Chapter 15
A Framework for Designing Recommender System for Consumers Using Distributed Data Clustering

Flavius L. Gorgônio
Federal University of Rio Grande do Norte, Brazil

José P. Araújo Neto
University of Brasília, Brazil

Taciano M. Silva
Federal University of Rio Grande do Norte, Brazil

ABSTRACT
In the past, consumers looked for information about quality of products and services with family members, friends, vendors, and experts. Currently, this reality is changing, and the number of consumers using Internet to find this kind of information is increasing, but not only to obtain additional information about a specific product, but to compare its features with other similar products. However, Internet provides a considerable amount of information through high volume of commercial sites, making the search for really useful information costly and difficult. Recommender systems are a Web social based process, performed by ordinary people, where users want to describe their degree of appreciation about items (products, services or people) based on their personal experience. This chapter proposes a framework for designing Web recommender systems that combine a meta-search engine and a data clustering strategy for product evaluation, enabling consumers to decide which products should be chosen.

INTRODUCTION
In their attempt to purchase a new product, consumers often seek to obtain as many information as possible about it, that is capable of resolving all doubts and make sure they are doing a good purchase. Another justification for the acquisition of this information is using them to compare a product with similar products or the same category, choosing one that offers the best cost versus benefit. For instance, when a user has no knowledge or experience about a particular product, he or
she usually seeks more information about this product with family, friends, vendors and experts (Araújo Neto, 2009).

More recently, the Internet has become the most important source for obtaining such information, usually through a search engine available on the network. Recent research shows that consumers are increasingly using the internet to search for information about products, either through technical specifications available on vendor Web sites, or expert opinion on the simple interaction among users, very common in current social networks (Gonçalves, 1998) (Langerak et al., 2004) (Hess & Madansky, 2006) (Havenstein, 2008). These features help consumers both in their purchasing decisions, and in sharing their experiences, continuing the cycle of information exchange, also known as online word-of-mouth.

The Internet, particularly via the Web, offers users quick and free access to a considerable amount of information, thus becoming an excellent alternative to searching for specialized information. However, the big challenge for consumers is finding out which Web pages provide factual information and which of them are really reliable. In the Web context, besides the diversity of available sources that provide a particular piece of information, there is an endless amount of sites that can provide different representations of information to a user’s interest. Moreover, the information needed for decision making can be, for example, distributed in several communities and forums available on the Web.

The popularization of Web 2.0 and social media, facilitating the creation of forums, blogs, micro blogging and discussion groups, contributes to several online communities be able to be developed and evolve, creating private discussion forums. However, the vast majority of these communities have been focusing on a particular product or in a specific category of products, limiting the possibility of creating discussions on products that have direct link with the main category of the community. In a sense, the existence of these communities helps the user in making decisions because of the large amount of information that such communities provide them all together. However, it is not always possible to obtain specific information about a particular product from a particular community, because information is scattered in a large number of these communities. In addition, there are serious difficulties in monitoring and seeking new information, including sharing it, since most of these communities is restricted to registered users only (Araújo Neto, 2009).

These restrictions require the users interested in such information to make registration and attend a large number of virtual communities in order to carry out a survey, ask questions or share new information. This factor hinders the discovery of information, because the restricted access to community content prevents bots (also known as web crawlers, web spiders or web robots) from search engines to extract information relevant and of interest to a user community.

There are a variety of Internet services that intent to facilitate the search for products undertaken by the user. Some examples of such services available, in Brazil, are: i) BuscaPé (http://www.buscape.com.br); ii) Bondfaro (http://www.bondfaro.com.br); iii) EuComparo (http://www.eucomparo.com.br); and iv) JáCotei (http://www.jacotei.com.br). It is noticed that these services, as well as other similar services available via the Internet, have distinct focus on their search engines. The main goal is to find a product interesting to the consumer in a set of pre-registered premises, accessible via web service, and to allow the consumer purchase decision based on price and reliability that the store has with the search service, which is measured by feedback from users.

Another type of service is Google Product Search, internationally available for Google, which provides an engine for searching for a product offering results on its values. However,