Chapter 12
Social Search and Personalization Through Demographic Filtering

Kamal Taha
Khalifa University, UAE

Ramez Elmasri
University of Texas at Arlington, USA

ABSTRACT

Most existing personalized search systems do not consider group profiling. Group profiling can be an efficient retrieval mechanism, where a user profile is inferred from the profile of the social groups to which the user belongs. The authors propose an XML search system called DemoFilter which employs the concept of group profiling. DemoFilter simplifies the personalization process by pre-defining various categories of social groups and then identifying their preferences. Social groups are characterized based on demographic, ethnic, cultural, religious, age, or other characteristics. DemoFilter can be used for various practical applications, such as Internet or other businesses that market preference-driven products. In the ontology, the preferences of a social group are identified from published studies about the social group. They experimentally evaluate the search effectiveness of DemoFilter and compare it to an existing search engine.

INTRODUCTION

With the growth of massive information on the Web, it has become increasingly difficult to search for useful information. As one of the most promising approaches to alleviate this overload, recommender systems have emerged in domains such as E-commerce, digital libraries, and knowledge management. In general, recommendation systems suggest items or products by analyzing what users with similar tastes have chosen in the past (Herlocker et al., 2002). There are two prevalent approaches to formulate recommendations: collaborative filtering recommendation and content-based recommendation. They depend on the type of items to be recommended and on the

way that user models (Allen, 1990; Huang et al., 2004) are constructed. In collaborative filtering approach (Dalal, 2007; Herlocker et al., 2004; Hofmann, 2003), information is filtered for a larger group of users. The term collaborative filtering was coined by Goldberg et al. (1992). Collaborative-filtering algorithms aim to identify users that have relevant interests and preferences by calculating similarities and dissimilarities between user profiles. The idea behind this method is that it may be of benefit to one's search for information to consult the behavior of other users who share the same or relevant interests. A content-based approach (Carmine, 2003; Kalles, 2003; Meng, 2002) provides recommendations by comparing an item's content with the content that the user is interested in. In this approach, a model of user ratings is first developed. Then the filtering process is envisioned by computing the expected value of a user prediction given the user's ratings on other items. Content-based algorithms are principally used when documents are to be recommended, such as web pages, publications, or news. The agent maintains information about user preferences either by initial input about user's interests during the registration process or by rating documents. Recommendations are formed by taking into account the content of documents and by filtering in the ones that better match the user's preferences and logged profile.

In this chapter, we propose an XML-based recommender system, called DemoFilter (Demographic Filtering). It is a type of collaborative information filtering system. DemoFilter uses ontology-driven social networks, where nodes represent social groups. A social group is an entity that defines a group based on demographic, ethnic, cultural, religious, age, or other characteristics. For example, people of ethnic group E_x; people who follow religion R_y; and people who live in neighborhood N_y can all be considered to form various social groups. In the DemoFilter framework, query results are filtered and ranked based on the preferences of the social groups to which the user

belongs. If the user belongs to social group G_x, results will be filtered based on the preferences of G_x and the preferences of *each* ancestor social group of G_x in the social network. In social communities, it is commonly accepted that people who are known to share a specific background are likely to have additional connected interests (Herlocker et al., 2002). DemoFilter can be used for various practical applications, such as Internet or other businesses that market preference-driven products.

An individual user may belong to more than one social group. Therefore, DemoFilter outputs ranked lists of content items, taking into account not only the initial preferences of the user, but also the preferences of the user's various social groups. Consider for example a Mexican-American user. The user belongs to social groups Mexicans and Americans: the portion of Mexicans living in the USA. The results of a query submitted by this user will be *filtered* and *ranked* based on the union of the interests of social groups "Mexicans" and "Americans". The social groups to which a user belongs usually have *class-subclass* relationships. A subclass social group has its own properties while inheriting the properties of its superclass(es). For example, consider a user who belongs to the ethnic group "Berbers", who lives in the country of "Morocco", which is part of "North Africa". We could have the following representation of the hierarchical relationships between the three social groups: North Africans → Moroccans → Berbers. The Berbers may have their own concerns and preferences, while sharing the concerns and preferences of Moroccans, and more general concerns and preferences of North Africans. Thus, the user's query will be filtered and ranked based on the preferences of "Berbers", "Moroccans", and "North Africans".

In the framework of DemoFilter, the preferences of a social group are identified from published studies about the social group (the availability of such data has had a significant boost with the emergence of the World Wide Web).

19 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/social-search-personalization-through-demographic/61519

Related Content

Mining Customer Knowledge for a Recommendation System in Convenience Stores
Shu-Hsien Liao, Chih-Hao Wen, Pei-Yuan Hsian, Chien-Wen Liand Che-Wei Hsu (2014). *International Journal of Data Warehousing and Mining (pp. 55-86).*
www.irma-international.org/article/mining-customer-knowledge-for-a-recommendation-system-in-convenience-stores/110386

Hyperlink Structure Inspired by Web Usage
Pawan Lingras (2009). *Handbook of Research on Text and Web Mining Technologies (pp. 386-400).*
www.irma-international.org/chapter/hyperlink-structure-inspired-web-usage/21737

A Proposed Solution for Identifying Online Fake Reviews in the Research Process
Victor-Alexandru Briciu, Cristian-Laureniu Romanand Arabela Briciu (2021). *New Opportunities for Sentiment Analysis and Information Processing (pp. 188-205).*
www.irma-international.org/chapter/a-proposed-solution-for-identifying-online-fake-reviews-in-the-research-process/286911

Automated Integration of Heterogeneous Data Warehouse Schemas
Marko Banek, Boris Vrdoljak, A Min Tjoaand Zoran Skocir (2008). *International Journal of Data Warehousing and Mining (pp. 1-21).*
www.irma-international.org/article/automated-integration-heterogeneous-data-warehouse/1815

A Survey of Open Source Tools for Business Intelligence
Christian Thomsenand Torben Bach Pedersen (2009). *International Journal of Data Warehousing and Mining (pp. 56-75).*
www.irma-international.org/article/survey-open-source-tools-business/3896