Community-based Social Networks: Generation of Power Law Degree Distribution and IP Solutions to the KPP

By

Wentao Wu

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: DECISION SCIENCES AND ENGINEERING SYSTEMS

The original of the complete thesis is on file

In the Rensselaer Polytechnic Institute Library

Examining Committee:

Wai Kin (Victor) Chan, Thesis Adviser

Cheng Hsu, Member

Charles J. Malmborg, Member

Lei Chi, Member

Rensselaer Polytechnic Institute
Troy, New York

November 2012
(For Graduation December 2012)

ABSTRACT

The objective of this thesis is two-fold: (1) to investigate the degree distribution property of community-based social networks (CSNs) and (2) to provide solutions to a pertinent problem, the Key Player Problem.

In the first part of this thesis, we consider a growing community-based network in which the ability of nodes competing for links to new coming nodes depends on their fitness parameter. In particular, we consider a network with two communities, in which the fitness does not only depend on the degree of the present nodes in the network, but also the property of the incoming new nodes. By using a mean-field approximation, we demonstrate that in a CSN the community with stronger fitness will become richer. One direct result of this finding is a power law degree distribution with exponents being a function of the percentage and fitness of each community. Moreover, we show that the gap in the average degrees between the two communities narrows when the percentage of the rich one increases.

The second part of this thesis deals with a pertinent optimization problem in social networks (SNs)—the Key Player Problem (KPP). While a variety of node centrality have been studied extensively, measures of group importance (group centrality) have not received much attention. One important type of KPPs is the Key Player Problem–Positive (KPP-POS), which is to identify a set of k nodes from a SN of size n such that the number of nodes connected to these k nodes is maximized. The KPP is, therefore, associated with group importance. The KPP has applications in social diffusion and products adoption as it helps maximize information diffusion and impact. It has also been studied in small-scale non-community-based SNs. However, previous approaches are not designed for large-scale SNs or CSNs. This thesis aims to provide an efficient algorithm to solve the KPP-POS for both traditional SNs and CSNs. We first formulate the KPP as a binary integer program and then propose two semi-definite program (SDP)-based algorithms for KPP-POS. We test our algorithms on a real small network with 472 nodes as well as three types of network structures: random graph, scale-free networks, and scale-free networks with community structure. Computational results demonstrate that the SDP-based algorithms outperform previous heuristic methods in both efficiency

and accuracy. For example, one SDP-based algorithm runs 10 times faster than previous methods for large k. It is shown that CSNs allow more nodes to be reached for a given k than the scale-free and random networks.