Heuristic algorithms for a vehicle routing problem with simultaneous delivery and pickup and time windows in home health care

Ran Liu, Xiaolan Xie, Vincent Augusto, Carlos Rodriguez

To cite this version:


HAL Id: hal-00838253
https://hal.archives-ouvertes.fr/hal-00838253
Submitted on 8 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Heuristic algorithms for a vehicle routing problem with simultaneous delivery and pickup and time windows in home health care

Ran LIU$^{a,b}$, Xiaolan XIE$^{a,b}$, Vincent AUGUSTO$^a$, Carlos RODRIGUEZ$^a$

$^a$Ecole Nationale Supérieure des Mines
Centre for Health Engineering, CNRS UMR 6158 LIMOS-ROGI
158 cours Fauriel, 42023 Saint Etienne, France

$^b$Shanghai Jiao Tong University
Centre for Healthcare Engineering, Dept. Industrial Engr. & Logistics Management
800 Dongchuan Road, Min Hang District, 200240 Shanghai, China

Correspondence: Prof. Xiaolan XIE, Centre for Health Engineering (CIS), Ecole Nationale Superieure des Mines de Saint-Etienne (ENSM.SE), 158 cours Fauriel, 42023 Saint-Etienne cedex 2 France. Tel. +33(0)477426695; Fax +33(0)477420249; xie@emse.fr

Submitted on 10/04/2012, revised 27/03/2013

Abstract: This paper addresses a vehicle scheduling problem encountered in home health care logistics. It concerns the delivery of drugs and medical devices from the home care company's pharmacy to patients' homes, delivery of special drugs from a hospital to patients, pickup of bio samples and unused drugs and medical devices from patients. The problem can be considered as a special vehicle routing problem with simultaneous delivery and pickup and time windows, with four types of demands: delivery from depot to patient, delivery from a hospital to patient, pickup from a patient to depot and pickup from a patient to a medical lab. Each patient is visited by one vehicle and each vehicle visits each node at most once. Patients are associated with time windows and vehicles with capacity. Two mixed-integer programming models are proposed. We then propose a Genetic Algorithm (GA) and a Tabu Search (TS) method. The GA is based on a permutation chromosome, a split procedure and local search. The TS is based on route assignment attributes of patients, an augmented cost function, route re-optimization, and attribute-based aspiration levels. These approaches are tested on test instances derived from existing VRPTW benchmarks.

Keywords: Home health care logistics, vehicle routing, pickup and delivery, time windows, metaheuristics,
1 Introduction

This paper considers a special vehicle routing problem with simultaneous delivery and pickup and time windows in home health care. Home hospitalization organizations have been created for patients requiring long and regular health cares in order to provide quality health service at their home while reducing the bed requirements at hospitals. Home hospitalization initially focused on nursing cares and has been extended to complex and technical cares such as chronic cares, rehabilitation, end-of-life palliative cares, and home chemotherapy. Home health care services are provided in France by Home Health Care (HHC) companies. Each day, a HHC company has various logistic activities including delivering drugs and medical devices from its pharmacy (also called depot in this paper) to patients at their home. It also takes some special drugs, such as chemotherapy drugs and blood products, from hospitals to patients. On the other hand, the HHC also needs to pick up materials from patients and deliver to different locations. Blood samples of the patients are collected and delivered to a medical lab. Medical wastes, unused drugs and medical devices are collected and brought back to the HHC or the depot. As HHC companies are usually small but serve rather large number of patients with dispersed locations, it is crucial to carefully design the routes of the HHC vehicles in order to reduce its operating cost while improving the service quality to patients.

Since an HHC patient may be a delivery and a pickup client simultaneously and have both pickup and delivery demands, the design of HHC vehicle routes is related to the vehicle routing problem with simultaneous pickup and delivery and time windows (VRPSDPTW) introduced by (Hokey, 1989). The VRPSDPTW is a hard and challenging problem in the field of vehicle routing problem (VRP). It considers clients that require simultaneous pickup and delivery service. Some common constraints must be satisfied in both HHC’s vehicle scheduling problem and the VRPSDPTW. For example, each client must be visited and served in a given time window; the load on a vehicle must always be below the vehicle capacity. However, the problem faced by the HHC company is more complex than the classical VRPSDPTW. The first reason is the complexity of its logistic operations with different types of pickup and delivery demands of patients. According to the origins and destinations, both pickup and delivery demands can be divided into two subclasses. The pickup demands include: (i) picking up the material from patients’ homes and deliver to a lab, e.g., biological samples; (ii) picking up some materials from the patients’ homes and bring back to the depot, e.g., medical waste. Similarly, there are two subclasses of delivery demands required by the patients: (i) delivering the products from the company’s depot to patients; (ii) delivering some materials from a hospital to patients’ homes, e.g., special drugs for cancer treatment. In the classical VRPSDPTW all delivery goods are loaded at depot and all pickup goods have to be transported to depot. In our HHC vehicle scheduling problem besides the depot, goods can be transported from a hospital to patients and from the patients to a lab. Clearly, the composition of vehicles’ loads in our case is more complex than VRPSDPTW. Furthermore, different from the classical VRPSDPTW, each route of our problem must satisfy some precedence
constraints, e.g., for a patient who needs drugs provided by the hospital, the vehicle visiting the patient has to visit the hospital first. Such special constraints are similar to the pairing and precedence constraints in classical *pickup and delivery problem* (PDP), in which each customer request is defined by an origin location and a destination, the origin must visited before the destination by the same vehicle. However, the PDP is less complicated than our problem, since the origins as well as the destinations of transportation requests in the PDP are locations other than the depot, and a customer in the PDP only has either pickup or delivery request. Thus, in this paper the HHC vehicle scheduling problem is rather a special VRPSDPTW variant which has never been studied before. Since both the VRPSDPTW and PDP are NP-hard problems, our problem is more complex than these problems and is also NP-hard. To the best of our knowledge, we have not found any existing work dealing this special simultaneous pickup and delivery problem in HHC industry.

In this paper, we first perform a literature review, propose two mathematical formulations of our problem, and then develop two heuristic algorithms for this special vehicle scheduling problem. The rest of this paper is organized as follows. Section 2 introduces the relevant literature. Our problem is formally defined and two mathematical models are given in Section 3. Section 4 proposes a Genetic algorithm (GA) for our problem. Section 5 proposes a Tabu Search (TS) algorithm for solving the problem. Computational experiments are described in Section 6. Section 7 concludes the paper.

2 Literature review

As stated before, two main bodies of vehicle routing literature are relevant to our problem. The first is the *vehicle routing problem with simultaneous pickup and delivery and time windows*, in which goods are transported by a fleet of vehicles between the depot and customers within their time windows. The second one is the *pickup and delivery problem with time windows* (PDPTW) problem, in which goods are transported between *n* pickup and *n* delivery locations, and the vehicle visiting each location must be within an associated time window. We survey the literature in two parts.

The VRPSDPTW is an extension of the VRPSDP, and has been much less studied than the VRPSDP. The VRPSDP can be seen an extension of the *vehicle routing problems with backhauls* (VRPB). In the VRPB, the set of customers are divided in two subsets consisting of linehaul and backhaul custumers, where a linehaul customer requires a given quantity of product to be delivered from the depot, and a backhaul customer requires a given quantity of product to be picked up to the depot. In the VRPB, it is assumed that the vehicles only pick goods up (serve backhaul customers) after they have finished delivering their entire load (serve linehaul customers) (P. Toth and Vigo, 1997a) (Goetschalckx and Jacobs-Blea, 1989). One reason for this assumption is the difficulty to re-arrange delivery and pickup goods on the vehicles. The objective of the VRPB is to design a set of minimum cost routes so that on each route neither the total load of linehaul customers nor that of backhaul customers exceed the vehicle capacity. The VRPB is a NP-hard problem in strong sense and a number of algorithms
are proposed for this problem. Exact methods for the VRPB are proposed by (Yano, et al., 1987) (P. Toth and Vigo, 1997a) (Mingozzi, et al., 1999). Heuristics have been developed by (Goetschalckx and Jacobs-Blecha, 1989) (Paolo Toth and Vigo, 1999) (Osman and Wassan, 2002) (Tavakkoli-Moghaddam, et al., 2006) (Y Gajpal and PL Abad, 2009). If the linehaul and backhaul customers can be freely mixed within a route, the VRPB is transformed to the **vehicle routing problem with mixed backhauls** (VRPMB). Clearly, the vehicle capacity check in the VRPMB is more complicated than the VRPB. Exact solution methods for the VRPMB have only been developed for the single vehicle case (Eilam Tzoreff, et al., 2002) (Süral and Bookbinder, 2003) (Baldacci, et al., 2003). Heuristics for the VRPMB are given by (Nagy and Salhi, 2005) (Salhi and Nagy, 1999) (Wade and Salhi, 2004) (Reimann and Ulrich, 2006). Based on the VRPMB, if we allow customers to have both pickup quantity and delivery quantity, then, there exist two special problems: the **vehicle routing problem with divisible delivery and pickup** (VRPDDP), and the **vehicle routing problem with simultaneous delivery and pickup** (VRPSDP). The difference between the VRPDDP and VRPSDP is the number of times a customer is visited. In the VRPDDP customers do not have to be visited exactly once, i.e., a customer can be visited twice, once for pickup and once for delivery service. The VRPSDP requires that each customer is visited only once by a vehicle. The VRPDDP instances can be transformed to VRPMB by modeling every customer’s pickup and delivery service as two separate customers. One exact method for the VRPSDP was designed by (Dell’Amico, et al., 2006). (Dethloff, 2002) proposes an extension of the cheapest insertion heuristic to the VRPSDP. Several tabu search algorithms for VRPSDP were proposed in (Alfredo Tang Montané and Galvão, 2006) (Chen and Wu, 2005) (Bianchessi and Righini, 2007) (Crispim and Brandão, 2005). Recently, (Ai and Kachitvichyanukul, 2009) (Y. Gajpal and P. Abad, 2009) (Subramanian, et al., 2010) have proposed several metaheuristics to solve VRPSDP.

Contrary to the VRPB, VRPMB, and VRPSDP, only a fewer researchers consider the time window constraints in these problems, especially for the VRPSDP. For example, an exact algorithm was designed for the **VRP with time windows** in (Gélinas, et al., 1995). (Duhamel, et al., 1997) (REIMANN, et al., 2002) (Thangiah, et al., 1996) (Zhong and Cole, 2005) proposed heuristics for this problem. For the **VRPMB with time windows**, (Hasama, et al., 1998) (Kontoravdis and Bard, 1995) (Zhong and Cole, 2005) designed heuristics, with the primary objective of minimizing the number of the vehicles and the second objective of minimizing traveling distances. For the most complex one, **VRPSDPTW**, only (Angelelli and Mansini, 2002) proposed an exact method and (Mingyong and Erbao, 2010) and (Wang and Chen, 2012) proposed genetic algorithms.

Compared with VRPSDPTW, there is an abundant body of research on the second related problem **PDPTW**. The PDPTW originates from the basic **PDP** (M. W. P. Savelsbergh and Sol, 1995). In the PDP, a customer order consists of two parts: a pickup at one location and a delivery at another location. The PDP has been intensively studied in the past three decades. For survey on the PDP, the reader is referred to (M. W. P. Savelsbergh and Sol, 1995),
(Berbeglia, et al., 2007), and (S. Parragh, et al., 2008). For the PDPTW, several exact approaches have been designed. (Dumas, et al., 1991) (M. Savelsbergh and Sol, 1998) (Xu, et al., 2003) (Sigurd and Pisinger, 2004) used branch and price schemes for the PDPTW. (J.-F. Cordeau, 2006) (Ropke, et al., 2007) developed branch and cut approach for the PDPTW. (Ropke and Cordeau, 2009) introduced a new branch and cut and price algorithm for the PDPTW. Meanwhile, many heuristics have proposed for the PDPTW. (J.-F. Cordeau, 2006) (Ropke, et al., 2007) developed branch and cut approach for the PDPTW. (Ropke and Cordeau, 2009) introduced a new branch and cut and price algorithm for the PDPTW. Meanwhile, many heuristics have proposed for the PDPTW. (Jaw, et al., 1986), (Madsen, et al., 1995), (Diana and Dessouky, 2004), (Lu and Dessouky, 2006) presented various insertion-based heuristics for solving the PDPTW. (P. Toth and Vigo, 1997b) (Nanry and Wesley Barnes, 2000) (J.-F. Cordeau and Laporte, 2003) solved the PDPTW by means of tabu search heuristics. (Li and Lim, 2001), (Pankratz, 2005), (Ropke and Pisinger, 2006) and (S. N. Parragh, et al., 2010) designed simulated annealing, genetic algorithm, adaptive large neighborhood search heuristic, and variable neighborhood search heuristic for solving the PDPTW. Extensive review of the PDPTW literature is out of the scope of in this paper and interested readers are referred to the surveys of (Berbeglia, et al., 2007) and (S. Parragh, et al., 2008).

Although a large number of literatures have studied the VRPSDPTW and PDPTW, to our best knowledge, no existing literature consider the problem, which contains following two vehicle service strategies simultaneously: (i) the transportation of goods from the depot to linehaul customers and from backhaul customers back to the depot; (ii) goods are transported between pickup and delivery locations. Our HHC vehicle scheduling problem includes both of these two service strategies and can be seen as a special VRPSDPTW variant.

3 Mathematical Formulation

This paper addresses the daily scheduling problem of vehicles of a home health care company for delivery of drugs and medical devices and for pickup of biological samples and medical wastes or unused drugs. This section provides a formal description of the problem and then presents two equivalent mixed integer programming formulations of the problem that will serve to assess the efficiency of the heuristic methods of this paper.

The problem can be defined as follows. Let \( G = (V, A) \) be a directed graph with a set \( V = \{0, 1, \ldots, n, n+1\} \cup \{h, l\} \) of nodes and a set \( A = \{(i, j) : i, j \in V, i \neq j\} \) of arcs. Nodes 0 and \( n+1 \) represent the origin and destination depots which are in practice the pharmacy of the home health care company. Each vehicle starts at node 0 and ends at node \( n+1 \). Nodes \( N = \{1, \ldots, n\} \) correspond to patients’ homes. Node \( h \) and \( l \) represent the locations of a hospital and a medical lab.

Each patient \( i \in N \) has four types of delivery and pickup requirements: \( d_{i1}, d_{i2}, p_{i1} \) and \( p_{i2} \), where \( d_{i1} \) represents the amount of materials (drugs/medical devices) to deliver from the depot 0 to patient \( i \), \( d_{i2} \) the amount of materials (special drugs) to deliver from the hospital to patient \( i \), \( p_{i1} \) the amount of materials to pick up from patient \( i \) and bring back to the depot \( n+1 \), and \( p_{i2} \) the amount of materials (biological samples) to pick up from patient \( i \) and bring to the medical.
lab $h$. Each type of requirements is called a demand. Different materials are assumed to be compatible and can be loaded in the same vehicle. $D_1 \subseteq N$ denotes the set of patients needing type 1 delivery service, i.e. patients $i$ with $d_{i1} > 0$. Similarly, $D_2$, $P_1$, $P_2$ denote sets of patients needing type 2 delivery, type 1 and type 2 pickup services. A patient may require different types of demands. For example, for a patient $i \in D_2 \cap P_2$, the company has to pick up the quantity $p_{i2}$ from node $i$ and deliver to the lab and deliver the quantity $d_{i2}$ from the hospital to this node. For notation convenience, we set zero-demands for nodes $0$, $n+1$, $h$, $l$.

A time window $[a_i, b_i]$ is associated with each node $i \in V$, where $a_i$ and $b_i$ represent the earliest and latest time. A vehicle is allowed to arrive before $a_i$ and wait until the patient becomes available, but arrivals after $b_i$ are prohibited. The depot node also has a time window, representing the earliest and latest times when the vehicles may leave from and return to the depot. Each arc $(i, j) \in A$ is associated with a routing cost $c_{ij}$ and a travel time $t_{ij}$. The service time for a patient $i$ is assumed to be included in the travel time $t_{ij}$.

A fleet $K$ of identical vehicles, initially located at the depot, is available to serve the patients. Each vehicle has a capacity of $Q$.

The problem consists in determining a set of at most $K$ routes of minimal overall cost in order to serve all delivery and pickup demands of all patients under the obvious time window and vehicle capacity constraints and the following assumptions.

**Assumption A.** Each route starts and ends the depot and visits each location at most once;

**Assumption B.** Each patient is visited by exactly one vehicle for all its demands;

**Assumption C.** Each route makes a hospital visit before visits to its $D_2$-patients;

**Assumption D.** Each route makes a lab visit after all visits to its $P_2$-patients.

A typical route is as follows. The vehicle starts the depot with all materials for its $D_1$-patients, visits some patients for $D_1$-delivery and any pickup, visits the hospital to load all materials for its $D_2$-patients, visits other patients, then goes to the lab to deliver materials of all $P_2$-patients, visits other patients before returning to the depot with all materials of its $P_1$-patients.

In the following, we propose two three-index MIP formulations. The first mathematical formulation termed MIP1 is derived from the model of (Dell’Amico, et al., 2006) and (Ropke and Cordeau, 2009). Four types of decision variables are used.

- $x_{ij}^k$ binary variable equal to 1 if vehicle $k$ travels directly from node $i$ to node $j$;
- $B_i^k$ time at which vehicle $k$ begins to serve at node $i$;
- $y_{ij}^k$ quantity of $P_1$- and $P_2$-pickup carried along arc $(i, j)$ by vehicle $k$;
- $w_{ij}^k$ quantity of $D_1$- and $D_2$-delivery carried along arc $(i, j)$ by vehicle $k$.

**MIP1:**

$$
\text{Min } \sum_{i \in V} \sum_{j \in V} \sum_{k \in K} c_{ij} x_{ij}^k
$$  \hspace{1cm} (1)
Subject to:

\[
\sum_{k \in K} \sum_{j \in V} x^k_{ji} = 1 \quad \forall i \in N \tag{2}
\]

\[
\sum_{j \in V} x^k_{ji} \leq \sum_{j \in V} x^k_{ij} \quad \forall i \in P_2, k \in K \tag{3}
\]

\[
\sum_{j \in V} x^k_{ji} \leq \sum_{j \in V} x^k_{ij} \quad \forall i \in D_2, k \in K \tag{4}
\]

\[
\sum_{i \in V} x^k_{ii} \leq 1 \quad \forall k \in K \tag{5}
\]

\[
\sum_{i \in V} x^k_{ji, i+1} \leq 1 \quad \forall k \in K \tag{6}
\]

\[
\sum_{j \in V} x^k_{ji} = \sum_{j \in V} x^k_{ij} \quad \forall i \in N \cup \{l, h\}, k \in K \tag{7}
\]

\[
\sum_{j \in V} x^k_{ji} \leq 1 \quad \forall i \in \{l, h\}, k \in K \tag{8}
\]

\[
B^k_j \geq (B^k_j + t_{ij})x^k_{ij} \quad \forall i \in V, j \in V, i \neq j, k \in K \tag{9}
\]

\[
\sum_{i \in V} y^k_{ji} - \sum_{i \in V} y^k_{ij} = p_{j1} + p_{j2} \quad \forall j \in N \cup \{h\} \tag{10}
\]

\[
\sum_{i \in V} y^k_{ji} - \sum_{i \in V} y^k_{ij} = d_{j1} + d_{j2} \quad \forall j \in N \cup \{l\} \tag{11}
\]

\[
\sum_{i \in V} y^k_{li} - \sum_{i \in V} y^k_{il} = \sum_{i \in V} \sum_{j \in D_2} x^k_{ij}d_{i2} \quad \forall k \in K \tag{12}
\]

\[
\sum_{i \in V} y^k_{li} - \sum_{i \in V} y^k_{il} = \sum_{i \in V} \sum_{j \in P_2} x^k_{ij}p_{i2} \quad \forall k \in K \tag{13}
\]

\[
B^k_i \geq (B^k_i + t_{ji})\sum_{j \in V} x^k_{ij} \quad \forall i \in P_2, k \in K \tag{14}
\]

\[
B^k_i \geq (B^k_i + t_{hi})\sum_{j \in V} x^k_{ij} \quad \forall i \in D_2, k \in K \tag{15}
\]

\[
a_i \leq B^k_i \leq b_j \quad \forall i \in V, k \in K \tag{16}
\]

\[
y^k_{ij} + w^k_{ij} \leq Qx^k_{ij} \quad \forall i \in V, j \in V, k \in K \tag{17}
\]

\[
x^k_{ij} \in \{0, 1\} \quad \forall i, j \in V, i \neq j, k \in K \tag{18}
\]

The objective function (1) minimizes the total routing cost. Constraints (2) ensure that all the demands are satisfied. Constraints (3) guarantee that a vehicle visiting a $P_2$-patient also visits the lab. Constraints (4) are similar but for hospital visit. Constraints (5) and (6) force the route of each vehicle to start and end at the depot. Constraints (7) ensure the flow balance of the vehicles, i.e., if a vehicle visits a node it must leave this node. Constraints (8) indicate that each vehicle can only visit the hospital and the lab once. Constraints (9) impose the consistency of the visiting times. Constraints (10) and (11) are flow equations for pickup and delivery demands. Constraints (12) impose that all delivery demands of $D_2$ patients are loaded at the hospital. Constraints (13) impose the unloading of all pickup demands of $P_2$ patients at the lab. Constraints (14) ensure that each $P_2$ patient is visited before a lab visit; constraints (15)
ensure the hospital visit before any visit to a $D_2$ patient. Finally, constraints (16) and (17) impose the time window and vehicle capacity constraints. This MIP1 model is nonlinear because of constraints (9), (14) and (15) that can be linearized as follows:

$$B_i^k \geq B_i^k + t_{ij}^k - M(1-x_{ij}^k) \quad \forall i \in V, j \in V \setminus 0, k \in K$$ (19)

$$B_i^k \geq B_i^k + t_{ij}^k - M(1-\sum_{j \in V} x_{ij}^k) \quad \forall i \in P_2, k \in K$$ (20)

$$B_i^k \geq B_i^k + t_{ij}^k - M(1-\sum_{j \in V} x_{ij}^k) \quad \forall i \in D_2, k \in K$$ (21)

The second mathematical formulation termed MIP2 uses decision variables $x_{ij}^k$ and $B_i^k$ as in MIP1 plus the following new decision variables:

$z_{ij}^k$: quantity of $P_j$-pickup with $j \in \{1, 2\}$ carried by vehicle $k$ when leaving node $i$;

$v_{ij}^k$: quantity of $D_j$-delivery with $j \in \{1, 2\}$ carried by vehicle $k$ when leaving node $i$.

MIP2: $\min \sum_{i \in V} \sum_{j \in V} \sum_{k \in K} c_{ij} x_{ij}^k$

subject to constraints (2)-(9), (14)-(16), (18) and

$$z_{ij}^k \geq (z_{i1}^k + p_{j1})x_{ij}^k \quad \forall i \in V, j \in N \cup \{l, h\}, i \neq j, k \in K$$ (22)

$$z_{ij}^k \geq (z_{i2}^k + p_{j2})x_{ij}^k \quad \forall i \in V \setminus \{l\}, j \in N \cup \{l, h\}, i \neq j, k \in K$$ (23)

$$v_{ij}^k \geq (v_{j1}^k + d_{j1})x_{ij}^k \quad \forall i \in V, j \in N \cup \{l, h\}, i \neq j, k \in K$$ (24)

$$v_{ij}^k \geq (v_{j2}^k + d_{j2})x_{ij}^k \quad \forall j \in V \setminus \{h\}, i \in N \cup \{l, h\}, i \neq j, k \in K$$ (25)

$$z_{ih}^k + z_{ih}^k + v_{ih}^k + v_{ih}^k \leq Q \quad \forall i \in V, k \in K$$ (26)

Constraints (22) determine the vehicle loading for $P_1$-pickup. Constraints (23) track the vehicle loading for $P_2$-pickup and ensure that it becomes null after the lab visit at the optimum. Constraints (24) and (25) determine the vehicle loading for $D_1$ and $D_2$ delivery. Constraints (26) are vehicle capacity constraints. Again nonlinear constraints (22)-(25) can easily be linearized by standard reformulation techniques.

As explained in Sections 1 and 2, our problem is related to highly complex vehicle routing problems including VRPSDPTW and PDPTW problems. It can be easily proved that our problem is strongly NP-hard as it covers classical VRP problems as special case. It is expected that the above MIP formulation can only be used to solve small-size problems. This is evidence by numerical results of Section 6 in which we try to solve different test instances with the Cplex 12.3 solver. Even for instances of very small size with 30 patients and 40 demands, Cplex is not able solve the problem optimally. For some of the small size instances, it even cannot get a feasible solution in reasonable time (about 48 hours). For these reasons, we propose in the following a genetic algorithm and a tabu search method to address problems of large size.
4  A genetic algorithm

The genetic algorithm proposed in this paper combines the following features: a permutation chromosome, exact fitness computation by splitting, improvement by local search, diversity of the population and multi-start with partially replaced population. These ingredients were shown powerful in (Prins, 2004) in designing efficient GA for vehicle-routing like problems.

In this paper, the chromosome is a permutation of all patients that give order of visits in different routes. An exact split algorithm will be presented to split the permutation into sub-strings of patients to be visited by a vehicle and to insert hospitals and labs.

Algorithm 1: Outline of the GA

1: Generate an initial population $\Pi$ of chromosomes

Main GA exploration phase

2: Select two parents $P_1$ and $P_2$ by binary tournament;
3: Crossover $(P_1, P_2)$ by OX operator;
4: Evaluate the two resulting children by Splitting;
5: Repeat 2-4 if no child is feasible. Otherwise, select randomly a feasible child $C$;
6: Improve $C$ by Local Search, with probability $p_m$;
7: Insert $C$ in $\Pi$ to replace a randomly selected individual among the half worst of $\Pi$ if $C$ is not the current worst and $C$ has a distinct fitness value than those in $\Pi$;
8: Repeat 2-7 for $N_1$ iterations or till $N_2$ iterations without improving the current best;

End of the main GA phase

9: Restart the main GA phase 2-8 with a partially replaced population, for $N_3$ phases or till $N_4$ phases without improving the best solution.

The overall structure of the GA is illustrated in Algorithm 1. It starts with the generation of an initial population with insertion of good heuristic solutions to be presented. The central part of our GA is an incremental GA exploration phase in which only one chromosome is replaced at each iteration. It starts with the selection of two parents by binary tournament. The OX operator that was proved appropriate for VRP problems in (Prins, 2004) is then applied to generate two child chromosomes. These child chromosomes are evaluated by the exact split algorithm. The selection and crossover operations are repeated till obtaining a feasible child chromosome, i.e. a chromosome for which a feasible split solution exists. With some probability, this feasible child chromosome is further improved by Local Search to be presented. The new child chromosome is inserted in the current population under two conditions: (i) it is better than the current worst and (ii) it does not have identical fitness as an existing chromosome.

In our GA implementation, the main GA phase is performed for $N_1=3000$ iterations or till $N_2=1500$ iterations without improving the current best. The main GA phase is repeated for
\( N_3 = 10 \) phases or till \( N_4 = 5 \) phases without improving the current best. To restart the GA phase, we keep the \( \psi \) best chromosomes with \( \psi = 3 \) and replace the others with randomly generated chromosomes. Through some preliminary experiments, the proposed GA gives better solution when increasing the local search probability \( p_m \) and the size of the population \( |\Pi| \), especially when \( p_m \) is smaller than 0.8 and \( |\Pi| \) is less than 35. The negative side is the increasing running time of GA. To balance the accuracy and the speed of GA, the local search is applied with a probability \( p_m = 0.8 \), and the size of the population is kept fixed at \( |\Pi| = 30 \) chromosomes.

It has been proven that perverting the diversity of GA population can diminish the risk of premature convergence (Sörensen and Sevaux, 2006). A simple and stricter rule is imposed in this paper to keep the diversity of the population, i.e., the fitness of any two feasible chromosomes must be different. For this reason, a child chromosome \( C \) is inserted during each main GA phase only if it has a different fitness than existing individuals. Diversity is also checked in the generation of the initial population and the partially replaced population for restart.

The remaining of this Section is devoted the detailed presentation of the fitness evaluation, generation of initial solutions, and local search.

### 4.1 The chromosome and fitness evaluation

In our GA, a chromosome \( C \) is a permutation \((s_1, s_2, \ldots, s_n)\) of all patients \((1, 2, \ldots, n)\) and the fitness is the optimal criterion value of our problem such that each route visits a sub-string of patients of \( C \) and in the order of \( C \) with of course necessary visits to the hospital and lab. For example, the solution of a chromosome \((3, 5, 1, 2, 4)\) could be \((0, 3, 5, h, 1, n+1)\) and \((0, 2, 4, l, n+1)\).

Hereafter we propose a split procedure for fitness evaluation of our problem. It is a shortest path approach. It first builds an auxiliary graph \( H \) which contains nodes \((s_0, s_1, s_2, \ldots, s_n)\) with \( s_0 = 0 \) and in which each arc corresponds to a feasible route. There is an arc \((s_i, s_j)\) from node \( s_i \) to \( s_j \) with \( i < j \) if there exists a feasible route visiting all patients \((s_{i+1}, \ldots, s_j)\) in the given order. The length of the arc is the minimal total routing cost of such a route. The fitness of the chromosome is the shortest path from \( s_0 \) to \( s_n \) with at most \( K \) arcs.

In this approach, when considering an arc \((s_i, s_j)\), all possible and necessary insertions of the hospital \( h \) and the lab \( l \) in the route \((0, s_{i+1}, \ldots, s_j, n+1)\) are considered. For each insertion of \((h, l)\), the earliest visiting time of each node is determined, the corresponding time window constraint checked and the insertion is abandoned if the time window constraint is violated. The length \( l_{ij} \) of the arc \((s_i, s_j)\) is the smallest earliest visiting time of node \( n+1 \) among all feasible routes.

Once the auxiliary graph \( H \) is built, the fitness is determined by dynamic programming by determining iteratively the shortest path \( SP_{ik} \) from \( s_0 \) to \( s_i \) with at most \( k \) arcs. Clearly, the fitness is \( SP_{nk} \) determined by the following recursion:
\[
SP_{jk+1} = \min_{i \in J} SP_{ik} + l_{ij}, \forall k = 1, \ldots, K - 1
\]
\[
SP_{00} = 0
\]

Figure 1 is an illustrative example. The first part of Figure 1 shows a small instance with a depot(D), a hospital(h), a lab(l), three patients (a, b, c) and 2 identical vehicles with capacity Q=10. The chromosome is (a, b, c). The distances between any two nodes are given beside the arc, and the demand of each patient is shown in brackets. There are no time window constraints. In the second part of Figure 1, each arc in the auxiliary graph represents a possible vehicle route and its traveling distance. The route visiting patients a and b, i.e., arc (D, b), has two possible positions for visiting the hospital and the one leading to shortest distance is chosen to represent this arc. Based on the second part, is the shortest path with no more than 2 arcs is (D-a-c) and has a cost of 52. The corresponding solution consisting of two vehicle routes is given in the third part of Figure 1.
4.2 Constructing the initial population

The initial population is a combination of heuristic solutions and randomly generated chromosomes. It uses four simple constructive heuristics including three saving-based heuristics and a Nearest Neighbor (NN) heuristic. The heuristics make use of randomly set parameters but details will be given later.

The initial population is generated in two phases. Phase 1 starts with a population of the four heuristic solutions with randomly set parameters and randomly generated chromosomes. Phase 1 repeats till a feasible chromosome is obtained. Phase 2 tries to ensure the diversity of the population. It replaces two types of individuals with new random chromosomes: feasible chromosomes having the same fitness as another one and infeasible chromosomes. Phase 2 is repeated for $N_5 = 100$ times or till the diversity is ensured.

When restarting the GA exploration phase, the current population is partially replaced by keeping the $\psi$ best chromosomes, replacing the remaining ones with random chromosomes, and applying phase 2 to ensure the diversity.

We now give details of the four construction heuristics. The first simplest saving method, called $saving_1$, is as follows.

Step 1 assigns each patient to a separate route. Each patient of $P_1$ and $D_1$ is connected with the depot, forming a route beginning and ending at depot. Each $P_2$ patient $j$ is connected with the lab and the depot, forming a route (depot-$j$-l-depot). Similarly, each patient of $D_2$ is connected with the hospital and the depot;

Step 2 merges two routes associated with the maximal saving value, which is calculated as in the classical saving method by checking the time windows and vehicle capacity constraints. In the classical saving, when two routes (depot-$\ldots$-$i$-depot) and (depot-$\ldots$-$j$-depot) can be feasibly merged into a single route (depot-$\ldots$-$i$-$j$-$\ldots$-depot), the saving value is $c_{i0}+c_{0j}-\lambda c_{ij}$ where $\lambda$ is a parameter randomly sampled in $[0.8, 1]$. In our problem, the saving value is equal to the total distance of arcs to delete in the merged route minus the distances of arcs to add in the new giant route. If both two routes contain the hospital (lab), only the first hospital (the second lab) is kept in the new giant route. For example, the left part of Figure 2 show two routes (depot-$h$-$\ldots$-$i$-depot) and (depot-$h$-$j$-$\ldots$-depot) where $i$ and $j$ are $D_2$ patients and the others are $D_1$ or $P_1$ patients. When merging two routes into one giant route (shown in the right part), three arcs are deleted (represented by dotted line), and one arc $(i, j)$ is inserted into the new route. Thus, the saving value is $\Delta_i = c_{i0}+c_{0j}+c_{0h}-\lambda c_{ij}$. 
Step 3: repeat merging routes until no improvement is possible. The final chromosome is obtained by concatenation of the resulting routes in the linguistic order of the first patient in each route.

Two other saving heuristic $\text{saving}_2$ and $\text{saving}_3$ are also implemented with different one-step look-ahead criterion to improve $\text{saving}_1$. $\text{saving}_2$ combines the immediate saving value $\Delta_i$ of merging two routes and the maximum positive saving value $\Delta_{i, \text{max}}$ of this giant route with other remaining routes in the next step. $\text{saving}_2$ merges two routes with the largest $\Delta_2=\Delta_i+\lambda_i\Delta_{i, \text{max}}$ among all merging with positive $\Delta_i$ where $\lambda_i$ is the parameter, randomly sampled in $[0.1, 0.3]$. $\text{saving}_3$ is similar with $\Delta_2$ replacing by $\Delta_3=\Delta_i+\lambda_i\Delta_{i, \text{sum}}$ where $\lambda_i$ is a third parameter, randomly sampled in $[0.01, 0.03]$ $\Delta_{i, \text{sum}}$ is the sum of top-10 positive savings of merging the giant route and remaining routes in the next step.

In the last heuristic $\text{NN}$, we only consider the depot and patients and neglect the lab and hospital. First, $\text{NN}$ links the depot to its closest patient, i.e. the patient $i$ with the smallest routing cost from 0 to $i$. Then, the $\text{NN}$ adds a directed path connecting the last added node to the closest unvisited patient. The procedure repeats until all the patients are included in the path. The resulting sequence of patients is the chromosome of the $\text{NN}$ heuristic.

### 4.3 Improving a feasible chromosome by local search (LS)

This subsection starts with the split solution of a feasible chromosome and then improves it by several local moves including 1-1-exchange, 1-0 relocation, 2-Opt exchange, and 2-Opt* exchange. The first-accept strategy is used, i.e., once a new better neighboring solution is identified, it replaces the current solution. For each solution, it first tries 1-1-exchange and accept the first improving 1-exchange. If no improvement is possible with 1-exchange, it then tries the other local moves. The process repeats until no improvement is possible at all.

For 1-1-exchange and 1-0 relocation, both intra-route and inter-route movements are tried. The 1-1-exchange tries to exchange positions of any two patients. The 1-0 relocation is the operator for one patient and transfers a patient from its position in one route to another position either the same or a different route. The 2-Opt exchange is only executed on a single route and it tries to improve the route by replacing two of its edges by two other edges. The 2-Opt* exchange is an inter-route method: two edges are selected from two different routes,
respectively, and the end portions of two edges are exchanged, so as to generate two new routes.

Note that the inter-route 1-0 relocation and 2-Opt* may combine two routes into one and reduce the usage of the vehicles in the solution. In such case, the empty route is removed from the solution. Compared with classical VRPTW, in our work additional precedence constraints also must be checked during the LS procedure. During the search procedure, once a neighboring solution is identified, we checked whether it can satisfy the vehicle capacity, time windows, and precedence constraints. If not, it is aborted directly.

When the LS procedure stops, the resulting routes are concatenated in linguistic order to give the improved chromosome. The split procedure is applied to this new chromosome for evaluation of its fitness.

5 A tabu search algorithm

Tabu search (TS) is also one of the most powerful and competitive heuristics in the fields of VRP, VRPTW and other VRP variants (Gendreau, et al., 1994) (J. Cordeau, et al., 1997) (Hertz, et al., 2000) (Côté and Potvin, 2009) (Zachariadis, et al., 2009). Essentially, TS iteratively explores the solution space by moving from the current solution to another solution in its neighborhood. Since the current solution may deteriorate during the search, some anti-cycling strategies are used to help the search process explore a broad portion of the solution space.

In this section, we design a TS algorithm for our problem. It is based on the general TS framework of (J. F. Cordeau, et al., 2001) with similar attribute set and augmented criterion function for constraint violations. The general framework of our TS is given in Algorithm 2. It starts with the best feasible solution of the GA initial population given in Section 4 and searches on the set of solutions in which each patient belongs to a route. The neighborhood is defined by relocation of a patient from one route to another route or location exchange of patients in different routes. Our tabu search also includes the re-optimization of all modified routes after each local move. The TS stops after $L_1=5000$ iterations or after $L_2=2500$ iterations without improvement of the best solution. Our TS restarts from the second best and empty tabu list if there is no improvement after $L_3=1000$ iterations. Restarts bring TS to new search regions (Dell’amico, et al., 1999).

In the following, we give detailed presentation of the augmented criterion function, the neighborhood structure, the route re-optimization, the attribute set, tabu duration and aspiration criterion.

**Algorithm 2: Outline of the TS**

1: Determine an initial solution $Sol$;
2: Determine the best neighbor solution $Sol'$ that is not tabu or satisfies an aspiration criterion;
3: Re-optimize each modified route of Sol’ and set Sol := Sol’;
4: Update tabu list and aspiration levels;
5: Stop the TS if the stopping criterion is met;
6: Restart the TS from the second best if a restart criterion is met;
7: Go to 2.

### 5.1 Augmented criterion function

In our TS, both feasible and infeasible solutions are allowed. Each solution Sol partitions the set of patients into different routes and is completely represented by its set of routes \( k \) with each route represented by the sequence of nodes visited including the depot, the hospital and the lab.

For each route, the visiting times at different nodes can be easily determined by taking into account the visiting sequence and the earliest available time of each node. The vehicle load can be determined as follows. The vehicle starts at the depot with all demands of \( D_1 \) patients, drops all delivery demands \( d_{i_1} \) and \( d_{i_2} \) and loads all pickup demands \( p_{i_1} \) and \( p_{i_2} \) at each patient visit, loads all demands of \( D_2 \) patients of the route at the hospital, and drops all demands of \( P_2 \) patients of the route at the lab visit.

For each route \( k \) of a solution Sol, let \( w_k(Sol) \) be the total traveling distance (original objective value), \( d_k(Sol) \) be the total violation of vehicle capacity, \( e_k(Sol) \) be the total violation of time window, \( g_k(Sol) \) be the total violation of precedence constraints. The evaluation of \( w_k(Sol) \) is straightforward and the evaluation of the others is as follows:

\[
d_k(Sol) = \sum_{i \in R_k} (Load_{ik} - Q)^+
\]

\[
e_k(Sol) = \sum_{i \in R_k} (m_{ik} - b_i)^+
\]

\[
g_k(Sol) = N_{hk} + N_{lk}
\]

where \((x)^+ = \max(0, x)\), \( R_k \) is the set of nodes visited by vehicle \( k \), \( Load_{ik} \) is the vehicle load when leaving node \( i \), \( m_{ik} \) is the visiting time of node \( i \) in route \( k \) and \( b_i \) its latest visiting time, \( N_{hk} \) (\( N_{lk} \)) is the number of \( D_2 \) (\( P_2 \)) patients visited before a hospital visit (after a lab visit) in route \( k \).

The following augmented criterion function \( f(Sol) \) is used in our TS:

\[
f(Sol) = \sum_{i \in k} (w_i(Sol) + \alpha \cdot d_i(Sol) + \beta \cdot e_i(Sol) + \gamma \cdot g_i(Sol))
\]

where \( \alpha, \beta, \gamma \) are positive parameters to adjust the penalty of constraint violation. If Sol is a feasible solution, \( f(Sol) \) coincides with \( w(Sol) \). According to this definition, the TS search process will contain a mix of feasible and infeasible solutions, reducing the probability of becoming trapped in a local solution.

Our TS algorithm adjusts \( \alpha, \beta, \gamma \) dynamically to facilitate the exploration of the search
space. The TS algorithm starts or restarts from three initial parameters $\alpha_0, \beta_0$ and $\gamma_0$, set at respectively 1, 1, 100. We set three intervals $[\alpha_{\text{min}}, \alpha_{\text{max}}]$, $[\beta_{\text{min}}, \beta_{\text{max}}]$ and $[\gamma_{\text{min}}, \gamma_{\text{max}}]$ to limit these parameters during search process in the following respective ranges [0.01, 1000], [0.01, 1000] and [0.01, 5000]. At each TS iteration, penalty parameter for the vehicle capacity constraint $\alpha$ is modified as follows. If the solution generated after route re-optimization is feasible, $\alpha$ is divided by a factor $1+\varphi_1$. If the solution is infeasible and the vehicle capacity is violated, $\alpha$ is multiplied by a factor $1+\varphi_2$. If the solution is infeasible but the vehicle capacity is satisfied, parameter $\alpha$ is divided by a factor $1+\varphi_2^2$ $(0 < \varphi_2 \leq \varphi_1)$. And, parameters $\beta$ and $\gamma$ are adjusted in the same rules. The following parameters $\varphi_1 = 0.2$ and $\varphi_2 = 0.05$ are used. Note these parameters, e.g., $\alpha, \beta, \gamma, \varphi_1$ and $\varphi_2$, are tuned by the preliminary experiments, as well as some guidelines in previous studies on unified tabu search (J. Cordeau, et al., 1997). The sensitivity analyses on the parameters were performed sequentially, leaving the remaining parameters unchanged.

5.2 Neighborhood structure

Recall that each solution $Sol$ in our TS partitions the set of patients into different routes and is completely represented by its set of routes with each route represented by the sequence of nodes visited including the depot, the hospital and the lab.

The neighborhood of a solution $Sol$ is defined by two local moves. The first local move for a solution $Sol$ removes a patient from one route $k$ and inserts it into another nonempty route $k'$. The second local move consists in exchanging the route assignment of two patients $i$ and $j$ in different routes $k$ and $k'$. The relocation move can be denoted as $(i, k')$ and route exchange move denoted as $(i, j, k, k')$

More specifically, relocation move $(i, k')$ of a patient on route $k$ consists in (i) removing patient $i$ from route $k$; (ii) inserting it in route $k'$ at a position that minimize the cost of route $k'$, i.e.

$$w_i(Sol) + \alpha \cdot d_i(Sol) + \beta \cdot e_i(Sol) + \gamma \cdot g_i(Sol),$$

and (iii) removing the hospital (lab) in route $k$ if there is no more $D_2 (P_2)$ patient in route $k$ after the local move. Route exchange move $(i, j, k, k')$ of two patients on different routes $k$ and $k'$ consists in (i) removing patients $i$ and $j$ from their routes, (ii) inserting patient $i (j)$ in route $k' (k)$ at a position that minimize the cost of route $k' (k')$, and (iii) removing any unnecessary hospital and lab visit in routes $k$ and $k'$. The cost of the neighbor solution is the augmented cost function of the solution obtained after the local move.

Note that when inserting a patient in route $k$, we do not relocate or insert the hospital visit and the lab visit in the route even if it is necessary to satisfy the new patient. This is due to the relative long computation time needed to find the best position for inserting hospital or lab. Nevertheless, the relocation or insertion of hospital and lab visits are considered in the re-optimization of modified route once the next local move has been selected.
Other local moves such as *intra-route crossover* method and 2-opt* (Alfredo Tang Montané and Galvão, 2006) were also used in the literature for solving the VRPs and VRPSDP. We tested such local moves in our TS and did not observe significant improvement. Since most of our TS running time is spent by the neighborhood search, implement more local moves has a clear disadvantage of increasing very substantially the size of the neighborhood and the computation time. For this reason, we limit our TS to two local moves.

### 5.3 Re-optimization of a modified route

Each local move in our TS can modify two routes. A modified route might miss necessary hospital or lab visit, violate vehicle capacity, time window and precedence constraints. This subsection presents a method to rebuild and improve each modified route. The remaining part of this subsection concerns the re-optimization of a given modified route.

The re-optimization starts with a constructive method and then improves the route by local search. More specifically, it starts with a simple and effective constructive heuristic, the *nearest insertion* (NI) (Bentley, 1992). Our NI heuristic is a three-step method.

**Step1:** Delete all the existing trips in this route. Then create a partial route beginning and ending at the depot and visiting the farthest patient.

**Step2:** Insert the patient whose insertion generates the smallest increment on the cost \( w(Sol) + \alpha \cdot d(Sol) + \beta \cdot e(Sol) \), i.e., travel cost plus penalty of the vehicle capacity and time window violation. This step is repeated until all patients are inserted.

**Step3:** Insert the lab and hospital into the route at feasible positions with the smallest increment on the cost \( w(Sol) + \alpha \cdot d(Sol) + \beta \cdot e(Sol) \). Precedence constraints are satisfied and \( g(Sol) = 0 \) at this step, i.e. the hospital is visited before all \( D_2 \) patients and the lab is visited after all \( P_2 \) patients.

The NI route is not necessarily good enough or even feasible for our problem. It is further improved by local search comibing the first-accept strategy and three moves: 1-1exchange, 1-0 relocation, 2-Opt exchange. Before the local search, we check the feasibility of the NI route. If the NI route is feasible, infeasible solutions are not allowed in local search. If the NI route is not feasible, infeasible neighbor solutions are allowed in local search and the augmented criterion function \( f(Sol) \) is used.

### 5.4 Attribute set, Tabu list, Tabu duration and Aspiration criterion

In our TS, we associate with each solution \( Sol \) an attribute set \( At(Sol) = \{(i, k)|i \in N, k \in K\} \) indicating for each patient \( i \) the vehicle \( k \) serving it. Of course, the attribution set is only a partial characterization of the solution used to define the tabu list. The relocation move \((i, k')\) of a patient \( i \) in route \( k \) to a different route \( k' \) is equivalent to replace an attribute \((i, k)\) from \( At(Sol) \) by a new attribute \((i, k')\) with \( k \neq k' \). The route exchange of two patients in different routes can be seen as replacing two attributes in \( At(Sol) \).

TS utilizes adaptive memory, called tabu list and tabu duration, to implement a
diversification strategy. In our TS, when a patient \( i \) is removed from a route \( k \), we assign a tabu status to the attribute \((i, k)\), and set a tabu duration \( \theta \) to this attribute. That is to say, in the next \( \theta \) iterations, inserting patient \( i \) back into route \( k \) is forbidden when we perform the neighborhood search. The size of the tabu list \( \theta \) takes its values in \([\theta_{\min}, \theta_{\max}]\) and starts from \( \theta_0 \). Tabu list \( \theta \) is also a self-adjusting parameter, and dynamically modifies during the TS search. After each improvement of the current best solution \( S_{\text{best}} \), we set parameter \( \theta \) equal to \( \theta_{\min} \). After \( \theta_\phi \) consecutive times unimproved iteration, parameter \( \theta \) is updated to be \( \min(\theta + 1, \theta_{\max}) \). In the preliminary experiments, we find \( \theta_0 = 7 \), \( \theta_\phi = 30 \) suit most test instances. It is not necessary to increase the value of \( \theta_\phi \) with the instance size (customer number). And, we set \([\theta_{\min}, \theta_{\max}] = [5, 15]\) in our TS implementation.

The tabu status of a local move can be overridden by an aspiration criterion in our TS. We define an aspiration value for each attribute, which is equal to the cost of the best feasible solution found with that attribute. For each feasible solution found after the re-optimization of modified routes, we update the aspiration value of each attribute of this feasible solution. A relocation move \((i, k')\) is considered at an iteration if \((i, k')\) is not in the tabu list, or the neighbor solution is feasible and its cost is smaller than the aspiration value of \((i, k')\). A route exchange move \((i, j, k, k')\) is considered if \((i, k')\) and \((j, k)\) are not in the tabu list, or the neighbor solution is feasible and its cost is smaller than the smallest aspiration value of \((i, k')\) and \((j, k)\).

6 Computational experiments

This section reports the results of a series of computational experiments for comparison of the genetic algorithm, the tabu search, and application of the commercial solver Cplex 12.3 for the two mathematical formulations MIP1 and MIP2 of Section 3, and other methods of the literature for some special cases of our problem.

To the best of our knowledge, this paper is the first study of this special vehicle scheduling problem in the home health care industry. There are no benchmark instances to evaluate the performances of our heuristic approaches. Therefore, we construct some test instances based on existing VRPTW benchmarks. Further, as our problem with only \( P_1 \) or \( D_1 \) patients reduces to the classic VRPMBTW, our approaches are also compared with existing VRPMBTW approaches on existing benchmark instances for the VRPMBTW.

As our problem is highly combinatorial, the performance of Cplex solver strongly depends on its parameterization. We tried different Cplex parameters including default settings, strong branching, depth-first search, MIP emphasis feasibility or optimality. No parameter setting led to satisfactory performance. For most of the small-size test instances of this Section, Cplex cannot find a feasible solution after 48 hours. It does not make sense to compare our heuristic solutions with direct implementation of MIP models in Cplex. Instead, we use in this section \( \text{MIPstart} \) strategy by letting Cplex to start with the initial solution of our TS, i.e. the best initial solution of our GA. Cplex installs it as the incumbent and initial
solution of its branch and cut procedure, which allows Cplex to eliminate portions of the search space and results in smaller branch and cut trees. Cplex starts from this initial solution and goes on solving the problem based on this solution, until exhausting the memory or predetermined maximum computation time. In our preliminary experiments, we also find that Cplex with MIP2 formulation always outperforms Cplex with MIP1. For this reason, this section limits to Cplex with MIP2 formulation. Further, the Cplex lower bounds are very poor and are not given in this paper.

All the algorithms of this paper are implemented in C. All heuristic algorithms (GA, TS,...) are carried out on a 3.2 GHz Dual Core computer with a 2 GB memory under Linux. We set a time limit of 72 hours and a memory limit of 20 GB for Cplex for each instance. After some preliminary experiments, the parameters of the proposed GA and TS have been set to the values reported earlier in this paper. For fair comparison with GA, our tabu search algorithm is implemented with two stopping criteria: (i) maximum number of iterations and maximum number of iterations without improvement, and (ii) same computation time as GA. The first tabu search is denoted TS1 and the second one TS2. GA, TS1 and TS2 run 10 times for each instance. The best results, the average results and average running time are used to assess the efficiency of these algorithms.

6.1 Test Instances from VRPTW benchmarks

We first derive test instances from existing VRPTW benchmarks of (Solomon, 1987) and (Gehring and Homberger, 1999). Eighteen Solomon VRPTW instances are selected to generate our test instances. Each Solomon instance contains 100 customers over a service region defined on a 100×100 grid. These VRPTW instances are divided into three classes that differ by the geographical distribution of the customers: they are clustered in the C type instances, randomly located in the R type instances, and partly clustered, partly randomly located in the RC type instances. Meanwhile, each class is divided into two series: in the 100-series instances time windows are tighter, and in the 200-series instances time windows are wider. To test different characteristics of instances, we select 6 C type instances, 6 R type instances and 6 RC type instances. Among 6 instances of each type, both the 100-series instances and 200-series ones exist.

For each Solomon instance, we derive 6 new instances for our problem with 4Z demands as follows. First, we randomly choose Z customers from the Solomon instance as the $P_1$ patients in our new instance, each of which has a demand equal to 50% of the customer’s demand given in the Solomon instance. Then, Z $P_2$, Z $D_1$ and Z $D_2$ patients are randomly selected from the Solomon instance. Clearly, one patient may be selected more than once and the number of patients is less than then number of requirements (4Z). If the basic Solomon instance is C101 and 37 patients and 40 demands exist in our new instance, it is denoted as C101-37-40. The coordinates of the depot is inherited in our instances, and the locations of lab and hospital are (10, 15) and (40, 50). For each patient, the time window in the Solomon instance is used directly. Time windows for the depot, lab and hospital are selected as follows
to avoidance infeasible solutions. The depot’s time window of the Solomon instance is multiplied by 1.2 and assigned to the depot, lab and hospital in our instance. If ‘violative’ patients still exist in our instance, the new time window is repeatedly multiplied again by 1.2 until all ‘violative’ patients are eliminated. In the preliminary experiment, we find that all ‘violative’ patients disappear after two tries.

For each Solomon instance, this constructing procedure are repeated 6 times, generating 2 small (40 demands), 2 moderate (80 demands), and 2 large (120 demands) instances. Concerning the vehicle capacity and vehicle number, their values have been reduced compared to the ones considered for the VRPTW, because they are loose for our problem. The detailed information about these two characteristics is illustrated in Tables 4-7.

Besides the Solomon instances, we also create the largest instances from VRPTW instances of (Gehring and Homberger, 1999). These instances are similar but larger than Solomon instances and have hundreds of customers. We choose 12 instances from this benchmark, each of which has 400 customers, and undergoes the procedure described above to generate 24 new instances for our study. Each of these instances contains 200 demands, and also named as the VRPTW instance’s label with patients number, and demands number.

### 6.2 Computational Results on VRPTW-based instances

In this subsection, Tables 1 and 2 summarize the results obtained from the GA and TS on all VRPTW-based instances. The detailed computational results obtained on the small instances (containing 40 demands), moderate instances (80 demands), large instances (120 demands), and the largest instances (200 demands) are presented in four Tables 4-7 in the Appendix, respectively.

<table>
<thead>
<tr>
<th>Demand #</th>
<th>Type</th>
<th>Cplex</th>
<th>GA</th>
<th>TSI</th>
<th>TS2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Best</td>
<td>AVG</td>
<td>Worst</td>
</tr>
<tr>
<td>50</td>
<td>C</td>
<td>1065.7</td>
<td>881.9</td>
<td>890.5</td>
<td>895.6</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>1051.0</td>
<td>875.3</td>
<td>889.0</td>
<td>901.2</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>1158.3</td>
<td>954.4</td>
<td>966.4</td>
<td>977.2</td>
</tr>
<tr>
<td>80</td>
<td>C</td>
<td>1776.8</td>
<td>1424.3</td>
<td>1457.4</td>
<td>1480.7</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>1680.7</td>
<td>1409.1</td>
<td>1441.1</td>
<td>1463.9</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>1867.8</td>
<td>1577.6</td>
<td>1617.5</td>
<td>1648.9</td>
</tr>
<tr>
<td>120</td>
<td>C</td>
<td>1965.6</td>
<td>1503.9</td>
<td>1542.1</td>
<td>1565.1</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>2012.8</td>
<td>1505.4</td>
<td>1553.5</td>
<td>1575.9</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>2181.7</td>
<td>1691.3</td>
<td>1736.4</td>
<td>1764.8</td>
</tr>
<tr>
<td>200</td>
<td>C</td>
<td>9744.1</td>
<td>8267.5</td>
<td>8387.2</td>
<td>8485.5</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>10759.9</td>
<td>9219.3</td>
<td>9380.8</td>
<td>9499.1</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>9773.1</td>
<td>8580.1</td>
<td>8706.5</td>
<td>8808.7</td>
</tr>
<tr>
<td>AVG</td>
<td></td>
<td>3176.8</td>
<td>2654.6</td>
<td>2704.0</td>
<td>2739.6</td>
</tr>
</tbody>
</table>
Table 2. Percentage of best solutions found of GA and TS on VRPTW-based instances

<table>
<thead>
<tr>
<th>Demand#</th>
<th>Type</th>
<th>Cplex</th>
<th>GA</th>
<th>TS1</th>
<th>TS2</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>C</td>
<td>0%</td>
<td>83%</td>
<td>92%</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0%</td>
<td>58%</td>
<td>67%</td>
<td>58%</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>0%</td>
<td>75%</td>
<td>75%</td>
<td>67%</td>
</tr>
<tr>
<td>80</td>
<td>C</td>
<td>0%</td>
<td>42%</td>
<td>42%</td>
<td>58%</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0%</td>
<td>42%</td>
<td>17%</td>
<td>58%</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>0%</td>
<td>50%</td>
<td>25%</td>
<td>42%</td>
</tr>
<tr>
<td>120</td>
<td>C</td>
<td>0%</td>
<td>50%</td>
<td>42%</td>
<td>67%</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0%</td>
<td>58%</td>
<td>17%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>0%</td>
<td>42%</td>
<td>25%</td>
<td>42%</td>
</tr>
<tr>
<td>200</td>
<td>C</td>
<td>0%</td>
<td>13%</td>
<td>63%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0%</td>
<td>25%</td>
<td>38%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>0%</td>
<td>13%</td>
<td>63%</td>
<td>25%</td>
</tr>
<tr>
<td>AVG</td>
<td></td>
<td>0%</td>
<td>48%</td>
<td>46%</td>
<td>53%</td>
</tr>
</tbody>
</table>

Table 1 shows the average routing costs by grouping problem instances according to the number of demands and the type of the instance. The results are obtained from 10 independent runs for each problem instance of the three approaches (GA, TS1, and TS2) plus the one obtained with CPLEX. Column ‘Best’ is the average over all relevant problem instances of the best solutions among 10 independent runs of an approach for each instance. Column ‘AVG’ is the average travel cost over all problem instances and over all runs of the approach. Similarly, column ‘Worst’ represents the average value of the worst solution costs among relevant test instances. We calculate the standard deviation of 10 runs of each test instance, and Column ‘STD’ gives the average standard deviation among all relevant test instances. Column ‘CPU’ is the average CPU time in seconds of one run among relevant test instances. The last line in Table 1 provides the average values for all the test instances. Table 2 shows the percentage of the best solutions obtained by GA, TS1 and TS2 on each type of test instances.

Several conclusions can be drawn from these experimental results. First, the proposed GA and TS (both TS1 and TS2) perform well for test instances of different types and different sizes. For each combination of types and sizes, our heuristic algorithms significantly dominate the Cplex solver. For all 132 test instances, the Cplex solution costs and best GA solution costs deviate on average by 16.4%; the Cplex solution costs and TS1, TS2 best solution costs deviate on average up to 17.0% and 17.1%, respectively. Recall that both Cplex and TS start from the best solution of the initial GA population built by simple constructive heuristics, random generated solutions and the optimal split procedure. This implies that GA and TS can significantly improve the solutions of these heuristic solutions and our algorithmic approaches are highly competitive.

Concerning the solution quality of the GA and TS approaches, the performances of all three heuristics are satisfactory. For example, as shown in Table 2, out of 36 small size test
instances, GA, TS1 and TS2 can get 72%, 78%, 72% best solutions, respectively. For 17 instances out of all 36 small size instances, GA, TS1 and TS2 all find the same solution (referred to Tables in Appendix). Among the total 132 test instances, GA, TS1 and TS2 are able to find 48%, 46%, 53% best solutions, respectively. Note that Cplex does not get any best solution and the Cplex solution is always far from the best solutions found by other approaches.

Meanwhile, we find that our approaches GA, TS1 and TS2 are robust and can find good solutions for different test instances in different runs. As shown in Table 1, for all the test instances, the average standard deviation of GA, TS1 and TS2 are 25.5, 24.8 and 24.1. Comparing GA and TS, we find that TS1 is slightly better the GA. For the total 132 test instances, the best solution costs of GA and TS1 deviate on average by 0.71%; their average and worst solution costs deviate by 0.79% and 0.97%, respectively. The superiority of TS1 is especially true for problem instances of the largest size. For example, as shown in Table 1, for the largest 24 instances, the best solution costs of GA and TS1 deviate on average by 1.29%; the deviation between their average and worst solution costs is 1.52% and 1.93%. Meanwhile, Table 2 shows that TS1 succeeds in finding 55% best solutions of the largest instances, while GA is able to find 17% best solutions. Nevertheless, for these instances, TS1’s improved solution quality is obtained at the cost of longer computation time. Over the 132 test instances, the CPU time of TS1 is on average 9.28% longer than that of GA.

With the same computation time between GA and TS2, the performance of TS2 is slightly better than that of the GA. Over 132 test instances, concerning their best solution, TS2 is slightly better than GA on the average best solution costs with a deviation of 0.76%. Meanwhile, TS2 can find more best solutions than GA, with 48% best solutions found by GA and 53% by TS2.

6.3 Test on VRPMBTW benchmarks

This subsection restrict our problem to only $P_1$ and $D_1$ patients and our problem reduces to the classic VRP with mixed backhauls and time windows (VRPMBTW). Our GA and TS are tested on the benchmarks for the VRPMBTW against existing best solutions. (Gélinas, et al., 1995) construct test data for the VRPMBTW against existing best solutions. They constructed VRPMBTW instances by randomly choosing 10%, 30% and 50% of the 100 total customers in Solomon VRPTW benchmark to be backhaul customers instead. (Hasama, et al., 1998) designed 15 VRPMBTW test problems from the data of (Gélinas, et al., 1995) by relaxing the linehaul-before-backhaul constraint, and proposed a simulated annealing (SA) heuristic method for the VRPMBTW. We test our two heuristics, GA and TS2, on Hasama’s problem instances.

The results obtained by our algorithms and those provided the SA algorithm of (Hasama, et al., 1998) are detailed in Table 3. Column ‘Name’ gives the label of the instance. Column ‘BH%’ presents the percentage of the backhaul among all customers. Columns 3 and 4 contain
the best solutions found by Hasama’s SA algorithm, where ‘Cost’ shows the travel distance and ‘V’ shows the number of the vehicles used. The next six columns contain the best, the average and the worst solution cost, the standard deviation, the average CPU time (for one run) and minimum vehicle number used in 10 GA runs for each instance. The number of vehicles used in the minimum cost solution is also given in parenthesis if it is larger than V. Similar results are given for TS2.

Table 3. Results on VRPMBTW test instances

<table>
<thead>
<tr>
<th>Instance</th>
<th>SA</th>
<th>GA</th>
<th>TS2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>BH%</td>
<td>Cost</td>
<td>V</td>
</tr>
<tr>
<td>R101</td>
<td>10</td>
<td>1686.5</td>
<td>21</td>
</tr>
<tr>
<td>R101</td>
<td>30</td>
<td>1693.4</td>
<td>20</td>
</tr>
<tr>
<td>R101</td>
<td>50</td>
<td>1696.4</td>
<td>21</td>
</tr>
<tr>
<td>R102</td>
<td>10</td>
<td>1753.4</td>
<td>20</td>
</tr>
<tr>
<td>R102</td>
<td>30</td>
<td>1583.0</td>
<td>19</td>
</tr>
<tr>
<td>R102</td>
<td>50</td>
<td>1602.7</td>
<td>19</td>
</tr>
<tr>
<td>R103</td>
<td>10</td>
<td>1250.6</td>
<td>15</td>
</tr>
<tr>
<td>R103</td>
<td>30</td>
<td>1259.1</td>
<td>16</td>
</tr>
<tr>
<td>R103</td>
<td>50</td>
<td>1495.5</td>
<td>17</td>
</tr>
<tr>
<td>R104</td>
<td>10</td>
<td>1078.2</td>
<td>11</td>
</tr>
<tr>
<td>R104</td>
<td>30</td>
<td>1125.3</td>
<td>12</td>
</tr>
<tr>
<td>R104</td>
<td>50</td>
<td>1177.4</td>
<td>12</td>
</tr>
<tr>
<td>R105</td>
<td>10</td>
<td>1479.4</td>
<td>16</td>
</tr>
<tr>
<td>R105</td>
<td>30</td>
<td>1417.4</td>
<td>15</td>
</tr>
<tr>
<td>R105</td>
<td>50</td>
<td>1464.5</td>
<td>15</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>1450.9</td>
<td>1358.7</td>
</tr>
</tbody>
</table>

From Table 3, our GA and TS2 algorithms dominate the SA algorithm of (Hasama, et al., 1998) in terms of solution quality. The primary objective of (Hasama, et al., 1998) is to minimize the number of vehicles used to serve the customers, and the second objective is to minimize the total travel distance. Therefore, we first compare the number of vehicles. Although our heuristics do not use this as their objective, the proposed GA still can reduces the number of vehicles needed for 4 out of the 15 instances, and TS2 reduces the number of vehicles for 6 instances. Concerning the travel distance, the solutions of our GA and TS2 are much better than the solution produced by (Hasama, et al., 1998). For each instance, GA and TS2 can get better solution. The average percentage deviations between the best solution costs of our GA and TS2 and Hasama’s SA are 6.35% and 6.27% respectively.

7 Conclusions and future research

This paper investigates a special simultaneous pickup and delivery problem with time windows in HHC, an extension of the classical VRPSDPTW. The problem is of interest because of its theoretical complexity and of the important applications in the home health care
industry. We formulate the problem as two integer programming models to minimize the total vehicle cost for serving all patients demands. We also propose two meta-heuristics, Tabu search and genetic algorithm, to solve this problem. As this problem is new and no benchmark exists, experiments are conducted by using a range of test instances, which are designed based on existing VRPTW benchmarks to reflect different realistic scenarios. In general, both TS and GA can provide good solutions in a reasonable time span, and TS requires relatively more computational time. Our proposed heuristic approaches are also tested on a set of VRPMBTW benchmarks against best known results. The results of our metaheuristics are clearly better than the best-known solutions in the existing literature.

This research can be extended in different directions. Firstly, the problem can be extended to a planning horizon of several days, e.g., a week, to combine planning of delivery/pickup and vehicle routing. In real-life applications, usually each patient requires a certain number of visits and services within this time horizon. The HHC has to choose the visiting days for each patient and to solve a vehicle scheduling problem for each day. In the vehicle scheduling problem of each day, some special constraints and conditions encountered in the HHC must be considered. This problem is a special periodic vehicle scheduling problem and even more complex than the problem studied in this paper. Meanwhile, in many real-world home health care logistic applications, some elements in the problem may not be known in advance, e.g., patients’ delivery/pickup demands, travel and service times of the vehicle. Such uncertain elements significantly affect the system performance. The vehicle scheduling problems in HHC under such uncertain conditions are also very important and interesting.

Acknowledgement

This research is partially funded by Saint Etienne Metropôle for providing a Post-Doc fellowship for the first author. The authors are also grateful to OIKIA home health care company for sharing information for home health care logistics.

References


---

25


