Finding Overlapping Communities in Networks Using Propositional Satisfiability

Said Jabbour *, Nizar Mhadhbi*
Badran Raddaoui** Lakhdar Sais *

*CRIL - CNRS UMR 8188, University of Artois
F-62307 Lens Cedex, France
{jabbour, mhadhbi, sais}@cril.fr
**SAMOVAR, Télécom SudParis, CNRS, Univ. Paris-Saclay
F-91011 Evry Cedex, France
badran.raddaoui@telecom-sudparis.eu

Résumé. Community detection is a fundamental issue for understanding the structure of large and complex networks such as social, biological and information networks. In this paper, we propose a new approach to detect overlapping communities in large complex networks. We first introduce a parametrized notion of a community, called *k-linked community*, allowing us to characterize node/edge centered k-linked community with bounded diameter. Such community admits a node or an edge with a distance at most $\frac{k}{2}$ from any other node of that community. Next, we show how the problem of detecting node/edge centered k-linked overlapping communities can be expressed as a Partial Max-SAT optimization problem. Then, we propose a post-processing strategy to limit the overlaps between communities. An extensive experimental evaluation on real-world networks shows that our approach outperforms several popular algorithms in detecting relevant communities.

1 Introduction

Many complex interactions can be represented by networks, which are set of nodes connected by edges. Such connections might represent different type of relations between individuals or entities. In social networks an edge represents some kind of social interaction, while in the world of information networks, an edge represent logical connections such as hyper links and citations. Nodes in networks can be organized into *communities*, which often correspond to groups of nodes that share common properties, roles or functionnalities, such as functionally related proteins, social communities, or topically related webpages.

One of the most important task when studying networks is that of identifying communities. Communities correspond to groups of nodes in a graph that share common properties or have similar roles. Indeed, detecting and analyzing communities is of great interest in several application domains, including clustering web clients who have similar interests, identifying clusters of customers in the network of customers-products purchase relationships of online