

Chapter IV

The Next Generation of Personalization Techniques

Gulden Uchyigit
Imperial College London, UK

ABSTRACT

Coping with today's unprecedented information overload problem necessitates the deployment of personalization services. Typical personalization approaches model user preferences and store them in user profiles, used to deliver personalized content. A traditional method for profile representation is the so called keyword-based representation, where the user interests are modelled using keywords which are selected from the contents of the items which the user has rated. Although, keyword based approaches are simple and are extensively used for profile representation they fail to represent semantic-based information, this information is lost during the pre-processing phase. Future trends in personalization systems necessitate more innovative personalization techniques that are able to capture rich semantic-based information during the representation, modelling and learning phases. In recent years ontologies (key concepts and along with their interrelationships) to express semantic-based information have been very popular in domain knowledge representation. The primary goal of this chapter is to present an overview of the state-of-the art techniques and methodologies which aim to integrate personalization technologies with semantic-based information.

INTRODUCTION

The advent of the Internet, personal computer networks and interactive television networks has lead to an explosion of information available online from thousands of new sources, a situation which is overwhelming to the end-user and

is likely to worsen in the future. Personalization technologies have emerged as specialized tools in assisting users with their information needs. Personalization can be defined as the process of enabling a system to tailor information to its user's needs and preferences.

Personalization technologies became popular in the early 90's, soon after the Web first came into existence. As the number of services and the volume of content continues to grow personalization technologies are more than ever in demand. Over the years they have been deployed in several different domains including entertainment and e-commerce domains, their applications ranging from electronic newspapers to online shops. As a general rule, personalization systems acquire *domain knowledge* along with user's *information need* before they are able to deliver any personalized information to the user.

In recent years developments into extending the Web with semantic knowledge in an attempt to gain a deeper insight into the meaning of the data being created, stored and exchanged has taken the Web to a different level. This has led to developments of semantically rich descriptions to achieve improvements in the area of personalization technologies (Pretschner and Gauch, 2004).

Traditional approaches to personalization include the *content-based* method ((Armstrong et al., 1995), (Balabanovic and Shoham, 1997), (Lieberman, 1995), (Mladenovic, 1996), (Pazzani and Billsus, 1997), (Lang, 1995)). These systems generally infer a user's profile from the contents of the items the user previously seen and rated. Incoming information is then compared with the user's profile and those items which are similar to the user's profile are assumed to be of interest to the user and are recommended. To be successful the content-based method needs to be capable of accurately predicting interest from the contents of other items. A traditional method for determining whether information matches a user's interests is through *keyword matching* (Smyth & Cotter, 2000). If a user's interests are described by certain keywords then the assumption is made that information containing those keywords should be of relevant and interest to the user. Such methods may match lots of irrelevant information as well as relevant information, mainly because any

item which matches the selected keywords will be assumed interesting regardless of its existing context. For instance, if the word *learning* exists in a paper about *student learning* (from the educational literature) then a paper on *machine learning* (from artificial intelligence literature) will also be recommended. In order to overcome such problems, it is important to model the semantic meaning of the data in the domain. In recent years ontologies have been very popular in achieving this.

Ontologies are formal explicit descriptions of concepts and their relationships within a domain. Ontology-based representations are richer, more precise and less ambiguous than ordinary keyword based or item based approaches (Middleton et al., 2002). For instance they can overcome the problem of *similar* concepts by helping the system understand the relationship between the different concepts within the domain. For example to find a job as a *doctor* an ontology may suggest relevant related terms such as *clinician* and *medicine*. Utilising such semantic information provides a more precise understanding of the application domain, and provides a better means to define the user's needs, preferences and activities with regard to the system, hence improving the personalization process.

The primary challenge of next generation of personalization systems is to effectively integrate semantic knowledge from domain ontologies into the various parts of the personalization method, including data preparation, user modelling and recommendation phases (Mobasher, 2005). This chapter will present a comprehensive overview in the following areas:

- **Data preparation:** *Ontology learning, extraction and pre-processing* - This combines research from natural language processing, statistical analysis and machine learning. Challenging aspects of this research is to automatically extract and learn domain ontologies and automatically define domain

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/next-generation-personalization-techniques/24471

Related Content

Business Graduates as End-User Developers: Understanding Information Literacy Skills Required

Sandra Barker (2008). *End-User Computing: Concepts, Methodologies, Tools, and Applications* (pp. 382-388). www.irma-international.org/chapter/business-graduates-end-user-developers/18193

The Technology Acceptance Model: A Meta-Analysis of Empirical Findings

Qingxiong Ma and Liping Liu (2008). *End-User Computing: Concepts, Methodologies, Tools, and Applications* (pp. 1088-1100). www.irma-international.org/chapter/technology-acceptance-model/18243

The Changing Demographics: The Diminishing Role of Age and Gender in Computer Usage

M. B. Knight and J. M. Pearson (2008). *End-User Computing: Concepts, Methodologies, Tools, and Applications* (pp. 2284-2298). www.irma-international.org/chapter/changing-demographics-diminishing-role-age/18296

The Effects of Environmental Factors on the Adoption and Diffusion of Telework

Cynthia Ruppel and Geoffrey S. Howard (1998). *Journal of End User Computing* (pp. 5-14). www.irma-international.org/article/effects-environmental-factors-adoption-diffusion/55757

Impact of Mobile Ad Wearout on Consumer Irritation, Perceived Intrusiveness, Engagement, and Loyalty: A PLS-SEM Analysis

Ahmad A.M. Alwreikat and Husam Rjoub (2021). *Journal of Organizational and End User Computing* (pp. 69-84). www.irma-international.org/article/impact-of-mobile-ad-wearout-on-consumer-irritation-perceived-intrusiveness-engagement-and-loyalty/276377