Towards an unified framework for routing and scheduling planning in a Integrated Continuous Care Unit

Maria Teresa Godinho[1,3] and Maria João Lopes[2,3]

[1] Polytechnic Institute of Beja, School of Technology and Management, Rua Pedro Soares, 7800-295 Beja, Portugal
mtgodinho@ipbeja.pt
[2] Iscte-IUL University Institute of Lisbon, Business School, Avenida das Forças Armadas, 1649-026 Lisboa, Portugal
maria.joao.lopes@iscte-iul.pt
[3] CMAFcIO, University of Lisbon, Campo Grande, 1749-016 Lisboa, Portugal

Abstract. We study a real world routing and scheduling problem arising in home health care context. Requisites defined by stakeholders are introduced and examined. Then, the problem is modeled as a Time-constrained Vehicle Routing Problem with Time Windows and a new Single Commodity Flow Formulation of it is proposed. Valid Inequalities to enhance the model are presented. Extensions of the problem, such as aggregating different types of care and of ensuring a threshold on the difference between the duration of the routes, are also discussed. A set of computational tests, performed on instances adapted from benchmark instances from the literature were executed and results discussed.

Keywords: Home Health Care · Routing and Scheduling · Mixed Integer Linear Formulations.

1 Introduction and Literature Review

The generic term "Home health care" covers several programs of care delivered at the patient's home. In this work, we address a real case study motivated by an Integrated Continuous Care (ICC) Unit routing and scheduling needs. Next, we review some contributions on the routing and scheduling problem in health care context. Our review is based on [6, 3, 7].

The home health care routing and scheduling problem (HHCRSP) consists of designing a set of routes that allow to provide planned care visits to a set of patients over a planning horizon, while optimizing some given criteria and respecting several constraints [3]. Thus, the HHCRSP can be seen as a VRP where side-constraints are used for modeling the different requirements of each application, including the scheduling requirements.

Side constraints can be grouped into i) patient constraints; ii) caregivers constraints and iii) treatment/visit constraints. Constraints type i) deal with the patient's preferences regarding both the visit's time-span and the caregiver;

constraints of type ii) tackle labour legislation and also the preferences of the caregivers; finally, iii) gathers constraints stemmed from specifications of the care and constraints that ensure compliance with temporal dependencies between different visits to the same patient, as each patient may require more than one visit per day. The objective functions also vary. The most frequent objectives are the minimization of the routes and/or of the staff related total cost. Other objectives present in the literature are the optimization of the quality of service and well-being at work. These objectives are handled mostly under multi-objective approach's, either by considering weighted sums or by approximating the Pareto Frontier and analysing trade-offs.

Both exact and heuristic methods can be found in the literature. In our work, we present a new Mixed Integer Linear Programming Model of this problem. As mentioned, the HHCRSP can be seen as a VRP, to which side constraints have been added. Thus, many MILP models of it, in the literature, are based on models originally proposed for the VRP or its variants. This is the case of [9, 1, 2]. While Gomes and Ramos [9] present a model adapted from a widely used formulation of the VRP with Time Windows proposed in [5], Cappanera and Scutellà and Cappanera et al [1, 2] present flow based models rooted in the work of Wong [12] and Claus [4] and of Gavish and Graves [8], respectively.

Pure MILP approaches are seldom seen in recent papers. The references cited above, with the exception of [1], use the proposed models combined with other approaches. For fairness, it should be said that the conclusions of [1] point in the same direction. Nevertheless, in this work, our model is solved by feeding it to CPLEX. In fact, we use the proposed model as an exploratory mean to gain a better understanding as to the many aspects of the problem, but we are fully aware that to solve medium size instances in reasonable times we need to go a step further.

The remainder of this work is structured as follows: in Section 2, the case study that motivated this work is described in detail; in Section 3, the problem is formally define, a MILP formulation of it is introduced and its extensions discussed; in Section 4, a computational study is presented and Section 5 concludes the text.

2 Case Study

As mentioned, this work tackles a real case study originated in a ICC Unit. In order to understand the requisites of the problem, the head of the transport department, the Head Nurse and the drivers were heard. Due to the existence of many conflicting aspects, it was decided that the Head Nurse requests should be given priority.

The Head Nurse pointed out that some patients have to be visited more than once on the same day and that while some treatments can be performed at any given time, others have to be carried out within a specific time-slot. She also indicated that the number of teams is fixed, but the composition of the teams may vary. The allocation of the personnel to the teams is done in compliance with their skills, ensuring that the teams are balanced, and also with work shifts and days off. As it is, any team can be assigned to any patient. Enquired about the importance of continuity of care (we say that there is when the same patient is treated by the same team), the Head Nurse said that as the teams role and as ICC teams deal with long term care and are composed of dedicated personnel, in the end of the day, all ICC members get know quite well all the patients. Thus, for the time being, she does not feel the need to introduce that type of constraints. The teams are out for the maximum duration of the shift and use a vehicle with a driver to travel to and between patients' homes. Finally, she elected as her objective for this project, the desire to increase the percentage of time effectively dedicated to treating the patients in each shift.

In conclusion, requisites are: minimization of wasted time, while meeting the time windows requirements and the bounds on both the duration of the routes and the number of teams.

3 Modeling approach

The problem described in the previous section can be modeled as a Time-constrained Vehicle Routing Problem with Time-windows. (`TCVRPTW`). The `TCVRPTW` seeks a set of routes of minimum cost in a given network, such as time-windows, the number of vehicles and duration of the routes requirements are met.

Recall that any team can carry out any treatment to any patient but there are treatments with and without time windows associated with it. Thus, we consider two types of treatment: those with time windows associated with it, designated by specific (s), and the remainders designated by general (g).

Consider the set $P = \{1, ..., p\}$ of the patients to treat on a given day and the set $C = \{s, g\}$ of the types of care the Unit can provide. Thus, p^c, $p \in P, c \in C$, define the tasks to be executed by the team. As patients may need more than one care on the same day, there may be more than one task associated with the same patient.

Let $G = (N, A)$ be a complete directed graph with $N = \{0, ..., n\}$. Node 0 represents the depot and the remaining nodes represent the tasks. We let N^s and N^g denote the subsets of the nodes associated with tasks type s and g, respectively. d_i indicates the time required to complete task $i \in N \backslash \{0\}$ and $[a_i, b_i]$ denotes the time window within which the task $i \in N^s$ has to be concluded. The arcs represent the trips. t_{ij} indicates the traveling time between the homes of the patients associated with tasks i and j, for all $(i, j) \in A$. When tasks i and j are associated with the same patient, t_{ij} is null. Finally, T denotes the maximum duration of a route, which is, as mentioned, given by the duration of the shift. The number of teams defines the size of the fleet, denoted by V. We assume that all parameters are non-negative integers (the time unit is minutes).

Consider, also, the decision variables:

- binary variables, $x_{ij} = 1$ if j is visited just after i; $x_{ij} = 0$, otherwise, for all $i, j \in N, i \neq j$;

- variables y_{ij} representing the arrival time at j just after visiting i, for all $i, j \in N, i \neq j$;
- w_{ij} representing the waiting time at j, after traveling from i to j, for all $i \in N, j \in N^s, i \neq j$.
- binary variables $u_{p^s,p'^g} = 1$, if care s is provided before care g to a patient requiring both types of care, $u_{p^s,p'^g} = 0$, otherwise, for all $p^s \in N^s$, $p'^g \in N^g$, with $p = p'$.

Then, the feasibility set of TCVRPTW can then be modeled as:

$$\sum_{i \in N} x_{ij} = 1 \qquad j \in N \backslash \{0\} \tag{1}$$

$$\sum_{j \in N} x_{ij} = 1 \qquad i \in N \backslash \{0\} \tag{2}$$

$$y_{0j} = t_{0j} x_{0j} \qquad j \in N \backslash \{0\} \tag{3}$$

$$\sum_{i \in N} (y_{ij} + w_{ij} + d_j x_{ij}) = \sum_{i \in N} (y_{ji} - t_{ji} x_{ji}) \qquad j \in N^s \tag{4}$$

$$\sum_{i \in N} (y_{ij} + d_j x_{ij}) = \sum_{i \in N} (y_{ji} - t_{ji} x_{ji}) \qquad j \in N^g \tag{5}$$

$$y_{j0} \leq T x_{j0} \qquad j \in N \backslash \{0\} \tag{6}$$

$$y_{ij} + w_{ij} \leq (T - d_j - min_l\{t_{jl}\}) x_{ij} \qquad i \in N, j \in N^s, i \neq j \tag{7}$$

$$y_{ij} \leq (T - d_j - min_l\{t_{jl}\}) x_{ij} \qquad i \in N, j \in N^g, i \neq j \tag{8}$$

$$\sum_{i \in N} (y_{ij} + w_{ij}) \leq b_j - d_j \qquad j \in N^s \tag{9}$$

$$\sum_{i \in N} (y_{ij} + w_{ij}) \geq a_j \qquad j \in N^s \tag{10}$$

$$\sum_{i \in N} (y_{ik^s} + w_{ik^s}) + d_k \leq \sum_{i \in N} (y_{il^g}) + M(1 - u_{k^s l^g}) \qquad k^s \in N^s, l^g \in N^g, k = l \tag{11}$$

$$\sum_{i \in N} y_{il^g} + d_l \leq \sum_{i \in N} y_{ik^s} + M u_{k^s l^g} \qquad k^s \in N^s, l^g \in N^g, k = l \tag{12}$$

$$\sum_{j \in N \backslash \{0\}} x_{0j} \leq V \tag{13}$$

$$x_{ij} \in \{0, 1\}, y_{ij} \geq 0 \qquad i, j \in N, i \neq j \tag{14}$$

$$w_{ij} \geq 0 \qquad i \in N, j \in N^s, i \neq j \tag{15}$$

$$u_{i^s,j^g} \in \{0,1\} \qquad i^s \in N^s, j^g \in N^g, i = j \tag{16}$$

Time is modeled as a flow system whose source and sink are node 0. The flow values increase as patients are visited and cared. The assignment constraints, (1) and (2), ensure that all the tasks are executed; the flow conservation constraints constraints, (3), (4) and (5), ensure the connectivity of the routes and, thus, time continuity; the linking constraints, (6), (7) and (8) ensure that the duration of a route does no exceed T and, together with the assignment and the flow conservation constraints, prevent the presence of sub-circuits; constraints (9) and (10), ensure the that the time windows are satisfied and constraints (11) and (12) prevent that patients are treated for the two types of care at the same time; constraints (13) ensure that the total number of vehicles is not exceeded and, finally, constraints (14), (15) and (16) define the variables domain.

As mentioned, we aim at minimizing "wasted" time. With that purpose, we consider two criteria: minimizing the total traveling time, (17), and minimizing the total waiting time, (18) that will be studied separately.

$$\min F_1 = \sum_{i,j \in N, i \neq j} t_{ij} x_{ij} \tag{17}$$

$$\min F_2 = \sum_{i \in N, j \in N^s, i \neq j} w_{ij} x_{ij} \tag{18}$$

The corresponding ILP models are:

$$\begin{aligned} &\min F_i \\ &\text{subject to } \ (1) - (16) \end{aligned}$$

with $i = 1, 2$.

3.1 Enhancing the Models

The linear programming relaxation of the models described above can be enhanced by adding valid inequalities to it. Consider the following set of inequalities:

$$(min_l\{t_{lj}\} + t_{ji} + d_j)x_{ji} \leq y_{ji} \qquad j \in N \backslash \{0\}, i \in N, i \neq j \tag{19}$$

Constraints (19) model the fact that if node i is not the first to be visited (that is if $x_{ji} = 1$ to some $j \neq 0$), then some time has already been consumed in the previous travels and treatments. That time cannot be inferior to $min_l\{t_{lj}\} + t_{ji} + d_j$ for $l \neq i, j$. Similarly, observing that the care associated with nodes in N^s cannot start outside their time window, we have:

$$(a_j + t_{ji} + d_j)x_{ji} \leq y_{ji} \qquad j \in N^s, i \in N, i \neq j \tag{20}$$

Thus for the nodes in N^s constraints (19) can be replaced by

$$[max(min_l\{t_{lj}\}, a_j) + t_{ji} + d_j]x_{ji} \leq y_{ji} \qquad j \in N^s, i \in N, i \neq j \quad (21)$$

Another set of valid inequalities:

$$y_{ji} \leq (b_j + t_{ji})x_{ji} \qquad j \in N^s, i \in N, i \neq j \quad (22)$$

Preliminary tests, showed that, in most cases, the improvements on the lower bounds and computational times associated with adding (19), (20), (21) and (22) to the models are substantial. Thus, in the following, we consider the enhanced version of the two formulations.

3.2 Model Extensions

We have considered two different model extensions:

- Aggregating general and specific treatments when applied to the same patient;
- Ensuring time work balance within the teams.

Firstly, we observe that any solution, of one of the models, in which a general care and a specific care associated with the same patient are carried out in the same visit, can never be more costly than a solution where the two cares are carried out in different visits. That is, the later will occur only when the former is unfeasible. Thus, a new version of the model, where the two cares are aggregated was written, by making just a few adjustments to the current formulations. As a result of the decrease of the size of the network and of the elimination of constraints (11) and (12) the time consumed to solve the model to optimality reduced most considerably. Consequently, the computational experiments related in Section 4 were performed using the new version of the formulations as the base model. If unfeasibility would occur in a given instance (which did not happened) then we would resort to the previous model to solve the instance.

Secondly, large differences in the duration of the routes were observed in some of the solutions obtained in the preliminary tests. In consequence, a new set of constraints imposing a threshold on the difference between the duration of the routes, was added to the model. We considered a threshold of 30 minutes:

$$D_{min} \leq y_{j0} + M(1 - x_{j0}) \qquad j \in N \backslash \{0\} \quad (23)$$

$$y_{j0} \leq D_{max} \qquad j \in N \backslash \{0\} \quad (24)$$

$$D_{max} - D_{min} \leq 30 \quad (25)$$

4 Computational Study

The models presented in the previous sections were analyzed empirically with the goal of characterizing the solutions produced in each case and the associated trade-offs. Experiments were carried out on a 64-bit PC with an Intel®Core™ i9-7700HK Quad core at 2.8GHz with 16GB of RAM. We have used IBM ILOG CPLEX version 22.1 as the ILP solver, with no time limit. Real instances are not available due to confidentiality issues, regarding patient's related data. Thus, the instances considered were adapted from the literature. We have used the first six instances of the `C2` set of the Solomon instances [11] and performed the following adjustments:

1- The size of the instances was reduced to 45: 30 of which would require specific care only; 10 would require general care only and 5 would require both. Patients in each category were randomly chosen among the first 45 individuals of the original instances;
2- The values of the traveling times were the rounded to the next integer;
3- The total number of vehicles was set to 13;
4- The maximum duration of a route was set to 540 minutes;
5- The duration of the cares was a random integer between 30 and 60 minutes;
6- The Time-window of each patient $i \in N^s$ was generated by: i) obtaining the center of the interval, generating a random integer between t_{0i} and $540 - t_{i0}$; ii) obtaining the width of the interval, generating a random integer between 60 and 120; iii) correcting unfeasibilities - if a_i is negative $-a_i$ is added to both bounds; if $t_{0i} + d_i < b_i$, b_i is perturbed until unfeasibility cease to exist.

The values introduced in adjustments 4 and 5 were supplied by the stakeholders; the procedure described in 6 is suggested in [10] for a similar problem (in the original set of instances, all time windows have the same amplitude, which is not realistic); the value proposed in 1 is arbitrary; the value proposed in 3 was changed accordingly. Results for this set of tests are reported in tables 1 to 3. In all these tables, MnRD, AvRD and MxRD denote the minimum, the average and the maximum routes duration; TTT and TWT the total traveling and waiting times, respectively, and NR the number of routes.

In the upper part of Table 1 we present the results obtained when the total traveling time was minimized. The number of routes was 8 for two instances and 9 for the remaining four. The routes duration ranged from 167 to 540 minutes. In addition, the maximum duration for all instances was greater than 500 minutes. The total waiting time, for each instance, was greater than the total traveling time. On the other hand, to meet time windows there were several routes with waiting time before the first care, because each route started at minute 0. This means that several routes can start after minute 0 and their duration should not include the previous waiting times.

In the second part of Table 1 we present the results obtained excluding the waiting times before the first care. The waiting times reduced but they are still significant. This correction led to routes duration ranging from 71 to 540 minutes. So, the exclusion of initial waiting times increased the bias on routes duration.

Table 1. Results for the minimization of the: i) total traveling time; ii) total traveling time excluding the waiting time before the first care; iii) total traveling time with routes duration balance; iv) total traveling time with routes duration balance and excluding the waiting time before the first care.

Instances	C201-45	C202-45	C203-45	C204-45	C205-45	C206-45
MnRD	426	236	429	173	167	255
AvRD	488.1	407.7	477.3	409.9	429.3	451.9
MxRD	540	540	502	540	540	535
TTT	685	681	649	699	716	718
TWT	1012	736	939	801	762	1068
NR	8	9	8	9	9	9
MnRD	309	76	309	171	71	89
AvRD	439.6	360.9	445.4	388	396.2	421.3
MxRD	540	540	498	528	506	535
TTT	685	681	649	699	716	718
TWT	624	315	684	604	464	793
NR	8	9	8	9	9	9
MnRD	476	464	472	467	471	496
AvRD	483.6	481.7	485.4	476.2	498.9	510.8
MxRD	506	491	502	497	501	526
TTT	693	719	656	743	726	737
TWT	968	1364	997	1354	807	1068
NR	8	9	8	9	8	8
MnRD	246	294	359	387	416	441
AvRD	405.4	398.4	457.3	466.0	471.0	494.0
MxRD	484	489	501	497	501	526
TTT	693	719	656	743	726	737
TWT	342	615	772	1262	656	934
NR	8	9	8	9	8	8

Table 2. Results for the minimization of the total waiting time with routes duration balance.

Instances	C201-45	C202-45	C203-45	C204-45	C205-45	C206-45
MnRD	463	469	464	458	467	481
AvRD	481.3	484.9	476.6	469.1	480.1	492.9
MxRD	493	499	493	488	488	504
TTT	1642	1613	1583	1552	1455	1659
TWT	0	14	0	12	0	3
NR	8	8	8	8	8	8

Table 3. Results for the minimization of total traveling time, constraining the total waiting time to the minimum, with routes duration balance, for the first 5 instances.

Instances	C201-45	C202-45	C203-45	C204-45	C205-45
MnRD	470	464	466	458	467
AvRD	480.4	475.0	476.9	463.3	475.3
MxRD	491	491	494	497	497
TTT	1155	1059	1108	1506	941
TWT	0	14	0	12	0
NR	7	7	7	8	7

In the third part of Table 1 we present the results when the total traveling time was minimized and a gap for routes duration of at most 30 minutes was imposed. For all instances but one the maximal gap was attained. The exception was a gap equal to 27 minutes. The number of instances with 8 routes increased to 4. The others kept the 9 routes. For each instance, the total waiting time was greater than the total traveling time and several routes had waiting time before the first care.

In the bottom part of Table 1 we present the results excluding the waiting times before the first care. In spite of the reduction on the total waiting time, for all instances, these times are still very high. In addition, the correction on routes duration showed that their gap are greater than 30 minutes. To try to overcome this disadvantage, the objective was changed to the minimization of the total waiting time.

Table 2 reports these results. The total waiting time reduced significantly. For three instances, in one of the routes there was a waiting times before the first care, but these times were at most equal to 12 minutes. These times had a little effect on routes duration gaps. The number of routes was equal to 8. Thus, there was a reduction for 2 instances. However, the total traveling time increased significantly, it exceeded twice the total traveling time presented in the upper and third parts of Table 1. On the other hand, for all instances but one the total traveling time reported in Table 2 was lower than the corresponding total time (traveling plus waiting times) represented in the third part of Table 1.

Notice that the computational results analyzed so far minimize one objective (traveling or waiting time) without constraining other objectives. To keep the advantages of the previous solutions, preliminary computational experiments were carried out. The objective was to minimize the total traveling time, constraining the total waiting time to the minimum, and excluding biased routes duration. The results obtained so far for the first 5 five instances are presented in Table 3. Notice that for the four first instances, we present data concerning the best feasible solution obtained. In general, the reduction on the total traveling time was very significant. Moreover, for four instances, the number of routes needed decreased to 7.

5 Conclusions and future work

We addressed the routing and scheduling problem of an ICC Unit, by means of a Single Commodity Flow Formulation (SCF). SCF models have the advantage of being simple to adapt to the many extensions of this problem and easy to solve.

Results permitted to analyze trade-offs as conditions varied. Results are now to be analyzed with the Head Nurse in order to understand what still needs to be accommodate in the model and in which direction work should proceed.

In the near future, we plan to generate more and more realistic instances. Namely, it is necessary to study the effect on the results of varying the percentages of each type of patient in the data. In the long range, we intend to introduce uncertainty in the model and to develop an hybrid approach that combines MILP with heuristic methods.

References

1. Cappanera, P., Scutellà, M.G.: Joint assignment, scheduling, and routing models to home care optimization: A pattern-based approach. Transp. Sci. **49**, 830–852 (2015)
2. Cappanera, P., Scutellà, M.G., Nervi, F., Galli, L.: Demand uncertainty in robust home care optimization. Omega **80**, 95–110 (2018). https://doi.org/https://doi.org/10.1016/j.omega.2017.08.012, https://www.sciencedirect.com/science/article/pii/S0305048316309008
3. Cissé, M., Yalçındağ, S., Kergosien, Y., Şahin, E., Lenté, C., Matta, A.: Or problems related to home health care: A review of relevant routing and scheduling problems. Operations Research for Health Care **13-14**, 1–22 (2017). https://doi.org/https://doi.org/10.1016/j.orhc.2017.06.001, https://www.sciencedirect.com/science/article/pii/S2211692316300157
4. Claus, A.: A new formulation for the travelling salesman problem. SIAM Journal on Algebraic Discrete Methods **5**(1), 21–25 (1984). https://doi.org/10.1137/0605004, https://doi.org/10.1137/0605004
5. Cordeau, J.F., Desaulniers, G., Desrosiers, J., Solomon, M.M., Soumis, F.: 7. VRP with Time Windows, pp. 157–193. https://doi.org/10.1137/1.9780898718515.ch7, https://epubs.siam.org/doi/abs/10.1137/1.9780898718515.ch7
6. Di Mascolo, M., Martinez, C., Espinouse, M.L.: Routing and scheduling in home health care: A literature survey and bibliometric analysis. Computers & Industrial Engineering **158**, 107255 (2021). https://doi.org/https://doi.org/10.1016/j.cie.2021.107255, https://www.sciencedirect.com/science/article/pii/S0360835221001595
7. Fikar, C., Hirsch, P.: Home health care routing and scheduling: A review. Computers & Operations Research **77**, 86–95 (2017). https://doi.org/https://doi.org/10.1016/j.cor.2016.07.019, https://www.sciencedirect.com/science/article/pii/S0305054816301848
8. Gavish, B., Graves, S.C.: The travelling salesman problem and related problems (1978)
9. Gomes, M.I., Ramos, T.R.P.: Modelling and (re-)planning periodic home social care services with loyalty and non-loyalty features. European Journal of Operational Research **277**(1), 284–299 (2019). https://doi.org/https://doi.org/10.1016/j.ejor.2019.01.061, https://www.sciencedirect.com/science/article/pii/S0377221719300967
10. Li, Y., Xiang, T., Szeto, W.Y.: Home health care routing and scheduling problem with the consideration of outpatient services. Transportation Research Part E: Logistics and Transportation Review **152**, 102420 (2021). https://doi.org/https://doi.org/10.1016/j.tre.2021.102420, https://www.sciencedirect.com/science/article/pii/S1366554521001861
11. Solomon, M.M.: Algorithms for the vehicle routing and scheduling problems with time window constraints. Operations Research **35**, 254–265 (1987)
12. Wong, R.T.: Integer programming formulation of the travelling salesman problem. Proceedings IEEE International Conference of Circuit and Computers pp. 149–152 (1980)