

Chapter 34

Web Semantics for Personalized Information Retrieval

Aarti Singh

Guru Nanak Girls College, Yamuna Nagar, India

Anu Sharma

ICAR-IASRI New Delhi, India

ABSTRACT

This chapter explores the synergy between Semantic Web (SW) technologies and Web Personalization (WP) for demonstrating an intelligent interface for Personalized Information Retrieval (PIR) on web. Benefits of adding semantics to WP through ontologies and Software Agents (SA) has already been realized. These approaches are expected to prove useful in handling the information overload problem encountered in web search. A brief introduction to PIR process is given, followed by description of SW, ontologies and SA. A comprehensive review of existing web technologies for PIR has been presented. Although, a huge contribution by various researchers has been seen and analyzed but still there exist some gap areas where the benefits of these technologies are still to be realized in future personalized web search.

INTRODUCTION

Current proliferation and innovations in web technologies has changed the face of web from information dissemination medium to knowledge provider. The web has become an important hub for providing information on almost all the aspects of human life from the most basic needs to highly specialized ones. Although, it has fascinated the masses but this abundance of information has led to many bottlenecks in accessing the right information at any point of time. This issue could be resolved by applying technologies like Data Mining (DM), Text Mining(TM) and Web Mining (WM). But, the present scenario has demanded the personalized face of web for an individual user to satisfy his information needs. This change has led to the development of a new area of research named as Web Personalization (WP).

WP can be defined as the ability of the system to provide a customized view of web by performing various actions for a single or group of users (Anand & Mobasher, 2005). It can deliver a wide variety of

DOI: 10.4018/978-1-5225-5191-1.ch034

facilities to the users in the form of greetings, bookmarking, granting personalized rights, modifying web site structure, tailored offers & services and adapted web search results. Search engines are an important mode of retrieving desired information from the web. Currently available search engines however do not consider the preferences of the user in account while displaying Web Search Results (WSR). So, it is evident that the current search engines are not able to resolve ambiguous queries and are unable to identify the user preference automatically. So, there is need to add personalized preferences and interest for web IR (Singh & Alhadidi, 2008).

Further, intelligent SW technologies namely ontology and SA (Wooldridge & Jennings, 1995) are found useful in the retrieval of useful knowledge oriented WSR. Intelligent SW technologies provides an important paradigm for use in internet applications (Ehlert, 2003). A number of agents works collaboratively to enable personalization by recognizing individual interests and then recommending the contents. Unification of SW with web IR will enhance the efficiency, scalability of PIR along with complete automation of tasks.

This chapter is organized into four sections. A brief overview of the topic is explained in this section. Section two describes the basic components of a PIR system and section three outlines the existing SW technologies. Section four provides a framework for using web semantics for PIR. Discussion and future research directions are presented in section five.

ANATOMY OF PERSONALIZED INFORMATION RETRIEVAL SYSTEM ON WEB

The main components of the system are personalized user query expansion, personalized techniques for user profiling and ranking of web search results followed by personalized filtering of search results based on long and short term interest. The anatomy of a web PIR system is shown in Figure 1.

These components are described in brief in next sub-sections.

Query Expansion Techniques

The most challenging issue to consider in web information retrieval is to resolve the ambiguity arising out of poorly defined queries. Information requirements may vary with the different search sessions. Query expansion techniques aims at reformulating the query to meet the user requirements. Mainly two types of techniques are available- global and local. Global methods utilizes the existing thesaurus /WorldNet, create thesaurus automatically or perform spell check. Local methods consist of recording the relevance, pseudo relevance and indirect relevance feedback. Relevance feedback is recorded either implicitly or explicitly. Implicit user feedback is calculated by observing their behavior on web. These methods are further divided into word co-occurrence, probabilistic methods, context and location based methods. Figure 2 represents schematically the classification of query expansion techniques.

Various methods for creating user profiles are given in next sub-section.

User Modelling

User Modelling (UM) is referred to as gathering and exploiting the information about preferences, interests and behavior of the individual user for creating user adaptive websites (Carmagnola et al., 2011). The prominent characteristics considered for user profiling are background information, interests, pref-

14 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/web-semantics-for-personalized-information-retrieval/198576

Related Content

An Improved Text Extraction Approach With Auto Encoder for Creating Your Own Audiobook

Shakkthi Rajkumar, Shruthi Muthukumar, Aparna S. S. and Angelin Gladston (2022). *International Journal of Information Retrieval Research* (pp. 1-17).

www.irma-international.org/article/an-improved-text-extraction-approach-with-auto-encoder-for-creating-your-own-audiobook/289570

An Algorithm of Pattern Match Being Fit for Mining Association Rules

Hong Shi and Ji-Fu Zhang (2004). *Intelligent Agents for Data Mining and Information Retrieval* (pp. 254-262).

www.irma-international.org/chapter/algorithm-pattern-match-being-fit/24168

Efficacious Hyperlink Based Similarity Measure Using Heterogeneous Propagation of PageRank Scores

Vasantha Thangasamy (2019). *International Journal of Information Retrieval Research* (pp. 36-49).

www.irma-international.org/article/efficacious-hyperlink-based-similarity-measure-using-heterogeneous-propagation-of-pagerank-scores/236655

An Online Measure of Discernment

Hazel C. V. Trauffer, Corné L. Bekker, Mihai C. Bocarnea and Bruce E. Winston (2013). *Online Instruments, Data Collection, and Electronic Measurements: Organizational Advancements* (pp. 254-270).

www.irma-international.org/chapter/online-measure-discernment/69745

Intelligent Information Retrieval for Reducing Missed Cancer and Improving the Healthcare System

Madhu Kumari and Prachi Ahlawat (2022). *International Journal of Information Retrieval Research* (pp. 1-25).

www.irma-international.org/article/intelligent-information-retrieval-for-reducing-missed-cancer-and-improving-the-healthcare-system/284488