Chapter 5.12
Identifying Users Stereotypes for Dynamic Web Pages Customization

Sandro José Rigo
Universidade Federal do Rio Grande do Sul (UFRGS), Brazil

José Palazzo M. de Oliveira
Universidade Federal do Rio Grande do Sul (UFRGS), Brazil

Leandro Krug Wives
Universidade Federal do Rio Grande do Sul (UFRGS), Brazil

ABSTRACT
Adaptive Hypermedia is an effective approach to automatic personalization that overcomes the difficulties and deficiencies of traditional Web systems in delivering the appropriate content to users. One important issue regarding Adaptive Hypermedia systems is the construction and maintenance of the user profile. Another important concern is the use of Semantic Web resources to describe Web applications and to implement adaptation mechanisms. Web Usage Mining, in this context, allows the generation of Websites access patterns. This chapter describes the possibilities of integration of these usage patterns with semantic knowledge obtained from domain ontologies. Thus, it is possible to identify users’ stereotypes for dynamic Web pages customization. This integration of semantic knowledge can provide personalization systems with better adaptation strategies.

INTRODUCTION
With the enormous quantity of documents that are now available on the Web, accessing and collecting the desired and relevant data has become a difficult task that produces low quality results. The Websites adaptation allows the minimization of this problem as an adaptive application generates Website content or the structure in accordance with a class of users. In fact, the personalization aspects are a critical factor for the successful user experience. As a personalization example, it is common now to find several customization options in an increasing number of Websites. The reasons for this are due to the diversity of users and its experience, intents, needs, preferences and even available equipment and
software. The design of a Website with thousands of daily visitors will face hard time to fulfill these very different expectations. The personalization resources available can help users to have a more personal interaction, by observing their needs and preferences.

There are many different definitions for “Web personalization” in the literature. In a more general sense, it is considered as a set of actions that fine-tune the results of some user interaction, regarding this user or a set of similar users (Mobasher, 2005). The practical personalization depends on the context. For an e-Commerce Website it may be related to the set of products that are shown to the user each time he logs on, but for other applications it may refer to the interface organization, the navigational structure and content options.

All approaches have their specific problems and some of them are hybrid, combining their better techniques (Middleton, 2004; Kleinberg 2004). It is important to notice that a superior result for the personalization requires not only an efficient approach to the analysis of the contents or users behaviors, but it is also dependent on the Website life cycle. The personalization application should be integrated with tasks such as content management, users profile management, adaptation strategies and interface generation. These tasks are well known in Adaptive Hypermedia initiatives.

Adaptive Hypermedia (Brusilovsky, 2004, 2001; De Bra, 1999) has as its objective the establishment of better user experiences by adapting hyper-documents and hypermedia to the users’ needs, preferences and goals. Usability improvement is achieved with the construction of models that represent the users’ objectives, preferences, previous knowledge, and skills. The use of these models, together with some complementary information as context, usage records or adaptation rules, allows the identification of possible topics of interest, restrictions and personalization options. In addition, domain information is very important in this process and drives the adaptation choices. This can be with respect to different aspects of a Website, such as its content or structure. Briefly put, this adaptation is based on the relationship between information concerning the application domain and information regarding the user profile.

One important topic in Adaptive Hypermedia systems research is the generation and maintenance of the users’ profiles. Some approaches create the user profile from data obtained at the registration process, others incorporate the results of interviews and some perform automatic acquisition of information tracking the resources usage. In general, the profile based on the user identification tends to generate information valid over long periods. In some circumstances, short-term information can also be very useful and this kind of profile relies almost exclusively on the user interaction.

Web Usage Mining originates in prior Data Mining research with the purpose of automatic or semi-automatic discovery of Websites users’ access patterns to generate information to be used by recommendation systems or by personalization systems (Mobasher, 2005). Analyzing the approaches to the generation of users’ profiles by Web Usage Mining, a general pattern is identified and involves several stages (Markelou, 2005; Woon, 2005) that are briefly cited. The first is the acquisition of usage data. The second stage is dedicated to the pre-processing of data and the identification of access sessions amongst other necessary adjustments due to the Web environment (proxy servers, cookies or access errors, for example). At the end of the second stage, data is organized in appropriate formats for patterns mining where association rules and clusters can be generated or frequent pathways indicated. The third and last stage deals with the analysis and handling of these patterns in specific applications or contexts.

From the analysis of the obtained patterns, clusters identified or validated association rules, it is possible to generate complementary information to support the adaptation stage of Adaptive