Organization of external cares using the Large Neighborhood Search optimization

Timon Roxard, Lorraine Trilling, Maria Di Mascolo, Thibaud Monteiro

Organization of external cares using the Large Neighborhood Search optimization

1st Timon Roxard, 3th Maria Di Mascolo
timon.roxard@etu.univ-grenoble-alpes.fr
maria.di-mascolo@grenoble-inp.fr
Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP,
38000 Grenoble, France

2nd Lorraine Trilling, 4th Thibaud Monteiro
lorraine.trilling@insa-lyon.fr
thibaud.monteiro@insa-lyon.fr
INSA Lyon, Université Lumière Lyon 2,
Université Claude Bernard Lyon 1,
Université Jean Monnet Saint-Etienne,
DISP UR4570,
Villeurbanne, 69621, France

***Abstract*—In this article, we address the organization of external care in multiple sites and, in particular, the planning of the accompaniment of a group of patients with disabilities by a service. To model this VRP-TW, a mixed integer linear programming model is proposed and allows to optimize the travel time of professionals between care sites (schools), taking into account different constraints. These constraints include respect of patient and professional availabilities (time-windows), matching between patient need and professional competencies, capacities of the facilities where the care takes place, as well as scheduling rules specific to this problem. High computation time of the MILP has been observed when the size of the problem increases. In this paper, Large Neighborhood Search is investigated to reduce model computation time and propose an efficient resolution method.**

***Index Terms*—home health care, large neighborhood search, vehicle routing problem**

I. Introduction

In France, healthcare systems and the medico-social sector are undergoing rapid changes that require greater efficiency and better control of resources. With the aim of inclusion, SESSAD-type structures ("Service d'Education Spéciale et de Soins A Domicile", or Special education and home care service) provide individualized support to children and adolescents with disabilities in their normal living and educational environment. Multidisciplinary teams are responsible for this support, including speech and language therapists (SLTs), physiotherapists, and psychologists.

Our aim in this paper is to propose a model enabling efficient scheduling of SLTs activities within a SESSAD. We focus on the accompaniment of children (named patients in the remainder of the paper) by SLTs which takes place only within the schools. Our objective is to minimize the travel time of the SLTs but also to minimize the waiting time before the start of a session while taking into account different constraints in terms of schedules, competencies, and admission capacities of the facilities.

This problem can be modeled as a Vehicle Routing Problem with Time Windows (VRP-TW). Each patient is represented by its location and requires treatment with a specific frequency and a minimum number of days between sessions, as well as a specific skill. SLTs are also represented by their location and each has daily work hours. Availability time slots are defined for each patient and each SLT according to their schedule. The time is then divided into slots of one hour. One SLT cares for one patient per slot. The main goal is to minimize the travel time of the SLTs by finding a match between the schedules and the skills of each SLT and taking into account the capacity of the schools.

We can find numerous references on planning and scheduling in the home healthcare sector, which includes our problem. Most articles consider different objectives and methods for solving the home healthcare planning problem. A summary of all these articles is presented in the following literature review [2].

Note that our problem is at the intersection of different kinds of VRP problems, namely the periodic vehicle routing problem (P-VRP), the multi-depot vehicle routing problem (MD-VRP), and the vehicle routing problem with multiple time windows (VRP-MTW), with the addition of a specific, rarely studied demand for skills, which is very similar to a multi-product VRP (studies of this kind of problem focus essentially on the distribution of products in the vehicle, and never take into account any maximum customer capacity). In addition to all this, we take into account the rest time between each treatment for patients, the maximum working time for professionals, and the treatment time typical of all home healthcare problems. To illustrate how specific our skill requirements are, note that in our problem, a patient may need two different treatments requiring different skills, each requiring a dedicated visit.

Based on [2], as well as on other recent articles dealing with problems with similarities, it can be seen that most of the techniques specific to VRP do not work in our case. Especially, using a genetic algorithm seems to be very difficult in our case. Indeed, due to the time division in slots and the small number of patient slots it is almost impossible to change a piece of a route by day or to shift it in the day, and, due to the consideration of required skills, it is difficult to change the professional performing a task, since the skills and requests

may not match. Also, as it will be illustrated later, clustering methods are quite limited in this problem, again due to the low number of available patient slots.

In general, the few techniques that consider slots are not suitable for this problem, either because of the number of days and/or rest time, the skill constraint, or the multiplicity of these slots.

We could find an algorithm in the literature that seems to be adaptable to the problem, namely the Inferior First Bidirectional Searching Algorithm (IFBSA), which is described in [4]. However, adjusting it to consider rest times seems tricky.

In this paper, we will rely on some of our previous works on the subject, and especially [3], in which we proposed a MILP (mixed integer linear programming) model to organize health care activities outside the home. It allows us to determine which professional is responsible for each action, where it takes place, and how to organize the schedule to reduce costs. This model proposes to optimize the travel time of professionals through a lexicographic approach while reducing the waiting time before the start of sessions, but it has the disadvantage of exponential complexity, which makes the calculation time very high above six employees in the instance. We thus proposed, in [1], some approximate resolution techniques based on different ways of creating clusters and improving these clusters by incrementally changing some of the patients in their clusters. We thus managed to find rather accurate results for real instances, but still with an excessive computation time.

The objective of this paper is thus to explore new methods, based on LNS (Large Neighborhood Search) to solve efficiently our specific problem, by providing accurate results within a reasonable computation time.

The remainder of this paper is organized as follows: Section II, gives an exact MILP model, which is an improvement of the one presented in [3], that enables solving small instances of the problem. In Section III, we describe the new methods we propose in this paper. These methods are then applied to real data and the experiments and results are presented in Section IV. Some concluding remarks and research issues are finally given is Section V.

II. Mathematical model

Here is the MILP model, which will be used in the reconstruction part of the LNS presented in Section III. It is an improvement of the one found in [3]. The main improvement relies on the inclusion of capacities facilities constraints (Constraints 22).

Data

P : Set of patient sites.
E : Set of starting points of professionals.
$R = P \cup E$: Set of all locations.
D : Set of days.
S : Set of care types.
W : Set of time slots considered.
$T_{r_1 r_2}$: Travel time between locations r_1 and r_2.
Dem_{ps} : Patient's p request in care s.
$Disp_{rdw}$: 1 if r is available at slot w on day d, 0 otherwise.
QE_{es} : 1 if professional e is qualfied for care s, 0 otherwise.
$TTime$: Processing time.
$WTime$: Total time of a slot.
$MTime$: Margin time for start a treatment.
K_p : Rest time of p, minimum number of days required between two treatment.
M : Big M, a number large enough to be considered infinite if not multiplied by 0.
$SQE_{S'}$: set of professionals who all have at least one skill in the S' set (Calculated from QE_{es}).
$C_{r,d,w}$: capacity of institution r, on day d, slot w.
$V_{r',r}$: equal to 1 if r and r' are in the same school, 0 otherwise

Variables

$X_{r_1 r_2 edw}$: 1 if the professional e move from r_1 to r_2 at day d on slot w.
Y_{erd} : time of arrival of professional e on day d at slot w.
Tf_{ed} : time of arrival of professional e at his home on the days d.

Objective

$$optTrav = \min \sum_{r_1 \in R} \sum_{r_2 \in R} \sum_{e \in E} \sum_{d \in D} \sum_{w \in W} X_{r_1 r_2 edw} * T_{r_1 r_2} \tag{1}$$

$$\begin{aligned} optWaitT = \min & \sum_{e \in E} \sum_{d \in D} (Tf_{ed} - Y_{eed}) \\ & - \sum_{r \in R} \sum_{s \in S} Dem_{rs} * TTime \\ & - \sum_{r_1 \in R} \sum_{r_2 \in R} \sum_{e \in E} \sum_{d \in D} \sum_{w \in W} X_{r_1 r_2 edw} * T_{r_1 r_2} \end{aligned} \tag{2}$$

Constraints

$$X_{e_1 r e_2 dw} = 0 \quad \forall e_1 \in E, r \in R, e_2 \in E, d \in D, w \in W \tag{3}$$

$$X_{r e_1 e_2 dw} = 0 \quad \forall r \in R, e_1 \in E, e_2 \in E, d \in D, w \in W \tag{4}$$

$$X_{rredw} = 0 \quad \forall r \in R, e \in E, d \in D, w \in W \tag{5}$$

$$\begin{aligned} & X_{rpedw} \leq Avai_{pdw} * Avai_{edw} \\ & \forall e \in E, r \in R, p \in P, d \in D, w \in W \end{aligned} \tag{6}$$

$$X_{eped0} \leq Avai_{pd0} * Avai_{ed0} \quad \forall e \in E, p \in P, d \in D \tag{7}$$

$$\sum_{r_2 \in R} \sum_{e \in E} \sum_{w \in W} X_{r_2 r_1 edw} \leq 1 \quad \forall r_1 \in R, d \in D \tag{8}$$

$$\sum_{r \in R} \sum_{w \in W} X_{eredw} \leq 1 \quad \forall e \in E, d \in D \tag{9}$$

$$\sum_{r \in R} \sum_{w \in W} X_{reedw} \leq 1 \quad \forall e \in E, d \in D \tag{10}$$

$$\sum_{r_1 \in R} \sum_{r_2 \in R} X_{r_1 r_2 edw} \leq 1 \quad \forall e \in E, d \in D, w \in W \tag{11}$$

$$\begin{aligned} Y_{p_1 ed} + TTime + T_{p_1 p_2} \leq \\ Y_{ep_2 d} + M * \sum_{w \in W} (1 - X_{p_1 p_2 edw}) \\ \forall e \in E, p_1 \in P, p_2 \in P, d \in D \end{aligned} \tag{12}$$

$$\begin{aligned} Y_{eed} + T_{ep} \leq Y_{epd} + M * (1 - \sum_{w \in W} X_{epedw}) \\ \forall e \in E, p \in P, d \in D \end{aligned} \tag{13}$$

$$\begin{aligned} Y_{epd} - T_{ep} \leq Y_{epd} + M * (1 - \sum_{w \in W} X_{epedw}) \\ \forall e \in E, p \in P, d \in D \end{aligned} \tag{14}$$

$$\begin{aligned} (w * \sum_{r \in R} X_{rpedw} * WTime) - MTime \leq Y_{epd} \\ \forall e \in E, p \in P, d \in D \end{aligned} \tag{15}$$

$$\begin{aligned} Y_{epd} \leq M * (1 - \sum_{r \in R} X_{rpedw}) w + \sum_{r \in R} X_{rpedw} * WTime \\ + MTime \quad \forall e \in E, p \in P, d \in D \end{aligned} \tag{16}$$

$$\begin{aligned} Y_{er_1 d} \leq M * \sum_{r_2 \in R} \sum_{w \in W} X_{r_2, r_1 edw} \\ \forall e \in E, d \in D \end{aligned} \tag{17}$$

$$\begin{aligned} \sum_{r_1 \in R} \sum_{w \in W} X_{r_2 r_1 edw} = \sum_{r_2 \in R} \sum_{w \in W} X_{r_1 r_2 edw} \\ \forall e \in E, r_1 \in R, d \in D \end{aligned} \tag{18}$$

$$\begin{aligned} M * (1 - \sum_{r \in R} \sum_{w_2 = w}^{|W|} X_{predw_2}) \leq M * (1 - \sum_{r \in R} X_{r_1 r_2 edw}) \\ \forall e \in E, p \in P, d \in D, w \in W \end{aligned} \tag{19}$$

$$\begin{aligned} \sum_{r_2 \in R - \{r_1\}} \sum_{e \in E} \sum_{f=d}^{d+K_{r_1}} \sum_{w \in W} X_{r_1 r_2 efw} \leq 1 \\ \forall r_1 \in R, d \in [0, ..., |D| - K_{r_1}] \end{aligned} \tag{20}$$

$$\begin{aligned} Y_{erd} + \sum_{w \in W} (X_{reedw} * (T_{re} + TTime)) \leq Tf_{ed} \\ \forall e \in E, r \in R, d \in D \end{aligned} \tag{21}$$

$$\begin{aligned} \sum_{r_1 \in R} \sum_{r_2 \in P} \sum_{e \in E} X_{r_1 r_2 edw} * V_{pr_2} \leq C_{pdw} \\ \forall p \in P, d \in D, w \in W \end{aligned} \tag{22}$$

$$\begin{aligned} \sum_{e \in SQE_{S'}} \sum_{r \in R} \sum_{d \in D} \sum_{w \in W} X_{rpedw} \geq \sum_{s \in S'} Dem_{ps} \\ \forall p \in P, S' \subseteq S \end{aligned} \tag{23}$$

$$X_{r_1 r_2 edw} \in \{0, 1\} \quad \forall r_1 \in R, r_2 \in R, e \in E, d \in D \tag{24}$$

$$Y_{erd} \in \mathbb{N} \quad \forall r \in R, e \in E, d \in D \tag{25}$$

$$Tf_{ed} \in \mathbb{N} \quad \forall e \in E, d \in D \tag{26}$$

Objective 1 corresponds to the minimization of the total travel time, while Objective 2 corresponds to the minimization of the total waiting time, i.e. the time when the professionals are neither at home, nor performing a treatment, nor on a journey.

Constraints 3 to 7 initialize some variables to 0 to prevent inconsistencies in the solution.

Constraints 8 to 11 make sure that each professional starts and finishes the round at his own place, does only one treatment per slot, and doesn't treat the same patient twice on the same day.

Constraints 12 to 17 are used to ensure that available slots are respected, taking into account both treatment and travel times.

Constraints 18 and 19 ensure the conservation of path flow.

Constraint 20 checks that the time between two treatments for the same patient is respected.

Constraint 21 correctly defines the professionals' end-of-work time.

Constraint 22 limits visits to patients in the same institution to the capacity of that institution. Capacities depend on the day and the slot, which gives our model flexibility without losing performance.

Contraint 23 constraint, which is rather atypical, checks that patient demand is met for each type of care, i.e. a type of care can be associated with each visit the patient receives, such that demand is met. It is inspired by Hall's theorem

and has the advantage of not introducing any new variables, but the disadvantage of being exponential depending on the number of different types of care required by a patient. As in our real-life case (and certainly in most concrete cases of Home Healthcare), patients don't require many types of care, and this constraint is quite advantageous.
Constraints 24 to 26 define the variable domain.

III. Solving method

The MILP model presented in the previous section has been tested on real data including patients and healthcare professionals. When the size of the instance increases, the computation time of the optimal solution increases exponentially. The computation time of the optimal solution becomes too high from 6 professionals and 40 patients. To overcome this problem, we propose to develop a heuristic based on Large Neighborhood Search (LNS) [7] [6] .

The principle of local search (LS) is to start from a single solution, and destroy it to rebuild it differently. This allows to visit a certain neighborhood of the first solution. The major risk of LS is to get stuck in local optimal and have very low diversity in the solutions created. In Large Neighborhood Search (LNS), a large part of the solution is destroyed and repaired. In one iteration, a transformation by a Destroy-Repair method can give many more different solutions than LS.

A. Destruction heuristics

In most LNS algorithms, some nodes (in this case visits) are destroyed based on a heuristic or completely at random. However, for the problem we tackle, this may not be sufficient, given the relatively small space of feasible solutions, due to low patient availability and skill constraints. Indeed, it is probable that most of the nodes destroyed will be rebuilt identically. To avoid this, we propose to destroy entire portions of a route, entire routes, or even the entire schedule of certain professionals, randomly. Another proposal to destroy a large part of the solution is to use a random based on gradient noise.

To perturb the solution, a heuristic that associates a score or weight to each arc or node is used. An arc is a link between two nodes of the graph (either a patient or a professional). We then destroy as many arcs as possible, so that the destroyed arcs all have a score strictly higher than all the arcs kept, however destroying fewer arcs than a certain limit.

Different heuristics have been studied to associate a weight to each node or arc. Those weigths are then used to select the part of the solution destroyed:

1) Random node: Associates a positive random weight to each patient, and a weight equal to 0 to each professional. The weight associated with each arc is the greatest weight associated with the nodes of the arc.

2) Random visit: Associates a positive random weight with each visit. The weight associated with each arc is the greatest weight associated with the patients concerned, 0 if it is an arc from a professional to himself, meaning an empty day.

3) Random edge: Associates a positive random weight with each arc.

4) Worst deviation: Associates with each visit the travel time "lost" because of the visit, i.e. the travel time passing through the visit minus the travel time skipping the visit. The weight associated with each arc is the greatest weight associated with the patients concerned, 0 if it is an arc from a professional to himself, meaning an empty day.

5) Longest tour: Associates with each arc the travel time for the tour in which it appears.

6) Random tour: Randomly selects one of the tours and removes it completely from the current solution.

7) Random cluster: Randomly selects a patient/professional and associates each patient/professional with the opposite of the distance to the selected patient. The weight associated with each arc is the greatest weight associated with the nodes of the arc.

8) Intersection: Associates with each patient the number of professionals who visit them, 0 to professionals. The weight associated with each arc is the greatest weight associated with the nodes of the arc.

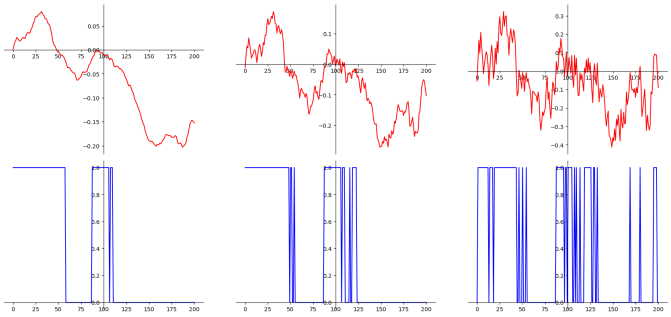
9) Random based on gradient noise: Associates a positive random weight to each arc, based on gradient noise in order to destroy "by portion" and not "by bit". This random-based gradient noise relies on Perlin theory [5], which introduces a method for generating pseudo-random and continuous functions based on gradient noise, definable on real spaces of any dimension. By summing several of those noises with different amplitudes and frequencies, we obtain new, more complex functions called fractal noises. We empirically choose to sum 20 noises, each with a frequency half that of the previous one and an amplitude a half that of the previous one, a being calculated as a function of a factor we call g granularity (the closer g is to 0, the closer the result is to Perlin noise, the larger g is, the closer the result is to a random function):

$$a = \begin{cases} \dfrac{1}{1+\dfrac{1}{g}} & \text{If } g > 0 \\ 0 & \text{If } g = 0 \end{cases}$$

This allows us to generate Boolean functions with varying degrees of consistency, as in Fig. 1. This Perlin noise acts as a filter to choose whether the arc is destroyed or not. In the experiments, we choose a granularity g that is equal to 1. This allows the destruction of a larger part of the solution.

B. Reconstruction

For the reconstruction, only an optimal reconstruction using the MILP model presented in section II was used. The main reason for this choice was based on the problem's complexity. The number of very different constraints is high, which makes it difficult to implement a heuristic that provides a correct solution relatively often. However, in future work, we could consider heuristics based on classic heuristics (cheapest insertion, optimal local insertion, etc.) with "preliminary steps".



(a) $g = 1$ (b) $g = 2$ (c) $g = 4$

Fig. 1: some examples of fractal noise of different granularity on the same seed (in red), and the associated Boolean function (in blue)

For example, if a request can only be met by a professional in a certain time slot, we add this visit to the solution. The benefit of using only one reconstruction method is to show the absolute potential of each destruction method, regardless of the reconstruction heuristic.

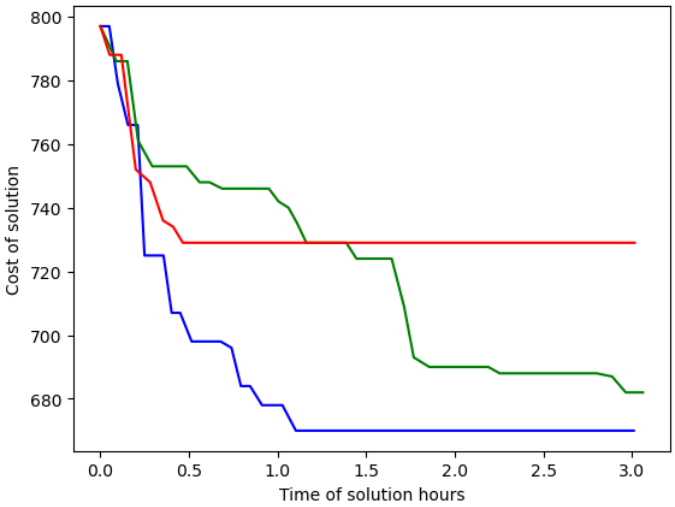
IV. Experiments and results

Using data from a SESSAD, we generated instances that took into account the actual schedules of speech therapists, the average duration of treatment (45 min), patient demands, and availabilities. Travel time values were calculated using queries made on the OSRM website (Open Source Routing Machine, http ://project-osrm.org/) using GPS coordinates of patients' schools and professionals' homes. We analyzed the schedules of schoolchildren, college students, and high-school pupils, to work out the availability of patients, targeting free periods.

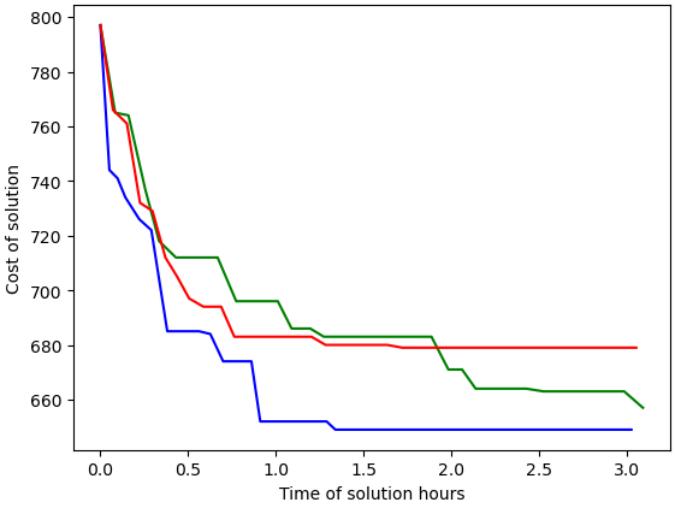
We aim to test the LNS's ability to find high-performance solutions within a reasonable computation time. Thus, and as the problem is NP-hard, it is impossible to guarantee finding the optimal solution for problems whose size corresponds to that of the real problem. So, our aim is to find solutions close to the optimal solution with a reduced computation time and respecting all constraints.

To determine which destruction operators are useful, we also wanted to test the impact of these various operators. This impact is both on the quality of the solution found, which essentially depends on the ability of the meta-heuristic to extract itself from a local optimum. The impact of operators can also be seen from the point of view of LNS convergence speed. This speed illustrates the ability to find a good solution in a reduced computation time.

Initially, we tested all 9 destruction operators. We found that operators with very little impact on the current solution were systematically dominated by operators with a wider impact. So in this paper, we have focused our presentation on the 3 most promising operators which are the ones that lead to a wider neighborhood: (1) Random based on gradient noise (red line in figures 2 and 3), (2) Random tour (green line), and (3) Random cluster (blue line).



(a) Instance with 60 patients



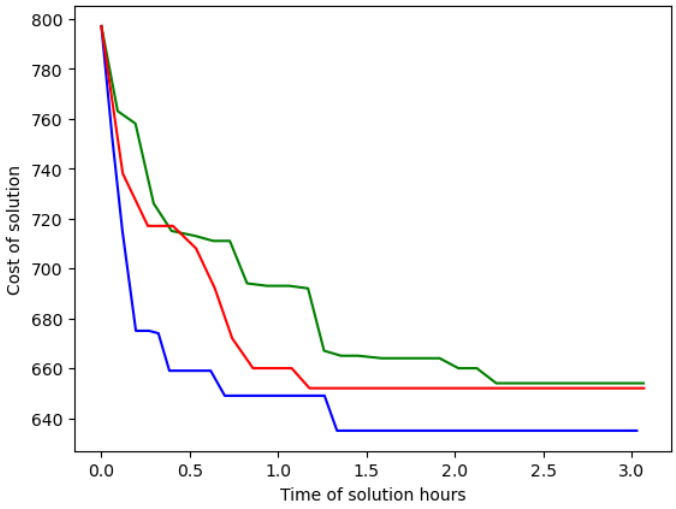
(b) Instance with 80 patients

Fig. 2: Convergence curves for instances with 60 and 80 patients - Random based on gradient noise (red line), Random tour (green line), Random cluster (blue line)

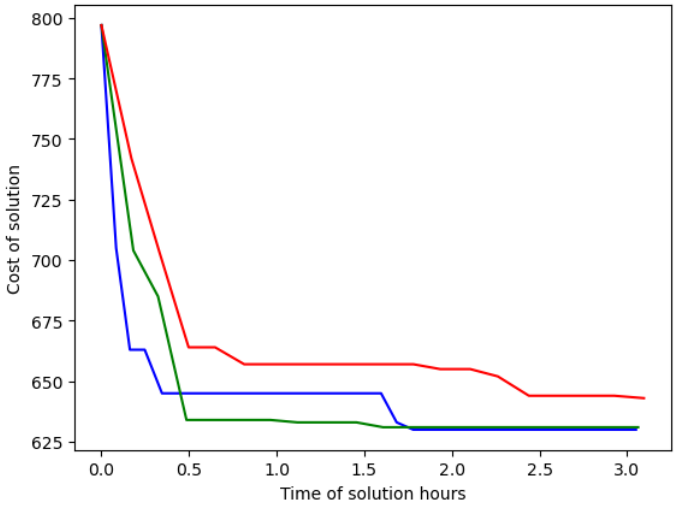
Figures 2 and 3 show convergence curves for instances ranging from 60 to 120 patients.

We can conclude that the LNS approach can propose solutions with a reasonable computation time. LNS is also able to handle larger problems than exact methods, which were restricted to problems with only 40 patients.

In terms of solution quality and in terms of convergence speed, the destruction operator that dominates the other ones is the Random cluster that consists of randomly deleting all patients of a professional. Random cluster is the operator that most extensively destroys the current solution. This operator, combined with an optimal reconstruction process, drives to better solutions. This approach allows a broader destruction of the current solution and facilitates the exit from the search for a local optimum.



(a) Instance with 100 patients



(b) Instance with 120 patients

Fig. 3: Convergence curves for instances with 100 and 120 patients - Random based on gradient noise (red line), Random tour (green line), Random cluster (blue line)

V. Conclusions

We have proposed a model for solving the problem of assigning and routing speech therapists to children with disabilities. As this problem is NP-hard, we used a heuristic approach to solve it in a reasonable computation time. The LNS meta-heuristic was tested and specific operators were proposed. Tests have demonstrated the ability of this approach to tackle problems whose size is similar to those of the initial real-life problem.

The operator that destroys the complete program of a speech therapist seems the most interesting if coupled with an exact reconstruction process.

A more in-depth study involving the combination of several operators would be useful to take advantage of the strengths of each.

Acknowledgment

For the purpose of Open Access, a CC-BY public copyright licence has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission.

References

[1] Estelle Collomb, Maria Di Mascolo, Lorraine Trilling, and Thibaud Monteiro. Optimization of the scheduling of medico-social activities at multiple sites: a clustering approach. In *24ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2023)*, Rennes, France, February 2023.

[2] Maria Di Mascolo, Clea Martinez, and Marie-Laure Espinouse. Routing and scheduling in Home Health Care: A Literature Survey and Bibliometric Analysis. *Computers & Industrial Engineering*, 158:1–48/107255, 2021.

[3] Alois Franzino, Trilling Lorraine, Di Mascolo Maria, and Monteiro Thibaud. Intérêt de la mutualisation des professionnels pour la planification d'activités médico-sociales. *Génie industriel et productique*, 5(Numéro spécial Gestion et ingénierie des systémes hospitaliers), 2022.

[4] Y H Li, H J Mao, and Y M Qin. Vehicle routing problem with multiple time windows and batch splitting based on inferior first bidirectional search algorithm. *Journal of Physics: Conference Series*, 1314(1):012116, oct 2019.

[5] Ken Perlin. An image synthesizer. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '85, page 287–296, New York, NY, USA, 1985. Association for Computing Machinery.

[6] David Pisinger and Stefan Røpke. *Large Neighborhood Search*, pages 399–420. Springer, 2 edition, 2010.

[7] Paul Shaw. Using constraint programming and local search methods to solve vehicle routing problems. pages 417–431. Springer Berlin Heidelberg, 1998.