Chapter II

Mining for Web Personalization

Penelope Markellou
University of Patras, Greece and
Research Academic Computer Technology Institute, Greece

Maria Rigou
University of Patras, Greece and
Research Academic Computer Technology Institute, Greece

Spiros Sirmakessis
University of Patras, Greece and
Research Academic Computer Technology Institute, Greece

ABSTRACT

The Web has become a huge repository of information and keeps growing exponentially under no editorial control, while the human capability to find, read and understand content remains constant. Providing people with access to information is not the problem; the problem is that people with varying needs and preferences navigate through large Web structures, missing the goal of their inquiry. Web personalization is one of the most promising approaches for alleviating this information overload, providing tailored Web experiences. This chapter explores the different faces of personalization, traces back its roots and follows its progress. It describes the modules typically comprising a personalization process, demonstrates its close relation to Web mining, depicts the technical issues that arise, recommends solutions when possible, and discusses the effectiveness of personalization and the related concerns. Moreover, the chapter illustrates current trends in the field suggesting directions that may lead to new scientific results.
INTRODUCTION

Technological innovation has led to an explosive growth of recorded information, with the Web being a huge repository under no editorial control. More and more people everyday browse through it in an effort to satisfy their “primitive” need for information, as humans might properly be characterized as a species of Informavores who “have gained an adaptive advantage because they are hungry for further information about the world they inhabit (and about themselves)” (Dennett, 1991, p. 181). Based on the observation that humans actively seek, gather, share and consume information to a degree unapproached by other organisms, Pirolli and Card (1999, p. 3) took the informavores approach one step further and introduced the Information Foraging Theory according to which, “when feasible, natural information systems evolve towards stable states that maximize gains of valuable information per unit cost” (see also Resnikoff, 1989). Under the information foraging assumption, people need information and the Web today provides open access to a large volume. Thus providing people with access to more information is not the problem; the problem is that more and more people with varying needs and preferences navigate through large and complicated Web structures, missing -in many cases- the goal of their inquiry. The challenge today is how to concentrate human attention on information that is useful (maximizing gains of valuable information per unit cost), a point eloquently made by H.A. Simon (as quoted by Hal Varian in 1995, p. 200), “…what information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.”

Personalization can be the solution to this information overload problem, as its objective is to provide users with what they want or need, without having to ask (or search) for it explicitly (Mulvenna et al., 2000). It is a multidiscipline area deploying techniques from various scientific fields for putting together data and producing personalized output for individual users or groups of users. These fields comprise information retrieval, user modeling, artificial intelligence, databases, and more. Personalization on the Web covers a broad area, ranging from check-box customization to recommender systems and adaptive Websites. The spectrum from customizable Websites (in which users are allowed, usually manually, to configure the site in order to better suit their preferences) to adaptive ones (the site undertakes to automatically produce all adaptations according to the user profile, recorded history, etc.) is wide and personalization nowadays has moved towards the latter end. We meet cases of personalization in use in e-commerce applications (product recommendations for cross-selling and up-selling, one-to-one marketing, personalized pricing, storefront page customization, etc.), in information portals (home page customization such as my.yahoo.com, etc.), in search engines (in which returned results are filtered and/or sorted according to the profile of the specific user or group of users), and e-learning applications (topic recommendations, student/teacher/administrator views, content adaptations based on student level and skills, etc.). And while recently there has been a lot of talking about personalization, one has to wonder whether it is hype or an actual opportunity. Doug Riecken, in his editorial article in the Communications of the ACM Special Issue on Personalization (2000), claims that it should be considered as an opportunity, but it must be defined clearly and it must be designed to be useful and usable, conditions that in the traditional HCI field
Related Content

Multi-Agent Based Dynamic E-Learning Environment
www.igi-global.com/chapter/multi-agent-based-dynamic-learning/53058?camid=4v1a

Mailing Lists and Social Semantic Web
www.igi-global.com/chapter mailing-lists-social-semantic-web/37678?camid=4v1a

eDAR Algorithm for Continuous KNN Queries Based on Pine
www.igi-global.com/chapter/edar-algorithm-continuous-knn-queries/5032?camid=4v1a
SWAMI: A Multiagent, Active Representation of a User's Browsing Interests
www.igi-global.com/article/swami-multiagent-active-representation-user/37586?camid=4v1a