

Chapter 15

Social Network Web Mining: Web Mining Techniques for Online Social Network Analysis

Balamurugan Balusamy
VIT University, India

Vegesna Tarun Sai Varma
VIT University, India

Sohil Sri Mani Yeshwanth Grandhi
VIT University, India

ABSTRACT

Today, social networks are major part of everyone's lives. They provide means to communicate with people across the globe with ease. As of July 2016, there are over 1.71 billion monthly active Facebook users. They generate significant amount of data, which if analysed well will provide us with valuable information. This can be done by analysing the log data collected at the respective social networking service. This chapter focuses on extraction and analysis of Facebook data since it is presently the most used social network. The result of analysis can be used in building decision support systems for an organization to help with the decision making process.

INTRODUCTION

Social Networks have been a major part of everyone's lives since the evolution of the web into Web 2.0 which emphasizes on user-generated content, usability and interoperability. A social network can formally defined as a platform to build social relations among people who share similar interests, backgrounds or real life connections. According to a survey conducted by PewResearchCenter (2015), 72% of American adult internet users use Facebook, as indicated in Table 1. This accounts to about 62% of the entire American adult population.

DOI: 10.4018/978-1-5225-1877-8.ch015

Social Network Web Mining

Table 1. Percentage of Social Network Users among American adult Internet Users

Social Network	Internet Users
Facebook	72.00%
Pinterest	31.00%
Instagram	28.00%
LinkedIn	25.00%
Twitter	23.00%

Source: Pew Research Center, March 17-April 12

As there is a huge number of users for Social Networks, there is a lot of data generated. Extracting knowledge from this data can give us a lot of useful information. This is done through social web mining algorithms and techniques. Social Network Mining is a hot research topic since it combines two very interesting research topics: Web Data Mining and Social Network Analysis. Social Network Mining discusses a lot more disciplines than discussed above such as Machine Learning, Network Analysis, Sociology, Ethnography, Statistics and may more. Before we discuss any more Social Network Mining let us see a brief introduction of Web Data Mining.

BACKGROUND: WEB DATA MINING

Overview

Web Mining refers to applying data mining techniques used to extract useful knowledge from patterns on the World Wide Web documents and services. Although Web Mining has deep roots into Data Mining, it converges from Data Mining in a lot of directions. Web Mining also digs into fields like Information Retrieval, Artificial Intelligence, Psychology and Statistics.

Web Mining can be decomposed into the following sub-tasks (Wang, Y., 2000):

- **Resource Discovery:** Refers to retrieving the raw data from the particular web document.
- **Information Extraction:** The task of pre-processing specific information from retrieved Web resources.
- **Generalization:** Discovering general patterns in information retrieved from the target Web document.
- **Analysis:** Analyzing the pattern mined from the document.

Web involves three types of data: data on the Web document (Content), Web user-logs and server-side data (Usage Data) and the HTML tag data (Structure Data). Hence, Web Mining can also be categorized into web content mining, web usage mining and web structure mining respectively.

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/social-network-web-mining/173836

Related Content

Strategy, Decision Making, and Cognition: An Integrative Perspective

Radha Appan, Dorcia E. Bolton and Sreedhar Madhavaram (2017). *International Journal of Strategic Decision Sciences* (pp. 78-98).

www.irma-international.org/article/strategy-decision-making-and-cognition/189079

SIDE: A Decision Support System Using a Combination of Swarm Intelligence and Data Envelopment Analysis

Rashmi Malhotra (2014). *International Journal of Strategic Decision Sciences* (pp. 39-58).

www.irma-international.org/article/side/111159

Asset Management for Buildings within the Framework of Building Information Modeling Development

Antonio Jesús Guillén López, Adolfo Crespo Márquez, Jose A. Sanz, Khairy A. H. Kobbacy, Samir M. Shariff, Etienne Le Page and Vicente González-Prida (2017). *Decision Management: Concepts, Methodologies, Tools, and Applications* (pp. 133-150).

www.irma-international.org/chapter/asset-management-for-buildings-within-the-framework-of-building-information-modeling-development/176754

Application of Markowitz Portfolio Theory by Building Optimal Portfolio on the US Stock Market

Martin Šírek and Lukáš Ken (2017). *Tools and Techniques for Economic Decision Analysis* (pp. 24-42).

www.irma-international.org/chapter/application-of-markowitz-portfolio-theory-by-building-optimal-portfolio-on-the-us-stock-market/170894

The Tetrad Influences: A Case Study of an Adaptable Software Configuration Management Process

Usman Khan Durrani, Zijad Pita and Joan Richardson (2014). *International Journal of Strategic Decision Sciences* (pp. 30-42).

www.irma-international.org/article/the-tetrad-influences/114627