MASTER'S THESIS

Overlapping community detection exploiting direct dependency structures in complex networks

Liang, Fengfeng

Date of Award:
2017

Abstract

Many important applications in the social, ecological, epidemiological, and biological sciences can be modeled as complex systems in which a node or variable interacts with another via the edges in the network. Community detection has been known to be important in obtaining insights into the network structure characteristics of these complex systems. The existing community detection methods often assume that the pairwise interaction data between nodes are already available, and they simply apply the detection algorithms to the network. However, the predefined network might contain inaccurate structures as a result of indirect effects that stem from the nodes' high-order interactions, which poses challenges for the algorithms upon which they are built. Meanwhile, existing methods to infer the direct interaction relationships suffer from the difficulty in identifying the cut point value that differentiates the direct interactions from the indirect interactions. In this thesis, we consider the overlapping community detection problem with determination and integration of the structural information of direct dependency interactions. We propose a new overlapping community detection model, named direct-dependency-based nonnegative matrix factorization (DNMF), that exploits the Bayesian framework for pairwise ordering to incorporate the structural information of the underlying network. To evaluate the effectiveness and efficiency of the proposed method, we compare it with state-of-the-art methods on benchmark datasets collected from different domains. Our empirical results show that after the incorporation of a direct dependency network, significant improvement is seen in the community detection performance in networks with homophilic effects.

Table of Contents