

ToTeM: une méthode de détection de communautés adaptées aux réseaux d'information

David Combe*, Christine Largeron*
Előd Egyed-Zsigmond**, Mathias Géry*

*Université de Lyon, F-42023, Saint-Étienne, France,
CNRS, UMR 5516, Laboratoire Hubert Curien, F-42000, Saint-Étienne, France
Université de Saint-Étienne, Jean-Monnet, F-42000, Saint-Étienne, France
{david.combe, christine.largeron, mathias.gery}@univ-st-etienne.fr

**Université de Lyon
UMR 5205 CNRS, LIRIS
7 av J. Capelle, F 69100 Villeurbanne, France
elod.egyed-zsigmond@insa-lyon.fr

Résumé. Alors que les réseaux sociaux s'attachaient à représenter des entités et les relations qui existaient entre elles, les réseaux d'information intègrent également des attributs décrivant ces entités ; ce qui conduit à revisiter les méthodes d'analyse et de fouille de ces réseaux. Dans cet article, nous proposons une méthode de classification des sommets d'un graphe qui exploite d'une part leurs relations et d'autre part les attributs les caractérisant. Cette méthode reprend le principe de la méthode de Louvain en l'étendant de façon à permettre la manipulation d'attributs continus d'une manière symétrique à ce qui existe pour les relations.

1 Introduction

L'objectif de la détection de communautés dans les graphes, ou encore dans les réseaux sociaux, est de créer une partition des sommets, en tenant compte des relations qui existent entre ces sommets dans le graphe, de telle sorte que les communautés soient composées de sommets fortement connectés (Fortunato (2010)). Ainsi, les principales méthodes de détection de communautés proposées dans la littérature se concentrent sur la structure des liens, en ignorant les propriétés des sommets. Or dans de nombreuses applications, les réseaux sociaux peuvent être représentés par des graphes dont les sommets ont des attributs qui peuvent être pris en compte pour détecter plus efficacement les communautés. Ceci a conduit à revisiter cette problématique afin d'opérer cette détection non seulement à partir des relations décrites par le graphe, mais aussi à partir d'attributs caractérisant les sommets et cela a donné lieu récemment à l'introduction de méthodes qui exploitent ces deux types de données (Moser et al. (2007); Zhou et al. (2009); Li et al. (2008); Cruz Gomez et al. (2011); Combe et al. (2012); Dang et Viennet (2012)).

Dans cet article, nous proposons ToTeM, une méthode de classification de graphes à vecteurs d'attributs qui reprend le principe de la méthode de Louvain, basée sur l'optimisation