Chapter 2.17
Mobile Information Filtering

Witold Abramowicz
The Pozañ University of Economics, Poland

Krzysztof Banaśkiewicz
The Pozañ University of Economics, Poland

Karol Wieloch
*The Pozañ University of Eco*nomics, Poland

Paweł Żebrowski
The Pozañ University of Economics, Poland

INTRODUCTION

Information filtering techniques have been continuously developed to meet challenges arisen from new requirements of Information Society. These techniques gain even much more on importance in the facet of greater mobility of people. One of the most dynamic and compelling areas is the environment of wireless and mobile devices. Just recently, information filtering and retrieval have begun to take into consideration circumstances in which they are being used. As information needs of mobile users are highly dynamic, this points out the necessity of considering additional set of attributes describing user situation—context. This article presents an information filtering system for mobile users (mobileIF) being developed in the Department of Management Information Systems at The Pozna University of Economics. Architecture of the mobileIF is a result of research done in the field of contexts, their taxonomies and influence on information relevance in dynamic user's environment. The paper shows our approach to contexts, discusses time perspective on filtering systems and finally, describes mobileIF architecture and basic data flow within it. At last, we present our current research in fields related to the mobileIF system.

BACKGROUND

The process of providing a user with relevant information can be viewed in two different ways. On the one hand, it can be described as the process where single query is performed on a set of documents (information retrieval—IR). On the other hand, it can be understood as applying a set of queries to a single document (information filtering—IF). Although, the aim of both methods is serving users with relevant documents, the way of processing content in information retrieval and information filtering systems significantly differs from each other (Belkin & Croft, 1992). What is more, queries performed in IR represent short-term information needs, whereas profiles, representing information needs in filtering, stand for relatively constant interests in a particular subject (Baeza-Yates & Ribeiro-Neto, 1999). There are many different applications of IR and IF in various areas, however, majority of them utilizes similar techniques such as Boolean model, vector space, and probabilistic models, as well as some brand new ones, like neural network or Bayesian models. Baeza-Yates et al. (1999) provide exhaustive comparison of those techniques.

In the central point of our interests is information filtering domain, that could be divided itself into several additional subdomains according to methods used. The most important ones are content-based (cognitive) filtering and social (collaborative) filtering. The idea that stands for the content-based filtering is to select the right information (relevant to user) by comparing representations of information being searched to representations of user profiles' contents (Oard & Marchionini, 1996). This method of IF has turned out in many systems to be very effective, especially in dealing with textual objects. The latter one overcomes some limitations of content-based filtering (such as problems with filtering multimedia objects, difficult to use for novices, etc.). The collaborative filtering improves results of IF by taking advantage of judgements of multiple users who have similar interests on the read documents (Shardanand & Maes, 1995). Basis for this technique is the assumption that users who judged the same documents in the similar way to others in the group, will most probably proceed like that in future, while judging new documents.

Both of those methods have many specific advantages as well as some drawbacks. The natural way of evolution is combining these techniques in order to achieve better results of filtering. Claypool, Gokhale, Miranda, Murnikov, Netes, and Sartin (1999) and Li and Kim (2003) proposed some hybrid methods.

There are many definitions of context provided in literature. Among them, several deserve special attention as stimulus to our further considerations. In one of the earliest definitions Schilit (1995) distinguishes the following types of context: computing context (network capacity, connectivity, communication costs, and available devices), user context (user's profile, location, people nearby, and social situation), and physical context (lightning, noise level, temperature, and traffic conditions). According to Schmidt, context is divided into two categories, namely: human factors (information of the user, social environment, and user's tasks) and physical environment (location, infrastructure, and conditions) (Schmidt, Beigl & Gellersen, 1999). Both presented definitions try to identify context by simple division of some characteristics into several groups of potentially distinct attributes. However, neither of them is suitable for inferring more aggregated and complex information. This inconvenience is reduced in the definition by Chen and Kotz (2000) who distinguish low-level and high-level context. The former group contains raw contextual information such as location, temperature etc. (mainly acquired from physical sensors), whereas the latter one is specified on the basis of supplied low-level contexts. More formal definition is provided by Dey and Abowd (1999) who argue "context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."

Systems that take into account context changes and adapt to them (to some degree) are defined as context-aware (Pascoe, 1999). Such adapta-

Related Content