Dynamics of Continuous, Discrete and Impulsive Systems
Series **B:** *Applications & Algorithms* **13 (2006)** 395-400

MINING REPRESENTATIVE NODES IN SCALE-FREE NETWORKS

Bin Liu[1], Bao-hua Cao[2], Bing Li[2] and De-yi Li[3]

[1] Computers School, Wuhan University, Wuhan, 430072, China

[2] State Key Laboratory of Software Engineer,Wuhan University, Wuhan, 430072, China

[3] Beijing Institute of Electronic System Engineering, Beijing 100039, China

Emails: hurricane_liubin@yahoo.com, ziqinLi@public2.bta.net.cn

Abstract. Many real-world networks have been shown to have a scale-free property-vertex degrees follow power law distributions, vertices tend to clusters, and average diameters are small. It is a useful to find those important nodes that can represent the network. We analyze two methods of mining those representative nodes and describe their advantage and disadvantage. By analyzing the relation of betweenness and degree, we put forward a method based on degree and approximate betweenness. We also demonstrate how our method can be used in quickly mining representative nodes in a scale-free network.

Keywords. scale-free, representative node, betweenness, UML, mining

1 Introduction

With the compute penetrating our daily lives, different kinds of computing systems have been established, including communication systems, the Internet, etc. carrying on a huge amount of information with different forms, in the form of phone-calls, or data packets [1], [2], [3]. Complex systems grow and evolve to reveal intricately networked organizations. The intensive research on complex networks, during recent years has provided deep insights into the topological properties of the underlying networks showing that these complex networks share many "scale-free" and "small-world" properties-vertex degrees follow power law distributions, vertices tend to clusters, and average lengths of paths are small. Because of the scale-free property a few nodes have high degrees, and at the same time many nodes have low degrees. The core problem of graph mining arising in many important problem domains is to find those representative nodes in a network, for example, a surge of interest within the network reduction in the properties of networks of many kinds, including the Internet, the world wide web, citation networks, transportation networks, software call graphs, email networks, food webs, social and biochemical networks, and viral marketing [1], [2], [3].

Our focus is on scale-free networks. Networks of this type are distinguished by three primary characteristics. First, they are highly clustered;