Chapter 6

Analyzing Social Networks to Mine Important Friends

Carson K.-S. Leung
University of Manitoba, Canada

Irish J. M. Medina
University of Manitoba, Canada

Syed K. Tanbeer
University of Manitoba, Canada

ABSTRACT

The emergence of Web-based communities and social networking sites has led to a vast volume of social media data, embedded in which are rich sets of meaningful knowledge about the social networks. Social media mining and social network analysis help to find a systematic method or process for examining social networks and for identifying, extracting, representing, and exploiting meaningful knowledge—such as interdependency relationships among social entities in the networks—from the social media. This chapter presents a system for analyzing the social networks to mine important groups of friends in the networks. Such a system uses a tree-based mining approach to discover important friend groups of each social entity and to discover friend groups that are important to social entities in the entire social network.

INTRODUCTION

Due to advances in technology, Web-based communities and social networking sites have emerged. They have facilitated collaboration and information sharing between users, and have led to a vast volume of social media data. Intuitively, social media can be considered as forms of electronic communication (e.g., Web sites such as Twitter for social networking and micro-blogging) through which users create online Web-based communities to share information, ideas, personal messages, and other contents such as images and videos. More formally, social media (Kaplan & Haenlein, 2010) refer to groups of electronic communications (e.g., for Web-based or mobile-based applications [Matera, 2009]) that (a) build on the ideological and technological foundations of Web 2.0 and (b) allow the creation and exchange of user-generated content. Common social media
Analyzing Social Networks to Mine Important Friends

include (a) blogs or tweets (e.g., Blogger, Twitter [Kelsey, 2010]), (b) collaborative projects (e.g., Wikipedia), (c) content communities (e.g., Flickr, SlideShare, YouTube [Lacy, 2008]), (d) social networking sites (e.g., Facebook, Google+, LinkedIn, MySpace), (e) virtual game worlds (e.g., multiplayer online role-playing games like EverQuest and its successors, as well as World of Warcraft), and (f) virtual social worlds (e.g., Second Life [Rymaszewski, et al., 2007]). Embedded in these social media data are rich sets of meaningful knowledge. This leads to social media mining, which aims to extract, represent, and exploit the rich sets of meaningful knowledge from vast volumes of social media data ranging from those in digital textual forms to those in rich multimedia formats.

Social media mining can also be considered as an interaction between data mining and social computing. Data mining (Frawley, et al., 1991) refers to non-trivial extraction of implicit, previously unknown, and potentially useful information from data (e.g., social media data); social computing intersects social behaviour and computing systems in the sense that it computationally facilitates social studies and human-social dynamics in social networks, creates social conventions through the use of computer software, and designs information and communication technologies to deal with social context. An important social media-mining task is to discover meaningful knowledge about the social networks residing in the social media data. A social network (Schwagereit & Staab, 2009) is a structure made of social entities (e.g., individuals, corporations, collective social units, or organizations) that are linked by some specific kinds of interdependency (e.g., friendship, kinship, common interest, beliefs, or financial exchange). A social entity is connected to another entity as his friend, next-of-kind, collaborator, co-author, classmate, co-worker, team member, or business partner. As such, social network analysis (Wasserman & Faust, 1994)—which often refers to a systematic method or process for identifying useful social information like interdependency among social entities (i.e., social relationships and connections) from the social networks—is in demand.

Nowadays, various social computing applications such as blogs, email, instant messaging, social bookmarking, social networking, and wikis have been widely popularized so that people could interact socially via computing space. For instance, a Facebook user can create a personal profile, add other users as friends (who can be further categorized into different customized lists such as close friends, acquaintances, or family), exchange private messages, post messages on friends’ walls, and join common-interest user groups. Similarly, a LinkedIn user can create a professional profile, establish connections to other users (who can be further annotated with tags corresponding to overlapping categories such as colleagues, classmates, business partners, and friends), exchange messages, recommend other users, and join common-interest user groups. A Google+ user can create a profile and add other users in one or more of his circles (e.g., circles for his friends, family, acquaintances, and followers). He can also share some posts, photos, or videos among users in the same circles.

Regardless which of the above social networking sites are used by users, there are some commonalities among the social experience of these users. Specifically, a user may have many friends in his social network. The number of friends may vary from one user to another. It is not uncommon for a user to have hundreds or thousands of friends. From the ego-centric prospective of a single user, some of his friends are more important than some others. It is desirable to discover his important groups of friends. To a further extent, from the socio-centric prospective of all users in the social networks, some of their groups of friends are important to some others. It is also desirable to discover groups of friends that are important in the networks. Hence, an objective of this chapter is to propose a tree-based system that (a) discovers
Related Content

**Influence of Social Networks in the Decision to Vote: An Exploratory Survey on the Ecuadorian Electorate**
[www.igi-global.com/article/influence-of-social-networks-in-the-decision-to-vote/139778?camid=4v1a](www.igi-global.com/article/influence-of-social-networks-in-the-decision-to-vote/139778?camid=4v1a)

**Community Detection Approaches in Real World Networks: A Survey and Classification**
[www.igi-global.com/article/community-detection-approaches-in-real-world-networks/116011?camid=4v1a](www.igi-global.com/article/community-detection-approaches-in-real-world-networks/116011?camid=4v1a)

**Acceptance, Use, and Influence of Political Technologies among Youth Voters in the 2008 US Presidential Election**
[www.igi-global.com/article/acceptance-use-influence-political-technologies/47197?camid=4v1a](www.igi-global.com/article/acceptance-use-influence-political-technologies/47197?camid=4v1a)

**Multimodal Mapping of a University’s Formal and Informal Online Brand: Using NodeXL to Extract Social Network Data in Tweets, Digital Contents, and Relational Ties**
[www.igi-global.com/chapter/multimodal-mapping-of-a-universitys-formal-and-informal-online-brand/203343?camid=4v1a](www.igi-global.com/chapter/multimodal-mapping-of-a-universitys-formal-and-informal-online-brand/203343?camid=4v1a)