Knowledge-Based Systems 117 (2017) 101–110

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Social network pruning for building optimal social network: A user perspective

Sumith N.[a,*], Annappa B.[a], Swapan Bhattacharya[b]

[a] Dept. of CSE, National Institute of Technology Karnataka, Surathkal, India
[b] Dept. of CSE, Jadavpur University Kolkata, India

ARTICLE INFO

Article history:
Received 31 March 2016
Revised 27 September 2016
Accepted 22 October 2016
Available online 22 October 2016

Keywords:
Social networks
Pruning
Small-world properties
Influence maximization
Viral marketing

ABSTRACT

Social networks with millions of nodes and edges are difficult to visualize and understand. Therefore, approaches to simplify social networks are needed. This paper addresses the problem of pruning social network while not only retaining but also improving its information propagation properties. The paper presents an approach which examines the nodal attribute of a node and develops a criterion to retain a subset of nodes to form a pruned graph of the original social network. To authenticate feasibility of the proposed approach to information propagation process, it is evaluated on small world properties such as average clustering coefficient, diameter, path length, connected components and modularity. The pruned graph, when compared to original social network, shows improvement in small world properties which are essential for information propagation. Results also give a significantly more refined picture of social network, than has been previously highlighted. The efficacy of the pruned graph is demonstrated in the information diffusion process under Independent Cascade (IC) and Linear Threshold (LT) models on various seeding strategies. In all size ranges and across various seeding strategies, the proposed approach performs consistently well in IC model and outperforms other approaches in LT model. Although, the paper discusses the problem with the context of information propagation for viral marketing, the pruned graph generated from the proposed approach is also suitable for any application, where information propagation has to take place reasonably fast and effectively.

1. Introduction

Social networks are becoming a pedestal for applications such as viral marketing, poll analysis, recommendation system and so on. In this paper, the social network is discussed with the context of viral marketing application. This application depends on the information propagation process seen in social networks. However, huge size of social network is a challenge to understand the information propagation process. Social networks with millions of users and connections among them will be difficult to visualize and understand. Moreover, its disconnected structure may result in numerous outliers. As a result, an information initiated at random node[1] may not possibly reach target nodes. In such situation, a reduced social network with the best properties still retained, can be more helpful than the original large social network.

Various network simplification concepts are already available [1,2]. However, when these are used to simplify the social network, it risks in losing certain properties that may makes simplified social network unsuitable for application of interest. Moreover, impact of removing users may result in disturbing graph properties, those that are essential for information propagation. The aim of the paper is to develop an approach to reduce the size of the social network to an optimal consumer network. Such a pruned graph should have small world properties [3] such as higher average clustering coefficient, lower diameter, lower average path length, fewer number of connected components and lower modularity for effective information propagation. The proposed approach also supports the claim earlier made by Faust [4] and Butt [5], which reinforced the theory that social networks have little or no social structure. The term social structure is explained in terms of the presence of properties that are not explained by nodal tendencies. The proposed approach demonstrates that nodal attributes such as user interactions can be used to simplify social networks.

The contributions in this paper are as follows:

1. Simplification of social network on the nodal attribute by retaining the small world properties.

* Corresponding author.
E-mail addresses: s.nireshwalya@gmail.com (Sumith N.), annappa@ieee.org (Annappa B.), bswapan2000@yahoo.co.in (S. Bhattacharya).

[1] The words node and user are used interchangeably in this article.

http://dx.doi.org/10.1016/j.knosys.2016.10.020
0950-7051/