Chapter 2.18
Semantic Knowledge Mining Techniques for Ubiquitous Access Media Usage Analysis

John Garofalakis
University of Patras, Greece

Theodoula Giannakoudi
University of Patras, Greece

Yannis Panagis
University of Patras, Greece

Evangelos Sakkopoulos
University of Patras, Greece

Athanasios Tsakalidis
University of Patras, Greece

ABSTRACT

In this chapter, an information acquisition system is proposed which aims to provide log analysis, dealing with ubiquitous access media, by use of semantic knowledge. The lately emerging figure of the semantic Web, the ontologies, may be used to exalt the Web trails to a semantic level so as to reveal their deeper usage information. The proposed architecture, which is presented in detail, intends to overcome mobile devices' trail duplicates problems and detect semantic operations similarity of server Web services, which are often composed to provide a function. The references that supplement the chapter provide publications that discuss mainly log file mining and analysis and semantic similarity. Useful technology-used URL resources are also provided.

INTRODUCTION

User driven access to information and services has become more complicated, and can sometimes be tedious for users with different goals, interests, levels of expertise, abilities, and preferences. The Boston Consulting Group announced that a full of 28% of online purchasing transactions failed

and 1/3 of them stopped shopping online due to usability difficulties (Boston Consulting Group, 2000). This problem is crucial in online sales systems with thousands of products of different kinds and categories. It is obvious that typical search methods are becoming less favourable as information increases, resulting in money losses.

In user-centred applications, two parameters affect usability:

- **Orientation and navigation strategy:** Users are frequently uncertain as how to reach their goals. Because users have different states of knowledge and experience, information presentation may be too redundant for some of them and too detailed for others.
- **Quality of search results:** Users cannot locate efficiently the information they need (results must be relevant and be presented quickly).

Moreover, with the unprecedented growth of the Internet usage, Web sites are being developed in an uncontrollable, ad-hoc manner, a fact frequently reflected to unpredictable visit patterns. Thus, a critical task for a Web site maintainer is to use enumerable metrics in order to identify substructures of the site that are objectively popular.

Web usage knowledge acquisition has emerged as a method to assist such a task. The fundamental basis for all knowledge manipulation operations entails processing Web server access log files. In its most simplified approach, usage trails' management entails registering absolute page visits or identifying popular paths of information inside a Web site, by the means of log file analysing software solutions such as Web trends (http://www.webtrends.com), and Analog (http://www.analog.cx). When the goal is to detect popular structural Web site elements, more elaborate techniques have been devised. Some representative work is presented hereafter.

In this work we propose knowledge acquisition techniques and annotation procedures of knowledge objects deriving from the Web usage through any kind of infrastructure, typical browsers, or new media (handheld devices and mobile phones). The key aim of our work is to enhance business intelligence by manipulation and collection of knowledge from ubiquitous access to Web data.

In this chapter we are going to analyze techniques to acquire and manage knowledge emanating from three different sources and corresponding categories:

- Web site business/academic context
- Academic ontologies forming educational Semantic Web sites
- (Semantic) Web usage objects

The proposed novel mechanisms enable bilateral knowledge acquisition from any source of the above to any resulting knowledge category.

We will present a promising prototype that validates our approach as effective. The evaluation of our study is quite encouraging, thus enabling efficient knowledge acquisition.

BACKGROUND

Significant work on converting server log files to valuable sources of access patterns has been conducted by Cooley (2000). Apart from analysing log files, it is important to use analysis as input, and determine which changes, if any, to bring to the Web site structure. Chen, Park, and Yu (1998) describe efficient algorithms to infer access patterns corresponding to frequently traversed Web site paths. Srikant and Yang (2002) infer path traversal patterns and use them to indicate structural changes that maximize (or minimize) certain site-dependent criteria. Finally, Christopoulou, Garofalakis, Makris, Panagis, Sakkopoulos, and Tsakalidis (2002, 2003), and

Related Content

Artificial Intelligence in Higher Education: A New Horizon

Fatima Al Husseiny (2023). *Handbook of Research on AI Methods and Applications in Computer Engineering (pp. 295-315).*

www.irma-international.org/chapter/artificial-intelligence-in-higher-education/318070

Exploring Cryptocurrency Sentiments With Clustering Text Mining on Social Media

Jiwen Fang, Dickson K. W. Chiuand Kevin K. W. Ho (2021). *Intelligent Analytics With Advanced Multi-Industry Applications (pp. 157-171).*

www.irma-international.org/chapter/exploring-cryptocurrency-sentiments-with-clustering-text-mining-on-social-media/272783

Ambulatory EEG Data Management System for Home Care Epileptic Patients: A Design Approach

Amol Pardhiand Suchita Varade (2022). *International Journal of Ambient Computing and Intelligence (pp. 1-15).*

www.irma-international.org/article/ambulatory-eeg-data-management-system-for-home-care-epileptic-patients/311500

A Multi Criteria Decision Making Method for Cloud Service Selection and Ranking

Rakesh Ranjan Kumarand Chiranjeev Kumar (2018). *International Journal of Ambient Computing and Intelligence (pp. 1-14).*

www.irma-international.org/article/a-multi-criteria-decision-making-method-for-cloud-service-selection-and-ranking/204345

Combining Supervised Learning Techniques to Key-Phrase Extraction for Biomedical Full-Text

Yanliang Qi, Min Song, Suk-Chung Yoonand Lori deVersterre (2011). *International Journal of Intelligent Information Technologies (pp. 33-44).*

www.irma-international.org/article/combining-supervised-learning-techniques-key/50484