Chapter 2
Data Mining Techniques for Social Network Analysis

Vijayaganth V.
Bannari Amman Institute of Technology, India

ABSTRACT

Social networks have increased momentously in the last decade. Individuals are depending on interpersonal organizations for data, news, and the assessment of different clients on various topics. These issues often make social network data very complex to analyze manually, resulting in the persistent use of computational means for analyzing them. Data mining gives a variety of systems for identifying helpful learning from huge datasets and a wide range of techniques for detecting useful knowledge from massive datasets like trends, patterns and rules. This chapter discusses different data mining techniques used in mining social networks.

INTRODUCTION

Interpersonal organization is a term used to portray online administrations that enable people to make an open/semi-open profile inside an area to such an extent that they can informatively associate with different clients inside the system (Chen, 2009). Interpersonal organization has enhanced the idea and innovation of Web 2.0, by empowering the arrangement and trade of User-Generated Content (Kaplan & Haenlein, 2010). Basically, informal community is a diagram comprising of hubs and connections used to speak to social relations on interpersonal organization destinations. The hubs incorporate elements and the connections between them frames the connections (Borgatti, 2009). The nodes include entities and the relationships between them forms the *links*.

Social networks are important sources of online interactions and contents sharing (Thompson, 2013; Chelmis & Prasanna, 2011), subjectivity (Asur & Huberman, 2010), assessments (Kim & Hsu, 2013), approaches (Korda & Itani, 2013), evaluation (Kaur, 2013), influences (Bakshy & Hofman, 2011), observations (Chou & Hunt, 2009), feelings (Kaplan & Haenlein, 2010), opinions and sentiments expres-

DOI: 10.4018/978-1-5225-7522-1.ch002

sions (Pang & Lee, 2008) borne out in text, reviews, blogs, discussions, news, remarks, reactions, or some other documents (Liu, 2011). The exercises on social network as of late appear to have changed the World Wide Web (www) into its proposed unique creation. Social network stages empower fast data trade between clients paying little heed to the area. Numerous associations, people and even legislature of nations now take after the exercises on social network. Data mining techniques have been found to be capable of handling the three dominant disputes with social network data namely; size, noise and dynamism. The voluminous idea of social network datasets requires robotized data preparing for dissecting it inside a sensible time. Strikingly, information mining procedures likewise require immense informational collections to mine momentous examples from information; social network locales give off an impression of being ideal destinations to mine with information mining instruments (Cortizo & Carrero, 2009). This structures an empowering factor for cutting edge indexed lists in web crawlers and furthermore helps in better comprehension of social information for inquire about and hierarchical capacities (Aggarwal, 2011).

SOCIAL NETWORK BACKGROUND

Amid the most recent decade social network have turned out to be prominent as well as reasonable and all-around acclaimed correspondence implies that has flourished in making the world a global village. Social network locales are regularly known for data spread, individual exercises posting, item audits, online pictures sharing, proficient profiling, ads and supposition/conclusion articulation. News alerts, breaking news, political debates and government policy are also posted and analyzed on social network sites (Pang & Lee, 2008). It is watched that more individuals are getting to be occupied with and depending on the social network for data continuously. Clients once in a while settle on choices in light of data posted by new people on social network expanding the level of dependence on the validity of these locales. Social network has prevailing with regards to changing the way unique substances source and recover profitable data independent of their area. Social network has likewise given clients the benefit to give feelings with next to no or no limitation.

Social Network: Power to the Users

Social sites have undoubtedly bestowed unimaginable privilege on their users to access readily available never-ending uncensored information. Twitter, for example, permits its users to post events in real time way ahead the broadcast of such events on traditional news media. Also, social network allows users to express their views, be it positive or negative (Aggarwal, 2011). Organizations are now conscious of the significance of consumers' opinions posted on social network sites to the patronage of their products or services and the overall success of their organizations. On the other hand, important personalities such as celebrities and government officials are being conscious of how they are perceived on social network. These entities follow the activities on social network to keep abreast with how their audience reacts to issues that concerns them (Castellanos & Dayal, 2011; Chen & Lee, 2011; Hoffman, 2008). Considering the enormous volume of data being generated on social network, it is imperative to find a computational means of filtering, categorizing, classifying and analyzing the social network contents.