

# Guidage des métaheuristiques par machine-learning, application au transport d'enfants en situation de handicap

Timothée Chane-Hai<sup>1</sup>, Samuel Vercraene<sup>1</sup>, Céline Robardet<sup>2</sup>, Thibaud Monteiro<sup>1</sup>

<sup>1</sup> Univ Lyon, INSA Lyon, Université Claude Bernard Lyon 1, Univ Lumière Lyon 2, DISP, EA4570,  
69621 Villeurbanne, France

{timothee.chane-hai, samuel.vercraene, thibaud.monteiro}@insa-lyon.fr

<sup>2</sup> Univ Lyon, INSA Lyon, CNRS, LIRIS UMR 5205, F-69621, Lyon, France  
celine.robardet@insa-lyon.fr

**Mots-clés** : *transport à la demande, transport sanitaire, machine-learning, recherche guidée*

## 1 Motivations

Ces travaux de recherche ont pour but d'améliorer le transport sanitaire en région Auvergne-Rhône-Alpes en France. Le transport de personnes en situation de handicap, de leur domicile vers les institutions médicales, est un problème de transport à la demande, nommé *dial-a-ride problem* (DARP) en recherche opérationnelle. Durant les travaux précédents sur le sujet, la méthode développée dans [3] a permis de résoudre des problèmes réels comprenant jusqu'à 600 usagers. Le but de ces travaux est de réduire le temps de calcul en exploitant les caractéristiques du problème.

Habituellement, les chercheurs s'attellent à construire des heuristiques robustes et correctement paramétrées pour résoudre les problèmes de transport et obtenir des résultats satisfaisants sur une grande diversité d'instances. Mais dans [1], une nouvelle approche a été proposée. L'accent a été mis sur l'étude des caractéristiques structurelles du problème en amont de la construction de l'heuristique. En connaissant les différences structurelles entre les bonnes et les mauvaises solutions, il est possible de construire des heuristiques efficaces qui intègrent ce savoir dans leur fonctionnement. Cette méthode de recherche guidée leur permet de résoudre de très grandes instances de *vehicle routing problem* (VRP), allant jusqu'à 30000 usagers en un temps de calcul raisonnable.

Le but de ces travaux est d'améliorer les performances obtenues dans [3] sur les problèmes réels en élargissant la méthode utilisée dans [1] du VRP au DARP.

## 2 Méthode

La méthode pour construire une heuristique de recherche guidée se décompose en trois étapes : la génération des données, l'analyse des données, l'intégration des opérateurs de recherche guidée dans l'algorithme d'optimisation.

Durant l'étape de génération des données, la base de données qui servira à l'apprentissage des caractéristiques structurelles du problème est créée. Tout d'abord, un nombre important d'instances est généré. Comme le nombre d'instances réelles est limité, il est nécessaire de générer des instances artificielles mais crédibles. Ensuite, pour chaque instance, deux types de solutions sont générées et sauvegardées : une proche de l'optimalité (bonne solution) et une éloignée de l'optimalité (mauvaise solution). Le but est d'apprendre les différences entre ces deux types de solutions. Pour représenter la structure d'une solution, plusieurs indicateurs sont générés pour une solution et l'instance correspondante. Ainsi, une entrée de la base de données correspond aux indicateurs de la solution, aux indicateurs de l'instance et à la qualité de la solution (bonne/mauvaise).

Lors de l'étape d'analyse de données, des algorithmes de classification sont entraînés pour prédire la qualité d'une solution en se basant sur ses caractéristiques. Dans un premier temps, les classificateurs sont entraînés en utilisant tous les indicateurs. Dans un second temps, ils sont entraînés en utilisant un seul indicateur de la solution à la fois. Dans les deux cas, si plus de la moitié des prédictions sont correctes, alors cela montre qu'il est possible de prédire correctement la qualité d'une solution en se basant sur un ou plusieurs indicateurs. En d'autres termes, les caractéristiques exprimées par ces indicateurs permettent de différencier les bonnes des mauvaises solutions.

Finalement, l'heuristique est construite à partir de l'analyse de données. Les nouveaux opérateurs de recherche guidée vont cibler les faiblesses des solutions en se basant sur les valeurs des indicateurs.

Prenons l'exemple suivant pour illustrer la méthode. Lors de l'étape d'analyse de données, l'indicateur "nombre de passagers moyen par route" permet de prédire correctement la qualité (bonne/mauvaise) de 75% des solutions. Une analyse plus détaillée sur cette indicateur révèle que les bonnes solutions ont tendance à présenter une charge moyenne plus importante que les mauvaises solutions. Il est donc possible de créer un opérateur qui cible les routes à faible charge moyenne et favorise l'insertion d'utilisateurs dans ces routes.

### 3 Résultats préliminaires et perspectives

Avant de travailler sur des instances réelles, nous réalisons une preuve de concept sur des instances artificielles. Les instances artificielles sont générées avec la méthode issue de [2]. Des résultats prometteurs sont obtenus.

Des instances contenant entre 20 et 80 usagers sont créées. Le dépôt peut être positionné au centre, sur un bord ou aléatoirement. Trois groupes de solutions avec des niveaux d'optimalité différents sont générés : 0% (groupe de référence), 2%, 4%. Les algorithmes de classification utilisés sont les arbres de décisions et les machines à vecteurs de support (SVM) linéaires. Différents indicateurs sont créés et testés collectivement et individuellement. Utilisés collectivement, le taux de bonnes prédictions atteint jusqu'à 75% pour les arbres de décisions avec le groupe de solutions à 4%. Utilisés individuellement, deux indicateurs se distinguent : le coût moyen de l'arc par route et le rang de connexion moyen. Dans une solution, si un usager est connecté à son plus proche voisin en terme de coût, son rang de connexion est de 1. Si il est connecté à son deuxième plus proche voisin, son rang de connexion est de 2, et ainsi de suite. Sur le groupe de solutions à 4%, les résultats obtenus sont donc de 73% et 63% de bonnes prédictions respectivement pour l'indicateur sur le coût moyen par route et l'indicateur sur le rang de connexion moyen.

Si la suite de ces travaux de recherche confirme ces résultats initiaux, la méthode sera adaptée au transport réel d'enfants en situation de handicap en utilisant les données issue du projet NOMAd.

### Références

- [1] Florian Arnold and Kenneth Sørensen. What makes a VRP solution good? The generation of problem-specific knowledge for heuristics. *Computers & Operations Research*, 106 :280–288, June 2019.
- [2] Jean-François Cordeau. A Branch-and-Cut Algorithm for the Dial-a-Ride Problem. *Operations Research*, 54(3) :573–586, June 2006. Publisher : INFORMS.
- [3] Oscar Tellez, Samuel Vercraene, Fabien Lehuédé, Olivier Péton, and Thibaud Monteiro. The fleet size and mix dial-a-ride problem with reconfigurable vehicle capacity. *Transportation Research Part C : Emerging Technologies*, 91 :99–123, June 2018.