# Problème de couverture par ensembles : une approche mémétique

Maxime Pinard[1,2], Laurent Moalic[1,2], Mathieu Brévilliers[1,2], Julien Lepagnot[1,2], Lhassane Idoumghar[1,2]

[1] Université de Haute-Alsace, IRIMAS UR 7499, F-68100 Mulhouse, France

[2] Université de Strasbourg, France

maxime.pin@live.fr

{laurent.moalic,mathieu.brevilliers,julien.lepagnot,lhassane.idoumghar}@uha.fr

**Mots-clés** : *métaheuristiques, algorithme mémétique, problème de couverture par ensembles*

## 1 Problème de couverture par ensembles

Soit un ensemble univers $U = \{u_1, \dots, u_m\}$ et une famille $S = \{s_1, \dots, s_n\}$ de sous-ensembles de $U$, le problème de couverture par ensembles (Unicost Set Cover Problem, USCP) consiste à trouver la plus petite sous-famille de $S$ permettant de couvrir chaque élément de $U$ au moins une fois. Une instance du problème est définie par une matrice d'incidence $A = (a_{i,j})$ de taille $m \times n$ avec :

$$\forall i \in \{1, \dots, m\},\ \forall j \in \{1, \dots, n\},\ a_{i,j} = \begin{cases} 1 & \text{si } u_i \in s_j \\ 0 & \text{sinon} \end{cases}$$

Une solution est définie par un vecteur $n$-dimensionnel $x = (x_j)$ avec :

$$\forall j \in \{1, \dots, n\},\ x_j = \begin{cases} 1 & \text{si } s_j \text{ fait partie de la solution} \\ 0 & \text{sinon} \end{cases}$$

Une solution est valide si et seulement si :

$$\forall i \in \{1, \dots, m\}\ ,\ \sum_{j \in \{1,\dots,n\}} a_{ij} x_j \geq 1$$

La fonction objectif à minimiser est :

$$f(x) = \sum_{j \in \{1,\dots,n\}} x_j$$

Ce problème est un cas particulier du problème de couverture par ensembles pondérés (Weighted Set Cover Problem, WSCP), le USCP et le WSCP sont des problèmes d'optimisation NP-difficiles, et NP-complets dans leur forme décisionnelle. Le USCP est généralement considéré comme plus difficile à résoudre que le WSCP [1] car les poids des sous-ensembles sont unitaires et n'apportent donc aucune information pour guider un algorithme dans la recherche d'une solution optimale.

## 2 Algorithme mémétique à deux individus

Un algorithme mémétique très efficace, mêlant une recherche locale et un algorithme génétique avec une population de seulement deux individus, a été récemment proposé pour résoudre le problème de coloration de graphe [2]. Le but du travail présenté ici est d'adapter cet algorithme pour résoudre le USCP. La recherche locale choisie pour cette hybridation est la Row Weighting Local Search (RWLS) [3], qui s'avère être à ce jour l'algorithme le plus performant pour résoudre les instances de benchmark classiques du SCP [4]. La principale caractéristique de RWLS réside dans un schéma de pondération des éléments de $U$, qui permet de donner plus de poids aux éléments difficiles à couvrir, et de favoriser les sous-ensembles qui couvrent ces éléments.

L'hybridation proposée a comme paramètres le nombre de générations et le nombre d'étapes de RWLS à réaliser par solution pour chaque génération. La population est initialisée avec 2 solutions aléatoires et 2 vecteurs de poids de valeur 1. Ensuite, pour chaque génération on applique RWLS aux solutions de la population et on génère les solutions de la génération suivante en croisant les solutions et les vecteurs de poids utilisés par RWLS. Le croisement des parents permet d'obtenir une solution fille de façon "gloutonne", en prenant alternativement dans une solution puis l'autre le sous-ensemble qui augmente le plus le nombre d'éléments couverts par la solution. Ce même croisement est appliqué en commençant par le 2$^{ème}$ parent pour générer le deuxième enfant. Notons que les poids des colonnes des deux enfants sont obtenus en effectuant la moyenne des poids des deux parents.

# 3 Résultats expérimentaux

L'algorithme proposé a été comparé a RWLS sur les instances de benchmark de référence pour le USCP et le WSCP de la OR-Library [5]. Les paramètres utilisés sont 333 générations de 150.000 étapes pour chacun des 2 individus pour l'algorithme mémétique et $333 \times 150.000 \times 2 \simeq$ 100.000.000 étapes pour RWLS. Le tableau 1 contient un extrait des résultats sur les instances `CYC`, les deux premières colonnes contiennent le nom de l'instance ainsi que la meilleure valeur connue à ce jour (Best Know Solution, BKS), puis pour chaque algorithme, la meilleure valeur trouvée, le nombre de fois que cette valeur a été atteinte et en combien d'étapes de RWLS en moyenne elle est atteinte. Une valeur en gras indique que l'algorithme a atteint la BKS et une valeur précédée d'une étoile indique que l'algorithme a amélioré la BKS.

| Inst. | BKS | RWLS | | | Memetic | | |
|---|---|---|---|---|---|---|---|
| | | Best | #Best | Total steps | Best | #Best | Total steps |
| CYC8 | 342 | **342** | 100/100 | 311 506 | **342** | 100/100 | 157 522 |
| CYC9 | 772 | **772** | 29/100 | 30 286 518 | **772** | 24/100 | 11 821 254 |
| CYC10 | 1798 | ***1794** | 2/100 | 27 241 299 | ***1792** | 8/100 | 11 963 321 |
| CYC11 | 3968 | **3968** | 3/100 | 18 871 638 | **3968** | 2/100 | 15 287 639 |

TAB. 1 – Extrait des résultats obtenus

L'algorithme proposé permet d'atteindre les meilleures solutions connues à ce jour obtenues par RWLS en moins d'étapes et même d'améliorer BKS pour l'instance `CYC10`.

# Références


[1] B. Yelbay, Ş. İ. Birbil et K. Bülbül, "The set covering problem revisited : An empirical study of the value of dual information", *Journal of Industrial & Management Optimization*, t. 11, p. 575, 2015, issn : 1547-5816. doi : 10.3934/jimo.2015.11.575.

[2] L. Moalic et A. Gondran, "Variations on memetic algorithms for graph coloring problems", *Journal of Heuristics*, t. 24, nº 1, p. 1-24, 2018-02, issn : 1572-9397. doi : 10.1007/s10732-017-9354-9.

[3] C. Gao, X. Yao, T. Weise et J. Li, "An efficient local search heuristic with row weighting for the unicost set covering problem", *European Journal of Operational Research*, t. 246, nº 3, p. 750-761, 2015, issn : 0377-2217. doi : 10.1016/j.ejor.2015.05.038.

[4] J. Kritter, M. Brévilliers, J. Lepagnot et L. Idoumghar, "On the optimal placement of cameras for surveillance and the underlying set cover problem", *Applied Soft Computing*, t. 74, p. 133-153, 2019. doi : 10.1016/j.asoc.2018.10.025.

[5] J. Beasley. (1990-2019). OR-Library. A collection of test data sets for a variety of Operations Research (OR) problems., adresse : http://people.brunel.ac.uk/~mastjjb/jeb/info.html.