Chapter 43
Semantics and Agents Oriented Web Personalization:
State of the Art

Aarti Singh
Maharishi Markandeshwar University, Mullana, India

Anu Sharma
Maharishi Markandeshwar University, Mullana, India

Nilanjan Dey
Jadavpur University, India

ABSTRACT

Advent of technologies like semantic web, multi-agent systems, web mining has changed the internet as knowledge provider. Web personalization offers a solution to the information overload problem in current web by providing users a personalized experience, considering their interest, behavior, context and emotions. Semantic web technology is based on use of software agents, ontologies and reasoning to add meaning to web information. An important technology for achieving personalization is the use of independent intelligent software agents. This work reviews, web personalization in the light of semantic web and software agent technology. A comparative study of recent works in the domain of web personalization has been carried out for this purpose. This review highlights ample scope for application of intelligent agents in the web personalization domain for solving many existing issues like personalized content management, user profile learning, modeling and adaptive interactions with users.

1. INTRODUCTION

Recent years have seen an exponential increase in size of World Wide Web (WWW) and thereby lead to many bottlenecks in accessing the required and relevant material from the pool of available information. With the advent of latest technologies like Semantic Web (SW), multi-agent systems, data mining, the nature of the Web has been changed from information provider to a knowledge resource. This change has led to the emergence of Web Mining (WM) and Web Personalization (WP)

(Singh & Mishra,2013). WM may be defined as the application of data mining techniques to discover patterns from the Web. WM is divided into three sub-categories namely web content, web structure and web usage mining (Singh & Mishra,2012).

WM has been successfully applied in various areas from e-commerce, e-banking, scientific research, e-learning, social networking, Web communities, blogs and personalized pages (these are discussed in detail in section 3). There has been tremendous growth and evolution in e-business which have necessitated the need for personalized information delivery from the web based on user interests and context. WP can be defined as any set of actions that can tailor the web experience for a particular user or set of users (Anand & Mobasher, 2005). It can provide a wide range of services to the users in the form of greetings, bookmarking, personalized access rights, personalized web site layout, personalized offers & services and personalized web search results (Pierrakos, Paliouras, Papatheodorou & Spyropoulos, 2003). WP relies on the collection, analysis and modeling of various types of web data including web access log files, client side information in the form of cookies, browser cache, proxy server logs, click-stream data, site structure, domain ontology, explicit user profiles and location specific information. WP is achieved through the implementation of all the phases of a typical data mining cycle.

The four main phases (Anand & Mobasher, 2005) of a WP system are user profile extraction, pattern discovery, recommendations and evaluation. User profile extraction is concerned with collection of explicit and implicit data to learn about the type, characteristics, preferences, context and cognitive aspects of the system users. User profiles are generally represented as sets of weighted keywords, semantic networks, or weighted concepts, or association rules. Some of the important issues related to web profiling are accuracy, inferencing/reasoning of user interest, scalability, privacy and user model interoperability.

During pattern discovery phase various statistical and data mining techniques are applied on collected and available data to mine useful and novel patterns. Web log usage data is used to build user model offline by applying various techniques like clustering, classification, association rule mining, sequential pattern discovery and markov models etc. This model is used in real time for comparing the current interaction of the user with the stored model and generating recommendations.

Recommendations phase consist of utilizing various approaches for generating recommendations. Pierrakos, Paliouras, Papatheodorou & Spyropoulos (2003) has classified WP into four generic approaches manual decision rule systems, content-based filtering systems, social or collaborative filtering systemsm and hybrid approaches. Manual decision rule systems personalized a web-based service via manual intervention of its designer and usually with the cooperation of the user. Typically, static user models are obtained through a user registration procedure and a number of rules are specified manually concerning the web contents that are provided to users with different models.

Content based techniques aims at studying the similarity between user profile and web contents and then applying machine learning methods to predict the similar items of interest. This approach has been found very suitable for a specific area. But its effectiveness decrease with the dynamic changes in the user's behavior. Also, this requires the application of natural language processing and text mining techniques which are difficult to apply as compared to other approaches.

Collaborative filtering approach is to study the preferences and activities of larger number of similar people and then making suggestions using collective features. This approach is very useful as it is able to make recommendation without considering the web contents. But on other hand, recommendations are difficult to make for a new item which has not been rated earlier by

Related Content

Multiple Attributes Decision Making Algorithms for Vertical Handover in Heterogeneous Wireless Networks
Enrique Stevens-Navarro, José D. Martínez-Moralesand Ulises Pineda-Rico (2012). *Wireless Multi-Access Environments and Quality of Service Provisioning: Solutions and Application (pp. 52-71).*
www.irma-international.org/chapter/multiple-attributes-decision-making-algorithms/61836

A Binary Search Algorithm to Determine the Minimum Transmission Range for Minimum Connected Dominating Set of a Threshold Size in Ad Hoc Networks
Natarajan Meghanathan (2020). *International Journal of Wireless Networks and Broadband Technologies (pp. 1-16).*
www.irma-international.org/article/a-binary-search-algorithm-to-determine-the-minimum-transmission-range-for-minimum-connected-dominating-set-of-a-threshold-size-in-ad-hoc-networks/257776

A Mobility-Based Routing Protocol for CR Enabled Mobile Ad Hoc Networks
Yan Sun, Jingwen Bai, Hao Zhang, Roujia Sunand Chris Phillips (2015). *International Journal of Wireless Networks and Broadband Technologies (pp. 81-104).*
www.irma-international.org/article/a-mobility-based-routing-protocol-for-cr-enabled-mobile-ad-hoc-networks/125820

Utilizing Augmented Reality in Library Information Search
Robert Gibson (2016). *Mobile Computing and Wireless Networks: Concepts, Methodologies, Tools, and Applications (pp. 1278-1289).*
www.irma-international.org/chapter/utilizing-augmented-reality-in-library-information-search/138330

4G Access Network Architecture
Young-June Choi (2010). *Fourth-Generation Wireless Networks: Applications and Innovations (pp. 46-60).*
www.irma-international.org/chapter/access-network-architecture/40696