-time and scheduled services. *Networks*, Forthcoming.

Pacheco, J. and Martí, R. (2006). Tabu search for a multi-objective routing problem. *Journal of the Operational Research Society*, 57(1):29–37.

Pallottino, S. and Scutellà, M. G. (1998). Shortest path algorithms in transportation models: Classical and innovative aspects. In Marcotte, P. and Nguyen, S., editors, *Equilibrium and Advanced Transportation Modelling*, pages 245—281. Kluwer, Norwell.

Paquette, J., Cordeau, J.-F., and Laporte, G. (2009). Quality of service in dial-a-ride operations. *Computers & Industrial Engineering*, 56(4):1721 – 1734.

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

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press, Cambridge.

Savas, E. S. (1978). On equity in providing public services. *Management Science*, 24(8):800–808.

Sen, A. (1973). *On Economic Inequality*. Oxford University Press, Oxford.

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

Tzeng, G.-H. and Shiau, T.-A. (1988). Multiple objective programming for bus operation: A case study for Taipei city. *Transportation Research Part B: Methodological*, 22(3):195–206.

Ziliaskopoulos, A. K. and Mahmassani, H. S. (1996). A note on least time path computation considering delays and prohibitions for intersection movements. *Transportation Research Part B: Methodological*, 30(5):359–367.