Ann Oper Res (2010) 175: 213–245
DOI 10.1007/s10479-009-0650-0

Exact algorithms for routing problems under vehicle capacity constraints

Roberto Baldacci · Paolo Toth · Daniele Vigo

Published online: 1 November 2009

Abstract The solution of a vehicle routing problem calls for the determination of a set of routes, each performed by a single vehicle which starts and ends at its own depot, such that all the requirements of the customers are fulfilled and the global transportation cost is minimized. The routes have to satisfy several operational constraints which depend on the nature of the transported goods, on the quality of the service level, and on the characteristics of the customers and of the vehicles. One of the most common operational constraint addressed in the scientific literature is that the vehicle fleet is capacitated and the total load transported by a vehicle cannot exceed its capacity.

This paper provides a review of the most recent developments that had a major impact in the current state-of-the-art of exact algorithms for vehicle routing problems under capacity constraints, with a focus on the basic Capacitated Vehicle Routing Problem (CVRP) and on *heterogeneous* vehicle routing problems.

The most important mathematical formulations for the problem together with various relaxations are reviewed. The paper also describes the recent exact methods and reports a comparison of their computational performances.

Keywords Capacitated vehicle routing problem · Heterogeneous vehicle routing problem · Exact algorithms · Survey

This is an updated version of the paper “Recent advances in vehicle routing exact algorithms” that appeared in 4OR, 5(4), 269–298 (2007).

R. Baldacci · D. Vigo (✉)
DEIS, University of Bologna, Via Venezia 52, 47023 Cesena, Italy
e-mail: daniele.vigo@unibo.it

R. Baldacci
e-mail: r.baldacci@unibo.it

P. Toth
DEIS, University of Bologna, Viale Risorgimento 2, 40136, Bologna, Italy
e-mail: paolo.toth@unibo.it

1 Introduction

The Vehicle Routing Problem (VRP) is one of the most studied combinatorial optimization problems and is concerned with the optimal design of routes to be used by a fleet of vehicles to serve a set of customers. Since it was first proposed by Dantzig and Ramser (1959), hundreds of papers were devoted to the exact and approximate solution of the basic version of the VRP, known as the *Capacitated* Vehicle Routing Problem (CVRP), in which a homogeneous fleet of vehicles is available and the only considered constraint is the vehicle capacity. An important variant of the basic CVRP, which received a great attention in the scientific literature, arises when the vehicle fleet is characterized by different capacities and costs. This problems is generally known as *heterogeneous* VRP.

An interesting survey covering early exact methods for the CVRP is given by Laporte and Nobert (1987). The book edited by Toth and Vigo (2002) provides a comprehensive overview of exact methods for the CVRP and other variants proposed up to the end of the twentieth century. This work was updated by the survey of Cordeau et al. (2007).

Many different heuristics are proposed in the literature for the CVRP and its variants. Among the various surveys on heuristic algorithms for the CVRP, we mention the surveys of Laporte and Semet (2002) and of Gendreau et al. (2002) in the book edited by Toth and Vigo (2002). A specific survey on heterogeneous VRPs can be found in Baldacci et al. (2008a) which also covers lower bounds for heterogeneous VRPs.

In this paper, an updated version of the survey paper published in Baldacci et al. (2007), we provide a review of the most recent developments in the exact solution of the CVRP and of the heterogeneous VRPs on undirected graphs that were not covered in the previous survey works and that had a major impact in the current state-of-the-art of exact algorithms for this problem family. In particular, we present the different mathematical formulations used in the literature and discuss their interrelations, combinations and properties that were exploited in the most successful recent exact approaches. The structure of such algorithms is also discussed and a comparative analysis of their performance on the solution of well-known test instances from the literature is given. As previously mentioned, we concentrate here on the undirected version of the problem which is the most widely studied. As far as we know, the most recent exact approaches specifically designed for the directed CVRP are those proposed by Fischetti et al. (1994) and by Pessoa et al. (2008).

This paper is organized as follows. In the following section we formally describe the CVRP, the different variants of the heterogeneous VRP and introduce the notation that will be used throughout the paper. Section 3 reviews mathematical formulations for the CVRP and its heterogeneous variants. In Sect. 4, we describe the valid inequalities proposed for the different mathematical formulations together with a comparison of various relaxations. Sections 5 to 7 review the most recent exact algorithms, while a comparison of their computational performances is reported in Sect. 8. Finally, the last section offers conclusions and suggestions for future research.

2 Problems definition and basic notation

In this section we give a formal definition of the CVRP and of the variants of the heterogeneous VRP studied in the literature. In this section, we introduce as well the relevant notation required to define suitable models used by the exact approaches.

In the CVRP all customers correspond to deliveries, the demands are deterministic, known in advance and may not be split among different routes. The vehicles are identical, all based at a single depot, and only the capacity restrictions for the vehicles are imposed. The objective is to minimize the total cost of the routes (i.e., their length or travel time) needed to serve all the customers. More precisely, the CVRP may be described as the following graph theoretic problem.

Let $G = (V, E)$ be a complete and undirected graph where $V = \{0, \dots, n\}$ is the vertex set and E is the edge set. Vertex set $V_c = \{1, \dots, n\}$ corresponds to n *customers*, whereas vertex 0 corresponds to the depot. A nonnegative *cost*, d_{ij}, is associated with each edge $\{i, j\} \in E$ and represents the *travel cost* spent to go from vertex i to vertex j. In several practical cases the cost matrix $[d_{ij}]$ satisfies the *triangle inequality*. Each customer $i \in V_c$ is associated with a known nonnegative *demand*, q_i, to be delivered (the depot having a fictitious demand $q_0 = 0$). A set of m identical vehicles, each with capacity Q, is available at the depot. Without loss of generality we assume that $q_i \leq Q$ for each $i \in V_c$.

In the variant of CVRP that is generally considered in the literature and that we study here, all available vehicles must be used, each performing exactly one route. We also assume that m is not smaller than $m_{\min}$, which is the minimum number of vehicles needed to serve all the customers. The value of $m_{\min}$ may be determined by solving the *Bin Packing Problem* (BPP) associated with the CVRP, which calls for the determination of the minimum number of bins (i.e., vehicles), each with capacity Q, required to load all the n items (i.e., customers), each with nonnegative weight q_i, $i \in V_c$.

The CVRP consists of finding a collection of exactly m simple *cycles* or *routes* with minimum cost, defined as the sum of the costs of the edges belonging to the routes, and such that:

(i) each route visits the depot vertex;
(ii) each customer vertex is visited by exactly one route;
(iii) the sum of the demands of the vertices visited by a route does not exceed the vehicle capacity Q.

The CVRP is known to be $\mathcal{NP}$-hard (in the strong sense), and generalizes the well-known *Traveling Salesman Problem* (TSP), which calls for the determination of a minimum cost simple cycle visiting all the vertices of G (Hamiltonian cycle), and arising when $Q \geq \sum_{i \in V_c} q_i$ and $m = 1$.

In the heterogeneous VRP, the vehicle fleet is composed of m different vehicle types, with $M = \{1, \dots, m\}$. For each type $k \in M$, m_k vehicles are available at the depot, each having a capacity equal to Q_k. Each vehicle type is also associated with a nonnegative fixed cost, equal to F_k, modelling, e.g., rental or capital amortization costs. In addition, for each edge $\{i, j\} \in E$ and for each vehicle type $k \in M$ a nonnegative routing cost, d_{ij}^k is given. A route is defined as the pair (R, k), where $R = (i_1, i_2, \dots, i_{|R|})$, with $i_1 = i_{|R|} = 0$ and $\{i_2, \dots, i_{|R|-1}\} \subseteq V_c$, is a simple cycle in G containing the depot, and k is the type of vehicle associated with the route. A route (R, k) is feasible if the total demand of the customers visited by the route does not exceed the vehicle capacity Q_k (i.e., $\sum_{h=2}^{|R|-1} q_{i_h} \leq Q_k$). The cost of a route corresponds to the sum of the costs of the edges forming the route, plus the fixed cost of the vehicle associated with it (i.e., $\sum_{h=1}^{|R|-1} d_{i_h i_{h+1}}^k + F_k$).

The most general version of the heterogeneous VRP consists of designing a set of feasible routes with minimum total cost, and such that:

(i) each customer is visited by exactly one route;
(ii) the number of routes performed by vehicles of type $k \in M$ is not greater than m_k.

Several variants of this general problem were presented in the literature, depending on the available fleet and the type of considered costs. In particular, the following problem characteristics were considered:

(i) the vehicle fleet is composed by an *unlimited* number of vehicles for each type, i.e., $m_k = +\infty$, $\forall k \in M$;
(ii) the fixed costs of the vehicles are *not considered*, i.e., $F_k = 0$, $\forall k \in M$;
(iii) the routing costs are *vehicle-independent*, i.e., $d_{ij}^{k_1} = d_{ij}^{k_2}$, $\forall k_1, k_2 \in M$, $k_1 \neq k_2$, and $\forall \{i, j\} \in E$.

Related problems that received attention in the literature are the Site-Dependent VRP (SDVRP) and the Multi-Depot VRP (MDVRP).

In the SDVRP there is a limited heterogeneous fleet available for the service, no vehicle fixed costs are considered, routing costs are vehicle-independent, and each customer may include restrictions on the vehicle types that may visit it. It may be observed that SDVRP is a special case of the general heterogeneous VRP described above, where the routing cost d_{ij}^k of all edges incident to customer j is set to infinity for all vehicles types k that are incompatible with customer j.

The MDVRP is an extension of the CVRP where a customer can be served by an unlimited fleet of identical vehicles of capacity Q, located at p depot. Let $[\widehat{d}_{ij}]$ be a $(n+p) \times (n+p)$ symmetric cost matrix, where $\widehat{d}_{n+k\ i}$ is the travel cost for going from depot $k = 1, \dots, p$ to customer $i \in V_c$. Any MDVRP instance can be converted into an equivalent heterogeneous VRP instance generating $m = p$ different vehicle types and setting for each vehicle type $k \in M$:

$$Q_k = Q, \quad m_k = n, \quad F_k = 0 \quad \text{and}$$

$$d_{ij}^k = \begin{cases} \widehat{d}_{n+k\ j}, & \text{if } i = 0, \\ \widehat{d}_{ij}, & \text{otherwise,} \end{cases} \qquad \forall \{i, j\} \in E.$$

Table 1 summarizes the different problem variants that were actually considered in the literature. The different problem variants have been referred in the literature using different names. However, there is a certain homogeneity towards calling heterogenous VRPs the variants with limited number of vehicles, and Fleet Size and Mix those with unlimited number of vehicles. Therefore, we adopted the unified naming convention introduced by Baldacci et al. (2008a), that uses two acronyms (HVRP and FSM) and adds them two letters indicating whether fixed or routing costs are considered: “F” for fixed costs and “D” for vehicle-dependent routing costs, respectively.

Thus, we will refer to the problem variants as follows (see Table 1):

(a) Heterogeneous VRP with Fixed Costs and Vehicle-Dependent Routing Costs (HVRPFD). This variant corresponds to the most general variant described above;

Table 1 Heterogeneous VRP: problem variants presented in the literature

Problem variant	Fleet size	Fixed costs	Routing costs
HVRPFD	Limited	Considered	Dependent
HVRPD	Limited	Not considered	Dependent
SDVRP	Limited	Not considered	Site-dependent
FSMFD	Unlimited	Considered	Dependent
FSMD	Unlimited	Not considered	Dependent
MDVRP	Unlimited	Not considered	Depot-dependent
FSMF	Unlimited	Considered	Independent

(b) Heterogeneous VRP with Vehicle Dependent Routing Costs (HVRPD);
(c) Fleet Size and Mix VRP with Fixed Costs and Vehicle Dependent Routing Costs (FSMFD);
(d) Fleet Size and Mix VRP with Vehicle Dependent Routing Costs (FSMD);
(e) Fleet Size and Mix VRP with Fixed Costs (FSMF).

As to SDVRP and MDVRP we kept the original acronyms that are consistently adopted in the literature. All the problems described above are $\mathcal{NP}$-hard as they are generalizations of the CVRP.

3 Mathematical formulations

In this section we describe the mathematical formulations of the CVRP and its heterogeneous variants that are used as a base for the most recent exact solution approaches. In particular, we examine two and three-index vehicle flow, commodity flow and set partitioning formulations.

In addition to the notation already introduced, for a subset $S \subseteq V_c$, let $\overline{S} = V_c \setminus S$ be the *complement* of S and let $\delta(S)$ be the *cutset* defined by S (i.e., $\delta(S) = \{\{i, j\} \in E : i \in S, j \notin S$ or $i \notin S, j \in S\}$. Moreover, we denote by $r(S)$ and by $q(S)$ the optimal solution value of the BPP associated with customer set S and the total demand of the customers in S, respectively. Also, let $E(S)$ denote the set of edges in G with both end-vertices in S and, given two disjoint vertex sets S_1, S_2, let $E(S_1 : S_2)$ denote the set of edges crossing from S_1 to S_2 (i.e., $E(S_1 : S_2) = \delta(S_1) \cap \delta(S_2)$). Finally, let $\mathcal{S} = \{S : S \subseteq V_c, |S| \geq 2\}$.

3.1 Vehicle flow formulations

The so-called vehicle flow formulations use binary decision variables to indicate if a vehicle travels between a pair of vertices in G. In three-index models there is a specific variable for each vehicle and vertex pair, whereas two-index models aggregate this information over the different vehicles.

The *two-index vehicle flow* formulation of the CVRP was originally proposed by Laporte et al. (1985) and is as follows. Let x_{ij} be an integer variable which may take value $\{0, 1\}$, $\forall \{i, j\} \in E \setminus \{\{0, j\} : j \in V_c\}$ and value $\{0, 1, 2\}$, $\forall \{0, j\}$, $j \in V_c$, with $x_{ij} = 1$ when edge $\{i, j\}$ is traveled and $x_{0j} = 2$ when a route including the single customer j is selected in the solution. Then, CVRP can be formulated as the following integer program.

$$\min \sum_{\{i,j\} \in E} d_{ij} x_{ij} \tag{1}$$

$$\text{s.t.} \quad \sum_{\{i,j\} \in \delta(\{h\})} x_{ij} = 2 \quad (\forall h \in V_c), \tag{2}$$

$$\sum_{\{i,j\} \in \delta(S)} x_{ij} \geq 2\lceil q(S)/Q \rceil \quad (\forall S \in \mathcal{S}), \tag{3}$$

$$\sum_{j \in V_c} x_{0j} = 2m, \tag{4}$$

$$x_{ij} \in \{0, 1\} \quad (\forall \{i, j\} \in E \setminus \{\{0, j\} : j \in V_c\}), \tag{5}$$

$$x_{0j} \in \{0, 1, 2\} \quad (\forall \{0, j\}, j \in V_c). \tag{6}$$

Constraints (2) are the degree constraints for each customer. Constraints (3) are the so-called *Rounded Capacity* (RC) inequalities which, for any subset S of customers that does not include the depot, impose that at least $\lceil q(S)/Q \rceil$ vehicles enter and leave it. Constraint (4) states that m vehicles must leave and return to the depot, while constraints (5) and (6) are the integrality constraints.

Golden et al. (1977) formulated the CVRP using three-index binary variables x_{ij}^k as vehicle flow variables to indicate whether vehicle k travels directly from customer i to customer j ($x_{ij}^k = 1$ if vehicle k travels directly from customer i to customer j, 0 if not). In addition, for each vertex and each vehicle there is an assignment variable y_i^k ($i \in V$; $k = 1, \dots, m$) that takes value 1 if customer i is served by vehicle k in the optimal solution, and 0 otherwise. The *three-index vehicle flow* formulation for the CVRP is:

$$\min \sum_{\{i,j\}\in E} d_{ij} \sum_{k=1}^{m} x_{ij}^k \tag{7}$$

$$\text{s.t.} \quad \sum_{\{i,j\}\in\delta(\{h\})} x_{ij}^k = 2y_h^k \quad (\forall h \in V_c, k = 1, \dots, m), \tag{8}$$

$$\sum_{\{i,j\}\in\delta(S)} x_{ij}^k \geq 2y_h^k \quad (\forall S \in \mathcal{S}, h \in S, k = 1, \dots, m), \tag{9}$$

$$\sum_{k=1}^{m} y_i^k = 1 \quad (\forall i \in V_c), \tag{10}$$

$$\sum_{k=1}^{m} y_0^k = m, \tag{11}$$

$$\sum_{i\in V_c} q_i y_i^k \leq Q \quad (k = 1, \dots, m), \tag{12}$$

$$x_{ij}^k \in \{0, 1\} \quad (\forall \{i, j\} \in E \setminus \{\{0, j\} : j \in V_c\}, k = 1, \dots, m), \tag{13}$$

$$x_{0j}^k \in \{0, 1, 2\} \quad (\forall \{0, j\}, j \in V_c, k = 1, \dots, m), \tag{14}$$

$$y_i^k \in \{0, 1\} \quad (\forall i \in V, k = 1, \dots, m). \tag{15}$$

Degree constraints (8) state that each customer served by a vehicle k has exactly two incident edges traveled by k. Constraints (9) are the subtour elimination constraints which prohibit subtours not containing the depot. Constraints (10) impose that each customer is assigned to exactly one vehicle and constraint (11) ensures that all vehicles are used. Finally, constraints (12) model the demand limitations imposed by the capacity Q of each vehicle.

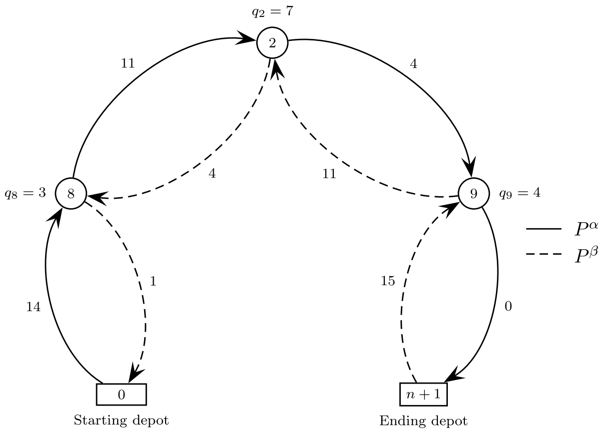
The *three-index vehicle flow* formulation for the CVRP can be extended to model the HVRPFD as follows:

$$\min \sum_{\{i,j\}\in E} \sum_{k=1}^{m} d_{ij}^k x_{ij}^k + \sum_{k=1}^{m} F_k y_0^k \tag{16}$$

$$\text{s.t.} \quad \sum_{i\in V_c} q_i y_i^k \leq Q_k \quad (k = 1, \dots, m), \tag{17}$$

(8), (9), (10), (11), (13), (14) and (15).

Fig. 1 Flow paths for a route of three customers



In the above model, the objective function (16) states to minimize the total routing cost plus the total fixed cost of the vehicles used. Similarly to constraints (12), constraints (17) impose the demand limitations imposed by the heterogeneous vehicle fleet.

3.2 Two-commodity flow formulation

The two-commodity flow formulation of the CVRP proposed by Baldacci et al. (2004) is based on the two-commodity flow formulation introduced by Finke et al. (1984) for the TSP.

To model single customer routes, this formulation requires the extended graph $\overline{G} = (\overline{V}, \overline{E})$ obtained from G by adding vertex $n+1$ which is a copy of depot vertex 0. Thus $\overline{V} = V \cup \{n+1\}$, $V_c = \overline{V} \setminus \{0, n+1\}$, $\overline{E} = E \cup \{\{i, n+1\}, i \in V_c\}$ and $d_{in+1} = d_{0i}$, $\forall i \in V_c$. Note that graph $\overline{G}$ has the same family of subsets $\mathcal{S}$ of graph G. In graph $\overline{G}$, a route $R = (i_1, i_2, \ldots, i_{|R|})$ is a simple path from vertex $i_1 = 0$ to vertex $i_{|R|} = n+1$.

This formulation uses two flow variables, y_{ij} and y_{ji}, to represent an edge $\{i, j\} \in \overline{E}$ of a feasible CVRP solution along which the vehicle carries a combined load of Q units. If a vehicle travels from i to j then the flow y_{ij} represents the load of the vehicle and the flow y_{ji} represents the empty space on the vehicle (i.e., $y_{ji} = Q - y_{ij}$). The flow variables y_{ij}, $i, j \in \overline{V}, i \neq j$, define two flow paths for any route of a feasible solution: one path from vertex 0 to vertex $n+1$ is given by the flow variables representing the vehicle load, while the second path from vertex $n+1$ to vertex 0 is defined by the flow variables representing the empty space on the vehicle.

Figure 1 shows an example of a three-customer route for a vehicle of capacity $Q = 15$ and the two paths P^α and P^β represented by the flow variables $\{y_{ij}\}$ defining the route. Path P^α is given by the variables representing the vehicle load: $(y_{08}, y_{82}, y_{29}, y_{9\,n+1})$. For example, the flow $y_{08} = 14$ indicates that the vehicle leaves the depot with a load equal to the total demand of the three customers. Path P^β is defined by the variables representing the empty space on the vehicle: $(y_{n+1\,9}, y_{92}, y_{28}, y_{80})$. Note that $y_{n+1\,9} = 15$ indicates that the vehicle arrives empty at the depot. Note also that for every edge $\{i, j\}$ of the route we have $y_{ij} + y_{ji} = Q$.

Let x_{ij} be a binary variable equal to 1 if edge $\{i,j\} \in \overline{E}$ is in the solution, 0 otherwise. The *two-commodity flow* formulation for the CVRP is as follows:

$$\min \sum_{\{i,j\}\in\overline{E}} d_{ij}x_{ij} \tag{18}$$

$$\text{s.t.} \quad \sum_{j\in\overline{V}} (y_{ji} - y_{ij}) = 2q_i \quad (\forall i \in V_c), \tag{19}$$

$$\sum_{j\in V_c} y_{0j} = q(V_c), \tag{20}$$

$$\sum_{j\in V_c} y_{j0} = mQ - q(V_c), \tag{21}$$

$$\sum_{j\in V_c} y_{n+1j} = mQ, \tag{22}$$

$$\sum_{\{i,j\}\in\delta(\{h\})} x_{ij} = 2 \quad (\forall h \in V_c), \tag{23}$$

$$y_{ij} + y_{ji} = Qx_{ij} \quad (\forall \{i,j\} \in \overline{E}), \tag{24}$$

$$y_{ij} \geq 0,\ y_{ji} \geq 0 \quad (\forall \{i,j\} \in \overline{E}), \tag{25}$$

$$x_{ij} \in \{0,1\} \quad (\forall \{i,j\} \in \overline{E}). \tag{26}$$

Constraints (19)–(22) and the nonnegative constraints (25) define a feasible flow pattern from the source vertices 0 and $n+1$ to sink vertices in $V_c \cup \{0\}$. The *outflow* at source vertex 0 (20) is equal to the total customer demand, while the *inflow* at source $n+1$ (22) corresponds to the total capacity of the vehicle fleet. Equation (19) states that the *inflow* minus the *outflow* at each customer $i \in V_c$ is equal to $2q_i$, while the *inflow* at vertex 0 (21) corresponds to the residual capacity of the vehicle fleet. Constraints (24) define the edges of a feasible solution and constraints (23) force any feasible solution to contain two edges incident to each customer.

Note that the two-commodity formulation can be rewritten in terms of variables y_{ij} only, once variables x_{ij} are substituted with variables y_{ij} using (24). In this case, the integrality constraints (26) should be replaced by $y_{ij} + y_{ji} \in \{0, Q\}$, $\forall \{i,j\} \in \overline{E}$.

The *two-commodity flow* formulation for the CVRP has been recently extended by Baldacci et al. (2009, to appear) to model the FSMF. Other FSMF mathematical formulations have been proposed by Yaman (2006), who described six mathematical formulations, called $HVRP_1$ to $HVRP_6$. The first four formulations are based on Miller-Tucker-Zemlin constraints, while the last two formulations use three-index vehicle flow variables to indicate if a vehicle of a specific type travels between a given pair of customers. In addition, Yaman (2006) made use of commodity-flow variables to eliminate subtours: model $HVRP_5$ used aggregated commodity-flow variables, whereas model $HVRP_6$ disaggregated such variables over the different vehicle types.

3.3 Set partitioning formulation

The set partitioning formulation of the CVRP was originally proposed by Balinski and Quandt (1964) and associates a binary variable with each feasible route.

Let $\mathcal{R}$ be the index set of all feasible routes and let $a_{i\ell}$ be a binary coefficient that is equal to 1 if vertex $i \in V$ belongs to route $\ell \in \mathcal{R}$ and takes the value 0 otherwise (note that $a_{0\ell} = 1, \forall \ell \in \mathcal{R}$). In the following we will use R_ℓ to indicate the subset of vertices (i.e., $R_\ell = \{0, i_1, i_2, \ldots, i_h\}$) visited by route $\ell \in \mathcal{R}$. Such a route represents the trip of one vehicle leaving the depot, delivering the demands of the customers in $R_\ell \setminus \{0\}$, and returning to the depot. Each route $\ell \in \mathcal{R}$ has an associated cost c_ℓ, that is equal to the optimal solution cost of the TSP instance defined by R_ℓ.

Let ξ_ℓ be a (0–1) binary variable that is equal to 1 if and only if route $\ell \in \mathcal{R}$ belongs to the optimal solution. The *set partitioning* formulation for the CVRP is as follows:

$$\min \sum_{\ell \in \mathcal{R}} c_\ell \xi_\ell \tag{27}$$

$$\text{s.t.} \quad \sum_{\ell \in \mathcal{R}} a_{i\ell} \xi_\ell = 1 \quad (\forall i \in V_c), \tag{28}$$

$$\sum_{\ell \in \mathcal{R}} \xi_\ell = m, \tag{29}$$

$$\xi_\ell \in \{0, 1\} \quad (\forall \ell \in \mathcal{R}). \tag{30}$$

Constraints (28) specify that each customer $i \in V_c$ must be covered by one route and constraint (29) requires that m routes are selected.

This model is valid for any type of cost matrix $[d_{ij}]$. However, if the cost matrix $[d_{ij}]$ satisfies the triangle inequality, then equality constraints (28) can be written as a '$\geq$' inequality, thus obtaining the so-called *set covering* formulation that preserves the optimal objective function value (see Bramel and Simchi-Levi 2002). The set partitioning formulation is very general and can take into account several route constraints (e.g., time windows, precedences, route lengths) since route feasibility is implicitly considered in the definition of set $\mathcal{R}$.

The set partitioning formulation described above remains valid if the set of routes $\mathcal{R}$ is enlarged with the set $\widehat{\mathcal{R}} \supset \mathcal{R}$ containing, in addition to elementary routes, *non-elementary* routes, which are routes in which the vehicle is permitted to visit customers more than once. In this case, coefficient $a_{i\ell}$ is a general integer coefficient that is equal to the number of times customer i is visited by route ℓ. Note that the overall integer programming formulation remains valid, since constraints (28) ensure that the variables representing non-elementary routes will be automatically eliminated when ξ is binary.

The above *set partitioning* formulation can be extended to model the HVRPFD as follows. Let $\mathcal{R}_k$ be the index set of all feasible routes for a vehicle of type $k \in M$. Each route $\ell \in \mathcal{R}_k$ has an associated cost c_ℓ^k. Let $\mathcal{R}_i^k \subset \mathcal{R}_k$ be the index subset of the routes for a vehicle of type k covering customer $i \in V_c$.

Let ξ_ℓ^k be a binary variable that is equal to 1 if and only if route $\ell \in \mathcal{R}_k$ belongs to the optimal solution. The set partitioning model is as follows:

$$\min \sum_{k \in M} \sum_{\ell \in \mathcal{R}_k} c_\ell^k \xi_\ell^k \tag{31}$$

$$\text{s.t.} \quad \sum_{k \in M} \sum_{\ell \in \mathcal{R}_i^k} \xi_\ell^k = 1 \quad (\forall i \in V_c), \tag{32}$$

$$\sum_{\ell \in \mathcal{R}_k} \xi_\ell^k \leq m_k \quad (\forall k \in M), \tag{33}$$

$$\xi_\ell^k \in \{0, 1\} \quad (\forall \ell \in \mathcal{R}_k, \forall k \in M). \tag{34}$$

Constraints (32) specify that each customer $i \in V_c$ must be covered exactly by one route and constraints (33) require that at most m_k routes are selected for a vehicle of type $k \in M$.

4 Relaxations and valid inequalities

Valid lower bounds on the VRP can be derived from the LP relaxations of the different mathematical formulations described in the previous section. Some of the resulting LP programs cannot be solved directly, even for moderate size VRPs, since either the number of variables or constraints is exponential in the problem size. Thus, the lower bounds are usually computed using cutting plane and column generation techniques. In addition, to strengthen the relaxations, a variety of valid inequalities have been described in the literature for the different formulations.

In this section, we review the most effective valid inequalities described in the literature together with a comparison of various CVRP relaxations.

4.1 Valid inequalities for the two-index formulation

Valid inequalities for the vehicle flow formulations have been described by Laporte and Nobert (1984), Cornuéjols and Harche (1993), Fischetti et al. (1995), Augerat et al. (1995) and Letchford et al. (2002) and a description of some of them for the two-index formulation can be found in Naddef and Rinaldi (2002). Some of these inequalities are rather intricate, and not all of them have been successfully implemented within a branch-and-cut algorithm. Thus, in this section, we will only provide a summary of the main known results. Conditions under which some of these inequalities induce facets of the CVRP polytope are discussed in Cornuéjols and Harche (1993).

Valid inequalities for the two-index formulation similar to the RC inequalities, can be obtained according to the way in which the right-hand side of constraints (3) is computed. *Fractional Capacity* (FC) inequalities take the form:

$$\sum_{\{i,j\}\in\delta(S)} x_{ij} \geq 2\frac{q(S)}{Q}, \quad \forall S \in \mathcal{S}. \tag{35}$$

Note that as $q(S)$ can be smaller than Q, the following *Subtour Elimination* (SE) inequalities are not dominated by the FC inequalities:

$$\sum_{\{i,j\}\in\delta(S)} x_{ij} \geq 2, \quad \forall S \in \mathcal{S}. \tag{36}$$

Given the set of all feasible routes $\mathcal{R}$, let $\mathcal{P}$ denote the index set of all feasible m-partitions $P = \{\ell_1, \dots, \ell_m\}$ of V_c. Let $R(S)$ be the minimum number of vehicles needed to satisfy all customer demands in S in a feasible m-partition. For any nonempty set $S \subseteq V_c$, the value $R(S)$ can be computed as

$$R(S) = \min_{P\in\mathcal{P}}\{|\{\ell_i \in P : R_{\ell_i} \cap S \neq \emptyset\}|\}. \tag{37}$$

For a given set S, we have $R(S) \geq r(S) \geq \lceil q(S)/Q \rceil \geq \max\{1, q(S)/Q\}$.

A CVRP example, proposed by Cornuéjols and Harche (1993), having eight customers and four vehicles, each one with capacity $Q = 7$, where the customer demands are $q_1 = 5$, $q_2 = q_3 = q_4 = 3$, $q_5 = q_6 = q_7 = 4$, $q_8 = 2$, and S is given by the first four customers, shows that $R(S)$ can be different from $r(S)$. Indeed, for the example, we have $R(S) = 4$, $r(S) = 3$, and $\lceil q(S)/Q \rceil = 2$.

Using the definition of $R(S)$, Augerat et al. (1995) introduced the *Generalized Capacity* (GC) inequalities. Let $\mathcal{B} = \{S_1, \ldots, S_t\}$ be a collection of t (> 1) disjoint subsets of V_c, then

$$\sum_{h=1}^{t} \sum_{\{i,j\}\in\delta(S_h)} x_{ij} \geq 2 \min_{P\in\mathcal{P}} \left\{ \sum_{h=1}^{t} |\{\ell_i \in P : R_{\ell_i} \cap S_h \neq \emptyset\}| \right\} \tag{38}$$

is a valid inequality for the CVRP.

A weak version, and more tractable valid inequalities than GC can be obtained as follows. Let H be a subset of V_c containing all the subsets in $\mathcal{B}$ and assume that $q(S_h) \leq Q$ holds for $h = 1, \ldots, t$. Then, define $r(H|S_1, S_2, \ldots, S_t)$ to be the optimal solution value of a bin packing problem with one item of size q_i for each $i \in H \setminus \bigcup_{h=1}^{t} S_h$, where all items in each subset $S \in \mathcal{B}$ are constrained to stay together in the same bin. Then, if $H = V_c$, the *Weak Generalized Capacity* (WGC) inequality is defined as

$$\sum_{\{i,j\}\in\delta(V_c)} x_{ij} + \sum_{h=1}^{t} \sum_{\{i,j\}\in\delta(S_h)} x_{ij} \geq 2t + 2r(V_c|S_1, S_2, \ldots, S_t), \tag{39}$$

or, equivalently, since $\sum_{\{i,j\}\in\delta(V_c)} x_{ij} = 2m$, as

$$\sum_{h=1}^{t} \sum_{\{i,j\}\in\delta(S_h)} x_{ij} \geq 2t + 2(r(V_c|S_1, S_2, \ldots, S_t) - m). \tag{40}$$

If in inequality (40) set V_c is replaced by any of its subsets H containing all the sets S_h, $h = 1, \ldots, t$, then the resulting inequality

$$\sum_{\{i,j\}\in\delta(H)} x_{ij} + \sum_{h=1}^{t} \sum_{\{i,j\}\in\delta(S_h)} x_{ij} \geq 2t + 2r(H|S_1, S_2, \ldots, S_t) \tag{41}$$

is a valid CVRP inequality, called *Framed Capacity* (FrC) inequality, which generalizes inequality (40).

Araque et al. (1990) defined the *multistar inequalities* for the CVRP with unit demands. These inequalities take the form:

$$\alpha \sum_{\{i,j\}\in E(N)} x_{ij} + \beta \sum_{\{i,j\}\in E(N:S)} x_{ij} \leq \gamma, \tag{42}$$

where $N \subset V_c$ is the so-called *nucleus*, $S \subseteq V_c \setminus N$ is the set of so-called *satellites*, and α, β, γ are constants which depend on $|N|$ and $|S|$. Araque et al. also introduced the *partial multistar* inequalities, which take the form:

$$\alpha \sum_{\{i,j\}\in E(N)} x_{ij} + \beta \sum_{\{i,j\}\in E(C:S)} x_{ij} \leq \gamma, \tag{43}$$

where $C \subset N$ is the set of so-called *connector* vertices, and, again, α, β and γ are constants which depend on $|N|$, $|S|$ and $|C|$.

Letchford et al. (2002) generalized the multistar and partial multistar inequalities to the CVRP, yielding the so-called *homogeneous* multistar and partial multistar inequalities. The same authors also described a procedure called *polygon* procedure, that can be used to generate all known homogeneous inequalities in the literature, along with some new ones.

Gouveia (1995) proposed a related, but not identical, set of valid inequalities, called *Generalized Large Multistar* (GLM) inequalities:

$$\sum_{\{i,j\}\in\delta(S)} x_{ij} \geq \frac{2}{Q}\left(q(S) + \sum_{j\in\overline{S}} q_j \left(\sum_{\{i,h\}\in E(S:\{j\})} x_{ih}\right)\right), \quad \forall S \in \mathcal{S}. \tag{44}$$

It is easy to see that the GLM inequalities dominates the FC inequalities. These inequalities have been called by Letchford et al. (2002) *inhomogeneous*, since the coefficients of the edges in $E(S:\{j\})$ vary depending on the vertex $j \in \overline{S}$ involved. Letchford et al. introduced additional valid inhomogeneous inequalities, the *Knapsack Large Multistar* (KLM) inequalities, which generalize the GLM inequalities:

$$\sum_{\{i,j\}\in\delta(S)} x_{ij} \geq \frac{2}{\beta}\left(\sum_{i\in S}\alpha_i + \sum_{j\in\overline{S}}\alpha_j \sum_{\{i,h\}\in E(S:\{j\})} x_{ih}\right), \quad \forall S \in \mathcal{S}, \tag{45}$$

where $\alpha, \beta > 0$ are such that the inequality $\sum_{i\in V_c}\alpha_i y_i \leq \beta$ is valid for the 0–1 knapsack polytope $KP(Q,q) = \text{conv}\{y \in \{0,1\}^{|V_c|} : \sum_{i\in V_c} q_i y_i \leq Q\}$. Their validity follows from the fact that as $\sum_{i\in V_c}\alpha_i y_i \leq \beta$ is a valid inequality for $KP(Q,q)$, then any feasible CVRP solution must also be feasible for a modified CVRP instance in which the vector of the demands q and the vehicle capacity Q are replaced by α and β, respectively. Inequality (45) is then the GLM inequality for the modified CVRP instance.

Given an edge subset $F \subset E$, assume that there is no feasible CVRP solution which uses only edges in F. Then, at least one edge belonging to set $E \setminus F$ must be in the solution. Thus, the following inequality, called *hypotour* (HI) inequalities (see Augerat 1995), is a valid CVRP inequality:

$$\sum_{\{i,j\}\in\delta(E\setminus F)} x_{ij} \geq 1. \tag{46}$$

Other classes of valid inequalities have been presented for the integer polytope associated with the two-index formulation. In particular, some of the other valid inequalities are based on the successful results of polyhedral combinatorics developed for the TSP by Chvátal (1973) and by Grötschel and Padberg (1979, 1985). These include, among others, *comb* inequalities, *path-bin* inequalities, etc. (see Naddef and Rinaldi 2002).

4.2 Valid inequalities for the two-commodity formulation

All the valid inequalities known for the CVRP are clearly valid inequalities for the two-commodity formulation. These inequalities are expressed in terms of variables x_{ij} defined for the two-index formulation, but can be added to the two-commodity formulation once variables x_{ij} are replaced with variables y_{ij} using equations $y_{ij} + y_{ji} = Qx_{ij}$, $\forall\{i,j\} \in \overline{E}$. In addition, Baldacci et al. (2004) have proposed the following valid inequalities, called *commodity flow* (CF) inequalities:

$$\left.\begin{aligned}(Q - q_j)y_{ij} - q_j y_{ji} \geq 0\\ (Q - q_i)y_{ji} - q_i y_{ij} \geq 0\end{aligned}\right\}, \quad \forall\{i,j\} \in \overline{E}. \tag{47}$$

Baldacci et al. (2004) have shown that the LP relaxation of the two-commodity formulation with the addition of CF inequalities (47) satisfies the GLM inequalities.

Baldacci et al. (2009, to appear) introduced new classes of valid inequalities that explicitly take into account the heterogenous fleet and that are used to strengthen the linear programming relaxation of the two-commodity formulation used to model the FSMF. Yaman (2006) investigated the use of different valid inequalities to strengthen the linear relaxation of the different mathematical formulations proposed for the FSMF. The following families of valid inequalities were added by Yaman (2006) to possibly improve the lower bounds: *covering*, SE and GLM inequalities. Baldacci et al. (2009, to appear) described also new covering-type inequalities, that generalize those proposed by Yaman (2006).

4.3 Valid inequalities for the three-index formulation

In the case of the three-index formulation, it is quite simple to see that any solution x of the three-index formulation can be transformed into a feasible solution of the two-index formulation by using the setting:

$$x_{ij} = \sum_{k=1}^{m} x_{ij}^k, \quad \forall \{i, j\} \in E. \tag{48}$$

Then, any valid inequality in the form $\alpha x \leq \beta$ for the two-index formulation can be transformed into a valid inequality for the three-index formulation as follows:

$$\sum_{\{i,j\} \in E} \alpha_{ij} \sum_{k=1}^{m} x_{ij}^k \leq \beta. \tag{49}$$

The resulting families of valid inequalities have been called by Letchford and Salazar González (2006) *aggregated* inequalities. In addition, any valid inequality for the so-called *multiple knapsack polytope*, $\text{conv}\{(y^1, \ldots, y^m) \in \{0,1\}^{n \times m} :$ (10), (12) and (15) hold$\}$, yields a valid inequality for the three-index formulation that only involves variables y. Moreover, if we relax the equations (10) in, say, a Lagrangean fashion, the three-index formulation decomposes into m independent identical subproblems. Each subproblem corresponds to finding a capacitated circuit (i.e., a circuit passing through a set of vertices whose total demand does not exceed Q) passing through the depot. Therefore, any valid inequality for the capacitated circuit polytope yields a valid inequality for the three-index formulation which involves only a single $k \in \{1, \ldots, m\}$. These two last families of valid inequalities have been called *Multiple Knapsack* (MK) inequalities and *One-Vehicle* (OV) inequalities, respectively (see Letchford and Salazar González 2006).

4.4 Valid inequalities for the set partitioning formulation

Valid inequalities for the set partitioning formulation can be derived by noting that any solution ξ of the set partitioning formulation can be transformed into a solution x of the two-index formulation by setting:

$$x_{ij} = \sum_{\ell \in \mathcal{R}} \eta_{ij}^{\ell} \xi_{\ell}, \quad \forall \{i, j\} \in E, \tag{50}$$

where the coefficients η_{ij}^{ℓ} are defined as follows:

- if ℓ is a single customer route covering customer h, then $\eta_{0h}^{\ell} = 2$ and $\eta_{ij}^{\ell} = 0$, $\forall \{i, j\} \in E \setminus \{0, h\}$;

– if ℓ is not a single customer route, then $\eta_{ij}^{\ell} = 1$ for each edge $\{i, j\}$ covered by route ξ and $\eta_{ij}^{\ell} = 0$ otherwise.

Then, any CVRP valid inequality designed for the two-index formulation can be transformed into a valid inequality for the set partitioning formulation. The family $\mathcal{F}$ of these inequalities can be expressed in a general form as:

$$\sum_{\{i,j\}\in E} \alpha_{ij}^t x_{ij} \geq \beta^t, \quad t \in \mathcal{F}, \tag{51}$$

and using (50), inequalities (51) become the following valid inequalities for the set partitioning formulation:

$$\sum_{\ell\in\mathcal{R}} \alpha^t(R_\ell)\xi_\ell \geq \beta^t, \quad t \in \mathcal{F}, \tag{52}$$

where $\alpha^t(R_\ell) = \sum_{\{i,j\}\in E} \alpha_{ij}^t \eta_{ij}^{\ell}$.

Results on which inequalities are implied by the LP relaxation of the set partitioning formulation have been reported by Baldacci et al. (2004) and by Letchford and Salazar González (2006). It can be easily shown that any fractional set partitioning solution satisfies the degree equations (2) and the bound inequalities $x_{ij} \leq 1, \forall\{i, j\} \in E \setminus \{\{0, j\} : j \in V_c\}$ and $x_{0j} \leq 2, \forall\{0, j\}, j \in V_c$.

In particular, Baldacci et al. (2004) have shown that the LP relaxation of the set partitioning formulation implies the FC, SE and GLM inequalities. More precisely, let us consider for each set $S \in \mathcal{S}$ the surrogate constraint obtained by adding (28) corresponding to the customers in S after having multiplied the equation associated with $i \in S$ by q_i:

$$\sum_{\ell\in\mathcal{R}(S)} b_\ell(S)\xi_\ell = q(S), \quad \forall S \in \mathcal{S}, \tag{53}$$

where $b_\ell(S) = \sum_{i\in S} q_i a_{i\ell}$, $\forall \ell \in \mathcal{R}$, and $\mathcal{R}(S) = \{\ell \in \mathcal{R} : R_\ell \cap S \neq \emptyset\}$. Since $b_\ell(S) \leq \min[Q, Q - b_\ell(\overline{S})]$, (53) imply that the LP relaxation of the set partitioning formulation satisfies the following inequalities:

$$\sum_{\ell\in\mathcal{R}(S)} \xi_\ell \geq \max\left[1, \frac{q(S)}{Q} + \frac{1}{Q}\sum_{\ell\in\mathcal{R}(S)} b_\ell(\overline{S})\xi_\ell\right], \quad \forall S \in \mathcal{S}. \tag{54}$$

Note that the right-hand-side of (54) can be smaller than $\lceil q(S)/Q\rceil$ for some $S \in \mathcal{S}$, however, Baldacci et al. (2004) showed that $\sum_{\ell\in\mathcal{R}(S)} \xi_\ell \geq \lceil q(S)/Q\rceil$ for any subset $S \in \mathcal{S}$ such that $b_\ell(S) \leq \frac{1}{2}Q, \forall \ell \in \mathcal{R}(S)$.

By noting that any route $\ell \in \mathcal{R}(S)$ contains at least two edges, one having an ending vertex in S and the other in $\overline{S} \cup \{0\}$, the following inequality holds for any $S \in \mathcal{S}$:

$$\sum_{\ell\in\mathcal{R}} \sum_{\{i,j\}\in\delta(S)} \eta_{ij}^{\ell}\xi_\ell \geq 2\sum_{\ell\in\mathcal{R}(S)} \xi_\ell, \tag{55}$$

and as

$$b_\ell(\overline{S}) \geq \sum_{j\in\overline{S}} q_j \sum_{\{i,h\}\in E(S:\{j\})} \eta_{ih}^{\ell}\xi_\ell, \tag{56}$$

from inequalities (54) and expression (50) we obtain the following valid inequalities:

$$\sum_{\{i,j\}\in\delta(S)} x_{ij} \geq 2\max\left\{1, \frac{q(S)}{Q} + \frac{1}{Q}\sum_{j\in\overline{S}} q_j \sum_{\{i,h\}\in E(S:\{j\})} x_{ih}\right\}, \tag{57}$$

for all $S \in \mathcal{S}$, thus the LP relaxation of the set partitioning formulation satisfies SE, FC and GLM inequalities. From the definition of KLM inequalities, it is easy to note that the LP relaxation of the set partitioning formulation implies also the KLM inequalities. Moreover, Letchford and Salazar González (2006) have shown that the LP relaxation of the set partitioning formulation implies by projection some *hypotour-like* inequalities.

It is quite easy to find examples of fractional set partitioning solutions violating the following inequalities:

$$\sum_{\ell\in\mathcal{R}(S)} \rho_\ell(S)\xi_\ell \geq 2\lceil q(S)/Q\rceil, \quad \forall S \in \mathcal{S}, \tag{58}$$

where $\rho_\ell(S) = \sum_{\{i,j\}\in\delta(S)} \eta^\ell_{ij}$, which are the projections of the RC inequalities in the set partitioning space. Using inequalities (54), Baldacci et al. (2008b) have observed that inequalities (58) can be strengthened as follows:

$$\sum_{\ell\in\mathcal{R}(S)} \xi_\ell \geq \lceil q(S)/Q\rceil, \quad \forall S \in \mathcal{S}. \tag{59}$$

Indeed, inequalities (59) are a lifting of inequalities (58). The following example, due to Baldacci et al. (2008b), show a fractional set partitioning solution satisfying inequalities (58) but violating inequalities (59), for a given set S.

Consider a subset of customers $S = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ having demands equal to $q_{i_1} = q_{i_2} = q_{i_4} = q_{i_5} = 20$, $q_{i_3} = q_{i_6} = 70$, while the vehicle capacity is $Q = 100$. Let ξ be a fractional solution where the variables corresponding to the five routes $R_1, R_2, \ldots, R_5$ have values $\xi_1 = \xi_2 = \cdots = \xi_5 = 0.5$ and where $R_1 \cap S = \{i_1, i_2, i_4, i_5\}$, $R_2 \cap S = \{i_2, i_3\}$, $R_3 \cap S = \{i_1, i_6\}$, $R_4 \cap S = \{i_5, i_6\}$ and $R_5 \cap S = \{i_4, i_3\}$. Thus, $\lceil q(S)/Q\rceil = \lceil 220/100\rceil = 3$ and $\rho_1(S) = 4$, $\rho_2(S) = \cdots = \rho_5(S) = 2$. This solution satisfies constraint (58) for S since:

$$4\xi_1 + 2\xi_2 + 2\xi_3 + 2\xi_4 + 2\xi_5 = 6 = 2\lceil q(S)/Q\rceil,$$

but violates the corresponding constraint (59) since:

$$\xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 = 2.5 \leq \lceil q(S)/Q\rceil.$$

Baldacci et al. (2008b) have considered other valid inequalities which have been designed for the general Set Partitioning problem. Let $H = (\mathcal{R}, \mathcal{E})$ be the *conflict graph* where each vertex corresponds to a route and the edge set $\mathcal{E}$ contains every pair $\{\ell, \ell'\}, \forall \ell, \ell' \in \mathcal{R}$, $\ell < \ell'$, such that $R_\ell \cap R_{\ell'} \neq \{0\}$. Let $\mathcal{C}$ be the set of all cliques of H. Then, the following inequalities, called *clique* inequalities:

$$\sum_{\ell\in C} \xi_\ell \leq 1, \; C \in \mathcal{C}, \tag{60}$$

are valid inequalities for the set partitioning formulation.

Inequalities (59) and clique inequalities (60) can be easily extended to the HVRPFD case (see Baldacci and Mingozzi 2009).

4.5 Comparison of various CVRP relaxations

Before the works of Baldacci et al. (2004) and Letchford and Salazar González (2006) no systematic study was performed on how the different CVRP formulations presented in Sect. 3 relate to each other. The aim of this section is to briefly review the main results concerning dominance relations among the different LP relaxations of the various CVRP formulations.

A first result, due both to Baldacci et al. (2004) and Letchford and Salazar González (2006), shows that the lower bound obtained by the LP relaxation of the two-index formulation with the RC inequalities replaced by the GLM inequalities coincides with the lower bound obtained from the LP relaxation of the two-commodity formulation plus the CF inequalities. Note that the number of variables and constraints of the LP relaxation of the two-commodity formulation plus the CF inequalities increases polynomially with the size of the problem, whereas the LP relaxation of the two-index formulation with (3) replaced by (44) is exponential in size. On the other hand, the separation problem for the GLM inequalities can be solved in polynomial time (see Letchford et al. 2002).

Concerning a comparison between the two-index formulation and the three-index formulation, Letchford and Salazar González (2006) have shown that if we consider for the three-index formulation aggregated inequalities, OV inequalities and MK inequalities only, one can construct a cutting plane algorithm based on the two-index formulation which gives lower bounds of the same quality. What is more, from a computational point of view, a cutting plane method based on the two-index version, having far fewer variables, is likely to run much faster in practice.

If we compare the set partitioning formulation and the two-index formulation, from the observations made in the previous section, the lower bound obtained from the LP relaxation of the set partitioning formulation is at least as good as the one obtained by using degree equations, bounds and KLM inequalities in the two-index formulation.

Finally, allowing non-elementary routes in the set partitioning formulation weakens the formulation. Indeed, it is quite easy to find feasible solutions of the LP relaxation of the set partitioning formulation with non-elementary routes which do not satisfy even the SE inequalities.

5 Branch-and-cut methods

The first sophisticated branch-and-cut algorithm based on the two-index formulation for the CVRP was proposed by Augerat et al. (1995). The algorithm used a variety of valid inequalities, such as the RC, GC and HI inequalities described in the previous section. These last inequalities led to significant improvements in the quality of the lower bound. A detailed description of the work of Augerat et al. (1995) can be found in Naddef and Rinaldi (2002).

Based on the original work of Augerat et al. (1995), other branch-and-cut methods based on the two-index formulation have been recently proposed by Ralphs et al. (2003) and Lysgaard et al. (2004). Ralphs et al. (2003) described a branch-and-cut algorithm based on the two-index formulation and on the addition of the RC inequalities in a cutting plane fashion. Lysgaard et al. (2004) used a variety of valid inequalities, including the RC, FrC, strengthened comb, multistar, partial multistar, extended HI inequalities, and classical Gomory mixed integer cuts.

Baldacci et al. (2004) described a branch-and-cut algorithm based on the two-commodity formulation of the CVRP, where the RC and CF inequalities are used in a cutting plane fashion to strengthen the lower bound obtained by the LP relaxation of the two-commodity formulation.

Below we briefly review two of the most important aspects of any branch-and-cut algorithm: the separation algorithms and the branching strategies.

5.1 Separation algorithms

The most important aspect of any branch-and-cut algorithm is designing exact or heuristic algorithms that effectively separate a given fractional point from the convex hull of the integer solutions. Given a (fractional) point x^*, the separation problem associated with a given family $\mathcal{F}$ of valid inequalities consists of finding a member $\alpha x \geq \beta$ of $\mathcal{F}$, such that $\alpha x^* < \beta$.

Several heuristic and exact separation algorithms have been designed for the separation problems associated with the valid inequalities described in Sect. 4. A detailed description of some of these algorithms can be found in Augerat et al. (1995, 1998), Naddef and Rinaldi (2002) and Lysgaard et al. (2004).

The FC, SE, GLM and KLM inequalities can be separated in polynomial type. Indeed, if we consider the KLM inequalities (45), which include the FC and GLM ones, since

$$\sum_{\{i,j\}\in\delta(S)} x_{ij} = \sum_{\{i,j\}\in E(\{0\}:S)} x_{ij} + \sum_{\{i,j\}\in E(S:\overline{S})} x_{ij}$$

and

$$\sum_{\{i,h\}\in\delta(\{j\})} x_{ih} = 2 = x_{0j} + \sum_{\{i,h\}\in E(\{j\}:S)} x_{ih} + \sum_{\{i,h\}\in E(\{j\}:\overline{S})} x_{ih},$$

the inequalities can be rewritten as:

$$\begin{aligned}\sum_{\{i,j\}\in E(\{0\}:S)} x_{ij} &+ \sum_{\{i,j\}\in E(S:\overline{S})} x_{ij}\\ &\geq \sum_{j\in S}\frac{\alpha_j}{\beta}\Bigg(x_{0j} + \sum_{\{i,h\}\in E(\{j\}:S)} x_{ih} + \sum_{\{i,h\}\in E(\{j\}:\overline{S})} x_{ih}\Bigg)\\ &\quad + \frac{2}{\beta}\sum_{j\in\overline{S}}\alpha_j \sum_{\{i,h\}\in E(S:\{j\})} x_{ih}, \quad \forall S\in\mathcal{S},\end{aligned} \tag{61}$$

or, equivalently, as

$$\begin{aligned}&\sum_{j\in S}\left(1-\frac{\alpha_j}{\beta}\right)x_{0j} + \sum_{j\in\overline{S}}\Bigg(\sum_{\{j,h\}\in\delta(\{j\})}\frac{\alpha_h}{\beta}x_{jh} + \sum_{\{h,j\}\in\delta(\{j\})}\frac{\alpha_h}{\beta}x_{hj}\Bigg)\\ &\quad + \sum_{\{i,j\}\in E(S:\overline{S})} (1-(\alpha_i+\alpha_j)/\beta)x_{ij}\\ &\geq \sum_{j\in V_c}\Bigg(\sum_{\{j,h\}\in\delta(\{j\})}\frac{\alpha_h}{\beta}x_{jh} + \sum_{\{h,j\}\in\delta(\{j\})}\frac{\alpha_h}{\beta}x_{hj}\Bigg), \quad \forall S\in\mathcal{S}.\end{aligned} \tag{62}$$

The term at the right-hand-side of the above inequality does not depend on S and if $\alpha_i \leq \beta$, $\forall i\in V_c$, and $x_{ij}=0$ if $\alpha_i+\alpha_j > \beta$, then all coefficients are nonnegative. Then, given α_i and β and a fractional solution x^*, the most violated KLM inequality can be found by computing a minimum s–t cut on an undirected capacitated graph $\overline{G} = (\overline{V}, \overline{E})$ with $\overline{V} = V_c \cup \{s, t\}$. The edge set is $\overline{E} = E^* \cup \{\{s, j\} : j \in V_c\} \cup \{\{j, t\} : j \in V_c\}$. Every edge $\{s, j\}$ is associated with a capacity $(1 - \alpha_j/\beta)x^*_{0j}$, while a capacity $\sum_{\{j,h\}\in\delta(\{j\})} \frac{\alpha_h}{\beta} x^*_{jh} + \sum_{\{h,j\}\in\delta(\{j\})} \frac{\alpha_h}{\beta} x^*_{hj}$ is assigned to each edge $\{j, t\}$. The remaining edges $\{i, j\}$ are associated with capacities $\{(1 - (\alpha_i + \alpha_j)/\beta)x^*_{ij}\}$. Let $(S, \overline{V} \setminus S)$ be the minimum s–t cut of $\overline{G}$ and assume that $t \in S$. One can see that if the cut capacity is strictly smaller than $\sum_{j\in V_c}(\sum_{\{j,h\}\in\delta(\{j\})} \frac{\alpha_h}{\beta} x^*_{jh} + \sum_{\{h,j\}\in\delta(\{j\})} \frac{\alpha_h}{\beta} x^*_{hj})$ then vertex set $S = S \setminus \{t\}$ defines the most violated KLM inequality. Otherwise, no such violated inequality exists.

Although the SE, FC, GLM and KLM inequalities can be separated in polynomial time, their use tends to be ineffective in the context of branch-and-cut algorithms. Valid inequalities which play a crucial role in the development of a cutting plane algorithm for the CVRP are the RC inequalities. However, the separation problem for these constraints was shown to be strongly $\mathcal{NP}$-hard (see Naddef and Rinaldi 2002). Thus, a number of heuristic algorithms have been designed for their separation (see Augerat et al. 1995, 1998; Lysgaard et al. 2004), and one of the most simple and effective techniques is the so-called *greedy randomized algorithm*. The greedy randomized algorithm is an iterative procedure that is applied to a number of customer subsets $\overline{\mathcal{S}} \subset \mathcal{S}$ generated a priori. At each iteration, the following procedure is repeated for each $S \in \overline{\mathcal{S}}$.

1. Let $i^* \in V_c \setminus S$ be the customer such that:

$$\sum_{\{j,h\}\in\delta(E(S:\{i^*\}))} x^*_{ih} = \max_{i\in V_c\setminus S}\left\{\sum_{\{j,h\}\in\delta(E(S:\{i\}))} x^*_{ih}\right\}.$$

2. If the current solution x^* violates the RC inequalities corresponding to the subset $S' = S \cup \{i^*\}$ then this inequality is added to the LP program, S is updated as $S = S'$ and the procedure is repeated until S contains all the customers in V_c.

The initial family $\overline{\mathcal{S}}$ is usually built by randomly generating a certain number of customer subsets.

Fukasawa et al. (2006) implemented exact separation procedures for the RC inequalities based on mixed-integer models. They observed that the heuristic procedures typically decrease the separation times by two orders of magnitude without any significant loss in terms of bound quality. Ralphs et al. (2003) described a decomposition-based separation methodology for the RC inequalities that takes advantage of the ability to solve small instances of the TSP efficiently. Specifically, when standard separation heuristic procedures fail to separate a candidate point, they attempt to decompose it into a convex combination of TSP tours; otherwise the tours present in this decomposition are examined for violated capacity constraints; if not, the Farkas' lemma provides a hyperplane separating the point from the TSP polytope.

Letchford et al. (2002) described separation heuristic algorithms for the multistar and partial multistar inequalities, while separation algorithms for the other classes of inequalities, including the comb, FrC and HI inequalities, have been described by Augerat et al. (1995) and Lysgaard et al. (2004).

The complete set of the separation routines used by Lysgaard et al. (2004) can be found in the software package CVRPSEP (see Lysgaard 2003), which is publicly available.

5.2 Branching strategies

Two branching strategies have been used by the different branch-and-cut algorithms for the CVRP proposed in the literature: the *branching on sets* strategy and the *branching on variables* strategy.

The branching on sets strategy has been introduced by Augerat et al. (1995). This strategy consists of choosing a set S, $S \subseteq V_c$, $|S| > 2$, for which $0 < p(S) < 2$, where $p(S) = \sum_{\{i,j\}\in\delta(S)} x_{ij} - 2\lceil q(S)/Q\rceil$, and creating two subproblems: one by adding the constraint $\sum_{\{i,j\}\in\delta(S)} x_{ij} = 2\lceil q(S)/Q\rceil$ and the other by adding the constraint $\sum_{\{i,j\}\in\delta(S)} x_{ij} \geq 2\lceil q(S)/Q\rceil + 2$.

The selection of the subset S is usually carried out in two steps: first, a candidate list of subsets is built heuristically and, second, one subset is selected from this list according to some criterion. Some of the strategies used to define the candidate list are the following:

– Select set S with maximum demand.
– Select set S which is farthest from the depot.
– Select set S such that $\sum_{\{i,j\}\in\delta(S)} x_{ij}$ is as close as possible to 3.
– Select set S such that $\sum_{\{i,j\}\in\delta(S)} x_{ij}$ is as close as possible to 2.75.

The set S to branch on is then selected by *LP testing* (see Applegate et al. 2006), i.e., by solving both linear programs induced by the two possible branches. The best set of the candidate list is then chosen as the one for which the minimum of the two objective function values is the largest.

When a suitable set S cannot be found, the branching on variables strategy is adopted. This involves the selection of an edge $\{i, j\}$ having a fractional value of x_{ij}, and the generation of two subproblems: one by fixing $x_{ij} = 1$ and the other by fixing $x_{ij} = 0$. The edge $\{i, j\}$ is selected in such a way that x_{ij} is as close as possible to 0.5. Ties are broken by choosing the edge $\{i, j\}$ having maximum cost d_{ij}.

6 Branch-and-cut-and-price methods

Fukasawa et al. (2006) suggested an approach to combine both two-index and set partitioning formulations using equations (50) linking the two-index variables x with the set partitioning variables ξ. Let a q-route (see Christofides et al. 1981) be a not necessarily simple cycle covering the depot and a subset of customers, whose total demand is equal to q. Note that the set of q-routes contains the set of feasible CVRP routes. In the formulation considered by Fukasawa et al. (2006), the variables correspond to the set of q-routes, while the set of constraints is defined by the set partitioning formulation constraints plus feasible CVRP inequalities designed for the two-index formulation. In addition to the RC inequalities, they also used the FrC, strengthened comb, multistar, partial multistar, GLM and HI inequalities, all presented in Lysgaard et al. (2004). Since the resulting formulation has an exponential number of both columns and rows, column and cut generation procedures for computing the lower bound and a branch-and-cut-and-price algorithm for solving the CVRP were proposed.

More precisely, Fukasawa et al. (2006) considered the following relaxation of the CVRP:

$$\min \sum_{\ell \in \widehat{\mathcal{R}}} c_\ell \xi_\ell \tag{63}$$

$$\text{s.t.} \quad \sum_{\ell \in \widehat{\mathcal{R}}} a_{i\ell} \xi_\ell = 1 \quad (\forall i \in V_c), \tag{64}$$

$$\sum_{\ell \in \widehat{\mathcal{R}}} \xi_\ell = m, \tag{65}$$

$$\sum_{\ell \in \widehat{\mathcal{R}}} \alpha^t(R_\ell) \xi_\ell \geq \beta^t \quad (\forall t \in \mathcal{F}), \tag{66}$$

$$\xi_\ell \geq 0 \quad (\forall \ell \in \widehat{\mathcal{R}}), \tag{67}$$

where $\alpha^t(R_\ell) = \sum_{\{i,j\} \in E} \alpha^t_{ij} \eta^\ell_{ij}$ and the set of routes $\widehat{\mathcal{R}}$ corresponds to the set of q-routes. Note that in this case, the coefficient η^ℓ_{ij} is a general integer coefficient that is equal to the number of times edge $\{i, j\}$ is traversed by q-route ℓ. In the formulation, the set $\mathcal{F}$ represents the family of the different valid inequalities designed for the two-index formulation (i.e., the FrC, strengthened comb, multistar, partial multistar, GLM and HI inequalities).

Fukasawa et al. (2006) used a pricing and cut generation technique to solve the LP program above. At first, the LP program (i.e., *master problem*) includes only the degree constraints (64) and (65) and a restricted set of q-routes. The resulting lower bound has been integrated by Fukasawa et al. in an enumeration scheme to solve the CVRP to optimality. The enumeration scheme is based on the branching on sets strategy described in Sect. 5.2.

The cut generation is performed by converting a solution ξ of the LP program into a corresponding solution x using (50). If this solution is fractional, it is given as input to the CVRPSEP package in order to separate the different classes of valid inequalities. The violated inequalities found are then translated back into the ξ variables to be introduced in the LP program.

The pricing subproblem consists of finding q-routes with minimum reduced cost. Although this problem is $\mathcal{NP}$-hard, it can be solved in pseudo-polynomial time as follows. Let $u = (u_0, u_1, \ldots, u_n)$ be a vector of dual variables, where $u_i, i \in V_c$, and u_0 are associated with constraints (64) and (65), respectively. Moreover, let w_t, $t \in \mathcal{F}$, be the dual variables of constraints (66).

Given a dual solution $(\widehat{u}, \widehat{w})$, the reduced cost of any q-route $\ell \in \widehat{\mathcal{R}}$ is equal to:

$$\overline{c}_\ell = c_\ell - \sum_{i \in V} a_{i\ell} \widehat{u}_i - \sum_{t \in \mathcal{F}} \alpha^t(R_\ell) \widehat{w}_t. \tag{68}$$

Define the modified cost $\overline{d}_{ij}$ of edge $\{i, j\}$ with respect to the dual solution $(\widehat{u}, \widehat{w})$ as follows:

$$\overline{d}_{ij} = d_{ij} - \frac{1}{2}\widehat{u}_i - \frac{1}{2}\widehat{u}_j - \sum_{t \in \mathcal{F}} \alpha^t_{ij} \widehat{w}_t, \quad \forall \{i, j\} \in E. \tag{69}$$

It is easy to see that for a given $\ell \in \widehat{\mathcal{R}}$, $\overline{c}_\ell = \sum_{\{i,j\} \in E} \eta^\ell_{ij} \overline{d}_{ij}$. Then, the q-route having the most negative reduced cost can be computed as follows.

Let $f(q, i)$ be the cost of the least cost path, not necessarily simple, $P = (0, i_1, \ldots, i_k)$, with $i_k = i$, from the depot 0 to customer i with total load $q = \sum_{h=1}^{k} q_{i_j}$. Such a path is called q-path. A q-path with the additional edge $\{0, i\}$ is called q-route and has cost $f(q, i) + d_{0i}$. The q-path functions, with the additional restriction imposing that the path should not contain loops formed by three consecutive vertices (see Christofides et al. 1981), can be computed in pseudo-polynomial time with a complexity of $O(n^2Q)$. Given a dual solution $(\widehat{u}, \widehat{w})$, the q-route having the most negative reduced cost (if any) can be computed as follows:

(i) compute q-path functions $f(.)$ using the modified edge cost $\overline{d}$;
(ii) let δ_i, $i \in V_c$, be computed as $\delta(i) = \min_{q_i \le q \le Q}\{f(q, i) + \overline{d}_{0i}\}$ and let $\overline{q}_i$ be the value of q corresponding to the minimum;
(iii) compute $i^* = \text{argmin}_{i \in V_c}\{\delta(i)\}$.

Then, the most negative q-route (if any) has cost $f(\overline{q}_i, i^*) + \overline{d}_{0i^*}$.

As usual in column generation, Fukasawa et al. (2006) found to be computationally convenient to add at each iteration more than one q-route having negative reduced cost. This operation has been embedded in the dynamic programming procedure used to compute q-path functions $f(.)$. Additional features of the branch-and-cut-and-price algorithm of Fukasawa et al. (2006) are:

1. *Cycle elimination*. To strengthen the formulation, s-cycle-free q-route (with s up to 4) have been computed.
2. *Heuristic acceleration*. In order to speed up the lower bound computation, different heuristic procedures are used to produce negative reduced cost q-routes. When no negative q-route is found using the heuristics, the full dynamic programming algorithm is run to check if none exists.
3. *Dynamic selection of column generation*. The master problem is usually highly degenerate and degeneracy implies alternative optimal dual solutions. Consequently, the generation of new columns and their associated variables may not change the value of the objective function of the master problem, the master problem may become large, and the overall method may become slow computationally. Moreover, in some CVRP instances, the increase in the lower bound value with respect to the one achieved by the pure branch-and-cut method is very small and is not worth the computing time required by the additional column generation approach. Thus, the exact algorithm presented by Fukasawa et al. (2006) decides at the root node, according to the best balance between running time and bound quality, either to use the pure branch-and-cut method or the branch-and-cut-and-price strategy.

Choi and Tcha (2007) proposed lower bounds for the FSMF based on the set partitioning formulation, that were computed by using q-route relaxation and column generation techniques. These authors also described lower bounds for the FSMFD and the FSMD.

An exact method for the FSMF, FSMFD and FSMD variants was recently proposed by Pessoa et al. (2009, to appear). These authors extend to the FSMF, FSMFD and FSMD variants the branch-and-cut-and-price method proposed for the CVRP by Fukasawa et al. (2006) and described above.

7 Set partitioning with additional cuts

Baldacci et al. (2008b) proposed an exact method based on the following set partitioning formulation with additional cuts:

$$\min \sum_{\ell \in \mathcal{R}} c_\ell \xi_\ell \tag{70}$$

$$\text{s.t.} \quad \sum_{\ell \in \mathcal{R}} a_{i\ell}\xi_\ell = 1 \quad (\forall i \in V_c), \tag{71}$$

$$\sum_{\ell \in \mathcal{R}} \xi_\ell = m, \tag{72}$$

$$\sum_{\ell \in \mathcal{R}(S)} \xi_\ell \geq \lceil q(S)/Q \rceil \quad (\forall S \in \mathcal{S}), \tag{73}$$

$$\sum_{\ell \in \mathcal{R}} \alpha^t(R_\ell)\xi_\ell \geq \beta^t \quad (\forall t \in \mathcal{F}), \tag{74}$$

$$\sum_{\ell \in C} \xi_\ell \leq 1 \quad (\forall C \in \mathcal{C}), \tag{75}$$

$$\xi_\ell \in \{0, 1\} \quad (\forall \ell \in \mathcal{R}), \tag{76}$$

where $\alpha^t(R_\ell) = \sum_{\{i,j\} \in E} \alpha^t_{ij} \eta^\ell_{ij}$ and again the set $\mathcal{F}$ represents the family of the different valid inequalities designed for the two-index formulation. Constraints (73) and (75) correspond to the strengthened capacity constraints and clique inequalities, respectively, described in Sect. 4.4.

The method proposed by Baldacci et al. (2008b) is based on a bounding procedure that computes a lower bound on the CVRP by finding a near-optimal solution of the dual of the LP relaxation of the integer problem above. Let $u = (u_0, u_1, \ldots, u_n)$ be a vector of dual variables, where u_i, $i \in V_c$, and u_0 are associated with constraints (71) and (72), respectively. Moreover, let v_S, $S \in \mathcal{S}$, w_t, $t \in \mathcal{F}$, and g_C, $C \in \mathcal{C}$, be the dual variables of constraints (73), (74) and (75), respectively. The dual problem is as follows:

$$\max \sum_{i \in V_c} u_i + m u_0 + \sum_{S \in \mathcal{S}} \lceil q(S)/Q \rceil v_S + \sum_{t \in \mathcal{F}} \beta^t w_t + \sum_{C \in \mathcal{C}} g_C \tag{77}$$

$$\text{s.t.} \quad \sum_{i \in V} a_{i\ell} u_i + \sum_{S \in \mathcal{S}} b_\ell(S) v_S$$

$$+ \sum_{t \in \mathcal{F}} \alpha^t(R_\ell) w_t + \sum_{C \in \mathcal{C}_\ell} g_C \leq c_\ell \quad (\forall \ell \in \mathcal{R}), \tag{78}$$

$$u_i \in \mathbb{R} \quad (\forall i \in V), \tag{79}$$

$$v_S \geq 0 \quad (\forall S \in \mathcal{S}), \tag{80}$$

$$w_t \geq 0 \quad (\forall t \in \mathcal{F}), \tag{81}$$

$$g_C \leq 0 \quad (\forall C \in \mathcal{C}), \tag{82}$$

where $\mathcal{C}_\ell = \{C \in \mathcal{C} : \ell \in C\}$ and the coefficient $b_\ell(S)$ is equal to 1, $\forall \ell \in \mathcal{R}$, such that $R_\ell \cap S \neq \emptyset$ and $b_\ell(S) = 0$ otherwise.

Given a near optimal dual solution (u', v', w', g') of value z', the exact method of Baldacci et al. (2008b) consists of finding, by means of the integer programming solver CPLEX (2006), an optimal integer solution of the set partitioning formulation with additional cuts resulting from the following reductions:

1. the route set $\mathcal{R}$ is replaced with the subset $\mathcal{R}' \subset \mathcal{R}$ containing all the routes whose reduced cost, with respect to the dual variables (u', v', w', g'), is smaller than the gap $z_{UB} - z'$, where z_{UB} is a valid upper bound on the CVRP;
2. the set of constraints (73), (74) and (75) is replaced by the set of active constraints which are generated during the lower bound computation.

The effectiveness of the method is based on having an efficient procedure for generating a near optimal dual solution (u', v', w', g'). As the dual solution gets better, the reduced costs of the routes of an optimal CVRP solution get smaller and, hopefully, the size of subset $\mathcal{R}'$ that must be generated to find an optimal solution gets smaller.

Baldacci and Mingozzi (2009) presented an exact algorithm for the HVRPFD that generalizes the bounding procedures and the exact method for the CVRP described above. They introduced new bounding methods that are particularly effective when the vehicle fixed cost contribution to the total cost is relevant. It is worth noting that the exact algorithm proposed for the HVRPFD by Baldacci and Mingozzi is able to solve all the heterogeneous VRP variants listed in Table 1.

Below we briefly review two of the key ingredients of the method used by Baldacci et al. (2008b) and by Baldacci and Mingozzi (2009): the procedure used for generating a near optimal dual solution and the method used to generate feasible CVRP routes.

7.1 Bounding procedure

The bounding procedure used to compute a near-optimal solution of the dual problem is an additive bounding method that computes a lower bound on the *CVRP* as the sum of the solution costs obtained by three heuristics, called H^1, H^2 and H^3 for solving the dual problem. The procedure is based on the additive method of Fischetti and Toth (1989) for combinatorial optimization problems and it extends the bounding procedure proposed by Baldacci et al. (2006) for the asymmetric VRP on a multi-graph. The three heuristic procedures are used in sequence and do not require the a priori generation of the entire route set $\mathcal{R}$.

Both Procedure H^1 and H^2 consider capacity inequalities (73) and are based on the following theorem.

Theorem 1 (Baldacci et al. 2008b) *Associate penalties* $\lambda_i \in \mathbb{R}$, $i \in V_c$, *with constraints* (71), $\lambda_0 \in \mathbb{R}$ *with constraint* (72) *and* $\sigma_S \geq 0$, $S \in \mathcal{S}$, *with constraints* (73). *Define*

$$b_i = q_i \min_{\ell \in \mathcal{R}_i} \left\{ \frac{c_\ell - \lambda(R_\ell) - \sigma(R_\ell)}{\sum_{i \in V_c} a_{i\ell} q_i} \right\}, \quad \forall i \in V_c, \tag{83}$$

where $\mathcal{R}_i \subset \mathcal{R}$ *is the index subset of the routes covering customer* $i \in V_c$, $\lambda(R_\ell) = \sum_{i \in V_c} a_{i\ell}\lambda_i$ *and* $\sigma(R_\ell) = \sum_{S \in \mathcal{S}} b_\ell(S)\sigma_S$.

A feasible solution (u, v, w, g) *of the dual problem of cost* $z(\lambda, \sigma)$ *is given by setting* $w = 0$, $g = 0$ *and by computing* u *and* v *according to the following expressions*:

$$u_i = b_i + \lambda_i, \quad i \in V_c, \quad u_0 = \lambda_0 \quad \textit{and} \quad v_S = \sigma_S, \quad S \in \mathcal{S}. \tag{84}$$

Note that expressions (83) cannot be used directly as they involve the entire route set $\mathcal{R}$.

Baldacci et al. (2008b) used procedures H^1 and H^2. In particular, H^1 is an extension of the bounding method proposed by Christofides et al. (1981) and it is based on the q-route relaxation of the CVRP route. H^2 is a column and cut generation procedure that considers feasible CVRP routes, but solves the master problem using Lagrangean relaxation. Both H^1 and H^2 use subgradient optimization procedures to solve $\max_{\lambda,\mu}\{z(\lambda, \mu)\}$.

The last procedure, procedure H^3, is a column and cut generation procedure based on the simplex method and that considers both capacity and clique inequalities. It is worth mentioning that Baldacci et al. (2008b) found it to be computationally convenient to consider capacity and clique inequalities only, as the improvement to the lower bound value given by the valid inequalities designed for the two-index formulation is not worth the extra computing time required.

In procedure H^3, the capacity constraints (73) are heuristically separated by converting a solution ξ of the master problem into a corresponding solution x using equations (50). If this solution is fractional, it is given as input to the CVRPSEP package in order to find any subset S whose RC inequality (3) is violated by x. Then, for any such S, the corresponding constraint (73) is added to the master problem.

The separation problem associated with the clique inequalities (75) is strongly $\mathcal{NP}$-hard as it corresponds to find the maximal weighted clique of the *conflict graph* associated with a given fractional solution (see Garey and Johnson 1990). Nevertheless, Baldacci et al. (2008b) found it to be computationally convenient to separate clique inequalities using the CLIQUER 1.1 package Niskanen and Östergård (2003), which is composed of a set of C routines for finding cliques in an arbitrary weighted graph based on the exact branch-and-bound algorithm developed by Östergård (2002).

Baldacci and Mingozzi (2009) proposed a new integer relaxation of the HVRPFD, called *RP*. The relaxation can provide a better lower bound than the linear programming relaxation of the set partitioning formulation of the HVRPFD for those HVRPFDs where the vehicle fixed cost contribution to the optimal cost is relevant or dominates the routing cost contribution.

RP involves two types of integer variables: $\xi_{ik} \in \{0,1\}$, $i \in V_c$, $k \in M$ and $y_k \in \mathbb{Z}^+$, $k \in M$. Variable ξ_{ik} is equal to 1 if and only if customer $i \in V_c$ is served by a vehicle of type $k \in M$. Variable y_k represents the number of vehicles of type k used in the solution.

Let β_{ik} be the *marginal routing cost* for covering customer $i \in V_c$ with a vehicle of type $k \in M$. We assume that the values β_{ik}, $i \in V_c$, $k \in M$, satisfy the following inequalities:

$$\sum_{i \in R_\ell^k} \beta_{ik} \le c_\ell^k, \quad \forall \ell \in \mathcal{R}^k,\ \forall k \in M. \tag{85}$$

It can be shown that the following integer problem, called *RP*, provides a valid lower bound on the HVRPFD for any solution β_{ik} of inequalities (85).

$$\min \sum_{k\in M}\sum_{i\in V_c} \beta_{ik}\xi_{ik} + \sum_{k\in M} F_k y_k \tag{86}$$

$$\text{s.t.} \quad \sum_{k\in M}\sum_{i\in V_c} q_i \xi_{ik} = q(V_c), \tag{87}$$

$$\sum_{i\in V_c} q_i\xi_{ik} \le Q_k y_k \quad (\forall k \in M), \tag{88}$$

$$y_k \le m_k \quad (\forall k \in M), \tag{89}$$

$$\xi_{ik} \in \{0,1\} \quad (\forall i \in V_c, \forall k \in M), \tag{90}$$

$$y_k \in \mathbb{Z}^+ \quad (\forall k \in M). \tag{91}$$

Relaxation *RP* is used by two bounding procedures, called DP^1 and DP^2, that correspond to two different methods for computing β_{ik} satisfying inequalities (85). DP^1 uses q-route relaxation while DP^2 uses column generation. Both procedures are based on a dual ascent procedure and solve problem *RP* by dynamic programming.

7.2 Route generation algorithm

The exact method of Baldacci et al. (2008b) and Baldacci and Mingozzi (2009) use a dynamic programming procedure, called GENROUTE, to generate feasible CVRP routes. Procedure GENROUTE is based on the observation that any route R can be decomposed, for every $i \in R$, into two paths P_i and $\overline{P}_i$ going from depot 0 to vertex i. These two paths are internally disjoint and their total customer demand is less than or equal to Q. GENROUTE consists of two phases: in the first phase, it generates the set of simple paths ending at vertices in V_c, while, in the second phase, the paths are combined to generate feasible CVRP routes.

Consider the reduced cost $\widehat{c}_\ell$ of route ℓ with respect to the four dual vectors $\widehat{u}$, $\widehat{v}$, $\widehat{w}$ and $\widehat{g}$ and the sets $\overline{\mathcal{S}} \subset \mathcal{S}$, $\overline{\mathcal{F}} \subset \mathcal{F}$ and $\overline{\mathcal{C}} \subset \mathcal{C}$:

$$\widehat{c}_\ell = c_\ell - \sum_{i \in V} a_{i\ell}\widehat{u}_i - \sum_{S \in \overline{\mathcal{S}}} b_\ell(S)\widehat{v}_S - \sum_{t \in \overline{\mathcal{F}}} \alpha^t(R_\ell)\widehat{w}_t - \sum_{C \in \overline{\mathcal{C}}_\ell} \widehat{g}_C, \tag{92}$$

where $\overline{\mathcal{C}}_\ell = \{C \in \overline{\mathcal{C}} : C \ni \ell\}$.

Given a dual solution $(\widehat{u}, \widehat{v}, \widehat{w}, \widehat{g})$, sets $\overline{\mathcal{S}} \subset \mathcal{S}$, $\overline{\mathcal{F}} \subset \mathcal{F}$ and $\overline{\mathcal{C}} \subset \mathcal{C}$ of valid inequalities and two user-defined parameters γ and Δ, GENROUTE produces the largest subset $\mathcal{B}$ of the route set $\mathcal{R}$ satisfying the following conditions:

$$\left.\begin{array}{l} \max_{\ell \in \mathcal{B}}\{\widehat{c}_\ell\} \le \min_{\ell \in \mathcal{R}\setminus\mathcal{B}}\{\widehat{c}_\ell\}, \\ |\mathcal{B}| \le \Delta, \\ \max_{\ell \in \mathcal{B}}\{\widehat{c}_\ell\} \le \gamma. \end{array}\right\} \tag{93}$$

The parameters γ and Δ allow GENROUTE to generate the route subsets required by procedures H^2, H^3 and by the exact method for generating the final route $\mathcal{R}' \subset \mathcal{R}$.

8 Computational experiments

In this section we report a summary of the computational results for the CVRP and for the different variants of the heterogeneous VRP obtained by the most effective exact methods proposed in the literature.

8.1 Computational experiments for the CVRP

Both analytical and computational results on the CVRP can be found in Augerat et al. (1995), Naddef and Rinaldi (2002), Ralphs et al. (2003), Baldacci et al. (2004, 2008b) Letchford et al. (2002), Lysgaard et al. (2004) and Fukasawa et al. (2006). However, extensive computational results over a common set of instances taken from the literature have been performed only by the last three of these exact methods. Moreover, according to the results reported, the last three exact methods represent the most effective exact methods currently available for the CVRP, both for the quality of the lower bounds produced and for the number of instances solved to optimality. Thus, in this section, we report a comparison of the results obtained by three exact algorithms of Lysgaard et al. (2004), Fukasawa et al. (2006) and Baldacci et al. (2008b).

Six classes of test instances called A, B, E, M, P and F are usually adopted to perform computational results on the exact algorithms for the CVRP. Classes A, B and P were proposed by Augerat (1995). Instance class M was proposed by Christofides et al. (1979) while classes E and F were produced by Christofides and Eilon (1969) and Fisher (1994), respectively. The data of all instances including the best upper bounds known and the solutions of the instances solved to optimality can be found at the URL address http://branchandcut.org/VRP/data. For all the instances considered, the computational results presented in the literature use integer-valued distances obtained by rounding to the nearest integer the Euclidean distance between each pair of vertices. More precisely, the cost of edge $\{i, j\}$ is set to $d_{ij} = \lfloor e_{ij} + 0.5 \rfloor$, where e_{ij} is the corresponding Euclidean distance between vertices i and j.

Classes A, B, E and P contain different instances with number of customers between 12 and 100 and different number of vehicles (up to 14 vehicles), whereas class M contains instances with number of customers between 100 and 199 with up to 17 vehicles. Class F contains three instances, with 44 customers and 4 vehicles, 71 customers and 4 vehicles, and 134 customers to serve with 7 vehicles.

Table 2 reports a summary of the computational results over the common set of instances of the five classes A, B, E, M and P considered by Lysgaard et al. (2004), Fukasawa et al. (2006) and Baldacci et al. (2008b).

Column *#Inst* of the table reports the total number of instances in the corresponding class. For each exact method and for each class, the table reports the following data:

#Opt: number of instances solved to optimality by the method;
%LB: average percentage ratio of the lower bound with respect to the optimal solution value;
t_{LB}: average running time in seconds for computing the lower bound;
t^1: average running time in seconds of the exact method.

For the exact method of Fukasawa et al. (2006), column *#B&C* reports the number of instances, for the corresponding class, for which column generation was not used. In this case, the exact algorithm of Fukasawa et al. (2006) becomes the exact branch-and-cut algorithm of Lysgaard et al. (2004).

The averages of the running times of the exact methods (column t^1) are computed over all the instances solved to optimality by all methods. For the exact methods of Fukasawa et al. (2006) and Baldacci et al. (2008b), column t^2 reports the averages of the running times (in seconds) of the two exact methods over all the instances solved to optimality by both methods. The last lines of the table report summation and averages over all the classes.

The computational results of Lysgaard et al. (2004) have been performed on a machine equipped with an Intel Celeron running at 700 MHz, while Baldacci et al. (2008b) and Fukasawa et al. (2006) have performed their computational results on machines equipped with Pentium 4 processors running at 2.6 GHz and 2.4 Ghz, respectively. According to computational benchmarks of the different machines mentioned above, the machine used by Baldacci et al. (2008b) is about ten percent faster and at least five times faster than the machines used by Fukasawa et al. (2006) and by Lysgaard et al. (2004), respectively. A time limit of eight hours was imposed by Lysgaard et al. (2004) to their branch-and-cut algorithm except for instance E-n76-k7 that was solved to optimality in about 33 hours. No time limits were imposed to the exact methods of Fukasawa et al. (2006) and Baldacci et al. (2008b). In particular, the method of Baldacci et al. (2008b) terminates without finding the optimal solution whenever either CPLEX or procedure GENROUTE (see Sect. 7.2) run out of memory.

Table 2 indicates that, on the set of instances considered, the lower bound produced by Baldacci et al. (2008b) is on average superior to the other lower bounds in all the classes.

Table 2 Summary of the computational experiments for the CVRP

Class	*#Inst*	Lysgaard et al. (2004)[a]				Fukasawa et al. (2006)[b]						Baldacci et al. (2008b)[c]				
		#Opt	*%LB*	t_{LB}	t^1	*#Opt*	*#B&C*	*%LB*	t_{LB}	t^1	t^2	*#Opt*	*%LB*	t_{LB}	t^1	t^2
A	22	15	97.9	60.0	6637.9	22	2	99.2	182.7	402.7	1961.2	22	99.8	40.9	35.4	118.1
B	20	16	99.4	28.6	783.0	20	14	99.5	83.8	134.3	4763.0	20	99.8	119.2	66.0	422.1
E-M	9	3	97.7	262.2	39591.7	9	2	99.0	956.6	15568.0	42614.5	8	99.6	192.6	1143.4	1025.1
P	23	16	97.3	38.3	12019.0	23	8	99.2	139.7	430.6	2466.3	21	99.7	34.5	180.1	191.8
Sum.	74	50				74	26					71				
Avg.			98.1	97.3	14757.9			99.2	340.7	4133.9	12951.3		99.7	96.8	356.2	439.3

[a]Computing time in seconds of an Intel Celeron 700 MHz

[b]Computing time in seconds of a Pentium 4 2.4 GHz

[c]Computing time in seconds of a Pentium 4 2.6 GHz

The method of Fukasawa et al. (2006) solved to optimality all the 74 instances, 26 of them were solved without using the column generator. Taking the computers used by the different authors into account in examining the computational results, Table 2 indicates that on the instances solved to optimality by all the methods, the algorithm of Baldacci et al. (2008b) is on average faster than both the algorithms of Lysgaard et al. (2004) and of Fukasawa et al. (2006), while the method of Fukasawa et al. compares favorably with that of Lysgaard et al. In addition, as shown by columns t^2, the algorithm of Baldacci et al. is on average faster than the method of Fukasawa et al. on the set of 71 instances solved to optimality by both methods.

Table 3 reports detailed results on instance classes E and M, which represent difficult CVRP instances. In the table, column labelled "z^*" reports the cost of the optimal solution of the corresponding instance. For each exact method and for each instance (whose name includes the number of vertices and the number of vehicles), the table reports the following data:

- $\%LB$: percentage ratio of the lower bound with respect to the optimal solution value;
- t_{LB}: running time in seconds for computing the lower bound;
- t_E: running time in seconds of the exact method; for the exact method of Lysgaard et al. (2004) "–" denotes that the time limit has been reached while for the exact method of Baldacci et al. (2008b) "–" denotes that the memory limit has been reached.

In addition, column labelled "s" reports the size of the cycles eliminated by the column generation procedure of Fukasawa et al. (2006) ("–" indicates that the column generation was not used).

Three instances, namely E-n101-k8, P-n76-k5 and P-n101-k4, cannot be solved to optimality by the exact method of Baldacci et al. (2008b). Table 3 shows that instance E-n101-k8 have been solved to optimality by the method of Fukasawa et al. (2006) in 801 963.0 seconds. Instances P-n76-k5 and P-n101-k4 have been solved to optimality by Lysgaard et al. (2004) and by Fukasawa et al. (2006) but without using column generation.

The results on these three instances, that are loosely constrained, testify the fact that set partitioning based methods may not work well on loosely constrained instances (i.e., instances where n is large and m is small), since the number of promising routes can be huge in such cases. On the other hand, as shown by Table 3, difficult CVRP instances, such as instance E-n76-k10, which cannot be solved to optimality by the branch-and-cut algorithm of Lysgaard et al., can be solved to optimality by both the algorithms of Fukasawa et al. and Baldacci et al. in 80 722.0 and in 174.4 seconds, respectively. Thus, set partitioning based methods complement branch-and-cut approaches, which tend to work better on loosely constrained instances.

Lysgaard et al. (2004) and Fukasawa et al. (2006) considered and solved to optimality three instances of the class F (namely, F-n45-k4, F-n72-k4 and F-n135-k7). These instances have not been considered by Baldacci et al. (2008b) due to the complexity involved in the computation of the q-path functions as they are characterized by large vehicle capacities. Indeed, Fukasawa et al. (2006) solved the instances without using column generation. In particular, instance F-n135-k7 has been solved to optimality by Lysgaard et al. (2004) in 3 092.0 seconds and by Fukasawa et al. (2006) in 7 065.0 seconds. Finally, it is worth mentioning that instance F-n135-k7 was solved to optimality also by Augerat et al. (1995) in 18 871.0 seconds on a Sun Sparc 10, and by Baldacci et al. (2004) in 6 599.0 seconds on a machine equipped with a Pentium III 933 MHz processor. To date, instance F-n135-k7 is the largest non-trivial CVRP instance solved to optimality in the literature.

Table 3 Computational experiments on classes E and M of the CVRP

Name	z^*	Lysgaard et al. (2004)[a]			Fukasawa et al. (2006)[b]				Baldacci et al. (2008b)[c]		
		%*LB*	t_{LB}	t_E	%*LB*	t_{LB}	*s*	t_E	%*LB*	t_{LB}	t_E
E-n51-k5	521	99.6	24.0	59.0	99.5	51.0	–	65.0	100.0	12.6	12.6
E-n76-k7	682	97.7	72.0	118683.0	98.2	264.0	2	46520.0	99.0	146.4	3370.6
E-n76-k8	735	97.7	136.0	–	98.8	277.0	2	22891.0	99.3	103.5	873.2
E-n76-k10	830	96.4	158.0	–	98.5	354.0	3	80722.0	99.5	59.6	174.4
E-n76-k14	1021	95.0	181.0	–	98.6	224.0	3	48637.0	99.6	16.8	44.9
E-n101-k8	815	98.5	222.0	–	98.8	1068.0	3	801963.0	99.0	249.9	–
E-n101-k14	1067	96.2	555.0	–	98.8	658.0	3	116284.0	99.7	153.5	1230.0
M-n101-k10	820	100.0	33.0	33.0	100.0	119.0	–	119.0	100.0	47.0	47.0
M-n121-k7	1034	98.4	979.0	–	99.7	5594.0	3	25678.0	99.8	943.7	2448.3

[a] Computing time in seconds of an Intel Celeron 700 MHz

[b] Computing time in seconds of a Pentium 4 2.4 GHz

[c] Computing time in seconds of a Pentium 4 2.6 GHz

8.2 Computational experiments for the heterogeneous VRPs

In this section, we report a summary of the computational results obtained by the exact methods of Baldacci and Mingozzi (2009) and of Pessoa et al. (2009, to appear) on the main sets of instances from the literature. The instances correspond to the different variants of the heterogeneous VRP listed in Table 1. The complete details of the instances can be found in Baldacci and Mingozzi (2009).

Pessoa et al. considered in their computational experiments only the FSMF, FSMFD and FSMD variants, while Baldacci and Mingozzi considered the whole set of instances corresponding to the problems listed in Table 1.

It is worth noting that computational results about lower bounds for the FSMF, FSMFD and FSMD can also be found in Choi and Tcha (2007), Yaman (2006) and Baldacci et al. (2009, to appear). The computational results show that the lower bounds obtained by these authors are dominated by the lower bounds obtained by Baldacci and Mingozzi and Pessoa et al.

The algorithm described in Baldacci and Mingozzi (2009) was run on a personal computer with an AMD Athlon 64 X2 Dual Core 4200+ processor at 2.6 GHz, while the algorithm described in Pessoa et al. (2009, to appear) was run on a personal computer with a Pentium Core 2 Duo at 2.13 GHz.

Table 4 summarizes the results obtained over all the variants considered by the two methods. This table reports the following columns: the total number of instances in each class (#*Inst*), the number of instances solved to optimality by each method (#*Opt*), the average percentage ratio of the lower bound with respect to the optimal solution value (%*LB*), the average running time in seconds for computing the lower bound (t_{LB}), the average running time in seconds of the exact method computed over all the instances solved to optimality by all methods (t_{TOT}). In the table, the two classes MDVRP1 and MDVRP2 correspond to the two sets of MDVRP instances proposed by Cordeau et al. (1997) and by Baldacci and Mingozzi (2009), respectively.

Table 4 shows that the exact method of Pessoa et al. (2009, to appear) solved to optimality 29 out of 36 instances considered. Five more instances were solved to optimality by the exact method of Baldacci and Mingozzi (2009) with respect to the instances solved by the method of Pessoa et al. Overall, the exact method of Baldacci and Mingozzi (2009) was able to solve to optimality 74 out of 86 instances considered. The table shows that the lower bound computed by Baldacci and Mingozzi on average dominates the lower bound computed by Pessoa et al.

9 Conclusions

In the last few years some innovative exact approaches for vehicle routing problems under capacity constraints were proposed, producing a significant improvement on the size of the instances that can be solved to optimality. Indeed, these algorithms have brought above one hundred the number of customers that may be handled, thus doubling this limit with respect to the best available methods.

The key factor of the success of these approaches is the effective combination of set partitioning and two-index vehicle flow formulations of the problem. In particular, the inclusion of known families of cuts into column generation based algorithms, significantly improved the quality of the resulting lower bounds that are now very close to the optimal solution values.

Table 4 Summary of the computational experiments for the heterogeneous VRPs

Variant	*#Inst*	Lower bounds				Exact methods			
		Pessoa et al. (2009, to appear)[a]		Baldacci and Mingozzi (2009)[b]		Pessoa et al. (2009, to appear)		Baldacci and Mingozzi (2009)	
		%LB	t_{LB}	*%LB*	t_{LB}	*#Opt*	t_{TOT}	*#Opt*	t_{TOT}
HVRPFD	12	–	–	99.6	224.8	–	–	10	259.9
HVRPD	8	–	–	99.2	128.9	–	–	7	564.8
SDVRP	13	–	–	99.1	183.9	–	–	9	880.6
FSMFD	12	99.7	243.8	99.7	143.5	10	963.46	11	172.9
FSMD	12	99.2	330.5	99.5	81.9	10	2309.0	12	281.1
MDVRP1	9	–	–	99.2	310.5	–	–	7	875.3
MDVRP2	8	–	–	99.7	189.5	–	–	7	4788.6
FSMF	12	99.6	229.5	99.8	147.0	9	4741.2	11	125.4
Sum.	86					(29)	74(34)		

[a]Computing time in seconds of a Pentium Core 2 Duo at 2.13 GHz

[b]Computing time in seconds of an AMD Athlon 64 X2 Dual Core 4200+ at 2.6 GHz

Nevertheless, some space for further improvement remains since pure branch-and-cut approaches are still the better ones on loosely constrained instances, i.e., where the average number of customers per vehicle is larger than 30.

On the other hand the set partitioning based approaches proved quite general. In fact, they are able to easily incorporate additional characteristics of the problem arising in practical applications, such as time windows, precedence constraints and a heterogeneous fleet. There is therefore a large room for research activities in this specific field whose achievement may also give important insight in the heuristic solution of real-world vehicle routing problems.

Acknowledgement Work partially supported by Italian MiUR under PRIN research project.

References

Applegate, D. L., Bixby, R. E., Chvátal, V., & Cook, W. J. (2006). *The traveling salesman problem: A computational study*. Princeton: Princeton University Press.

Araque, J. R., Hall, L., & Magnanti, T. (1990). *Capacitated trees, capacitated routing and associated polyhedra* (Technical Report Discussion Paper 9061). CORE, Louvain La Nueve.

Augerat, P. (1995). *Approche polyèdrale du problème de tournées de véhicules*. PhD thesis, Institut National Polytechnique de Grenoble.

Augerat, P., Belenguer, J. M., Benavent, E., Corberán, A., Naddef, D., & Rinaldi, G. (1995). *Computational results with a branch and cut code for the capacitated vehicle routing problem* (Technical Report 1 RR949-M). ARTEMIS-IMAG, Grenoble, France.

Augerat, P., Belenguer, J. M., Benavent, E., Corberán, A., & Naddef, D. (1998). Separating capacity constraints in the CVRP using tabu search. *European Journal of Operational Research*, *106*, 546–557.

Baldacci, R., & Mingozzi, A. (2009). A unified exact method for solving different classes of vehicle routing problems. *Mathematical Programming*, *120*(2), 347–380.

Baldacci, R., Hadjiconstantinou, E., & 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.

Baldacci, R., Bodin, L., & Mingozzi, A. (2006). The multiple disposal facilities and multiple inventory locations rollon-rolloff vehicle routing problem. *Computers and Operations Research*, *33*(9), 2667–2702.

Baldacci, R., Toth, P., & Vigo, D. (2007). Recent advances in vehicle routing exact algorithms. *4OR: A Quarterly Journal of Operations Research*, *5*(4), 269–298.

Baldacci, R., Battarra, M., & Vigo, D. (2008a). Routing a heterogeneous fleet of vehicles. In B. L. Golden, S. Raghavan, & E. Wasil (Eds.), *The vehicle routing problem: latest advances and new challenges* (Vol. 43, pp. 3–27). Berlin: Springer.

Baldacci, R., Christofides, N., & Mingozzi, A. (2008b). An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming Ser. A*, *115*(2), 351–385.

Baldacci, R., Battarra, M., & Vigo, D. (2009, to appear). Valid inequalities for the fleet size and mix vehicle routing problem with fixed costs. *Networks*. DOI: 10.1002/net.20331

Balinski, M., & Quandt, R. (1964). On an integer program for a delivery problem. *Operations Research*, *12*, 300–304.

Bramel, J., & Simchi-Levi, D. (2002). Set-covering-based algorithms for the capacitated VRP. In P. Toth & D. Vigo (Eds.), *SIAM monographs on discrete mathematics and applications: Vol. 9. The vehicle routing problem* (pp. 85–108). Philadelphia: SIAM.

Choi, E., & Tcha, D. W. (2007). A column generation approach to the heterogeneous fleet vehicle routing problem. *Computers and Operations Research*, *34*, 2080–2095.

Christofides, N., & Eilon, S. (1969). An algorithm for the vehicle dispatching problem. *Operational Research Quarterly*, *20*, 309–318.

Christofides, N., Mingozzi, A., & Toth, P. (1979). The vehicle routing problem. In N. Christofides, A. Mingozzi, P. Toth, & C. Sandi (Eds.), *Combinatorial optimization* (pp. 315–338). New York: Wiley. Chap. 11.

Christofides, N., Mingozzi, A., & Toth, P. (1981). Exact algorithms for the vehicle routing problem based on spanning tree and shortest path relaxation. *Mathematical Programming*, *10*, 255–280.

Chvátal, V. (1973). Edmonds polytopes and weakly Hamiltonian graphs. *Mathematical Programming*, *5*, 29–40.

Cordeau, J. F., Gendreau, M., & Laporte, G. (1997). A tabu search heuristic for periodic and multi-depot vehicle routing problems. *Networks*, *30*, 105–119.

Cordeau, J. F., Laporte, G., Savelsbergh, M. W. P., & Vigo, D. (2007). Vehicle routing. In C. Barnhart & G. Laporte (Eds.), *Transportation, handbooks in operations research and management science* (Vol. 14, pp. 367–428). Amsterdam: North-Holland.

Cornuéjols, G., & Harche, F. (1993). Polyhedral study of the capacitated vehicle routing. *Mathematical Programming*, *60*, 21–52.

CPLEX. (2006). *ILOG CPLEX 9.0 callable library*. ILOG.

Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management Science*, *6*(1), 80–91.

Finke, G., Claus, A., & Gunn, E. (1984). A two-commodity network flow approach to the traveling salesman problem. *Congressus Numerantium*, *41*, 167–178.

Fischetti, M., & Toth, P. (1989). An additive bounding procedure for combinatorial optimization problems. *Operational Research*, *37*(2), 319–328.

Fischetti, M., Toth, P., & Vigo, D. (1994). A branch-and-bound algorithm for the capacitated vehicle routing problem on directed graphs. *Operational Research*, *42*, 846–859.

Fischetti, M., Salazar González, J. J., & Toth, P. (1995). Experiments with a multi-commodity formulation for the symmetric capacitated vehicle routing problem. In *3rd meeting of the EURO working group on transportation Barcelona* (pp. 169–173).

Fisher, M. L. (1994). Optimal solution of vehicle routing problems using minimum K-trees. *Operational Research*, *42*, 626–642.

Fukasawa, R., Longo, H., Lysgaard, J., de Aragão, M.P., Reis, M., Uchoa, E., & Werneck, R.F. (2006). Robust branch-and-cut-and-price for the capacitated vehicle routing problem. *Mathematical Programming (A)*, *106*, 491–511.

Garey, M. R., & Johnson, D. S. (1990). *Computers and intractability; A guide to the theory of NP-completeness*. New York: Freeman.

Gendreau, M., Laporte, G., & Potvin, J.-Y. (2002). Metaheuristics for the capacitated VRP. In P. Toth & D. Vigo (Eds.), *SIAM monographs on discrete mathematics and applications: Vol. 9. The vehicle routing problem* (pp. 129–154). Philadelphia: SIAM.

Golden, B. L., Magnanti, T. L., & Nguyen, H. Q. (1977). Implementing vehicle routing algorithms. *Networks*, *7*, 113–148.

Gouveia, L. (1995). A result on projection for the vehicle routing problem. *European Journal of Operational Research*, *85*, 610–624.

Grötschel, M., & Padberg, M. W. (1979). On the symmetric traveling salesman problem: I and II. *Mathematical Programming*, *16*, 265–280.

Grötschel, M., & Padberg, M. W. (1985). Polyhedral theory. In E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, & D. B. Shmoys (Eds.), *The traveling salesman problem: A guided tour of combinatorial optimization* (pp. 231–305). Chichester: Wiley.

Laporte, G., & Nobert, Y. (1984). Comb inequalities for the vehicle routing problem. *Methods of Operations Research*, *51*, 271–276.

Laporte, G., & Nobert, Y. (1987). Exact algorithms for the vehicle routing problem. *Annals of Discrete Mathematics*, *31*, 147–184.

Laporte, G., & Semet, F. (2002). Classical heuristics for the capacitated VRP. In P. Toth & D. Vigo (Eds.), *SIAM monographs on discrete mathematics and applications: Vol. 9. The vehicle routing problem* (pp. 109–128). Philadelphia: SIAM.

Laporte, G., Nobert, Y., & Desrochers, M. (1985). Optimal routing under capacity and distance restrictions. *Operational Research*, *33*, 1058–1073.

Letchford, A. N., & Salazar González J. J. (2006). Projection results for vehicle routing. *Mathematical Programming*, *105*(2–3), 251–274.

Letchford, A. N., Eglese, R. W., & Lysgaard, J. (2002). Multistars, partial multistars and the capacitated vehicle routing problem. *Mathematical Programming*, *94*, 21–40.

Lysgaard, J. (2003). *CVRPSEP: A package of separation routines for the capacitated vehicle routing problem* (Technical Report). Dept. of Mgt. Science and Logistics, Aarhus School of Business.

Lysgaard, J., Letchford, A. N., & Eglese, R. W. (2004). A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming*, *100*(2), 423–445.

Naddef, D., & Rinaldi, G. (2002). Branch-and-cut algorithms for the capacitated VRP. In P. Toth & D. Vigo (Eds.), *SIAM monographs on discrete mathematics and applications: Vol. 9. The vehicle routing problem* (pp. 53–81). Philadelphia: SIAM.

Niskanen, S., & Östergård, P. R. J. (2003). *Cliquer user's guide* (Technical Report 48). Helsinki University of Technology Communications Laboratory.

Östergård, P. R. J. (2002). A fast algorithm for the maximum clique problem. *Discrete Applied Mathematics*, *120*(1–3), 197–207.

Pessoa, A., de Aragão, M. P., & Uchoa, E. (2008). Robust branch-cut-and-price algorithms for vehicle routing problems. In B. L. Golden, S. Raghavan, & E. Wasil (Eds.), *The vehicle routing problem: Latest advances and new challenges* (Vol. 43, pp. 297–325). Berlin: Springer.

Pessoa, A., & Uchoa, E. de Aragão, M.P. (2009, to appear). A robust branch-cut-and-price algorithm for the heterogeneous fleet vehicle routing problem. *Networks*. DOI: 10.1002/net.20330

Ralphs, T. K., Kopman, L., Pulleyblank, W. R., & Trotter, L. E. (2003). On the capacitated vehicle routing problem. *Mathematical Programming (B)*, *94*, 343–359.

Toth, P., & Vigo, D. (2002). *SIAM monographs on discrete mathematics and applications*: *Vol*. 9. *The vehicle routing problem*. Philadelphia: SIAM.

Yaman, H. D. (2006). Formulations and valid inequalities for the heterogeneous vehicle routing problem. *Mathematical Programming Ser. A*, *106*, 365–390.