

Chapter 11

Methodologies and Techniques of Web Usage Mining

T. Venkat Narayana Rao

Sreenidhi Institute of Science and Technology, India

D. Hiranmayi

Sreenidhi Institute of Science and Technology, India

ABSTRACT

Web usage mining attempts to discover useful knowledge from the secondary data obtained from the interactions of the users with the Web. It is the type of Web mining activity that involves the automatic discovery of out what users are looking for on the Internet. In this chapter methodology of web usage mining explained in detail which are data collection, data preprocessing, knowledge discovery and pattern analysis. The different Web Usage Mining techniques are described, which are used for knowledge and pattern discovery. These are statistical analysis, sequential patterns, classification, association rule mining, clustering, dependency modeling. Pattern analysis is needed to filter out uninterested rules or patterns from the set found in the pattern discovery phase.

INTRODUCTION

Web mining is the technique of data mining which is used to discover patterns from web. It contains three types of techniques namely web usage mining, web content mining, web structure mining. Web usage mining is the process of extracting information about server logs. This technique is used to extract the information about the users' access. By using web mining we are able to find out for what users are looking about in the internet. This type of mining is used for the collection of web access information about the web pages. This usage data provides the paths leading to access web pages. Web server stores this information automatically in the log files.

Web usage mining is used for companies to produce the information about products and their future business analysis based on the present productive information. This usage data provides the companies to increase their sales. Web usage mining is also useful for e-business of the companies. The use of this type of data mining is used to find the information about customer visiting sites. This helps to know

DOI: 10.4018/978-1-5225-0613-3.ch011

about company's in-depth logging information. This web mining also enables Web based businesses to provide the best access routes to services or other advertisements. When a company advertises for services provided by other companies, the usage mining data allows for the most effective access paths to these portals. In addition, there are typically three main uses for mining in this fashion.

Web usage mining mainly having three main uses:

1. It is used to complete pattern discovery, and used for processing.
2. Content processing which means that converting of web information like text, images to useful forms.
3. Structure processing which means that analyzing the structure of each page in the web site.

Recently, millions of electronic data are included on hundreds of millions data that are previously on-line today. With this significant increase of existing data on the Internet and because of its fast and disordered growth, the World Wide Web has evolved into a network of data with no proper organizational structure. In addition, survival of plentiful data in the network and the varying and heterogeneous nature of the web, web searching has become a tricky procedure for the majority of the users. This makes the users feel confused and at times lost in overloaded data that persist to enlarge. Moreover, e-business and web marketing are quickly developing and significance of anticipate the requirement of their customers is obvious particularly. As a result, guessing the users' interests for improving the usability of web or so called personalization has turn out to be very essential. Web personalization can be depicted as some action that builds the web experience of a user personalized according to the user's interest.

The ease and speed with which business transactions can be carried out over the Web has been a key driving force in the rapid growth of electronic commerce. Specifically, ecommerce activity that involves the end user is undergoing a significant revolution. The ability to track users' browsing behavior down to individual mouse clicks has brought the vendor and end customer closer than ever before. It is now possible for a vendor to personalize his product message for individual customers at a massive scale, a phenomenon that is being referred to as mass customization. The scenario described above is one of many possible applications of Web Usage mining, which is the process of applying data mining techniques to the discovery of usage patterns from Web data, targeted towards various applications.

BACKGROUND

Day by day the data in the web increasing in rapid way. To analyze and process this data is a complex task. The number of pages available on the Web is currently around 1 billion and is increasing at the rate of approximately 1.5 million per day. The Web-based business has been a key driving force for this rapid growth of the Web. Retailers on the Web need the ability to track users' browsing behavior history, which can increase the sale and build a strong customer relationship[8]. This ability also can personalize the retailer's Web pages for different individual customers. Although Web log mining is a relatively new field, it has generated a lot of interest and research in the past ten years. As a sub research field of Web Usage Mining, Web log mining is the process of applying data mining technologies to discover usage patterns from the Web data. One important source to discover such patterns is the Web log data that contains users Web browsing history. Most of Web log data is generated automatically by Web servers.

20 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/methodologies-and-techniques-of-web-usage-mining/162899

Related Content

A Model-Based Approach for Diagnosing Fault in Web Service Processes

Yuhong Yan, Philippe Dague, Yannick Pencoléand Marie-Odile Cordier (2010). *Web Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 1970-1993).

www.irma-international.org/chapter/model-based-approach-diagnosing-fault/37725

Ranking Web Services using Web Service Popularity Score

Selwa Elfirdoussi, Zahi Jarirand Mohamed Quafafou (2014). *International Journal of Information Technology and Web Engineering* (pp. 78-89).

www.irma-international.org/article/ranking-web-services-using-web-service-popularity-score/115936

How Culture May Influence Ontology Co-Design: A Qualitative Study

Linda Anticoliand Elio Toppano (2011). *International Journal of Information Technology and Web Engineering* (pp. 1-17).

www.irma-international.org/article/culture-may-influence-ontology-design/55380

Towards Automated Bypass Testing of Web Applications

J. Miller, L. Zhang, E. Ofuonyeand M. Smith (2010). *Web Engineering Advancements and Trends: Building New Dimensions of Information Technology* (pp. 212-229).

www.irma-international.org/chapter/towards-automated-bypass-testing-web/40429

Social Network Analysis: Basic Concepts, Tools, and Applications

Soufiana Mekouar (2018). *Handbook of Research on Contemporary Perspectives on Web-Based Systems* (pp. 388-415).

www.irma-international.org/chapter/social-network-analysis/203433