

Clusters dans les réseaux sociaux : intersections entre liens conceptuels fréquents et communautés

Erick Stattner, Martine Collard

Laboratoire LAMIA
Université des Antilles et de la Guyane, France
{estattne, mcollard}@univ-ag.fr

Résumé. La recherche de liens conceptuels fréquents (FCL) est une nouvelle approche de clustering de réseaux, qui exploite à la fois la structure et les attributs des noeuds. Bien que les travaux récents se soient déjà intéressés à l'optimisation des algorithmes de recherche des FCL, peu de travaux sont aujourd'hui menés sur la complémentarité qui existe entre les liens conceptuels et l'approche classique de clustering qui consiste en l'extraction de communautés. Ainsi dans ce papier, nous nous intéressons à ces deux approches. Notre objectif est d'évaluer les relations potentiellement existantes entre les communautés et les FCL pour comprendre la façon dont les motifs obtenus par chacune des méthodes peuvent correspondre ou s'intersecter ainsi que la connaissance utile résultant de la prise en compte de ces deux types de connaissance. Nous proposons pour cela un ensemble de mesures originales, basées sur la notion d'homogénéité, visant à évaluer le niveau d'intersection des FCL et des communautés lorsqu'ils sont extraits d'un même jeu de données. Notre approche est appliquée à deux réseaux et démontre l'importance de considérer simultanément plusieurs types de connaissance et leur intersection.

1 Introduction

L'identification de groupes est une des tâches les plus courantes dans le domaine de la fouille de données. En effet, dans de nombreux systèmes naturels ou sociaux, les agents ont souvent tendance à s'organiser en différents groupes. L'identification de ces groupes est devenu un domaine de recherche actif pour comprendre les potentielles relations impliquant les interactions sous-jacentes et le rôle au sein du système.

Pourtant, si de nombreux travaux ont été menés pour concevoir de nouveaux algorithmes d'identification de groupes, ou des adaptations d'algorithmes existants, nous observons que ces travaux sont souvent menés séparément, sans prendre en compte la complémentarité avec les groupes extraits par d'autres méthodes. C'est par exemple le cas des méthodes qui s'intéressent à la recherche de groupes au sein de réseaux sociaux. Alors que les méthodes traditionnelles de clustering dédiées aux réseaux exploitent uniquement la structure du réseau pour extraire des **communautés** (Fortunato, 2009; Blondel et al., 2008), les approches récentes se sont, elles, intéressées à la fois à la structure du réseau et aux attributs des noeuds pour identifier des **liens**