Chapter 6
Applications of Data Mining in Dynamic Social Network Analysis

Manish Kumar
IIIT, Allahabad, INDIA

ABSTRACT
Social Networks are nodes consisting of people, groups and organizations growing dynamically. The growth is horizontal as well as vertical in terms of size and number. Social network analysis has gained success due to online social networking and sharing sites. The accessibility of online social sites such as MySpace, Facebook, Twitter, Hi5, Friendster, SkyRock and Beb offer sharing and maintaining large amount of different data. Social network analysis is focused on mining such data i.e. generating pattern of people’s interaction. The analysis involves the knowledge discovery that helps the sites as well as users in terms of usage and business goals respectively. Further it is desired that the process must be privacy preserving. This chapter describes the various mining techniques applicable on social networks data.

INTRODUCTION
Social network is a network made up of a set of nodes consisting individuals, groups, organizations and systems that share relationship such as contacts, friends, group participation and many others. Social networks are self-organizing and complex with the incremental property in terms of size such that it can generate a global pattern from the local interaction which constitutes system (Newman et al., 2006). Due to the increase in network size, patterns become more apparent. The research components in social networks involves at two levels; micro-level and macro level. At the micro level, research associated with small group of people related with particular social context whereas macro level focuses on outcomes of group interaction and transfer of resources over large groups. The social network creates platform to analyze the structure of entire complex social entities (Wasserman et al.,1994). These structures generate local and global patterns.

DOI: 10.4018/978-1-4666-4213-3.ch006
from social entities and it also examines the social network dynamics. Social networks analysis is an interdisciplinary research activity which involves social psychology, information science, biology, political science, statistics, data mining and graph theory (Wasserman et al., 1994; Scott, 2000). Social network analysis becomes now a major paradigm in social and formal sciences. However, a global network analysis is very much difficult and contains so much uninformative information. Mathematically, social network (Nam et al., 2011) is modeled as a graph $G = (V, E)$ where $V$ denotes vertices of individuals and $E$ denotes edges of interaction. Dynamic social networks deals with dynamic interactions i.e. changes over time are considered for the mining purposes. To model the dynamic social network (Jiangtao et al., 2011; Takafoli et al., 2010), let $G_d = (V_d, E_d)$ be a graph network depending on time i.e. instances are recorded at time $t$. Let $V_e$ and $E_e$ be the sets of vertices and edges to be inserted or deleted at time $t$ and hence $G_e = (V_e, E_e)$ is the modified graph after the changes at time $t$. The sequence of network changes over time: $G_0, G_1, G_2, \ldots, G_m$ creates a dynamic social network. The application of data mining in social networks is a major area of research which involves identification of different pattern of online community. Dynamic Social networks are represented as complex networks which require modeling and new techniques to evaluate the system and methods to interpret the information from the networks. Dynamic social network provides important methods for analyzing the friendship graphs (Catanese et al., 2010) i.e. relationship dynamics. Various models and statistical algorithms have been proposed for identifying the actor nodes, groups and relationships for friendship graph networks. Now, social networks are useful to predict the organizational relationships and interpersonal relationship. The Figure 1 depicts the different modes of dynamic social network.

**Clustering**

Clustering is unsupervised learning that forms clusters of the data points. Clustering may be applied on numerical as well as categorical data with transformation in suitable data types. The clustering in social network analysis is different

Figure 1. Social networks (source: Jiangtao et al., 2011)
Related Content

Automatic Syllabus Classification Using Support Vector Machines
[www.irma-international.org/chapter/automatic-syllabus-classification-using-support/21717/](www.irma-international.org/chapter/automatic-syllabus-classification-using-support/21717/)

Data Intensive Cloud Computing: Issues and Challenges
[www.irma-international.org/chapter/data-intensive-cloud-computing/150186/](www.irma-international.org/chapter/data-intensive-cloud-computing/150186/)

A Mathematical Database to Process Time Series
[www.irma-international.org/article/a-mathematical-database-to-process-time-series/208690/](www.irma-international.org/article/a-mathematical-database-to-process-time-series/208690/)

Mining Progressive Confident Rules in Sequence Databases
Wynne Hsu, Mong Li Lee and Junmei Wang (2008). *Temporal and Spatio-Temporal Data Mining* (pp. 87-111).
[www.irma-international.org/chapter/mining-progressive-confident-rules-sequence/30263/](www.irma-international.org/chapter/mining-progressive-confident-rules-sequence/30263/)

A New Approach for Supervised Dimensionality Reduction
Yinglei Song, Yongzhong Li and Junfeng Qu (2018). *International Journal of Data Warehousing and Mining* (pp. 20-37).
[www.irma-international.org/article/a-new-approach-for-supervised-dimensionality-reduction/215004/](www.irma-international.org/article/a-new-approach-for-supervised-dimensionality-reduction/215004/)