1 INTRODUCTION

The RS (Recommender systems) became to an independent emerged research area in the mid-1990s. And their foundations are based on work done in the fields of cognitive science, information retrieval, approximation theory, forecasting theories, consumer choice modelling in marketing and also have links to management science (Adomavicius, G., Tuzhilin, A., 2005).

According to (Deshpande and Karypis, 2004), the recommender systems are personalized information filtering technology used to either predict whether a particular user will like a particular item (prediction problem) or to identify a set of N items that will be of interest to a certain user (top-N recommendation problem).

The problem of prediction and recommendation is increased, because users have their own culture, expectations, commitments and beliefs, behaving differently when they act alone or in an organized way.

In order to capture user’s interests, knowledge, background, skills, goals, behaviour interaction preferences, individual characteristics and the user’s context it is necessary a representation of information about an individual user (i.e. user profile) that is essential for the (intelligent) application we are considering (Schiaffino and Amandi, 2009).

Another pressure that recommendation systems feel is related to a change in habits of users, now more than ever they use e-commerce to make their purchases, express their views (e.g. commenting, rating) on collaboration environments and maintaining links with friends and family; or simply users with the same preferences.

2 MAIN RS APPROACHES

Most of the researchers agree in the existence of two main approaches for recommendation systems – item-based analysis and user-based analysis – the first approach determines items that are related to a specific item (e.g. when a user likes a particular item, all of which are related are recommended), the second approach uses personal user information to suggest the best recommendations (e.g. based on profile information, user actions, and lists of the contacts or user's friends) (Patel and Balakrishnan, 2009).

In order to overcome some disadvantages of approaches exclusively item-based or content-based...
and pure user-based methods of recommendation, hybrid methods have been created based upon collaborative filtering and content-based, which maintains user profiles based on content analysis, and directly compares these profiles to determine similar users for collaborative recommendation (Balabanovic and Shoham, Y., 1997).

To capture the context in which recommendations are made the multidimensional approach to recommendations extended to support additional dimensions capturing the context in which recommendations are made (Adomavicius and Tuzhilin, 2001).

All of the work is not only focused on recommending items to users and users to items, but takes into consideration additional contextual information such as time, place, the company of other people, and other factors affecting recommendation experiences (Adomavicius et al., 2005).

Although these different approaches there are three main recommendation methods, as can be seen in Figure 1, for finding similar items and similar users, and can be applied to each of them, techniques based on heuristics or on models for the rating estimation (Adomavicius and Tuzhilin, 2005).

![Figure 1: Methods used in recommendation systems - Adapted from: (Adomavicius and Tuzhilin, 2005).](image)

The RS now have the capability to capture context, through several methods of recommendation and use of techniques based on models or heuristics.

To unify user-based and item-based collaborative filtering approaches (Wang et al., 2006) uses the similarity fusion. This unification allows the estimation of final rating by fusing predictions from three sources: i) predictions based on ratings of the same item by other users; ii) predictions based on different item ratings made by the same user; iii) ratings predicted based on data from other but similar users rating other but similar items.

The user-based collaborative filtering has been proven to be the most successful technology for building recommender systems so far, and is extensively used in many commercial recommender systems, although content-based collaborative recommendation system solves many of the problems (Patel and Balakrishnan, 2009).

### 3 OBJECTIVES

The main objective is to develop a website for electronic commerce oriented to the publication of: (Scientific Magazines, Books, E-books and Articles).

Although the existence of several online e-commerce systems, the purpose is to develop a flexible website, better than most of the already existing systems giving each user a different experience with the online store based on his profile.

The focus is on RS, and should be considered the several approaches previously mentioned by focusing on user-based collaborative approach by applying filtering techniques based on models and heuristics, putting an emphasis on the user profile.

The differentiating aspect of this system is that it must be sufficiently intelligent to take the initiative and proactively recommending the user, content that we believe is of interest to him.

The user profile must contain essential information about an individual user and the motivation of building user profiles is that users differ in their interest, preferences, backgrounds and goals when using software applications, discovering these differences is essential to provide customized services (Schiaffino and Amandi, 2009).

For the content of the user profile this project will not only consider the content provided by the user explicitly, but must infer unobservable information about users from observable information about them (Zukerman and Albrecht, 2001).

![Figure 2: Predicted observations strategy - Source: (Oard and Kim, 1998).](image)

It should be combined inference and prediction to assist the user finding content that is of personal
interest. This combination helps using implicit feedback, because it is based on previous observations, which are used to predict user behaviour in response to new information, and then the inference phase seeks in order to estimate the value of information based on the predicted behaviour, as Figure 2, shows (Oard and Kim, 1998).

4 DISCOVERY OF THE FUNCTIONALITIES

To achieve the desired objectives it is necessary to study the e-commerce competitors and see what kind of recommendation technology they use.

A comparative study was conducted to analyze the market and see what other e-commerce sites have to offer. This comparative study was based on the direct experience of use.

High-level functionalities have been defined in order to compare these functionalities between some of the most popular e-commerce sites. The sites selected for this study were Amazon.com, FNAC, Pixmania and Barnes & Noble. Figure 3 shows the results of the comparison of the functionalities.

As our focus is on Amazon.com as a model to follow, we show in Figure 4 some of the applications used to interface recommendation, recommendation technology and how to find recommendations.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Recommendation Interface</th>
<th>Recommendation Technology</th>
<th>Finding Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customers who Bought</td>
<td>Similar item</td>
<td>Item to Item Correlation</td>
<td>Organic Navigation</td>
</tr>
<tr>
<td>Customer Comments</td>
<td>Average Rating</td>
<td>Aggregated Rating</td>
<td>Likert</td>
</tr>
<tr>
<td>Eyes</td>
<td>Email</td>
<td>Attribute Based</td>
<td>Keywords/freeform</td>
</tr>
<tr>
<td>Amazon.com Delivers</td>
<td>Email</td>
<td>Attribute Based</td>
<td>Selection Options</td>
</tr>
<tr>
<td>Book Matcher</td>
<td>Top N List</td>
<td>People to People Correlation</td>
<td>Request List</td>
</tr>
</tbody>
</table>

Figure 4: Amazon.com Recommender System - adapted from: (Schafer et al., 1999).

This research shows some of the applications used by Amazon.com.

These applications support the recommendation of books frequently purchased by customers who purchased the selected book, recommend authors whose books are frequently purchased by customers who buy books by the author of the book selected.

Such applications also do the notification of new books added to the catalog requests and provide recommendations based on research performed and persistent data.

They also offer the possibility of registered users receive text with recommendations based on the opinion of other registered users.

Beyond the implementation of these functionalities mentioned, our website will apply the implicit feedback techniques to avoid the effort of cognitive load in assigning precise ratings to large user populations and thus contributing to avoid the dispersion of data within these populations. These techniques seek to avoid this bottleneck inferred from observations that are available to the system, something similar to the ratings assigned by a user.

According to (Oard and Kim, 1998) in addition to explicit ratings were identified three major categories of potentially useful behavior observations: examination, retention and reference.

As Figure 5 shows, the category Examination extends beyond a single interaction between the user and the system it is characterized by repetition of the previous user behavior. The category of Retention aims to group these behaviors that suggest some future intention to use an object. Finally, the
Reference category is distinguished by the opportunity to direct observation of negative evaluations.

<table>
<thead>
<tr>
<th>Category</th>
<th>Observable Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination</td>
<td>Selection, Donation, Edit, Save, Repetition, Purchase</td>
</tr>
<tr>
<td></td>
<td>(object or subscription)</td>
</tr>
<tr>
<td>Retention</td>
<td>Save a reference or save an object (with or without annotation)</td>
</tr>
<tr>
<td></td>
<td>Post, Delete</td>
</tr>
<tr>
<td>Reference</td>
<td>Object-&gt;Object (forward, reply, post follow up)</td>
</tr>
<tr>
<td></td>
<td>Portion-&gt;Portion (hyperlink, citation)</td>
</tr>
<tr>
<td></td>
<td>Object-&gt;Portion (cut &amp; paste, quotation)</td>
</tr>
</tbody>
</table>

Figure 5: Observable behaviour for implicit feedback—source: (Oard and Kim, 1998).

For the contents of the user profile our website should have all the information cited above in Section 1.

Finally our website should not just allow the user to provide the content about their profile, but also should be able to infer the information which is not observable about their profile by using techniques based on Machine Learning. The user or customer profile is used to make personalized offers and to suggest or recommend products the user is supposed to like (Schiaffino and Amandi, 2009).

5 ISSUES

Taking into account the several approaches and the main recommendation methods as well as based techniques, the objectives defined in Section 3 and the functionalities mentioned in Section 4, many issues arise.

The companies that use Recommender Systems in their e-commerce are facing some issues such as: i) the need to have large amounts of data; ii) based on the preferences and previous behavior of the user; iii) the need for a large number of variables.

The first issue is related to the amount of data and the time required to provide effective recommendations by RS. It is necessary to save the data about the items, as well as all user behavior and profile. The lack of data could become a problem.

This problem is related to the occurrence of a new user, because it has to rate a sufficient number of items before a content-based RS can really understand the user’s preferences and present the user with reliable recommendations. Therefore, a new user, having very few ratings, would not be able to get accurate recommendations (Adomavicius and Tuzhilin, 2005).

The lack of data is also related to a new item added to RS, because it relies solely on users preferences to make recommendations. Therefore, until the new item is rated by a substantial number of users, the recommender system would not be able to recommend it (Adomavicius et al, 2005).

The second issue takes into account trends and user intentions. Trends are based on past behavior, and the user can change their tendencies; about the intentions of the user, it does not always have the same intentions when browsing a site. The problem is related to changes in data and user preferences.

The last issue is related to the need of adding contextual information, and will raise the problem of complexity because there are a larger number of variables.

This complexity problem can also be related to a large retailer that might have huge amounts of data, tens of millions of customers and millions of distinct catalog items, and can also be related to old customers who may have an excess of information based on thousands of purchases and ratings (Linden et al., 2003).

6 CONCLUSIONS AND FUTURE WORK

Although some major issues that have been mentioned, obviously our web site is by its nature limited by the peculiarities of the data and the recommendation domain.

We hope to achieve with this web site an improvement over existing approaches from the use of user profiles that contain information about user’s tastes, preferences, actions and needs.

The profiling information can be elicited from users explicitly, e.g., through questionnaires, or implicitly (e.g. learned from their transactional behavior over time).

We intend to use techniques for content-based recommendation such as Bayesian classifiers, and various machine learning techniques, including clustering, decision trees, and artificial neural networks.

The methods that we intend to use for searching similar items and similar users, should apply, techniques based on heuristics or on models for the rating estimation.

Even though is no more than an idea, our future work shall be to find some way of aggregating the best of the user-based collaborative filtering and
content-based collaborative filtering methodologies into a single hybrid methodology.

If the peculiarities of the site allow, this could be considered an approach based on multidimensional data model used for data warehousing and OLAP (Online Analytical Processing) applications in databases on hierarchical aggregation capabilities, and on user, item and other profiles defined for each of these dimensions (Adomavicius et al, 2005).

ACKNOWLEDGEMENTS

We would like to thank the Polytechnic Institute of Setubal, School of Technology of Setubal, for supporting the research work reflected in this paper, presented at ICE-B 2012.

REFERENCES


