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

Driving Behavior Evaluation Model Base on Big Data From Internet of Vehicles
Ruru Hao, Hangzheng Yangand Zhou Zhou (2019). *International Journal of Ambient Computing and Intelligence (pp. 78-95).*
www.irma-international.org/article/driving-behavior-evaluation-model-base-on-big-data-from-internet-of-vehicles/238055

Plant Diseases Concept in Smart Agriculture Using Deep Learning
Prachi Chauhan, Hardwari Lal Mandoria, Alok Negiand R. S. Rajput (2021). *Smart Agricultural Services Using Deep Learning, Big Data, and IoT (pp. 139-153).*
www.irma-international.org/chapter/plant-diseases-concept-in-smart-agriculture-using-deep-learning/264963

Fuzzy Joint Gaussian-Impulsive Noise Removal Using Joint Distribution Modelling in Sparse Domain
V. V. Satyanarayana Tallapragada, D. Venkat Reddyand Suresh Varma K. N. V. (2022). *International Journal of Fuzzy System Applications (pp. 1-15).*
www.irma-international.org/article/fuzzy-joint-gaussian-impulsive-noise-removal-using-joint-distribution-modelling-in-sparse-domain/312216

Inclusion of Children With Special Needs in the Educational System, Artificial Intelligence (AI)
Pradnya Mehta, Geetha R. Chillarge, Sarita D. Sapkal, Gitanjali R. Shindeand Pranali S. Kshirsagar (2023). *AI-Assisted Special Education for Students With Exceptional Needs (pp. 156-185).*
www.irma-international.org/chapter/inclusion-of-children-with-special-needs-in-the-educational-system-artificial-intelligence-ai/331738

Socio-Technical SIEM (ST-SIEM): Towards Bridging the Gap in Security Incident Response
Bilal AlSabbaghand Stewart Kowalski (2021). *Research Anthology on Artificial Intelligence Applications in Security (pp. 189-203).*
www.irma-international.org/chapter/socio-technical-siem-st-siem/270598