

Personalization in the Information Era

José Juan Pazos-Arias

University of Vigo, Spain

Martín López-Nores

University of Vigo, Spain

INTRODUCTION

We are witnessing the development of new communication technologies (e.g., DTV networks [digital TV], 3G [third-generation] telephony, and DSL [digital subscriber line]) and a rapid growth in the amount of information available. In this scenario, users were supposed to benefit extensively from services delivering news, entertainment, education, commercial functionalities, and so forth. However, the current situation may be better referred to as *information overload*; as it frequently happens that users are faced with an overwhelming amount of information. A similar situation was noticeable in the 1990s with the exponential growth of the Internet, which made users feel disoriented among the myriad of contents available through their PCs. This gave birth to *search engines* (e.g., Google and Yahoo) that would retrieve relevant Web pages in response to user-entered queries. These tools proved effective, with millions of people using them to find pieces of information and services. However, the advent of new devices (DTV receivers, mobile phones, media players, etc.) introduces consumption and usage habits that render the search-engine paradigm insufficient. It is no longer realistic to think that users will bother to visit a site, enter queries describing what they want, and select particular contents from among those in a list. The reasons may relate to users adopting a predominantly passive role (e.g., while driving or watching TV), the absence of bidirectional communication (as in broadcasting environments), or users feeling uneasy with the interfaces provided. To tackle these issues, a large body of research is being devoted nowadays to the design and provision of personalized information services, with a new paradigm of *recommender systems* proactively selecting the contents that match the interests and needs of each individual at any time. This article describes the evolution of these services, followed by an overview of the functionalities available in diverse areas of application and a discussion of open problems.

BACKGROUND

The development of personalized information services brings together diverse areas of technology, plus a significant body

of legislation. The following subsections group these aspects into five major topics.

User Modeling

To identify the most suitable contents for a user, it is necessary to handle profiles that capture the user's preferences and needs. Such profiles can be represented in many different ways: consumption histories, ontologies, neural networks, decision trees, and so on (Conati, McCoy, & Paliouras, 2007). Besides these, there exist many *de jure* or *de facto* application-dependent standards to manage data: *learner information packaging* (LIP) to track educational activities, TV-Anytime for TV programs, *Integrating the Healthcare Enterprise* (IHE) for health records, and so forth.

The initialization of the user profiles can be done manually, with the user entering a description of his or her interests, or automatically, with the program retrieving information from his or her interactions with other systems (e.g., commonly, from Web navigation histories). Whichever the approach, a recommender system must implement mechanisms to capture new data (*relevance feedback*) and discard obsolete information (*gradual forgetting*). For relevance feedback, some systems use explicit mechanisms that require the user to rate contents as interesting or not interesting; others provide implicit mechanisms that infer information from ongoing interactions (Montaner, López, & de la Rosa, 2003). In gradual forgetting, the common approach is to measure the obsolescence of data in the profiles as a function of time.

Context Awareness

Context awareness aims at acquiring information about physical and social situations to maximize the value of the information delivered to the user. Knowledge about context is added to that in the user's profile to drive the selection of contents. Regarding format and length, it is necessary to match the time the user will have to read or watch material, the size of the screens where it will be presented, the input mechanisms available, and so forth. As for the semantics, the goal is to identify the topics the user may welcome at a given moment.

The first possibility explored was to develop location-sensitive mobile applications that would display different contents following the users' moves in indoor environments (e.g., museums) or outdoors (e.g., city tours) (Baldauf, Dustdar, & Rosenberg, 2007). Other dimensions were progressively added, like the informational context (e.g., inferred from the words on screen), infrastructure (e.g., surrounding communication resources), and physical conditions (noise, light, etc.). Recent works on affective computing (Picard & Daily, 2005) bring the user's feelings (mood, stress, etc.) into consideration, too.

Characterization of Contents and Services

To enable automatic selection of contents, it is necessary to characterize the available resources regarding format, length, and semantics. The MPEG-7 standard (Manjunath, Salembier, & Sikora, 2002) here appears as a common umbrella for previously existing purpose-specific metadata specifications, such as the Dublin core for Web pages, TV-Anytime for programs, or *shared content object repository model* (SCORM) for educational resources. MPEG-7 provides descriptors for low-level audiovisual features (like color or texture) that can be annotated automatically, and for high-level characteristics of objects, events, and concepts. Furthermore, contents can be split into segments that can be handled separately: audio clips, video sequences, 3-D objects, and many others.

In the characterization of interactive services, the reference is the architecture of Web services (Cerami, 2002), which includes solutions for automatic discovery, invocation, and composition, covering practically the whole spectrum of applications over the Internet. At the core of most proposals, the Web service description language (WSDL) allows one to describe the operations offered by a service and the formats to follow in their invocation. Other initiatives, like the ontology Web language for services (OWL-S), focus on describing what the services do with a common conceptualization to reason about the functionalities delivered to the users. These solutions are now being extended to other platforms (e.g., mobile devices or DTV set-top boxes), where the services have been typically monolithic and purpose built.

Filtering

Personalization is achieved by matching the information in the user's profile with his or her context and the contents available. The first possibility explored in this regard was content-based filtering to make recommendations by looking at contents that gained the user's interest in the past (Balabanovic & Shoham, 1997). This approach is easy to

adopt, but the recommendations tend to be repetitive for considering that a user will always appreciate the same kind of content. Alternatively, collaborative filtering (Mobasher, Jin, & Zhou, 2003) evaluates not only the profile of the target user, but also those of users with similar interests (his or her neighbors). This approach solves the lack of diversity, but faces problems like sparsity when the number of contents is high (which makes it hard to find users with similar evaluations for the same contents) or problems in the treatment of users whose preferences are dissimilar to the majority. To neutralize these shortcomings, there exist hybrid approaches (Burke, 2002), such as recommending contents similar to the ones stored in the target user's profile, but considering two items similar if the users who show interest in the one tend to be interested in the other (item neighborhoods).

Regardless of the filtering strategy, the first recommender systems relied on syntactic matching techniques, comparing textual strings. This approach missed the ability to discover semantic relationships between preferences, context, and contents. Nowadays, research is focused on applying techniques from the Semantic Web (Antoniou & van Harmelen, 2004), which enable processes that gain insight into the meaning of words and sentences. The reasoning can be done directly over textual contents, and over metadata descriptions in the case of images, audiovisual material, or interactive programs.

Legal Aspects

Personalization is typically opposed to privacy. The effectiveness of the former depends on accumulating information about a user, while the latter aims at restricting the management of personal information by commercial or administrative entities. Since privacy is a right in itself, there is a legal framework that must be taken in account when implementing personalization. Globally, this is delimited by OECD guidelines, which lie at the core of more specific laws for different countries worldwide. These guidelines pose restrictions to the management of user data in such aspects as the necessary levels of the user's consent, the exchange and aggregation of data from different sources, the limits to granting access to third parties or to trade information, the requirement to keep the user informed of what his or her data will be used for, and so forth. In-depth details about these requirements can be found in Wang, Zhaoqi, and Kobsa (2006).

PERSONALIZED INFORMATION SERVICES

Personalized information services are emerging in diverse areas, considering a range of consumer devices and com-

4 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/personalization-information-era/14026

Related Content

Multimedia Impact on Human Cognition

Hayward P. Andres (2006). *Advanced Topics in Information Resources Management, Volume 5* (pp. 1-24). www.irma-international.org/chapter/multimedia-impact-human-cognition/4640

Bringing the User into the Project Development Process

Daniel W. Surry (2009). *Handbook of Research on Technology Project Management, Planning, and Operations* (pp. 105-113). www.irma-international.org/chapter/bringing-user-into-project-development/21629

AMERIREAL Corporation: Information Technology and Organizational Performances

Mo Adam Mahmood, Gary J. Mannand Mark Dubrow (2001). *Annals of Cases on Information Technology: Applications and Management in Organizations* (pp. 21-31). www.irma-international.org/article/amerireal-corporation-information-technology-organizational/44604

The Impact of the Mode of Data Representation for the Result Quality of the Detection and Filtering of Spam

Reda Mohamed Hamou, Abdelmalek Amineand Moulay Tahar (2017). *Ontologies and Big Data Considerations for Effective Intelligence* (pp. 150-168). www.irma-international.org/chapter/the-impact-of-the-mode-of-data-representation-for-the-result-quality-of-the-detection-and-filtering-of-spam/177392

Forecasting Coke's Price by Combination Semi-Parametric Regression Model

Jiaojiao Liand Linfeng Zhao (2022). *Information Resources Management Journal* (pp. 1-14). www.irma-international.org/article/forecasting-cokes-price-by-combination-semi-parametric-regression-model/308302