International Journal of Innovative
Computing, Information and Control
Volume **7**, Number **9**, September **2011**

ICIC International ©2011 ISSN 1349-4198
pp. 5163–5174

A SIMPLE BUT FAST AGGLOMERATIVE STRATEGY TO DETECT COMMUNITY STRUCTURE BASED ON VERTEX CLOSENESS

Zhiting Lin, Xiufang Jiang, Baohua Zhao and Yugui Qu

School of Computer Science and Technology
University of Science and Technology of China
No. 96, Jinzhai Road, Hefei 230026, P. R. China
ztlin@mail.ustc.edu.cn

Received April 2010; revised December 2010

Abstract. *The property of community structure is very important in social network analysis and has attracted particular attention of many researchers in this field. Recent years have seen the development of many community detection algorithms which can be mainly classified into two kinds: graph partition in computer science and hierarchical clustering in sociology. The vast majority of these methods, however, are computationally demanding, which largely limit their application to small, at most moderate scale networks. Here we present a very Simple Agglomerative Strategy (SAS) to detect community structure based on closeness among vertices, which not only yields excellent results but also has very low time complexity. The algorithm is tested on several real-world networks and proved to be high efficient and effective in community finding. The running time of the whole algorithm is only $O(m\overline{k})$ for sparse networks, where $\overline{k}$ is the mean vertex degree and m is the number of edges in the network.*

Keywords: Complex network, Community structure, Fast agglomerative strategy

1. **Introduction.** In the context of network theory, a complex network is a network (graph) with non-trivial topological features-features that do not occur in simple networks such as lattices or random graphs [1]. Many systems in nature and society can be described as complex networks. Examples include the World Wide Web, social and biological systems of various kinds [2-7]. Such networks exhibit community structure: the tendency of vertices to form communities (or modules) in which intra-community edges are denser than the edges between communities [8]. The automatic discovery of network communities is very useful because it can help throw light on the structures of networks which are far too large for humans to make sense of manually, even with the help of visualization techniques [9].

Many empirical algorithms have been presented to detect communities during the past few years. Generally, they can be mainly divided into two kinds of algorithms: graph partition in computer science and hierarchical clustering in sociology. One method introduced by Fiedler in [10] and the other by Kernighan and Lin in [11] are two famous algorithms in graph partition of computer science. Unfortunately, the spectral bisection method in [10] can only bisect the community each time while the efficient heuristic method in [11] will not turn out good results if the size of sub-communities is unknown. Besides, they have a common disadvantage – they need extra information on the expected number of communities because they give no indication when the bisection should terminate. As to hierarchical clustering in sociology, several typical algorithms have been proposed in [12-18], which can be classified into divisive and agglomerative methods. However, there are several disadvantages of these algorithms. The fatal flaw of divisive methods in [12-15] is that the computational demand is very high. With typical computer resources available at