Scatter search approach for solving a home care nurses routing and scheduling problem

Bouazza Elbenani¹, Jacques A. Ferland² and Viviane Gascon³*

¹ Département de mathématique et informatique, Faculté des sciences, Université Mohammed V-Agdal, B.P. 1014, Rabat, Morocco
² Département d’informatique et de recherche opérationnelle, Université de Montréal, P.O. Box 6128, Succursale Centre-ville, Montréal, Québec, Canada, H3C 3J7
³ Département des sciences de la gestion, Université du Québec à Trois-Rivières, P.O. Box 500, Trois-Rivières, Québec, Canada, G9A 5H7
* Corresponding author. Email: gascon@uqtr.ca; Phone: 819-376-5011, #3136; Fax: 819-376-5079

Abstract The home care nurses routing and scheduling problem consists in determining the sequences of visits performed by home care nurses to patients. Very often public medical clinics are responsible for this planning. The problem shows similarities with the vehicle routing problem with time windows (VRPTW) but it also includes additional constraints related to medical restrictions and to the continuity of care. In this paper, we describe the specific problem. We next present in details the solution methods used for the global problem. They are based on a scatter search approach. We present numerical results and comparisons with real life data provided by a medical clinic (CLSC Les Forges located in Trois-Rivières, Canada).

Keywords: routing and scheduling, home care nurses, mathematical programming, metaheuristics (Tabu).

1. Introduction

The ageing population and the urge to reduce costs entail health care services to be provided at home more often than they used to. Health care managers must solve new types of problems such as nurse scheduling, districting or routing problems. Routing home care nurses is a difficult task that we address by solving a mathematical programming model using a scatter search approach. In this paper we model a real problem and we analyse tools to specify nurse routes to visit their patients. This approach gives rise to a variant of the vehicle routing problem with time windows (VRPTW) including additional constraints related to medical restrictions and to the continuity of care. These additional constraints make the problem more complex and more difficult to solve.

In this paper, we illustrate the scatter search approach by solving a real problem faced by the CLSC Les Forges in Trois-Rivières, a public medical clinic covering a territory including several urban and rural sectors. Home care nurses visit patients each day. Each nurse and each patient are assigned to one sector. Usually a nurse can visit only patients from her sector. Hence, each sector has a set of nurses to provide service to its patients. Continuity of care is important in home care services. Patients prefer to always receive their treatments by the same nurse. Nurses also think that this is preferable to become better acquainted with their patients. Thus continuity of care is taken into account in our model. Moreover nurses are also responsible for taking blood samples for some patients. But then, for medical reason, each blood sample must be returned to the clinic within a time interval after being taken. For this reason, our model requires additional returns to the clinic that are more difficult to schedule because we do not know in advance the exact time of the day when a patient requiring blood sampling will be visited.
This paper is organized as follows. First, a literature review is presented in Section 2, in order to position our contribution in this field of research. Next, Section 3 depicts the home care routing and scheduling problem faced by the Centre Local de Services Communautaires (CLSC) Les Forges in Trois-Rivières (Canada, province of Québec). Section 4 is devoted to the mathematical programming model. Then in Section 5 we introduce the scatter search solution approach. Numerical results to compare the different solution approaches are presented in Section 6. We also compare the results obtained with the best solution approaches and the manual results provided by the clinic. To conclude, we summarize our work and we indicate possible extensions for future research.

2. Literature review

The literature includes only a few articles on routing and scheduling home care nurses. Those articles deal mainly with problems and cases related to specific clinics or companies. Bertels and Fahle (2006) develop a computerized system which solves the nurse scheduling and the routing home care nurses problems. Eveborn et al. (2005) introduce a scheduling problem for a variety of home care providers which is modeled as a set partitioning problem and solved with a repeated matching algorithm. Begur et al. (1997) developed a software enabling to visualize the patients’ locations. The objectives are to minimize the total traveling time and to balance nurse workloads. Cheng and Rich (1998) present the routing home care nurses problem as a vehicle routing problem with time windows and multiple depots. The problem is to determine optimal routes minimizing the total distance, the overtime worked by regular nurses, and the number of hours worked by part time nurses.

Several papers can be found in the literature on the vehicle routing and scheduling problem with time windows (VRPTW). Such a problem is underlying the routing and scheduling of home care nurses. We refer the reader to the papers of Solomon and Desrosiers (1988), Malandraki and Daskin (1992) and Bräysy and Gendreau (2005a, 2005b) focusing on heuristic methods. Papers of Solomon and Desrosiers (1988) and Malandraki and Daskin (1992) inspired us in defining our mathematical model. However routing and scheduling home care nurses with additional constraints and a dynamic aspect does not seem to have received much attention in the literature.

3. Problem overview

Nowadays, more patients receive their medical treatments at home. The medical public clinic CLSC Les Forges in Trois-Rivières (covering a territory divided into sectors) is accountable for planning nurse visits to patients. Each patient is assigned to a sector according to his home address (note that the masculine gender is used throughout the paper). Each nurse is also assigned to one sector, each sector includes several nurses. Even if a nurse should visits only patients from his sector, some nurses may however have to visit patients from other sectors in order to balance nurse workloads. If a patient has to be visited by a nurse from another sector, we try to choose the one from the nearest sector to reduce the total traveling time. The impact of this flexibility induces that the problem is not separable by sector.

When solving the problem, we assume that each nurse completes his route to return to the clinic no later than 12h00 PM even though in practice some flexibility is allowed to finish later. The rest of the working day is devoted to completing proper medical reports during the afternoon.

In home care services it is important to consider the human factor related to the continuity of care. The patients prefer receiving their treatments by the same nurse, and nurses also think that it is preferable to become better acquainted with their patients. But a follow-up nurse might not be available because of holidays and days off. Thus in our model, we account for the continuity of care requirements as soft
constraints. Moreover, to replace nurses during their holidays and days off, the administrators rely on a recall list including substitute nurses.

At the CLSC Les Forges, the rules specifying the time frames for returning the blood samples to the clinic are formulated as follows. If a blood sample is performed before 10h00 AM, the nurse must turn back the blood sample at the clinic no later than 10h00 AM (strictly speaking). If a blood sample is performed between 10h00 AM and 11h00 AM, the nurse must then turn back the blood sample no later than 11h00 AM. In our approach, we model these returns to the clinic as fictitious destination nodes to be visited before 10h00 AM and 11h00 AM, respectively. Hence the nurses must visit these fictitious nodes only if blood samples are completed.

The objective is to determine routes in a reasonable time frame satisfying all constraints and minimizing the objective function. The objective function to be minimized in our model is specified in terms of the following components:

- the traveling time of the nurses;
- the fixed cost (salary) for the nurses according to their types (regular or from the recall list);
- a penalty cost for each patient visited by a nurse different from his follow-up nurse;
- a penalty cost when a patient is visited by a nurse from a different sector.

Note that the third component is a way to account for the continuity of care requirements as a soft constraint. The objective function also includes an additional artificial component useful in formulating the blood sample constraints.

### 4. Mathematical model

The home care nurse routing problem is formulated as a vehicle routing problem with time windows (VRPTW) including additional constraints. In this model, we denote

- \( P = \{1, 2, ..., N\} \): the set of all patients.
- \( P^+ \): the set of patients requiring a blood sampling, \( P^+ \subset P \).
- \( I = I^r \cup I^l \cup I^f \): the set of all nurses including the sets \( I^r \), \( I^l \) and \( I^f \) of regular nurses, of nurses from the recall list, and of fictitious nurses, respectively.
- \( C_r, C_l \), and \( C_f \): the daily costs for the different nurse categories where \( C_r < C_l < C_f \).

Note that in addition to the regular nurses and to the nurses from the recall list categories, we introduce a third category of fictitious nurses that are assigned to patients in last resort indicating that these patients are not visited. Thus all patients are assigned to a route.

To specify the underlying network, we define the set of nodes \( V \) as follows:

- a node \( i \) is associated with each patient in \( P \);
- two different nodes 0 and \( D \) are associated with the clinic where 0 and \( D \) are used to denote the origin and the ending node of each route, respectively;
- two fictitious nodes \( p_{10} \) and \( p_{11} \) are also associated with the clinic; these nodes have to be visited whenever some nurse has to take back blood samples before 10h00 AM or 11h00 AM, respectively.

With each node \( i \in V \) we associate

- \( r_i \): the time required to complete the treatment of patient \( i \in P \).
- \( r_O = r_D = r_{p_{10}} = r_{p_{11}} = 0 \).
- \([e_i, f_i]\): the time window of the arrival time at node \( i \in V \).

More explicitly, the time windows for the different nodes are specified as follows:
where $H_8 < H_{10} < H_{11} < H_{12}$ are constant values associated with 8h00 AM, 10h00 AM, 11h00 AM, and 12h00 AM, respectively.

The set $A = \{(i, j) \in V \times V \mid e_i + r_i + t_{ij} \leq f_{ij}\}$ includes the admissible arcs; i.e. $(i, j) \in A$ if the time windows of nodes $i$ and $j$ allow sufficient time to move from $i$ to $j$.

The variables are denoted as follows:

$k_{ij}^k$: the time when nurse $k$ visits successively nodes $i$ and $j$.

$D_{ij}^k$: the time when nurse $k$ takes blood samples to the clinic before 10h00 AM (corresponding to the time when nurse $k$ arrives at the fictitious node $p_{i0}$).

$D_{ij}^{11}$: the time when nurse $k$ takes blood samples to the clinic for turning back blood samples before 11h00 AM but after 10h00 AM (corresponding to the time when nurse $k$ arrives at the fictitious node $p_{i1}$).

$y_i^k$: the additional modeling binary variables, $\forall i \in P^+, k \in I$

Note that for all nurses $k \in I$, the variables $b_i^k$ are initialized at the following values:

$H_8 < b_i^k < H_{11}$, $i \in P^+$

$H_8 < b_i^k < H_{12}$, $i \in V - P^+ \cup \{p_{i0}, p_{i1}\}$

$b_{i0}^k = H_{10}$

$b_{i1}^k = H_{11}$

The value of $b_i^k$ may be modified during the solution procedure whenever the nurse $k$ visits the node $i$.

To specify the objective function, denote by $t_{ij}$, the traveling time from node $i$ to node $j$. The cost associated with nurse $k$ moving from $i$ to $j$ is specified as follows:

$c_{ij}^k = \begin{cases} 
    t_{ij} & \text{if } k \in I^r \cup I^l \cup I^f \text{ and } i \neq 0 \\
    t_{ij} + C_r & \text{if } k \in I^r \text{ and } i = 0 \\
    t_{ij} + C_l & \text{if } k \in I^l \text{ and } i = 0 \\
    t_{ij} + C_f & \text{if } k \in I^f \text{ and } i = 0 
\end{cases}$

According to this notation, the traveling time is considered as a cost. Also when the nurse $k$ is leaving the clinic, we add the daily cost corresponding to his category. The cost structure is also modified in order to account for the requirement that a patient should be visited by a nurse of his sector whenever possible.

Denote by $s(k)$ and $y(j)$ the sectors of the nurse $k$ and of the node $j$, respectively. Note that $0, D, p_{i0}$ and $p_{i1}$
are nodes for every sector. For any pair of nodes $i$ and $j$ and for any nurse $k$, the cost is specified as follows:

$$c_{ij}^k = \begin{cases} c_{ij}^k & \text{if } \varsigma(j) = s(k) \text{ or } j = D_p, p_0, p_1 \\ c_{ij}^k + C_a & \text{if } \varsigma(j) \neq s(k) \end{cases}$$

where $C_a$ is a positive parameter of the model.

The standard constraints for the VRPTW constraints (see Solomon (1987)) are summarized as follows:

1. \[ \sum_{j \in V} x_{ij}^k = 1, \forall i \in P \] (1)
2. \[ \sum_{j \in V} x_{ij}^k - \sum_{j \in V} x_{ji}^k = 0, \forall i \in V \setminus \{0, D_p\}, k \in I \] (2)
3. \[ \sum_{j \in V} x_{0j}^k \leq 1, \forall k \in I \] (3)
4. \[ \sum_{j \in V} x_{jD}^k = \sum_{j \in V} x_{0j}^k, \forall k \in I \] (4)
5. \[ b_i^k + r_i + t_{ij} - b_j^k \leq M \times (1 - x_{ij}^k), \forall (i, j) \in A, k \in I \] (5)
6. \[ e_i \leq b_i^k \leq f_i, \forall i \in V, k \in I \] (6)

Now additional constraints are required for the blood samples and the continuity of care requirements.

**4.1 Blood sample related constraints**

The blood sample constraints are more complex to formulate. The constraints (7) specify that whenever a nurse $k$ performs a blood sample on at least one patient $i$ before 10h00 AM, then he must visit the fictitious destination node $p_{10}$ (i.e. he has to return to the clinic).

$$1 + \left( H10 - b_i^k \right) \leq M \times \left( \sum_{j \in P} x_{jp_{10}}^k + 1 - \sum_{j \in P \cup \{p_{10}\}} x_{ij}^k \right), \forall i \in P^+, k \in I$$

Here $M$ is a very large scalar.

To understand this constraint, first note that, if any patient $i \in P^+$ is visited by nurse $k$ before 10h00 AM, then it follows that $\sum_{j \in P \cup \{p_{10}\}} x_{ij}^k = 1$ and $b_i^k \leq H10$.

Hence $(H10 - b_i^k) \geq 0$ and $1 + (H10 - b_i^k) > 0$. Also, since $\left( 1 - \sum_{j \in P \cup \{p_{10}\}} x_{ij}^k \right) = 0$, it follows that

$$M \times \left( \sum_{j \in P} x_{jp_{10}}^k + 1 - \sum_{j \in P \cup \{p_{10}\}} x_{ij}^k \right) = M \times \left( \sum_{j \in P} x_{jp_{10}}^k \right).$$
Thus constraint (7) reduces to

$$0 < 1 + \left( H_{10} - b_{i}^{k} \right) \leq M \times \sum_{j \in P} x_{jp10}^{k}$$

inducing that $$\sum_{j \in P} x_{jp10}^{k} > 0$$ (i.e., forcing a visit to the fictitious destination node $$p_{10}$$).

Note that if nurse $$k$$ does not visit any patient $$i \in P^+$$ before 10h00 AM, then $$b_{i}^{k} > H_{10}$$ for all $$i \in P^+$$. Hence the constraints (7) are inactive for all $$i \in P^+$$ in the sense that they are satisfied for any value of $$\sum_{j \in P} x_{jp10}^{k}$$. Thus $$\sum_{j \in P} x_{jp10}^{k} = 0$$ since the costs associated with all arcs $$(j, p_{10})$$ are positive.

Similar arguments can be used to verify that the constraints (8) and (9) force a nurse $$k$$ performing any blood sample after 10h00 to visit the fictitious node $$p_{11}$$.

$$1 + \left( b_{i}^{k} - b_{p10}^{k} \right) \leq M \times \left( \sum_{j \in P} x_{jp11}^{k} + 1 - \sum_{j \in P \cup \{p_{11}\}} x_{ij}^{k} \right),$$

$$\forall i \in P^+, \forall k \in I$$  \hspace{1cm} (8)

$$1 + \left( b_{i}^{k} - H_{10} \right) \leq M \times \left( \sum_{j \in P} x_{jp11}^{k} + 1 - \sum_{j \in P \cup \{p_{11}\}} x_{ij}^{k} \right),$$

$$\forall i \in P^+, \forall k \in I$$  \hspace{1cm} (9)

The constraints (8) apply when the nurse $$k$$ had to turn back to the clinic before 10h00 and constraints (9) when he did not.

Now even if the constraints (7) guarantee that the nurse $$k$$ visits the fictitious destination node $$p_{10}$$ whenever he performs any blood sample before 10h00, nevertheless these constraints do not prevent the visit to $$p_{10}$$ taking place later than 10h00 AM (i.e., having $$b_{i}^{k} > H_{10}$$), nor they prevent this visit taking place before any blood sample has been performed. For these reasons, we include the following constraints requiring also additional binary modeling variables $$y_{i}^{k}, \forall i \in P^+$$, to be formulated. Note that these variables are useful to formulate the model but they have no particular interpretation.

First we introduce a set of constraints guaranteeing that $$y_{i}^{k} = 1$$ for exactly one patient $$i \in P^+$$ visited by nurse $$k$$ before 10h00 AM, if any. The constraints (10)

$$y_{i}^{k} \leq \sum_{j \in V - D} x_{ji}^{k}, \forall i \in P^+, \forall k \in I$$  \hspace{1cm} (10)

guarantee that $$y_{i}^{k}$$ can take value 1 only if nurse $$k$$ performs a blood sample on some patient $$i \in P^+$$, i.e. only if

$$\sum_{j \in V - D} x_{ji}^{k} = 1$$

The constraints (11)

$$H_{10} - b_{i}^{k} \leq M \times \sum_{j \in P^+} y_{j}^{k}, \forall i \in P^+, \forall k \in I$$  \hspace{1cm} (11)
guarantee that whenever a blood sample is performed on some patient \( i \in P^+ \) before 10h00 AM (i.e., \( b_i^k \leq H10 \)), then \( \sum_{j \in P^+} y_j^k > 0 \).

Moreover, in order to guarantee that \( \sum_{j \in P^+} y_j^k = 1 \) whenever \( \sum_{j \in P^+} y_j^k > 0 \), we introduce the term \( \sum_{k \in I} \sum_{j \in P^+} y_j^k \) in the objective function to be minimized.

Finally, it follows that the constraints
\[
b_{p_{10}}^k - b_{p_{10}}^k \leq M \times (1 - y_i^k), \forall i \in P^+, \forall k \in I
\] (12)
guarantee that \( b_{p_{0}}^k \leq H10 \), inducing that nurse \( k \) returns to the clinic no later than 10h00 AM whenever \( \sum_{j \in P^+} y_j^k = 1 \). Furthermore, it is easy to see that the constraints (12) are inactive when nurse \( k \) is not performing any blood sample before 10h00.

Finally to guarantee that nurse \( k \) returns to the clinic no later than 10h00 AM but only after performing at least one blood sample on a patient, constraints (13) are added to the model.
\[
b_{i}^k - b_{p_{10}}^k \leq M \times (1 - y_i^k), \forall i \in P^+, \forall k \in I
\] (13)
Indeed, since \( y_i^k = 1 \) for exactly one patient \( i \in P^+ \) visited by nurse \( k \) before 10h00 AM, it follows that the corresponding constraint (13) is active. Hence \( b_{i}^k - b_{p_{10}}^k \leq 0 \), and \( b_{i}^k \leq b_{p_{10}}^k \).

It is easier to deal with the case where the nurse \( k \) has to visit \( p_{11} \) (i.e., returning to the clinic before 11h00 AM) since this is the last return. On the one hand, to guarantee that \( p_{11} \) is visited no later than 11h00 AM, it is sufficient to fix properly the upper bound of the time window associated with \( b_{p_{11}}^k \) to the value 11h00. On the other hand, the constraints (14) guarantee that \( p_{11} \) is visited after the last blood sample has been performed.
\[
b_{p_{11}}^k \geq b_{i}^k, \forall i \in P^+, k \in I
\] (14)

4.2 Continuity of care constraints

In order to account for the continuity of care requirement, a set \( L_k \) of patients is assigned to each regular nurse \( k \in I' \) such that \( L_k \cap L_h = \emptyset \) for each pair of nurses \( k \) and \( h \). Also the set \( \bigcup L_k \) includes only patients requiring a follow-up because of medical decision based on the kind of patient treatment or because of his general health state. Since the continuity of care requirement is considered as a soft constraint, then it is modeled by including a penalty cost each time a nurse \( k \) visits a patient \( j \) requiring follow-up by nurse \( h \). The costs in the objective function are then modified accordingly: for all pairs of nodes \( i, j \in P \), and for all regular nurses \( k \in I' \).

\[
\tilde{c}_{ij}^k = \begin{cases} 
c_{ij}^k + C_{cc} & \text{if } j \in L_h \text{ with } k \neq h \text{ and } k \in I' \\
c_{ij}^k & \text{otherwise}
\end{cases}
\]

where \( C_{cc} \) is the penalty cost for bypassing the continuity of care requirement.
Consequently, the objective function to be minimized is specified as follows

\[ \sum_{k \in I} \sum_{i,j \in A} c_{ij}^k x_{ij}^k + \sum_{k \in I} \sum_{j \in P^+} y_j^k \]

including two terms related to the nurses and to the modeling binary variables \( y_j^k \), respectively.

5. Solution Approach to solve the problem

We use a metaheuristic approach based on the scatter search method (Glover et al. (2003)) relying on an adaptive memory (Rochat et al. (1995)) including pools of solutions generated using Tabu search. The approach is summarized in Figure 1. In the next sections, we describe different elements of the procedure.

<table>
<thead>
<tr>
<th>Initialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generate a set of feasible solutions for the problem.</td>
</tr>
<tr>
<td>Select a subset of solutions to generate a pool including the best solutions.</td>
</tr>
<tr>
<td>In the rest of the set, select a second pool including the most different ones from those in the first pool in order to explore more extensively the feasible domain.</td>
</tr>
</tbody>
</table>

**Step 1: Generating an offspring solution**

Select solutions from the pools according to some criterion.

Generate an offspring solution by recombining (according to some operator) the selected solutions.

Improve the offspring solution using a Tabu Search procedure.

**Step 2: Updating the pools**

Consider the offspring solution:

- if it is better than the worst solution in the first pool, then it replaces the later in the first pool;
- if not, and if it is more different from those in the first pool than any solution in the second pool, then it replaces the later in the second pool;
- otherwise, it is discarded.

**Step 3: Stopping criteria**

Repeat Step 1 for 200 iterations.

Figure 1. Scatter Search Approach

5.1 Initialization

Each feasible solution of the problem is generated with the purpose of complying with the objective that each nurse should visit patients of his sector in priority. Thus each sector problem is solved to determine the nurse routes, and then they are merged into a feasible solution for the problem.

Two solution methods based on a Tabu search approach are developed to solve the sector problem. The ultimate goals of the two methods are basically different. The first method relying on Lau et al. (2003) approach aims at reducing the number of nurses required to complete the visits in the sector. Hence, at each iteration an additional nurse (if needed) is introduced, and a Tabu search procedure is used to optimize the routing and scheduling of the current nurse. The set of visits not completed belongs to some holding list. The procedure iterates until all visits are assigned (or until the holding list is not
empty), even if fictitious nurses must be introduced. (Note that the visits assigned to a fictitious nurse are not completed in fact).

The purpose of the second approach is to reduce the total cost of the assignment. An initial solution is generated using Solomon (1987) heuristic, and then a Tabu search method is used to improve the solution.

A feasible solution is initiated as the union of the sector solutions generated with either Lau-Tabu method or Solomon-Tabu method. In a first step, we try to reduce the number of nurses working. In order to do so, we try to eliminate, if possible, nurses visiting fewer patients. These patients are then redistributed among the remaining nurses and inserted in their routes. If any patient cannot be inserted in one of the remaining nurse routes then the corresponding nurse route can not be eliminated. During this elimination process, the fictitious nurses are the first to be eliminated, followed by nurses belonging to the recall list, and finally, the regular nurses visiting only a few patients. This step may result in a worse solution according to the value of the objective function.

The process is repeated to generate an adaptive memory including two pools of solutions: a pool $BP$ of best solutions and a pool $DP$ of diverse solutions. Starting with a set $E$ of solutions for the global problem, the pool $BP$ is generated by selecting the $|BP|$ best solutions chosen from $E$. To generate the pool $DP$, $|DP|$ solutions are selected sequentially among the $E_{BP}$ solutions. The next solution is selected according to a dissimilarity measure with respect to the solutions in $BP$ and those already included in $DP$. The dissimilarity used in our implementation is due to Rego and Leao (2001). The measure $d_{ij}$ is equal to the number of arcs that are different in solutions $X_i$ and $X_j$: $d_{ij} = \left| (X_i \cup X_j) \setminus (X_i \cap X_j) \right|$. Therefore the $|DP|$ solutions are the ones showing the greatest diversity compared to the solutions already in $BP$ and $DP$.

### 5.2 Step 1: Generating a new solution from the pools

In order to account for the additional constraints related to the continuity of care and the blood sampling, we construct a new partial solution by selecting sequentially for each nurse a workload assigned to him in some solution of the pools. More specifically, consider each nurse individually. We first determine the pool where his workload will be chosen. The selection of the pool is made randomly: the pool $BP$ is selected with a probability $p_b$, otherwise the pool $DP$ is selected. Once the pool is specified, we select the solution $i$ from which the nurse workload is chosen according to one of the following operators:

- **Roulette selection operator** of Goldberg (1989) which consists in choosing a nurse workload proportionally to its fitness measured in terms of the number of patients visited. Let $f_i$ be the fitness of the nurse workload in solution $i$. The probability of choosing the workload in solution $i$ among the nurse workloads in the set of the $t$ solutions is equal to $p_i = \frac{f_i}{\sum_{i=1}^{t} f_i}$.

- **Tournament selection operator** of Michalewicz (1996) which consists in choosing randomly a set of workloads for the nurse among all solutions and in keeping the one with the best fitness.

- **Random selection operator** which consists in choosing randomly any candidate workload for the nurse among all solutions.

The partial solution generated according to this process needs to be repaired, in general, to obtain a new feasible solution. Indeed, some patients may be assigned to several nurses (and hence they will have to be removed from some nurse workloads) and some patients may not be assigned to any nurse.

We consider three different ways to remove a duplicated patient from some nurse workloads:
• **Mode 1**: consider the order in which the nurse workloads are generated; leave the patient in the first workload generated and eliminate him from the others.

• **Mode 2**: consider the solutions of the pools including the nurse workloads where the patient is assigned; leave the patient in the nurse workload belonging to the best of these solutions and eliminate him from the others.

• **Mode 3**: leave the patient in the workload of the nurse commissioned for his continuity of care.

Finally, a new feasible solution is obtained by including the remaining patients not yet assigned. They are included sequentially as follows:

- Include the patient in the workload of the nurse commissioned for his continuity of care, if possible.
- Otherwise, include the patient in the workload of a nurse from his sector.
- Otherwise, include the patient in the workload of a nurse from another sector.
- Otherwise, add a new nurse to visit the patient.

The preceding repair process generates a new feasible solution to be improved according to the following procedure:

i. First redefine the sectors as follows. Consider the nurses belonging to an original sector. The new sector associated with this set of nurses includes all the patients assigned to them in the new feasible solution.

ii. Improve the current solution of each new sector using the Tabu search.

iii. Starting with the union of the sector solutions, apply the approach defined in the Initialization step to generate a new improved global solution.

### 6. Numerical results

A first set of problems is generated randomly according to Solomon process (1987) adapted to our problem. This set is used to determine proper values for the parameters and to analyze the different selection strategies, the ways to remove a duplicated patient, and to include the remaining patients. The second set includes real world data problems provided by the CLSC Les Forges in Trois-Rivières.

#### 6.1 Analysis using the first set of random problems

We use Solomon process (1987) adapted to account for the continuity of care and the blood sampling constraints to generate problems with 50 and 100 patients. To generate the sectors, we used the K-means partitioning algorithm of MacQueen (1967). Problems with 50 and 100 patients include 2 and 4 sectors, respectively. Twenty four problems (twelve with 50 and twelve with 100 patients) are generated according to the same characteristics used for the sector problem generation.

Preliminary tests were performed to determine the parameter values. The parameters values used for the tests are \( C_r = 400 \), \( C_l = 2 \), \( C_f = 800 \), \( C_f = 3 \), \( C_{ce} = 1200 \), \( C_{cc} = 100 \), \( C_a = 60 \).

To compare the different selection operators, the pool size is fixed to 5, and we use the first removing mode. The numerical results indicate that the 3-tournament selection operator generates better results using smaller computational time than the two other operators. This seems to indicate that a more elitist selection is beneficial.

The different removing modes are also compared using a pool size equal to 5 and the 3-tournament selection operator. The numerical results indicate that mode 1 largely dominates the other two modes with respect to the quality of the solutions generated and the computational time used. This may be due to the
fact that mode 1 induces some diversification strategy to explore more extensively the feasible domains. Indeed this mode does not account for the quality of the solution including the workload of the follow-up nurse.

Finally, using the 3-tournament selection operator and the first removing mode, the results for three different pool sizes are compared: 5, 10 and 20. On the one hand, the quality of the solutions generated using pools of size 20 is slightly better. On the other hand, the computational time increases rapidly with the pool size. Therefore, we use a pool size of 5 to complete the comparison of the 4 variants.

On the one hand, we compare the solution approach when using the two different pools of solutions generated with Lau et al. (2003) heuristic (LMultiApproach) and Solomon (1987) heuristic (STMultiApproach), respectively. But on the other hand, we would also like to see the advantage of using a population based approach like Scatter Search by comparing these results with one of the solution generated during the Initialization. Denote by LMonoApproach and STMonoApproach the solutions generated with Lau heuristic and Solomon heuristic, respectively.

The numerical results to compare the four solution method variants are summarized in Table 1. Comparing the two mono solution approaches, the results indicate that the LMonoApproach generates solutions where the average cost is reduced by 11% at the expense of an increase of the computational time by a factor of 38% with respect to STMonoApproach. Furthermore the solutions generated with the LMonoApproach require an average number of nurses reduced by a factor of 12%, but as expected, this induces an increase of the number of patients without follow-up by their assigned nurse by a factor of 21%. Finally, note that for both approaches the average number of nurses required reduces when the number of visits with blood sampling reduces. This is not surprising since the number of returns to the clinic should increase with the number of visits with blood sampling.

When we compare the two multi solution approaches, we observe that the average cost of the solutions generated with the LMultiApproach is smaller by a factor of 5% with respect to the STMultiApproach, but the computational time required by the LMultiApproach is larger. However note that the difference between the two multi solution approaches is not as large as between the two mono solution approaches. In order to analyse more thoroughly the efficiency of the two multi solutions approaches, we refer to the curves in Figure 2 showing the evolution of the average cost over time. The average cost is evaluated every 20 seconds starting at 35 seconds after the procedures started. The curves in Figure 3 illustrate clearly the dominance of the LMultiApproach over the STMultiApproach.

<table>
<thead>
<tr>
<th></th>
<th>Average cost</th>
<th>Average number of nurses</th>
<th>Average number of patients without follow-up</th>
<th>Average computational time</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMonoApproach</td>
<td>9159.68</td>
<td>13.62</td>
<td>6.34</td>
<td>3.86</td>
</tr>
<tr>
<td>STMonoApproach</td>
<td>10179.03</td>
<td>15.28</td>
<td>5.23</td>
<td>2.80</td>
</tr>
<tr>
<td>LMultiApproach</td>
<td><strong>8120.03</strong></td>
<td><strong>12.19</strong></td>
<td>4.14</td>
<td><strong>404.68</strong></td>
</tr>
<tr>
<td>STMultiApproach</td>
<td>8546.37</td>
<td>12.93</td>
<td><strong>3.17</strong></td>
<td><strong>385.05</strong></td>
</tr>
</tbody>
</table>

Table 1. Comparison between the mono and multi solution approaches
Overall the multi solutions approaches generate better solutions where the average cost is reduced by a factor of 16% with respect to the mono solution approaches. But this improvement in the quality of the solutions is at the expense of a significant increase (by a factor of 117%) of the computational time.

### 6.2.2 Numerical results with real value data

The following tests are completed with data provided by the Centre Local de Services Communautaires (CLSC) Les Forges à Trois-Rivières. The total number of patients visited each day varies between 80 and 100. The territory covered by the clinic is divided into four sectors. 14 regular nurses and 12 nurses from the recall list are available to complete the home care visits. Each regular nurse is assigned to one sector but she can complete some visits in other sectors, if necessary. Each regular nurse also has a list of follow-up patients.

We use data available for three different days of September 2005. Table 2 summarizes the data for the three problems. The clinic provided us with an estimated time of service for each patient according to his treatment but it is worth noticing that the estimated times provided may be higher in practice. For instance, a patient may need to talk with the nurse more than usual. This may have an impact when comparing our results with those of the CLSC. Using each patient address, we established the traveling time matrix. These traveling times are adjusted to account for the time needed to find a parking space, to walk from the car to the house, to remove material from the car, etc.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Number of patients</th>
<th>Number of regular nurses</th>
<th>Number of nurses from the recall list</th>
<th>% of blood sample visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>84</td>
<td>9</td>
<td>4</td>
<td>34.5</td>
</tr>
<tr>
<td>Day 2</td>
<td>92</td>
<td>13</td>
<td>3</td>
<td>45.6</td>
</tr>
<tr>
<td>Day 3</td>
<td>82</td>
<td>9</td>
<td>6</td>
<td>23.1</td>
</tr>
</tbody>
</table>

Table 2. Real value data characteristics

Analyzing the solutions generated manually, we notice that some of the constraints are not always satisfied:
• Some blood sampling visits are completed after 11h00.
• Some regular visits end after noon.

This can be explained by the fact that in practice some flexibility with respect to these constraints is allowed to the nurses. But recall that the blood sampling and the time window constraints are considered as hard constraints in our models.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Manual cost</th>
<th>Number of blood sampling constraints unsatisfied</th>
<th>Number of regular visits ending after noon</th>
<th>Number of nurses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>9623</td>
<td>6</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Day 2</td>
<td>10673</td>
<td>5</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Day 3</td>
<td>11413</td>
<td>2</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Mean value (over the 3 problems)</td>
<td>10569.7</td>
<td>3</td>
<td>3.33</td>
<td>14.67</td>
</tr>
</tbody>
</table>

Table 3. Unsatisfied constraints by manual solutions

These test problems are solved 5 times with each of the four approaches LMonoApproach, STMonoApproach, LMultiApproach and STMultiApproach. The numerical results are summarized in Table 4 where the gap of a solution method measures, in percentage, the deviation of its average cost from the best average cost among all four solution methods. The numerical results are consistent with those obtained for randomly generated problems. The average cost is smaller for the solutions generated with the LMultiApproach even though the computational time increases by only a factor of 2% with respect to the STMultiApproach.

Comparing the solutions generated manually with those obtained with the LMultiApproach, the results in Tables 3 and 4 indicate that more nurses (12.8% more on the average where this percentage is evaluated by the difference between the average number of nurses obtained with the LMultiApproach solution (13) and the average number of nurses obtained with the manual solution (14.67) divided by the average number of nurses obtained with the LMultiApproach solution (13)) are required and that the average cost is higher in the former. However some reserve is in order. Indeed the solutions obtained manually are pieced together for the data available that do not include all the information gathered for our computerized methods. Hence some values have been evaluated in order to make comparisons. Moreover, the traveling times in our methods are only estimates, and in practice, last minute modifications occur which are difficult to model.

Another aspect to consider is that our objective function is modeled using different costs according to the priority given to the minimization of the total time, to the possibility for nurses to visit patients from other sectors, and to the possibility for patients to be visited by other nurses than their assigned one, etc. The weights associated with these priorities may not reflect precisely what the head nurse has in mind when he completes the planning in practice. Hence using other values for these weights may induce different solutions not as good as those presented here. Despite all that, we feel that our results are very interesting and might be encouraging for head nurses to use a computerized tool to complete their planning. Not only the quality of the solutions seems better, but it takes much less time to generate them. Moreover with a computerized program it becomes easier to solve the problem again whenever last minute changes arise. Indeed the numerical efficiency of the approach would allow the head nurse to complete several runs accounting for last minute modifications. Furthermore, the values of the different parameters in the global model can be adjusted to the satisfaction of the head nurse in order to obtain proper results for her specific context. The choice of these values allows accounting for the proper relative emphasis of the different problem components.
<table>
<thead>
<tr>
<th></th>
<th>Average cost</th>
<th>Average number of nurses</th>
<th>Average number of patients without follow-up</th>
<th>Gap</th>
<th>Average computational time</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMonoApproach</td>
<td>10287,13</td>
<td>13,80</td>
<td>2.27</td>
<td>7.9 %</td>
<td>5.65</td>
</tr>
<tr>
<td>STMonoApproach</td>
<td>10611,53</td>
<td>13,80</td>
<td>2.33</td>
<td>11.3 %</td>
<td>4.29</td>
</tr>
<tr>
<td>LMultiApproach</td>
<td>9531,27</td>
<td>13,00</td>
<td>0.60</td>
<td>0.0 %</td>
<td>549.78</td>
</tr>
<tr>
<td>STMultiApproach</td>
<td>9791.00</td>
<td>13.07</td>
<td>1.40</td>
<td>2.7 %</td>
<td>540.22</td>
</tr>
</tbody>
</table>

Table 4: Comparisons with real value data

7. Conclusion

In this paper, we introduce solution methods to deal with the home care routing problem. First, solution methods used for the sector problem are described since they are used to generate the initial solutions for the global problem. Two main solution approaches, a mono and a multi solutions approaches, are compared numerically on randomly generated data and real value data. The results are very encouraging because these procedures offer the possibility to obtain good quality solutions in very reasonable time.

We agree that more tests involving various cost parameters could allow generating different results. Also, modeling the problem over several consecutive days instead of one day at a time, would allow accounting more extensively for the continuity of care constraints. Finally, the districting problem was not addressed in this paper but we think that allowing a dynamic redrawing of the territory covered by the clinic could result in more balanced workloads for the nurses.

8. Acknowledgements

We are grateful to Mrs Ginette Gélinas of CLSC Les Forges in Trois-Rivières for her constant collaboration and for providing us with the data necessary to complete the tests. Support for this project was provided by NSERC (Canada) and we are thankful for this help.

References


