Chapter 12
Multi-Agent-Based Information Retrieval System Using Information Scent in Query Log Mining for Effective Web Search

Suruchi Chawla
Shaheed Rajguru College Delhi University, India

ABSTRACT

This chapter explains the multi-agent system for effective information retrieval using information scent in query log mining. The precision of search results is low due to difficult to infer the information need of the small size search query and therefore information need of the user is not satisfied effectively. Information Scent is used for modeling the information need of user web search session and clustering is performed to identify the similar information need sessions. Hyper Link-Induced Topic Search (HITS) is executed on clusters to generate the Hubs and authorities for web page recommendations to users who search with similar intents. This multi-agent system based on clustered query sessions uses query operations like expansion and recommendation to infer the information need of user search queries and recommends Hubs and authorities for effective web search.

INTRODUCTION

Information on the web is huge and the retrieval of web documents relevant to a user information need is a big challenge. Search engines are used for web information retrieval and retrieves large collection of documents for a given query. The user query issued for web search contains few keywords therefore difficult to infer information need of the user. This is because of small size user query, the search engine retrieves large collection of documents out of which few are relevant. There is the need for personalizing the web search to the information need of the user by retrieving more and more relevant documents in search results. The data mining techniques have been applied to web usage data to get knowledge of web user's browsing pattern for effective information retrieval. The search engines store the search history

of web users in query log which contains the search query and its associated clicked URLs (Broder, 2002; Jansen & Spink, 2006). There are four broad categorization of personalization techniques based on various factors such as context, behavior, location and history. In Saravanakumar and Deepa (2011), WordNet 3.1 is used to find different synonyms of query keywords to form alternate queries. In Gulati and Sharma (2010), the context of query is applied to retrieve the document containing the synonyms of user query which is otherwise missed in traditional search system. Personalization system based on capturing the behavior of user such as short term and long term interest has been proposed in Gao, Xi, and Im (2013). History based personalization has been proposed in Rastafari and Shamsuddin (2010), which uses the factors such as user query logs history, pages visited by user, action performed on that page, time spent on that page etc. Google generates the customized search result based on user search activity linked to cookie saved in browser. In Sethi and Dixit (2015), the personalized search system is proposed based on user previous history. The user behavior on the web like surfing pattern, search queries submitted and some explicitly collected information is used to re–rank the search results. In Li, Yang and Kitsuregawa (2009) and Morris, Teevan and Bush (2008), search engine search results are re-ranked based on user preferences. In Carman et al. (2010) and Lv et al. (2006), user implicit feedback is used to fine-tune the search results. In Stermsek, Strembeck, and Neumann (2007) and Ghosh and Dekhil (2009), web search personalization is based on user profile built from different resources, but it ignores the vocabulary problem and involves users in maintaining the profile. In location based personalization, a model for web personalization is proposed based on current location of users in (Mokbel et al., 2011). In Leung, Lee and Lee (2010); Bouidghaghen, Tamine, and Boughanem (2011) and Weber and Castillo (2010), personalized search techniques are proposed based on user's geographical location. In Moawad et al. (2012), multiagent system is proposed that uses profile and WordNet ontology for web search personalization. In Lieberman (1995) and Turner et al. (2001), software agent track the user browsing and generates the user profile in order to assist the user web search. In Menczer (2003) and Wei, Moreau and Jennings (2003), auction protocol and reward techniques are used for collaboration among agents in order to answer single and multiple queries. In Blanzieri et al. (2001); Chau et al. (2003) and Yu and Singh (2002), personal agents collaborate with another in order to improve the user browsing where users were asked to specify the areas of interest or analyze the similar search results. In Birukov, Blanzieri, and Giorgini (2005), agent system is used to generate the personalized search results using collaborative approach and generates the suggestions from the members of the community in addition to search engine results. In comparative analysis of personalized search techniques such as context, behaviour, location and history oriented. In context oriented technique synonyms and polysemy of keywords are used for context disambiguation with no user involvement. In behavior oriented, the user interest area is tracked using explicit or implicit involvement of user. In location oriented, geographic location is inferred automatically by the system but sometimes generates the misleading results. In history oriented, the user's clicks, action and time spent during web surfing is tracked implicitly and provides the relevant results. Google generates the personalized search results based on context, behavior, location and history. Yahoo uses the behavior approach and Bing generates the results based on user behavior and history (Mittal & Sethi, 2015). There are issues related to personalized search techniques such as there are many techniques but all are not applicable to all users at the same time. It is realized that hybrid of personalized search techniques can generate better results than any individual techniques. Thus the system proposed in this chapter provides the hybrid solution to various issues associated with effective information retrieval in one system. It uses query expansion and related queries recommendation for context of user input query and builds the user profile based on search behavior. At the same time web

24 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/multi-agent-based-information-retrieval-system-using-information-scent-in-query-log-mining-for-effective-web-search/198554

Related Content

A Hybrid User-Centric Approach for Efficient Web Service Selection
Neerja Negiand Satish Chandra (2020). *International Journal of Information Retrieval Research (pp. 1-20).*
www.irma-international.org/article/a-hybrid-user-centric-approach-for-efficient-web-service-selection/249698

Study on Brain Tumor Classification Through MRI Images Using a Deep Convolutional Neural Network
Kirti Sharma, Ketna Khanna, Sapna Gambhirand Mohit Gambhir (2022). *International Journal of Information Retrieval Research (pp. 1-19).*
www.irma-international.org/article/study-on-brain-tumor-classification-through-mri-images-using-a-deep-convolutional-neural-network/289610

Computational Intelligence Techniques Driven Intelligent Agents for Web Data Mining and Information Retrieval
Masoud Mohammadianand Ric Jentzsch (2004). *Intelligent Agents for Data Mining and Information Retrieval (pp. 15-30).*
www.irma-international.org/chapter/computational-intelligence-techniques-driven-intelligent/24153

The Use of E-Questionnaires in Organizational Surveys
Yael Brender-Ilanand Gideon Vinitzky (2013). *Online Instruments, Data Collection, and Electronic Measurements: Organizational Advancements (pp. 1-23).*
www.irma-international.org/chapter/use-questionnaires-organizational-surveys/69731

Context Window Based Co-Occurrence Approach for Improving Feedback Based Query Expansion in Information Retrieval
Jagendra Singhand Aditi Sharan (2018). *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications (pp. 1597-1613).*
www.irma-international.org/chapter/context-window-based-co-occurrence-approach-for-improving-feedback-based-query-expansion-in-information-retrieval/198616