A Matheuristic Algorithm Applied to the Home Health Care Problem

Author A, Author B, Author C

RESUMO

Este trabalho estuda a aplicação de uma *matheurística* para resolver o Roteamento e Escalonamento de Cuidadores para Pacientes em Domicílio. Este problema é similar ao Problema de Roteamento de Veículos com janela de tempo, e conta com restrições adicionais que modelam a sincronização espaço-temporal dos veículos durante a visita de alguns dos vértices do problema. O problema de atendimento em domicílio é uma alternativa a internações hospitalares de pacientes estáveis. Uma *matheurística* "fixa e otimiza" foi empregada, na qual são resolvidos iterativamente vários subproblemas que consideram a otimização de pares de rotas. Um banco de instâncias da literatura do problema foi utilizado nos experimentos computacionais. Resultados indicam que a *matheurística* é competitiva frente aos médotos de solução da literatura, com margens de ganho de até 17% em relação aos melhores resultados conhecidos.

PALAVRAS CHAVE. Roteamento de veículos, Janelas de Tempo, Interdependência entre rotas.

OC – Otimização Combinatória

ABSTRACT

This work study a *matheuristic* applied to the Home Health Routing and Scheduling Problem. The home care problem is an extension of the Vehicle Routing Problem with Time Windows. It considers additional constraints to model both spatial and temporal synchronization of caretakers during the routes. Home care is an alternative to hospital internment of stable patients. A fix and optimize *matheuristic* is applied to solve several subproblem iteratively, considering pairs of routes. An instance dataset of the literature of the problem is used in computational experiments. The results indicate that the *matheuristic* is competitive to solution methods of literature, achieving savings up to 17% on best known solutions from previous studies.

KEYWORDS. Vehicle routing. Time windows. Route interdependency.

OC – Combinatorial Optimization

1. Introduction

The increase in the elderly population is a worldwide observed phenomenon. Advances in agriculture that have allowed substantial increase in food supply and preventive medicine contributed to enlarge the life expectancy of 38 years to 50+ years in the last 100 years [Nasri, 2008]. Despite that, most countries have few strategies or regulations to manage and assist this increasing portion of the population. This aging effect also happens in Brazil. The projections of the Brazilian Institute of Geography and Statistics indicates that a quarter of the population will be above 60 at the end of the next half-century [IBGE, 2018].

It is natural that the elderly brings limitations to people. Depending upon how severe are the issues imposed by the advanced age, they might hamper a few day-to-day activities. Typically, the only option that provides the specialized care needed is nursing homes. Except for more critical situations, home care is a less disruptive alternative that enables the elderly to live normally with their family. The broader home care services are available, and the more population can take advantage of it. Moreover, not only the elderly might be benefited from home care services. Stable patients may proceed with their recovery out of hospitals, reducing their exposition to pathogens. Staying on home also helps to reduce psychological burdens. For hospital managers, home care increases the availability of beds as patients can be sent to continue, or either finish their treatments at home.

With home health care services, the population may receive specialized assistance at home. Going further the health services, home care can be used to assist with daily activities. Thus, the demand for home care services can be high. Cooking, for example, should be done daily. Meanwhile, laundry and housekeeping might be done once a week. The constant presence of professionals also helps to reduce the risk of accidents of patients. The constant presence of trusted professionals also brings tranquility to patients' family.

It will be problematic for governments to postpone discussions about home care services. The subtle peculiarities introduced by the home assistance require personnel with specific formation. Furthermore, the resources needed by home care services may vary abruptly over time, requiring efficient management from home care agencies. Another challenge is to have an efficient operation of home assistance services, what can be seen as an optimization problem. In this sense, physician and nurse rostering problems are inherent to this context. Additionally, the schedule problems must consider the distinct abilities of caretakers, as well as the demands of home care patients. Both optimization problems also have a logistic component, which handles the routing of caretaker to service users.

The remaining of this paper is structured as follows. Section 2 presents a brief literature review of home care problem, regarding relevant variations, and solution methods. Section 3 defines precisely the home care problem considered in this work through a mixed integer formulation. Section 4 presents a Fix and Optimize *matheuristic* to solve the problem. Section 5 presents the dataset used in computational experiments and compare the results found with the results of the literature of the problem. The paper finishes in Section 6. This section presents an overall review of the paper and highlights its major findings.

2. Related work

One of the earliest publications about home care appeared in the scientific literature 45 years ago. The work of Fernandez et al. [1974] was a result of joint efforts of three countries with a common problem: enable the access of population of a predominantly rural country to health services. The authors proposed a home care service to achieve this objective. To approach the problem, teams up to 5 nurses were considered. The geographical region was subdivided into districts, following the distribution of villages. Each team must be assigned to perform home attendances on a specific village, thus each village can be tackled as an isolated home care problem. Those simplifications were taken to work around the difficulty in solving the home care problem due to the lack of computational power and programming tools. Instead, the entire publication discussed estimation on distance performed by nurses using statistical frameworks.

In the following years, home care problems have been receiving increasing attention. More sophisticated and realistic approaches were considered thanks to the availability of powerful optimization tools. Eveborn et al. [2006] is the earliest publication on home care literature that resembles an operations research paper. The authors presented a set partitioning model to solve the assignment of nurses to patients. Due to the inefficiency of exact solvers of the time, the authors proposed a heuristic solving method based on repeated minimum cost matching algorithm to solve the assignment problem. Two objectives were considered, regarding the quality of routes and quality of service. The first objective seeks routes with minimum traveling distances. The second objective aims to reduce the variety of nurses that visit each patient, maximizing patient comfort. Computational experiments were conducted considering 28 caretakers and 150 patients in a study case using real data from a home care provider. The obtained solutions were compared to the one developed by the home care agency that provided the data. Both savings on planning time and traveling time were reported, without sacrificing the quality of service.

[Akjiratikarl et al., 2007] approached the home health care problem using a Particle Swarm Optimization metaheuristic (PSO) [Eberhart and Kennedy, 1995]. The authors were the first to establish a relation between the Home Care Problem (HCP) and the Vehicle Routing Problem with Time Windows (VRPTW). The former is defined as an extension to the VRPTW, with additional constraint and a weighted objective function. In VRPTW terminology, patients correspond to nodes, and caretakers correspond to vehicles. In the home care problem, an efficient route of each caretaker should be determined, starting on the depot, visiting a subset of patients, and returning to the depot at the end of operations. The assigned nodes define the scheduling of the caretakers, which should respect the time windows of patients. A constructive heuristic is used to obtain a feasible initial solution, which is later improved by a PSO procedure. Using real instances provided by a UK consortium of home care agencies, the authors reported savings up to 30% of travel distances of caretakers, when compared to solutions used in practice.

Motivated by the growing gap between the offer and the demand for home care in Austria, Trautsamwieser and Hirsch [2011] carried the research for practical solution methods to the home care problem. Such a method should be able to cope with a wide range of demands, considering language compatibility between nurses and patients, among other assignment preferences. The authors presented a sophisticated but realistic multiobjective MIP model, considering several practical soft constraints and regulatory rules. Such model considers client time windows (hard and target), contract workload, overtime and programmed breaks between shifts. The authors tested both MIP and a metaheuristic solving approaches on instances generated with data from Austrian Red Cross provider. Due to high memory and processing times demand, the model proved to apply only to small instances (up to 20 services types/jobs, 20 clients and 4 nurses). For larger instance sizes, an initial solution was created by a constructive heuristic. A VNS procedure then improves the solution. The authors reported results to instances up to 420 clients and 75 nurses, with a total request of 512 jobs. Note that some clients requested more than one job. The authors performed a comparison to the solutions from the health care provider and savings up to 45% of traveling distances were found. Results of sensitivity analysis indicated that the saved time can be used to enhance the service levels of patients.

Drexl [2012] introduce a survey of VRP problems with synchronization constraints. The survey considered VRP literature from early 1986 to 2010. A taxonomy was introduced to describe and classify VRP problems following their synchronization requirements. To cite some examples, some of the synchronization requirement found were: tasks synchronizations [Azi et al., 2010], operations synchronization [Mankowska et al., 2014], and synchronization during load and discharge [Desaulniers, 2010]. The task synchronization is the most common, and every VRP with multiple vehicles have it: the assignment of exactly one vehicle to each node of the problem implies a synchronization on the nodes. The author ends the discussion indicating that synchronization requirements should become more and more frequently on VRP literature since they are inherent on practice of several routing problems.

Relying on the home care problem with multiple service types and route interdependency, Mankowska et al. [2014] introduced the Home Health Care Routing and Scheduling Problem (HHCRSP). This problem consists of a VRPTW, in which each patient may request double services. In double services, a patient has two distinct service requests that should be performed in a specific order, or simultaneously, by two distinct caretakers. The former is called double services with precedence constraints, and the later is called double service with simultaneous attendance time. Such requirements implies in operations synchronization, following the terminology of Drexl [2012]. The authors proposed a dataset of 70 randomly generated instances, that are available on a web site[1]. The sizes of the instances range from 10 to 300 clients, and 3 to 40 caretakers. The authors considered a weighted objective function with three terms: minimizing travel distances, minimizing total tardiness (service starting before patient time window), and the largest tardiness observed between all clients. Within 10 hours of an execution, the MIP model was capable of producing optimal solutions only for the smaller instance subset (10 clients and three caretakers). To tackle instances of larger sizes, the authors proposed an Adaptive VNS metaheuristic (AVNS) that adjusts the sequence of neighborhood movements according to their previous contribution on improving the solution cost. Initial solutions were obtained using a similar approach to constructive approach of Akjiratikarl et al. [2007]. The metaheuristic solving time was limited to two hours. On small instances, the metaheuristic approach produced solutions with gaps of 1.3% to optimality. On larger sizes, the gap to the LP lower bound ranges substantially between 2% and 425%.

Several other variants and solving methods can be found on home health care literature, *cf.* [Fikar and Hirsch, 2017]. Most of the researches point to the advantages of optimizing solutions towards manual solution methods. Savings in traveling times and improvement on levels of services are the major reflects on using more efficient solutions. To the best of authors effort, there is just one scientific paper approaching home health care problems on entire South America [Gutiérrez and Vidal, 2013]. This study is an entry point to optimization tools on home health care services in Brazil. The Brazilian home health care services are in the early stages and are being tested by the SUS System in capitals of several Brazilian states [Ministério da Saúde, 2013]. Although the requirements are not quite set, the problem proposed by Mankowska et al. [2014] is a good start point to research. MIP models are the most flexible tools regarding changes in problem definition, and mathematical programming solvers can be applied ubiquitously.

3. Problem definition and formulation

The problem addressed in this work is the same proposed in Mankowska et al. [2014]. Let $\mathcal{C}$ be the set of patients, or clients, and $\mathcal{C}^0 = \mathcal{C} \cup \{0\}$ be the set of patients plus the depot node. The parameter $\mathrm{d}_{ij} \geqslant 0$ defines the travel distance between locations $i, j \in \mathcal{C}^0$. Let $\mathcal{V}$ the set of staff members/vehicles and $\mathcal{S}$ the set of service types offered by the home care provider. The parameter $\mathrm{r}_{is} \in \{0, 1\}$ indicates if patient $i \in \mathcal{C}$ requests a service type $s \in \mathcal{S}$. Similarly, the parameter

[1] https://prodlog.wiwi.uni-halle.de/forschung/research_data/hhcrsp/

$\mathrm{a}_{vs} \in \{0,1\}$ indicates if staff member $v \in \mathcal{V}$ can perform the service type $s \in \mathcal{S}$. Each service $s \in \mathcal{S}$ requested by a patient $i \in \mathcal{C}$ demands p_{is} units of time to be done. Each patient $i \in \mathcal{C}$ has a time window $[\mathrm{e}_i, \mathrm{l}_i]$ in which services can start. Services must begin not before the starting time e_i and, preferably, should start not after the ending time l_i, otherwise a tardiness is computed.

If a patient $i \in \mathcal{S}$ requires double services with precedence between them, there are minimum $\delta_i^{\min}$ and maximum $\delta_i^{\max}$ separation times between the two requested services. If double services with simultaneous attendance times are required, then in this case $\delta_i^{\min} = \delta_i^{\max} = 0$. The order of double services with precedence is defined globally. Supposing $\mathcal{S} = \{1, 2, \dots\}$, a global precedence defines that service type $i - 1$ must occur before service type i, $\forall i \in \mathcal{S} : i > 1$. The set $\mathcal{C}^{\mathrm{d}} \subset \mathcal{C}$ represents the subset of patients that require double services.

The decision variable $x_{ijvs} \in \{0,1\}$ indicates if staff member $v \in \mathcal{V}$ departs from $i \in \mathcal{C}^0$ to $j \in \mathcal{C}^0$ to perform the service type $s \in S$. The auxiliary variables $t_{ivs} \geqslant 0$ indicates in which time the attendance of patient $i \in \mathcal{C}$ by staff member $v \in \mathcal{V}$ begins for service type $s \in \mathcal{S}$. The auxiliary variable $z_{is} \geqslant 0$ holds the amount of tardiness when the service type $s \in \mathcal{S}$ starts before the time window ending e_i of patient $i \in \mathcal{C}$.

As mentioned earlier, the HHCRSP has three minimizing criteria, computed in variables $D, T, T^{\max} \geqslant 0$. The first criteria minimizes the length of routes performed by the staff members. The second criteria minimizes the total tardiness, and the largest tardiness observed over all patients, respectively. There are three multiplicative factors λ_1, λ_2 and λ_3 that combine these criteria in a single objective function.

The HHCRSP can be formally defined as follows.

$$\text{Minimize} \quad \lambda_1 D + \lambda_2 T + \lambda_3 T^{\max} \tag{1}$$

Subject to:

$$D = \sum_{v\in\mathcal{V}} \sum_{i\in\mathcal{C}^0} \sum_{j\in\mathcal{C}^0} \sum_{s\in\mathcal{S}} \mathrm{d}_{ij} x_{ijvs} \tag{2}$$

$$T = \sum_{i\in\mathcal{C}} \sum_{s\in\mathcal{S}} z_{is} \tag{3}$$

$$T^{\max} \geqslant z_{is} \qquad \forall i \in \mathcal{C}, s \in \mathcal{S} \tag{4}$$

$$\sum_{i\in\mathcal{C}^0} \sum_{s\in\mathcal{S}} x_{0ivs} = \sum_{i\in\mathcal{C}^0} \sum_{s\in\mathcal{S}} x_{i0vs} \qquad \forall v \in \mathcal{V} \tag{5}$$

$$\sum_{j\in\mathcal{C}^0} \sum_{s\in\mathcal{S}} x_{jivs} = \sum_{j\in\mathcal{C}^0} \sum_{s\in\mathcal{S}} x_{ijvs} \qquad \forall i \in \mathcal{C}, v \in \mathcal{V} \tag{6}$$

$$\sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{C}^0} \mathrm{a}_{vs} x_{jivs} = \mathrm{r}_{is} \qquad \forall i \in \mathcal{C}, s \in \mathcal{S} \tag{7}$$

$$t_{ivs_1} + \mathrm{p}_{is_1} + \mathrm{d}_{ij} \leqslant t_{jvs_2} + \mathrm{M}\,(1 - x_{ijvs_2}) \qquad \forall i \in \mathcal{C}^0, j \in \mathcal{C}, v \in \mathcal{V}, s_1, s_2 \in \mathcal{S} \tag{8}$$

$$t_{ivs} \geqslant \mathrm{e}_i \qquad \forall i \in \mathcal{C}, v \in \mathcal{V}, s \in \mathcal{S} \tag{9}$$

$$t_{ivs} \leqslant \mathrm{l}_i + z_{is} \qquad \forall i \in \mathcal{C}, v \in \mathcal{V}, s \in \mathcal{S} \tag{10}$$

$$\begin{aligned} t_{iv_2s_2} - t_{iv_1s_1} \geqslant \delta_i^{\min} \\ - \mathrm{M}\left(2 - \sum_{j\in\mathcal{C}^0} x_{jiv_1s_1} - \sum_{j\in\mathcal{C}^0} x_{jiv_2s_2}\right) \end{aligned} \qquad \forall i \in \mathcal{C}^{\mathrm{d}}, v_1, v_2 \in \mathcal{V}, s_1, s_2 \in \mathcal{S} : s_1 < s_2 \tag{11}$$

$$t_{iv_2s_2} - t_{iv_1s_1} \leqslant \delta_i^{\max} - \mathrm{M}\left(2 - \sum_{j \in \mathcal{C}^0} x_{jiv_1s_1} - \sum_{j \in \mathcal{C}^0} x_{jiv_2s_2}\right) \quad \forall i \in \mathcal{C}^{\mathrm{d}}, v_1, v_2 \in \mathcal{V}, s_1, s_2 \in \mathcal{S} : s_1 < s_2 \tag{12}$$

$$x_{ijvs} \in \{0, \mathrm{a}_{vs}\mathrm{r}_{js}\} \quad \forall i \in \mathcal{C}^0, v \in \mathcal{V}, s \in \mathcal{S} \tag{13}$$

$$t_{ivs},\, z_{is} \geqslant 0 \quad \forall i \in \mathcal{C}^0, v \in \mathcal{V}, s \in \mathcal{S} \tag{14}$$

$$D,\, T,\, T^{\max} \geqslant 0 \tag{15}$$

The objective function (1) minimizes the total distance performed by staff members, the sum of tardiness of patients, as well as the largest tardiness observed on a solution. The constraints (2–4) compute the indicators of solution quality. Constraints (5) indicate that each vehicle should depart and return to depot once. Constraints (6) model the flow balance of routes. Constraints (7) guarantee that the service requests of patients must be performed by a capable staff member. The Constraints (8) perform the subtour elimination by considering the distances between nodes and service processing times. The parameter M is a sufficiently large number. Constraints (9–10) model the time windows and compute the services start times on patients. Constraints (11–12) model the double services. Finally, the constraints (13–15) define the domain of auxiliary and decision variables of the problem.

4. A fix and optimize *matheuristic* to HHCRSP

Fix and Optimize (FO) *matheuristic* consists of iteratively solving subproblems of a large MIP model, seeking to improve the best known solution s. Subproblems are generated as follows. The decision variables of the problem are bound according to their values in s. Thus, a decomposition procedure selects a subset of decision variables to have their bounds restored, as specified in the MIP model. The restoring process is called relaxation. The resulting subproblem can be solved using a MIP solver. The solution s is updated once a better solution is found. This process repeats until either a stop criterion is met or a local minimum is achieved. FO was proposed by Sahling et al. [2009] to solve the capacitated lot-sizing problem. It was recently applied to other problems of OR literature [Toledo et al., 2013; Dorneles et al., 2014], for example.

According to decomposition criteria, FO *matheuristic* can be a powerful optimization framework. If the subproblems are easy to solve (tractable from an algorithmic point of view), FO can be understood as a local search procedure, in which a mathematical solver implicitly decides the neighborhood movements. If subproblems are hard to solve, FO effectively performs a very large neighborhood search [Chen, 2015].

FO relies on the mathematical formulation of the problem, and inherit one of its prominent advantages: the *matheuristics* remains valid when changes on problem definition are applied. However, more subtle changes can reduce the effectiveness of a previously good decomposition scheme. Regarding the HHCRSP, a possible decomposition relies on relaxing the decision variables that appear in the route of one caretaker. This decomposition scheme generates NP-*hard* subproblems (VRPTW), thus only a few routes can be relaxed at once to keep computational times short. In the proposed FO for HHCRSP, exactly two routes are relaxed in each iteration of *matheuristics*.

On the algorithm design phase, the systematic relaxation of all pairs of routes was proven to be ineffective. Due to time windows and double services, most time spent on the systematic approach is wasted since only a few pairs of routes have the potential to improve the current solution. Thus, two route selection approaches are proposed. The first approach selects two routes at random. The second approach exploits the structure of the solution, regarding both spatial and temporal characteristics of routes. Figure 1 depicts a solution with tree routes. Using the random approach, any pair from $\{(r_1, r_2), (r_1, r_3), (r_2, r_3)\}$ could be selected. The guided approach may select the pair (r_2, r_3), in which routes are close in space and time, and have greater chances to improve the current solution.

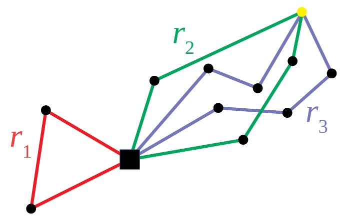

Figure 1: A routing solution to HHCRSP problem, considering 3 vehicles, and one node requiring double service with simultaneous attendance times, highlighted in yellow.

Algorithm 1 depicts a complete FO procedure applied to solve the HHCRSP. The *matheuristics* starts with the constructive solution of Mankowska et al. [2014] (line 1). Through lines (2–3), the model has its decision variables bounded according to s^*, and the MIP solver is called to generate the cost of the initial solution. Iteratively, two vehicles are chosen to have their routes optimized. The vehicles can be chosen using either random criteria (lines 8–9) or using the lateness of service times (guided criteria). The later criteria sorts the auxiliary variables t in decreasing order, according to their values in s^*. Then, two distinct vehicles are chosen randomly from the first half of the sorted list (lines 10–12), which holds the largest routes. Following the lines (lines 13–14), the decision variables x of the selected vehicles are relaxed. From lines (15–23), a MIP solver optimizes the generated subproblem up to a certain time limit $maxIterTime$. If the new solution $\bar{s}$ is better, or as good as s^*, the best known solution is then updated, and the stagnation flag $noImpr$ is set to 0 (lines 18–21). In line (21) the model has its decision variables bounded to the new best solution found. The fix and optimize loop (5–23) repeats until $\lfloor|\mathcal{V}|/2\rfloor$ iterations without improvement of the best known solution. The algorithm returns the best solution found and its objective value (line line 25).

5. Computational experiments

Computational experiments were performed on an Intel i7-3612QM machine, using up to 8GB of memory. The fix and optimize algorithm was implemented in C++, using the IBM CPLEX 12.9 as MIP solver. Parallel processing was disabled and the other solver parameters were kept defaults. Objective function weights are set to $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$, as used in Mankowska et al. [2014] for comparison purpose. FO parameter $maxIterTime$ was set to 25 seconds in all test cases.

The random instances dataset of Mankowska et al. [2014] was used in computational experiments. All instances consider 6 service types. Caretakers are split into two groups. Caretakers from the first group can perform up to 3 service types from subset $\{1, 2, 3\}$. Similarly, caretakers from the second group can perform up to 3 service types from $4, 5, 6$. Regarding double services, 15% of patients require simultaneous double service, and 15% require double services with precedence. The other patients require single services. Patients and depot were placed randomly inside a square region 100×100, and Euclidean distances were used. Each caretaker travel at 1 unit of distance per minute, thus, traveling time is the same as the distance between locations. For each patient $i \in \mathcal{C}$, the time value $\delta_i^{\min}$ was randomly generated from $[0, 60]$, represented in minutes. $\delta_i^{\max}$ was given by $\delta_i^{\min}$ added by a random number from $[0, 60]$. Processing times p_{is}, for patient $i \in \mathcal{C}$ and service type $s \in \mathcal{S}$ were generated from $[10, 20]$, in minutes. Time windows have a duration of 2 hours and were randomly generated inside a planning horizon of 10 hours. Seven subsets of instances

Algorithm 1: Fix and optimize matheuristic.

$s^* \leftarrow$ create a initial solution
fix bounds of decision variables x according their values in s^*
$z^* \leftarrow$ solve MIP (1–15) and get the cost of solution s^*
$noImpr \leftarrow 0$
repeat
  choose randomly a decomposition criteria m
  switch m **do**
    case random criteria **do**
      select v_1, v_2 randomly from $\mathcal{V}$, with $v_1 \neq v_2$
    case guided criteria **do**
      $cand \leftarrow$ sort t variables with decreasing values in s^*
      select $v_1 \neq v_2$ from the first half of list $cand$
  for $i, j \in \mathcal{C}^0, s \in \mathcal{S}$ **do**
    relax decision variables x_{ijvs} bounds to $\{0, \mathrm{a}_{vs}\mathrm{r}_{js}\}$
  $\bar{s} \leftarrow$ solve MIP model (1–15) up to $maxIterTime$ seconds
  $\bar{z} \leftarrow$ cost of solution $\bar{s}$
  if $\bar{z} \leqslant z^*$ **then**
    $s^* \leftarrow \bar{s}$
    $z^* \leftarrow \bar{z}$
    $noImpr \leftarrow 0$
    fix bounds of decision variables x according their values in s^*
  else
    $noImpr \leftarrow noImpr + 1$
until $noImpr \neq \lfloor |\mathcal{V}|/2 \rfloor$
return s^*, z^*

were generated, varying the number of caretakers and clients. Table 1 presents the characteristics for each instance subset. Column *Instance subset* identifies the configuration type, and columns $|\mathcal{V}|$ and $|\mathcal{C}|$ indicate the number of caretakers and patients for each configuration, respectively. For each configuration, 10 random instances were generated.

The model (1–15) considers a very large number of decision variables and constraints. A key strategy to scale such model to solve medium to large instances relies on the deletion of useless variables, following parameters a_{vs} and r_{is}. This way, if a patient does not require a service type, such variables must not be generated. Similarly, variables for a service type that a caretaker cannot perform must also not be generated. Table (1) compares the number of columns required to model the problem before and after the variables elimination. In instances with 75 nodes, the preprocessing eliminated about 920% the number of required variables. However, the preprocessing does not scale well on large instances (subsets E, F, and G). Noteworthy that the experiments could only be performed on instances up to 100 patients due to the memory limitation of the testing environment.

Instance subset	$\|\mathcal{V}\|$	$\|\mathcal{C}\|$	Cols before	Avg. cols after	Avg. reduction (%)
A	3	10	2,445	424.3	479.28
B	5	25	21,219	2,750.2	676.74
C	10	50	159,429	17,039	841.33
D	15	75	527,139	52,071	920.11
E	20	100	1,236,846	116,396.8	964.17
F	30	200	7,309,566	–	–
G	40	300	21,818,286	–	–

Table 1: Instance sizes and effect of variable elimination on model size.

Table 2 lists the published best known solutions and the best solutions found by the FO *matheuristic* for the dataset. The first column indicates the instance under testing. Four columns were presented for Mankowska et al. [2014]: LB, Best obj., Gap and Time. *LB* represents the best bound, a lower bound provided with 10 hours of CPLEX. Column *Best obj.* presents the best integer solution obtained, and column *Gap (%)* presents relative gaps of *Best obj.* to *LB*, according to the equation (Best obj. - LB)/LB $\times$ 100. Column *Time*$_{\rm b}$ *(sec.)* reports the times in which the best integer solutions were obtained (time to best). The stopping criteria used by Mankowska et al. [2014] is 2 hours for the heuristic solver, and up to 10 hours for the MIP solver. Results to subset A were solved to optimality using CPLEX. In Mankowska et al. [2014] the best results for instances B2, B3, B4, B8, and B9 were found by the solver. For the other instances, the best results were obtained using the heuristic. For the proposed *Fix and Optimize*, values of best integer solution found, the relative gap to *LB*, the running times, and the time to best integer solution are presented in columns *Best obj.*, *Gap (%)*, *Time*$_{\rm t}$ *(sec.)*, and *Time*$_{\rm b}$ *(sec.)*, respectively.

Instance	Mankowska et al. [2014]				Fix and Optimize			
	LB	Best obj.	Gap (%)	Time$_{\rm b}$ (sec.)	Best obj.	Gap (%)	Time$_{\rm t}$ (sec.)	Time$_{\rm b}$ (sec.)
A1	218.20	**218.20**	0.00	2.0	**218.20**	0.00	0.2	0.1
A2	246.60	**246.60**	0.00	5.0	246.63	0.01	0.2	0.1
A3	305.80	**305.90**	0.03	7.0	**305.90**	0.03	0.5	0.1
A4	186.90	**186.90**	0.00	8.0	**186.90**	0.00	1.8	0.3
A5	189.50	**189.50**	0.00	2.0	189.54	0.02	0.2	0.1
A6	200.10	**200.10**	0.00	2.0	**200.10**	0.00	0.1	0.0
A7	225.40	**225.40**	0.00	1.0	**225.40**	0.00	0.1	0.0
A8	232.00	**232.00**	0.00	4.0	232.05	0.02	0.2	0.0
A9	222.30	**222.30**	0.00	20.0	**222.30**	0.00	1.3	0.0
A10	225.00	**225.00**	0.00	1.0	225.01	0.00	0.1	0.0
Average		**225.19**	0.00		225.20	0.01		
B1	378.70	458.90	21.18	<1	**431.13**	13.85	58.1	16.1
B2	427.90	476.20	11.29	36000.0	**476.05**	11.25	9.1	5.3
B3	391.20	**399.20**	2.04	36000.0	421.87	7.84	155.9	94.9
B4	330.40	576.00	74.33	36000.0	**432.35**	30.86	17.4	7.3
B5	311.00	391.10	25.76	<1	**369.44**	18.79	86.9	46.2
B6	274.20	534.70	95.00	<1	**531.91**	93.98	96.5	76.1
B7	310.60	355.50	14.46	<1	**328.67**	5.82	17.9	17.3
B8	332.40	357.80	7.64	36000.0	**357.68**	7.61	67.4	65.5
B9	293.70	**403.80**	37.49	36000.0	446.62	52.07	225.2	94.2
B10	381.00	500.40	31.34	<1	**469.58**	23.25	147.1	86.7
Average		445.36	32.05		**426.53**	26.53		
C1	401.10	1123.60	180.13	<1	**1006.72**	150.99	984.7	906.0
C2	314.90	673.80	113.97	<1	**597.06**	89.60	185.9	140.2
C3	323.50	642.40	98.58	<1	**618.29**	91.12	298.5	256.3
C4	329.40	580.40	76.20	<1	**544.59**	65.33	81.1	75.4
C5	404.20	754.60	86.69	<1	**667.45**	65.13	158.7	114.1
C6	308.20	951.60	208.76	<1	**852.04**	176.46	790.0	599.5
C7	315.70	577.40	82.90	<1	**525.19**	66.36	421.7	383.2
C8	336.30	540.60	60.75	<1	**479.41**	42.55	394.4	367.6
C9	306.50	608.70	98.60	<1	**568.85**	85.60	465.6	392.1
C10	386.40	679.30	75.80	<1	**603.48**	56.18	321.5	240.5
Average		713.24	108.24		**646.31**	88.93		
D1	456.70	1321.80	189.42	5.0	**1301.01**	184.87	1649.4	1485.8
D2	336.80	892.70	165.05	4.0	**807.25**	139.68	764.8	659.5
D3	355.70	819.40	130.36	4.0	**675.33**	89.86	1563.5	1463.7

(Continues on next page.)

Table 2: Computational results for benchmark instances.

Instance	Mankowska et al. [2014]				Fix and Optimize			
	LB	Best obj.	Gap (%)	Time_b (sec.)	Best obj.	Gap (%)	Time_t (sec.)	Time_b (sec.)
D4	379.90	877.40	130.96	4.0	**817.83**	115.27	1331.1	1161.9
D5	372.20	872.10	134.31	5.0	**706.77**	89.89	823.3	716.7
D6	368.80	835.20	126.46	5.0	**762.33**	106.71	1058.3	988.8
D7	333.40	706.30	111.85	6.0	**671.70**	101.47	720.7	639.2
D8	373.30	811.40	117.36	4.0	**717.26**	92.14	750.2	690.9
D9	362.40	860.30	137.39	6.0	**739.63**	104.09	228.6	198.3
D10	434.40	**1306.60**	200.78	3.0	1330.39	206.26	712.5	602.0
Average		930.32	144.39		**852.95**	123.02		
E1	406.00	1604.90	295.30	17.0	**1349.19**	232.31	4864.0	4566.9
E2	411.90	1101.90	167.52	10.0	**941.35**	128.54	3654.2	3291.5
E3	437.20	986.40	125.62	14.0	**913.45**	108.93	2506.0	2316.0
E4	384.30	871.00	126.65	19.0	**851.40**	121.55	4332.1	3946.2
E5	392.40	1018.00	159.43	19.0	**905.41**	130.74	5035.1	4784.2
E6	390.20	1003.00	157.05	19.0	**878.79**	125.22	4074.0	3753.2
E7	374.40	921.10	146.02	20.0	**832.93**	122.47	2863.4	2531.9
E8	407.70	884.60	116.97	36.0	**832.73**	104.25	3765.3	3504.6
E9	422.10	1131.70	168.11	34.0	**1129.14**	167.51	1319.7	1139.7
E10	419.10	1053.60	151.40	11.0	**1026.83**	145.01	2961.8	2665.5
Average		1057.62	161.41		**966.12**	138.65		
F1	445.10	1721.40	286.70	889.0	–	–	–	–
F2	457.90	1763.80	285.20	909.0	–	–	–	–
F3	481.80	1549.60	221.60	868.0	–	–	–	–
F4	417.40	1420.40	240.30	1321.0	–	–	–	–
F5	452.30	1701.90	276.30	1145.0	–	–	–	–
F6	367.10	1639.70	346.70	836.0	–	–	–	–
F7	408.10	1384.30	239.20	1294.0	–	–	–	–
F8	454.30	1544.60	240.00	924.0	–	–	–	–
F9	426.80	1572.90	268.50	1642.0	–	–	–	–
F10	441.80	1581.00	257.90	1326.0	–	–	–	–
Average		**1587.96**	266.24		–	–		
G1	455.00	2248.00	394.10	7200.0	–	–	–	–
G2	463.10	2316.10	400.10	7200.0	–	–	–	–
G3	464.30	1885.30	306.10	7147.0	–	–	–	–
G4	461.50	2023.20	338.40	7200.0	–	–	–	–
G5	449.40	2247.60	400.10	7200.0	–	–	–	–
G6	471.50	2144.40	354.80	7200.0	–	–	–	–
G7	459.10	1971.50	329.40	6934.0	–	–	–	–
G8	472.40	1987.40	320.70	7200.0	–	–	–	–
G9	473.30	2415.50	410.40	7023.0	–	–	–	–
G10	451.60	2373.40	425.60	7003.0	–	–	–	–
Average		**2161.24**	367.97		–	–		

Table 2: Computational results for benchmark instances.

FO *matheuristic* outperformed the heuristic approach in 36 out of 40 test cases. FO seems to be more stable to achieve the best solutions for each instance subset. The results for instances B4, D3, and D5, hold the best improvements obtained by the FO, with savings up to 17% in comparison with the literature.

Figure 2 presents two scatter plots regarding the cost and relative gap to LB for the best integer solutions found by Mankowska et al. [2014] and FO. Each dot represents an instance of the dataset. A dot below the line indicates that FO obtained the best solution. A dot above the line indicates that the best solution was found by Mankowska et al. [2014]. The first graph indicates that the FO *matheuristic* outperformed results of the literature. Despite this, both solving methods provide solutions with costs within the same range of values (from 400 to 1200). This observation indicates that both approaches have similar performance, with a small advantage to the *matheuristic* approach. The scatter plot of the relative gaps indicates the same, with most of FO gaps ranging between 50% to 150%. The majority of gaps reported in the literature happen in a widening range of 50% to 200%.

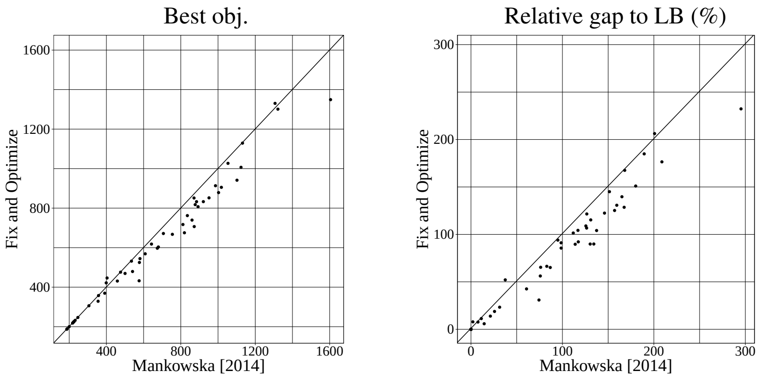

Figure 2: Scatter plots for best solutions and relative gap to lowerb bound.

6. Conclusion

This paper studied a *matheuristic* to solve the Home Care Routing and Scheduling Problem. This problem is relevant to public health services. The home health care service can also help to reduce the patients stay in hospitals, effectively increasing the availability of vacant beds. Such a service can also be offered to patients with chronic but stable health issues. The home environment can also enhance the patient recovery process.

A Fix and Optimize *matheuristic* was developed. Two decomposition criteria were proposed to exploit the structure of both solution space and local minima. Computational results were conducted on instances available in the literature of the problem. The experimental setup used an accessible computational environment, employing a standard MIP solver to solve subproblems iteratively. Regarding solver parameters, parallelism was disabled. Due to large memory requirements to solve big instances, a preprocessing was applied. It was observed a significant reduction in the number of decision variables used in the model.

Two decomposition schemes were proposed. The first scheme performs a random selection in the search space of the problem. The second scheme exploits the spatial and temporal characteristics of solutions. Preliminary results indicate that systematic decomposition is uninteresting. Overall, those simple decomposition schemes have proven to reduce the cost of an initial solution effectively. Since FO relies on a model, it inherits the robustness advantages of MIP against changes of problem definition, *w.r.t.* objective function and constraints.

Despite all the promising results of the proposed solution method, there are some points of the study that need more attention. A better route selection criteria can be employed to accelerate the convergence of the *matheuristic*. Other types of decomposition, like route segments, have a high potential to explore the neighborhood of a solution. A decomposition that exploits the structure of double services also has a great potential to improve solutions. As a further work, the solution approach should be adapted to solve the real home care problem found in the city of Porto Alegre (Rio Grande do Sul), and possibly other Brazilian cities, seizing the opportunity of Ministério da Saúde [2013].

References

Akjiratikarl, C., Yenradee, P., and Drake, P. R. (2007). Pso-based algorithm for home care worker scheduling in the UK. *Computers & Industrial Engineering*, 53(4):559–583.

Azi, N., Gendreau, M., and Potvin, J.-Y. (2010). An exact algorithm for a vehicle routing problem with time windows and multiple use of vehicles. *European Journal of Operational Research*, 202(3):756–763.

Bock, S. (2010). Real-time control of freight forwarder transportation networks by integrating multimodal transport chains. *European Journal of Operational Research*, 200(3):733–746.

Chen, H. (2015). Fix-and-optimize and variable neighborhood search approaches for multi-level capacitated lot sizing problems. *Omega*, 56:25–36.

Desaulniers, G. (2010). Branch-and-price-and-cut for the split-delivery vehicle routing problem with time windows. *Operations research*, 58(1):179–192.

Dorneles, Á. P., de Araújo, O. C., and Buriol, L. S. (2014). A fix-and-optimize heuristic for the high school timetabling problem. *Computers & Operations Research*, 52:29–38.

Drexl, M. (2012). Synchronization in vehicle routing—a survey of vrps with multiple synchronization constraints. *Transportation Science*, 46(3):297–316.

Eberhart, R. and Kennedy, J. (1995). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, p. 39–43. Ieee.

Eveborn, P., Flisberg, P., and Rönnqvist, M. (2006). Laps care–an operational system for staff planning of home care. *European journal of operational research*, 171(3):962–976.

Fernandez, A., Gregory, G., Hindle, A., and Lee, A. (1974). A model for community nursing in a rural county. *Journal of the Operational Research Society*, 25(2):231–239.

Fikar, C. and Hirsch, P. (2017). Home health care routing and scheduling: A review. *Computers & Operations Research*, 77:86–95.

Gutiérrez, E. V. and Vidal, C. J. (2013). Home health care logistics management problems: A critical review of models and methods. *Revista Facultad de Ingeniería Universidad de Antioquia*, (68):160–175.

IBGE (2018). Projeção da população do brasil e das unidades da federação. Available in https://www.ibge.gov.br/apps/populacao/projecao/. Last visited in February (2019).

Mankowska, D. S., Meisel, F., and Bierwirth, C. (2014). The home health care routing and scheduling problem with interdependent services. *Health care management science*, 17(1):15–30.

Ministério da Saúde (2013). Serviço de atenção domiciliar "melhor em casa". Available in http://portalms.saude.gov.br/acoes-e-programas/melhor-em-casa-servico-de-atencao-domiciliar/melhor-em-casa. Last visited in January (2019).

Nasri, F. (2008). O envelhecimento populacional no brasil. *Einstein*, 6(Supl 1):S4–S6.

Russell, R. A. and Morrel, R. B. (1986). Routing special-education school buses. *Interfaces*, 16(5):56–64.

Sahling, F., Buschkühl, L., Tempelmeier, H., and Helber, S. (2009). Solving a multi-level capacitated lot sizing problem with multi-period setup carry-over via a fix-and-optimize heuristic. *Computers & Operations Research*, 36(9):2546–2553.

Toledo, C. F. M., De Oliveira, R. R. R., and FrançA, P. M. (2013). A hybrid multi-population genetic algorithm applied to solve the multi-level capacitated lot sizing problem with backlogging. *Computers & Operations Research*, 40(4):910–919.

Trautsamwieser, A. and Hirsch, P. (2011). Optimization of daily scheduling for home health care services. *Journal of Applied Operational Research*, 3(3):124–136.