E-Commerce Recommendation Systems

Konstantinos Markellos
University of Patras, Greece

Penelope Markellou
University of Patras, Greece

Aristotelis Mertis
University of Patras, Greece

Angeliki Panayiotaki
University of Patras, Greece & Hellenic Ministry of Economy & Finance, Greece

INTRODUCTION

In the last decade electronic and wireless technologies have changed the way companies do business forever. E-commerce (electronic commerce) and e-business (electronic business) feature as extremely dynamic economic sectors and at the same time, as the most appealing ways of beginning or expanding a business activity. Successful companies today recognize these technologies and the Internet as mainstream to business success. Indeed, their future will continue to be promising to companies seeking means for cost cutting, enhanced productivity, improved efficiency, and increased customers' satisfaction. On the other hand, this networked economy is notably characterized by the impersonal nature of the online environment and the extensive use of IT (information technology), as opposed to face-to-face contact for transactions.

Since Internet technologies and infrastructures to support e-commerce are now in place, attention is turning to psychological factors that affect its acceptance by online consumers and their perceptions of online transactions. One such factor is *individuality* of e-customers, seen to be key to the proliferation of e-commerce. The demand for better products and services has been a pressing need. The question is how easily Internet users become e-customers and which are the internal "mechanisms" and external factors that participate in an e-purchase. The problem arises from the fact that the shoppers have varying needs, preferences, and background. At the same time, they are confronted with too many options and information that they have to deal with, the majority of which is often irrelevant to their needs and interests. In most cases, search engines are used for filtering pages according to explicit users' queries. However, their results are often poor since the produced lists are long, unmanageable, and contain irrelevant pages (Middleton, De Roure, & Shadbolt, 2004).

Currently, successful e-commerce strategies have focused on personalization technologies and opportunities. According to Personalization Consortium (2006):

"personalization is the combined use of technology and customer information to tailor electronic commerce interactions between a business and each individual customer". In other words, it means "gathering and storing information about web site visitors and analyzing this information in order to deliver the right content in a user-preferred form and layout." (Braynov, 2003)

Personalization is expected to be one means for pushing e-commerce and e-business forward. In this direction, the recent web technological advances help online companies to acquire individual customer's information in real-time and with low cost. Based on this information, they construct detailed profiles and provide personalized e-services. Thus, e-shops have now greater potential for increasing customer satisfaction, promoting customer loyalty, establishing one-to-one relationships, and consequently for return on investment.

The most popular forms of personalization are *recommendation or recommender systems (RSs)* (Adomavicius & Tuzhilin, 2005). They have emerged in the

middle of 1990's and from novelties used by a few Websites have changed to important tools incorporated to many e-commerce applications (e.g., Amazon.com, eBay.com, CDNow.com). Specifically, these systems take advantage of users' and/or communities' opinions/ratings in order to support individuals to identify the information or products most likely to be interesting to them or relevant to their needs and preferences.

In this chapter, we investigate the way RSs support e-commerce Web sites in their attempt to convert visitors to customers. We present the background field, compare the latest RSs and describe a general process for RSs. Finally, we illustrate the future trends and challenges and discuss the open issues in the field.

BACKGROUND

The dilemma of how to choose the most suitable products and services is a challenging research problem. Various approaches have been used to produce recommendations. Initial efforts were limited to *check-box personalization*, where portals allowed the users to select the links they would like on their "personal" pages, for example MyYahoo! (Manber, Patel, & Robinson, 2000). This has proved deficient since it depends on the users knowing in advance the content of their interest. Moving towards more intelligent approaches, *collaborative filtering (CF) or social filtering* (GoldBerg, Nichols, Oki, & Terry, 1992; Jung, Na, & Lee, 2003) typically based on product ratings explicitly defined by the users. The system recommends items that have been evaluated positively by similar user(s) whose ratings are in the strongest correlation with the current user. CF tries to model the way people take recommendations from friends, which would be the ideal situation. Its advantages include the capability of application in cross-domain recommenders and the simplicity compared to other recommendations techniques, as it does not need feature extraction, item representation, and so forth, but only ratings. Its drawbacks include the so-called "new-item" problem where a new item is not recommended unless a notable number of users rate it, the more common "new-user" problem, where a new user is registered and has not rated enough items so that the recommender can make accurate recommendations and the sparcity problem, where the number of available ratings are much larger than the ratings to be predicted by the recommender. CF is also vulnerable to attacks from vicious users for example promotions of an item or nuking.

Another approach is *content-based filtering* (Sarwar, Karypis, Konstan, & Riedl, 2000; Jung et al., 2003) that uses product features and recommends products to users that have similar features with those they rated highly during the past. It has some problems same with CF and some of its own. As it is based on item-to-item filtering, sometimes the recommendation results lack diversity. The feature extraction process when it comes, for example, to multimedia content can be rather difficult and a proper representation for the objects is not easy. The "new-user" problems still exist here. However, content-based filtering can work more easily with implicit feedback methods like keeping record of the purchase or search history, compared with CF, which most of the times needs explicit ratings of items.

Besides these two techniques, *demographic filtering* employs demographic data (e.g., gender, age, education, profession, etc.) to infer recommendation rules based on stereotypes. However, there are issues here about how the user information will be obtained and about the privacy policy of the company towards the customers. Privacy issues are common in demographic filtering and CF methods. Lastly, *statistics-based methods* (e.g., best-sellers) provide no-personalization but their results can be appreciated by users while they are easy to implement. Many systems incorporate them into their recommendation strategy as a stable technique. Due to the shortcomings of each recommendation method, *hybrid models* (Burke, 2002) have been proposed in order to combine the robustness and eliminate the drawbacks of the individual techniques. In the next tables a comparison of the most recent RSs is provided.

RECOMMENDATION SYSTEMS' PROCESS

Even though RSs are complex applications that combine several models, algorithms, and heuristics, most of them perform the basic steps of the knowledge discovery in databases (KDD) process. This process includes the following steps: data selection, data pre-processing and transformation, data analysis, interpretation and evaluation, and presentation (Geyer-Schulz & Hahsler, 2002), as depicted in Figure 1.