

PALM: Un algorithme parallèle pour extraire des clusters de liens dans les réseaux sociaux

Erick Stattner*, Reynald Eugenie*, Martine Collard*

* Université des Antilles

Laboratoire LAMIA

{erick.stattner, reynald.eugenie, martine.collard}@univ-antilles.fr

Résumé. Dans cet article, nous nous intéressons à l'optimisation du processus de recherche de clusters de liens. Nous proposons en particulier l'algorithme PALM (Stattner et al., 2017), qui vise à améliorer l'efficacité du processus d'extraction par l'exploration conjointe de plusieurs zones de l'espace de recherche. Ainsi, nous commençons par démontrer que l'espace des solutions forme un treillis de concepts. Nous proposons ensuite une approche qui explore en parallèle les branches de ce treillis tout en réduisant l'espace de recherche en s'appuyant sur différentes propriétés. Les bonnes performances de notre algorithme sont démontrées en le comparant avec l'algorithme d'extraction d'origine.

1 Introduction

Dans cet article, nous nous intéressons à l'optimisation d'une approche récente de clustering de réseau appelée la recherche de *liens conceptuels*. Il s'agit d'une nouvelle approche qui effectue des clusters de liens en exploitant la structure et les attributs des noeuds pour identifier les liens fréquents entre des groupes de noeuds au sein desquels les attributs sont communs. Ce travail est motivé par le fait que l'algorithme d'extraction d'origine effectue la recherche des clusters de liens de manière séquentielle, sans prendre en compte les parallélisations possibles.

Ainsi dans cet article, nous présentons PALM (Stattner et al., 2017), un algorithme parallèle qui vise à améliorer l'efficacité du processus d'extraction des liens conceptuels en explorant simultanément plusieurs zones de l'espace de recherche. Pour cela, nous démontrons que l'espace des solutions forme un treillis de concepts. Nous proposons ensuite une approche qui explore en parallèle les branches du treillis tout en réduisant l'espace de recherche en s'appuyant sur certaines propriétés des liens conceptuels. L'efficacité de l'algorithme est démontré en l'appliquant à un réseau de télécommunications et en comparant les performances avec l'algorithme d'extraction d'origine. Les résultats obtenus montrent un gain significatif sur le temps de calcul.

2 État de l'art

La recherche de clusters dans les réseaux sociaux, aussi appelée *clustering de réseaux*, est une des approches les plus répandues de la modélisation descriptive de réseaux. L'objectif est