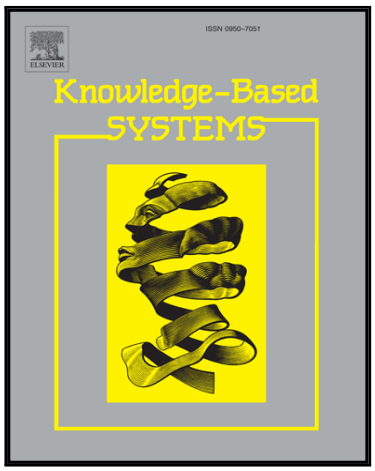
Accepted Manuscript



A Robust Two-stage Algorithm for Local Community Detection

Xiaoyu Ding, Jianpei Zhang, Jing Yang

PII: S0950-7051(18)30177-1
DOI: 10.1016/j.knosys.2018.04.018
Reference: KNOSYS 4300

To appear in: *Knowledge-Based Systems*

Received date: 25 September 2017
Revised date: 5 April 2018
Accepted date: 8 April 2018

Please cite this article as: Xiaoyu Ding, Jianpei Zhang, Jing Yang, A Robust Two-stage Algorithm for Local Community Detection, *Knowledge-Based Systems* (2018), doi: 10.1016/j.knosys.2018.04.018

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

A Robust Two-stage Algorithm for Local Community Detection

Xiaoyu Ding[1]

Nantong Street 145, Harbin

Xiaoyu Ding[a], Jianpei Zhang[a,], Jing Yang[a]*

[a]*College of Computer Science and Technology, Harbin Engineering University, Heilongjiang 150001, China*

Abstract

Local community detection addresses the efficiency problem faced by global community detection. Most existing local community detection algorithms take a seed as an initial community. They extend the community by running a greedy optimization process for a quality function. However, the quality of the detected community depends on the location of the seed. This leads to seed-dependent problem. Besides that, many local community detection algorithms cannot ensure the seed exists in the detected community. This leads to seed-invalid problem. This article proposes a robust two-stage local community detection algorithm (RTLCD) based on core detecting and community extension. To solve the seed-dependent problem, the core detecting stage replaces the seed with the core member of the target community. To solve the seed-invalid problem, the community extension stage takes the detected community core member as an initial community and extends the community based on *relation strength*. Experimental results on artificial and real-world networks show that RTLCD is more robust to the seed-dependent problem and the seed-invalid problem than earlier state-of-the-art algorithms. In addition, RTLCD has excellent performance in identifying more ground-truth community members.

Keywords: Local community detection, Seed-dependent problem, Seed-invalid problem, Core detecting, Community extension
2010 MSC: 00-01, 99-00

1. Introduction

The study of complex networks is one of the most active interdisciplinary research fields [1, 2, 3]. In the research of computer science and sociology, scientists model many complex networks as graphs composed of nodes and links [4]. The node represents the entity of interest; the link represents the association between the entities. Community structure is one of the most important properties of complex networks, and community detection is an effective way to study this property [1]. Communities are groups of nodes. The members within the same community link densely with each other, while, the links to the nodes out of the community are sparse [1, 2].

Previous studies have proposed a variety of community detection algorithms, and most of them focus on finding communities in a global sense [1, 2]. In the global sense, scientists describe communities regarding the entire network. Thus, global community detection algorithms are time-consuming. Besides that, it is difficult to get global information from large-scale networks. Sometimes people only want to know the community structure of a certain part of the network. Thus, scientists propose local community detection to detect the local community of a certain node.

Most existing local community detection algorithms take a seed as an initial community. They extend the community by running a greedy optimization process for a quality function. So far, scientists have

*Corresponding author
Email addresses: B615060001@hrbeu.edu.cn (Xiaoyu Ding), zhangjianpei@hrbeu.edu.cn (Jianpei Zhang), yangjing@hrbeu.edu.cn (Jing Yang)
[1]Since 1990.

Download English Version:

https://daneshyari.com/en/article/6861401

Download Persian Version:

https://daneshyari.com/article/6861401

Daneshyari.com