# Accepted Manuscript

Overlap community detection using spectral algorithm based on node convergence degree

Weimin Li, Shu Jiang, Qun Jin

PII: S0167-739X(17)31841-1
DOI: http://dx.doi.org/10.1016/j.future.2017.08.028
Reference: FUTURE 3623

To appear in: *Future Generation Computer Systems*

Received date : 30 September 2016
Revised date : 2 July 2017
Accepted date : 14 August 2017

Please cite this article as: W. Li, S. Jiang, Q. Jin, Overlap community detection using spectral algorithm based on node convergence degree, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.08.028

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

# *Overlap Community Detection Using Spectral Algorithm Based on Node Convergence Degree*

Weimin Li[1, 2], Shu Jiang[1], Qun Jin[3, 4 *]

[1]Computer Engineering and Technology, Shanghai University, Shanghai, China
[2]Department of Computer Science, UC Santa Barbara University, Santa Barbara, USA
[3]College of Information Engineering, China Jiliang University, Hangzhou, China
[4]Faculty of Human Sciences, Waseda University, Tokorozawa, Japan
e-mail: wmli@shu.edu.cn, jiangshu1990@shu.edu.cn, jin@waseda.jp
* Corresponding author

***Abstract*—Community structure is a typical feature of complex networks in cyberspace, and community detection is considered to be crucial to understanding the topology structure, network function and social dynamics of cyberspace. However, some particular nodes may simultaneously belong to several communities in cyberspace. Though there are many algorithms to detect the overlapping communities, most of them are based on the network structure without considering the attributes of the nodes. In this paper, we focus on the convergence characteristic of network and propose an overlap community detection algorithm based on the node convergence degree, which is defined as a combination of attribute convergence degree and structure convergence degree. It combines the network topology with the attributes of the nodes and considers both local and global information of a node. An improved PageRank algorithm is used to get the importance of each node in the global network, while the information of local network is used to measure the structure convergence degree. The overlap communities are thus identified by spectral cluster based on the node convergence degree. Finally, experiment results demonstrate the effectiveness and better performance of our proposed method.**

***Keywords—community structure; overlap; PageRank; node convergence degree.***

## 1. Introduction

With the rapid development of network technology and the Internet, many online social applications, such as Facebook and Twitter, become more and more popular. A large amount of social network data with a lot of potentially valuable information is emerged [1]. Community structure exists in all networks including both the real world and cyberspace, therefore it is significant to identify the community in the network. An online community is a community in cyberspace, such as a social network, that includes one or more groups of people as its members. We can understand the network structure and analyze the feature of the network from the view of the community. In the world-wide web, a set of web pages that have more links to other web pages in the community than outside. Identifying those communities would increase the precision of search engines and recall of the search results [2]. In a recommendation system, cold-start is the well-known issue, as the new node has little information, which leads to an inability to recommend items to this node, but we can recommend the item of similar nodes to the new node through finding the similar nodes based on community detection in social networks [3]. With the development of online financing services and activities, it is very hard to find suspicious activities among million numbers of normal and legal activities. Community detection in social networks can help detect the financial crimes [4]. Therefore, community discovery technologies have been widely used in many situations.

Another characteristic is the existence of local aggregation phenomenon for the node in the network. In the social network, the node degree distribution follows a power law which called a scale-free network. Some nodes have many links, which are called hubs while most nodes have fewer links. There are some different definitions for communities. For example, a community can be seen as a subgraph, in which the density of edges inside the subgraph is greater than the density of edges of the nodes between outside nodes. From this perspective, finding community structure in the network can be transformed into a graph partition problem.

Most community partition algorithms are based on the network topology, such as graph partitioning and hierarchical clustering [5, 6]. Graph partitioning can be used to identify the communities through adding or deleting edges of the network. It divides a graph into k subgraphs that are not connected to each other. Hierarchical clustering is a method to divide the node into the same community based on the similarity between nodes. By using these algorithms we can get community internal with high link density, but the attributes similarity of the node in the community is lower. In the cyber-physical system, the topology of the structure of social networks represents the relationships among the nodes, but the attributes of the nodes are the key factor for community detection and have an important effect on the dynamic network. The attributes of the node not only contain the information of node but have the characters of a community. Node attributes are used in some research work for community detection, such as SA-cluster [7] that defined a unified distance measurement by combining structure with node attribute, and then the cluster will be gained through the k-means clustering method. However, this method can only divide the nodes into the different community which cannot be used for finding overlap nodes.

Download English Version:

https://daneshyari.com/en/article/6873447

Download Persian Version:

https://daneshyari.com/article/6873447

Daneshyari.com