Élicitation Incrémentale combinée à la Recherche Locale et Recherche Gloutonne pour l'Optimisation de Matroïdes Pondérés

Nawal Benabbou, Cassandre Leroy, Thibaut Lust, Patrice Perny

Élicitation Incrémentale combinée à la Recherche Locale et Recherche Gloutonne pour l'Optimisation de Matroïdes Pondérés

Nawal Benabbou, Cassandre Leroy, Thibaut Lust, Patrice Perny
Sorbonne Université, CNRS, LIP6, F-75005 Paris, France, email : prénom.nom@lip6.fr

Mots-clés : *Elicitation incrémentale, matroïdes, recherche locale, algorithme glouton.*

1 Introduction

Dans ce travail, nous proposons deux méthodes d'élicitation incrémentale interactive pour l'optimisation de matroïdes pondérés. Les matroïdes apparaissent dans divers problèmes d'optimisation combinatoire tels que la sélection de sous-ensembles, l'ordonnancement des tâches, les arbres couvrants minimaux [3] avec diverses applications possibles en intelligence artificielle, par exemple dans le contexte du choix social et de la division équitable [2].

Un matroïde se définit de la manière suivante :

Définition 1 (Matroïde) *Soient $\mathcal{S}$ un ensemble fini non vide et $\mathcal{I}$ une collection non vide de sous-ensembles de $\mathcal{S}$. Le couple $\mathcal{M} = (\mathcal{S}, \mathcal{I})$ est un matroïde s'il vérifie les deux axiomes suivants :*
— (A_1) *si $Y \in \mathcal{I}$ et $X \subseteq Y$ alors $X \in \mathcal{I}$,*
— (A_2) *si $X \in \mathcal{I}$, $Y \in \mathcal{I}$ et $|Y| > |X|$ alors il existe $e \in Y \setminus X$ tel que $X \cup \{e\} \in \mathcal{I}$.*

L'axiome A_2 implique que tous les ensembles indépendants maximaux (par rapport à l'inclusion des ensembles) ont la même cardinalité. Un ensemble indépendant maximal est appelé base du matroïde.

Nous considérons en particulier l'optimisation de matroïdes pondérés. Plus précisément, étant donné un matroïde $\mathcal{M} = (\mathcal{S}, \mathcal{I})$, nous voulons calculer $\max_{X \in \mathcal{I}} w(X)$ où w est une fonction d'ensemble définie sur $2^{\mathcal{S}}$ mesurant le poids (ou l'utilité) de tout sous-ensemble de $\mathcal{S}$. Nous considérons ici le cas d'une fonction d'ensemble additive, caractérisée par le fait que $w(X) = \sum_{e \in X} w(e)$ pour tous les $X \subseteq \mathcal{S}$. La fonction w est donc entièrement caractérisée par les poids $w(e)$, $e \in \mathcal{S}$. Ces poids représentent les utilités des éléments de S et sont considérés comme strictement positifs. Sous cette hypothèse, w est monotone par rapport à l'inclusion d'ensembles, et donc tout sous-ensemble optimal indépendant est nécessairement une base du matroïde.

Trouver la base optimale dans un matroïde pondéré est un problème très général qui apparaît dans divers contextes pratiques. Les deux problèmes suivants donnent des exemples d'optimisation de matroïdes dans le contexte de l'optimisation combinatoire et de la théorie de la décision algorithmique :

Matroïde Uniforme. Ce matroïde est défini par $\mathcal{I} = \{X \subseteq \mathcal{S} : |X| \leq k\}$ pour un entier positif donné $k \leq n$. Tout ensemble de cardinalité k est une base. Cette structure apparaît en choix social lorsque le problème est de déterminer le meilleur comité de taille k.

Matroïde Graphique. Étant donné un graphe $G = (V, E)$ où V est l'ensemble des nœuds et E est l'ensemble des arêtes, les ensembles indépendants sont les sous-ensembles des arêtes qui ne contiennent aucun cycle (c'est-à-dire les forêts). Si le graphe G est connexe, toute base correspond à un arbre couvrant T du graphe.

2 Elicitation incrémentale

L'originalité de notre travail réside dans le fait que nous considérons que la fonction w n'est pas connue mais apprise en cours de résolution par des méthodes d'élicitation des préférences. L'élicitation des préférences est une problématique de recherche visant à concevoir des méthodes permettant d'aider le décideur dans sa prise de décision en collectant des informations sur ses préférences. L'approche classique considère en entrée une base de données contenant des informations sur les préférences du décideur, et consiste à déterminer les paramètres d'une fonction de scalarisation qui représente le mieux cette base de données. L'efficacité de cette approche dépend de la qualité de la base de données, et notamment de sa taille. Il faudrait une très grande base de données pour parvenir à déterminer précisément les paramètres représentant au mieux les préférences du décideur. Cependant, il est souvent difficile d'obtenir une base de données conséquente en pratique, parce que le nombre de questions que l'on peut poser au décideur est limité.

Afin de réduire le nombre d'interactions avec le décideur, des méthodes d'élicitation partielles ont été proposées, permettant de déterminer la meilleure solution pour le décideur sans chercher à apprendre précisément les paramètres du modèle. L'idée principale de cette approche, dite "incrémentale", est de poser des questions au décideur de manière itérative, de sorte à réduire efficacement l'espace des paramètres admissibles, jusqu'à être en mesure d'identifier la solution optimale. Plus récemment, il a été proposé d'utiliser le critère de décision minimax regret pour choisir les questions à poser au décideur afin de formuler rapidement une recommandation [1].

3 Recherche Gloutonne et Recherche Locale

Nous présentons deux algorithmes à garanties de performance permettant de résoudre tout problème d'optimisation combinatoire modélisé sous la forme d'un matroïde pondéré.

Nous proposons dans un premier temps une approche générale fondée sur la recherche gloutonne et l'élicitation incrémentale de préférences. En partant de l'ensemble vide, l'idée est de sélectionner itérativement un élément de poids maximal parmi ceux qui pourraient être insérés dans le sous-ensemble indépendant actuel sans perdre la propriété d'indépendance. Lorsque la fonction w n'est pas connue, nous pouvons collecter des préférences après du décideur (sous forme de question entre deux choix) afin de déterminer, son ordre de préférence sur les éléments de l'ensemble de base. Dans un souci d'efficacité, l'élicitation de préférences peut être effectuée pendant la recherche afin de focaliser l'élicitation sur des informations de préférence qui sont directement nécessaires pour mettre en œuvre l'algorithme. Pour cela, nous proposons une version interactive de l'algorithme glouton pour l'optimisation de matroïdes.

Une adaptation de la recherche locale est également proposée.

Nous avons testé nos algorithmes sur les matroïdes suivants : uniforme, ordonnancement et graphique. Nos algorithmes sont évalués selon leur temps d'exécution, le nombre de questions posées et l'écart à l'optimum. Nous avons pu mettre en évidence que nos algorithmes permettaient d'obtenir des solutions très proches de l'optimum (moins d'un pourcent) en posant un nombre relativement faible de questions (moins de 15 questions).

Références

[1] C. Boutilier, R. Patrascu, P. Poupart, and D. Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence*, 170(8-9) :686–713, 2006.

[2] L. Gourvès, J. Monnot, and L. Tlilane. Near fairness in matroids. In *proceedings of ECAI'14*, volume 263, pages 393–398, 2014.

[3] J. Lee. *A first course in combinatorial optimization*, volume 36. Cambridge University Press, 2004.