

Différentes Mesures de la diversité de la population d'un Algorithme Génétique pour l'optimisation du parcours patient.*

Gérard Olivier^{1,2}, Brisoux Devendeville Laure², Lucet Corinne², Darras Sylvain¹

¹ Laboratoire MIS (UR 4290), Université Picardie Jules Verne, France
{olivier.gerard, laure.devendeville, corinne.lucet}@u-picardie.fr

² Evolucare Technologies, France
{o.gerard, s.darras}@evolucare.com

Mots-clés : *algorithme génétique, planification, santé, optimisation.*

1 Introduction

L'épidémie de COVID-19 a mis en exergue la nécessité de moderniser le système de santé dans de nombreux pays. Parmi les axes d'amélioration possibles pour mieux prendre en charge les patients et faciliter le quotidien des soignants, nous nous intéressons dans cet article à la planification des rendez-vous médicaux. Ce problème NP-difficile appartient à la catégorie Resource Constrained Project Scheduling Problem (RCPSP) [1, 4] et a été le sujet de nombreuses études ces dernières décennies, particulièrement dans le domaine de la santé ces dernières années. Résoudre un problème RCPSP revient à trouver les meilleures affectations de ressources et les meilleures dates de début à affecter à un ensemble d'activités. Pour notre problème dans le domaine de la santé, cela consiste à affecter des dates de début et un ensemble de ressources médicales (personnel médical, salles et équipements) à un ensemble de rendez-vous nécessitant généralement la présence d'un patient spécifique. Dans cet article, nous proposons un algorithme génétique et différentes métriques nous permettant d'agir sur la diversité des populations pour résoudre notre problème.

2 Définition du problème

Notre problème de planification est défini comme suit : un ensemble A de rendez-vous doit être planifié sur un horizon H . Chaque rendez-vous $a \in A$ est caractérisé par sa durée, sa fenêtre de faisabilité $[ES_a, LS_a] \subseteq H$, et une quantité $qtreq_a^\pi$ de ressources de propriété π requise.

Chaque ressource r appartenant à l'ensemble fini des ressources renouvelables R peut posséder différentes propriétés $\pi \in \Pi$, qualifiant par exemple différentes compétences. À chaque rendez-vous sont également associés deux coefficients intervenant dans la priorité qui lui sera donnée pour sa planification : un coefficient d'importance $Essential_a$ et un coefficient d'urgence $Emergency_a$. Des contraintes de précedence s'appliquent entre les rendez-vous et certaines ressources peuvent être pré-affectées afin de représenter le suivi et les préférences des patients.

La qualité d'une solution Sol est évaluée par le nombre de rendez-vous $a \in \overline{Sol}$ qui n'ont pu être planifiés, pondéré par leur importance $Essential_a$ et par la somme des retards des rendez-vous $a \in Sol$ relativement à ES_a , pondéré par leur urgence $Emergency_a$. Notre objectif est de trouver une solution valide tout en minimisant la fonction objective f définie dans l'équation 1.

$$f = \sum_{a \in \overline{Sol}} Essential_a + \sum_{a \in Sol} \frac{t_a - ES_a}{LS_a - ES_a} \times Emergency_a \quad (1)$$

*Ce projet est soutenu par le projet LORH (CIFRE N° 2018/0425 entre Evolucare et le Laboratoire MIS)

3 Méthode de résolution

L'algorithme génétique est une méthode s'inspirant de la théorie de l'évolution de Darwin pour créer des populations de solutions, les faire évoluer et interagir pour voir émerger les solutions les plus adaptées au problème [3]. Cette métaheuristique est utilisée pour résoudre de nombreux problèmes, et s'avère être particulièrement efficace sur les problèmes de planification [5]. La population P de solutions valides est initialisée avec un algorithme glouton, puis passe par des processus de reproduction, de mutation et de sélection pour explorer l'espace de recherche et sélectionner les solutions pour chaque génération suivante. Ce cycle se répète jusqu'à ce qu'une solution satisfaisante émerge du processus d'évolution.

Lors de l'exécution d'un algorithme génétique, il arrive souvent que les individus de la population P convergent peu à peu vers des optimums locaux. Dans ces situations et sans mesures appropriées, il devient difficile de sortir de ces puits d'attractivité pour trouver les meilleures solutions. Pour mesurer et limiter ce phénomène, nous avons étudié et élaboré plusieurs métriques en nous inspirant des distances de Hamming [2], de Levenshtein [7], et de Kendall Tau [6]. Ces métriques nous permettent de calculer la distance entre deux individus Sol_1 et Sol_2 et donc la diversité de la population P . En mesurant la diversité au fil de l'exécution de l'algorithme, nous pouvons mettre en place des méthodes pour la faire varier en fonction de nos besoins d'exploration de nouvelles solutions ou d'exploitation des plus prometteuses. L'une de ces méthodes consiste à adapter le taux de mutation, un taux plus haut entraînant plus de modifications sur chaque individu et provoquant de fait une hausse de la diversité. L'autre est une méthode de sélection différente, prenant en compte la distance aux individus déjà sélectionnés en plus de la qualité des individus.

4 Expérimentations

Pour évaluer l'efficacité de notre algorithme génétique, nous avons construit de nombreuses instances à partir de scénarios élaborés avec des planificateurs de différents établissements de santé en France. Concernant les métriques, nous avons étudié et comparé les résultats obtenus avec chaque métrique sur différentes instances, pour voir lesquelles nous permettaient de mesurer efficacement la diversité de la population. Nous comparons les résultats obtenus avec l'algorithme génétique aux solutions obtenues avec un modèle de programmation linéaire 0-1 implémenté sous CPLEX, ainsi qu'aux résultats obtenus avec une méthode ALNS. L'ALNS et l'algorithme génétique obtiennent des solutions proches des solutions optimales fournies par CPLEX, et l'algorithme génétique fournit de meilleures solutions que l'ALNS. Les résultats seront présentés en détail lors de la conférence.

Références

- [1] P. Baptiste, P. Laborie, C. LePape, and W. Nuijten. Constraint-based Scheduling and Planning. *Foundations of artificial intelligence*, 2 : 761–799, 2006.
- [2] R.W. Hamming. *Coding and Information Theory*. Prentice-Hall, Inc., 1986.
- [3] J.H. Holland. Genetic Algorithms. *Scientific American*, 267(1) : 66–73, 1992.
- [4] D.S. Johnson and M.R. Garey. *Computers and Intractability : A Guide to The Theory of NP-completeness*. WH Freeman, 1979.
- [5] R.L. Kadri and F.F. Boctor. An Efficient Genetic Algorithm to Solve the Resource-Constrained Project Scheduling Problem with Transfer Times : The Single Mode Case. *European Journal of Operational Research*, 265(2) : 454–462, 2018.
- [6] M.G. Kendall. A New Measure of Rank Correlation. *Biometrika*, 30(1/2) : 81–93, 1938.
- [7] V.I. Levenshtein et al. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics Doklady*, 10(8) : 707–710, 1966.