ENHANCING RECOMMENDER SYSTEMS DEVELOPMENT WITH HUMAN, SOCIAL, CULTURAL AND ORGANIZATIONAL FACTORS

Andreas S. Andreou
Department of Electrical Engineering and Information Technology, Cyprus University of Technology
P.O.Box 50329, 3603 Limassol, Cyprus

Stephanos Mavromoustakos
Department of Computer Science and Engineering, European University Cyprus, P.O.Box 22006, 1516 Nicosia, Cyprus

Keywords: Web Recommender Systems, Human, Cultural, Social and organizational factors.

Abstract: Recommender Systems (RS) aim at suggesting filtered Web information adapted to the needs or interests of users by predicting their access behavior using a certain strategy or algorithm. The creation of RS is usually approached focusing mostly on user behavior modeling, while the recommendation engine often neglects critical, non-technical aspects of software systems development. Conceiving a RS primarily as a self-contained or part of a Web-application, the present paper utilizes the SpiderWeb methodology and takes into account important requirements that result from human, cultural, social and organizational factors (HSCO) so as to drive the RS development activities.

1 INTRODUCTION

Recommender Systems (RS) are a special class of software applications usually running on the Web to support users in making decisions while interacting with large spaces of information (Chen, Shtykh and Jin, 2009). RS recommend subjects or items of interest to users based on information that may be categorized in two main classes, the first being gathered from the users explicitly through interactive collection means (e.g. dialoguing or conversational modules) and the second being implied using various modeling and prediction techniques of user behaviour (Baraglia and Silvestri, 2007).

While there is a plethora of research works dealing with RS development the majority focuses on ways to implement the recommendation engine based purely on technical development factors. Nevertheless, a RS may be greatly enhanced if certain factors falling into a non-technical sphere are taken into account prior and during its creation. These factors mostly describe human, social, cultural and organizational (HSCO) aspects which may complete the picture for understanding the behavior, as well as the expectations of users working on the Web, thus contributing to better modeling certain behavioral characteristics and ultimately leading to finer information recommendations. In this context we will view RS as a specific-purpose (or part of a) Web application and we will provide a methodology for identifying and analyzing certain HSCO factors which will be incorporated in the RS thus enhancing its efficacy.

The methodology utilizes the SpiderWeb model proposed in (Andreou, Mavromoustakos and Schizas, 2002), which records critical factors that must be incorporated as functional or non-functional features in the RS under development.

2 THE SPIDERWEB MODEL AND THE RS INFORMATION GATHERING METHODOLOGY

The SpiderWeb model aims at visualizing and classifying valuable requirement components for the

S. Andreou A. and Mavromoustakos S..
ENHANCING RECOMMENDER SYSTEMS DEVELOPMENT WITH HUMAN, SOCIAL, CULTURAL AND ORGANIZATIONAL FACTORS.
DOI: 10.5220/0003553202110214
In *Proceedings of the 13th International Conference on Enterprise Information Systems* (ICEIS-2011), pages 211-214
ISBN: 978-989-8425-56-0

better identification of critical factors that will lead to the development of successful software systems running on the Web. In our case, success is assumed when the RS is able to provide information suggestions that are very close to the user interests, characteristics and expectations at the time of accessing particular portions of Web information. The model categorizes system requirements in three main axons: The Country Characteristics, the User Requirements, and the Application Domain axon (fig. 1).

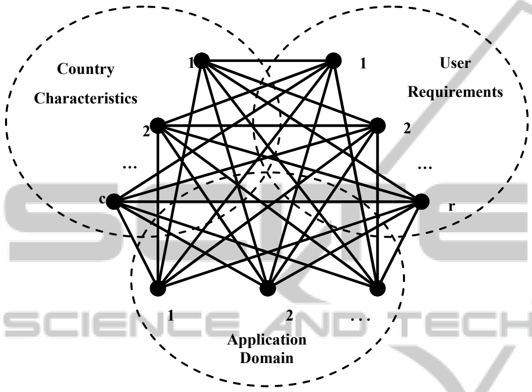

Figure 1: The SpiderWeb Model.

Each axon includes certain components, which are directly connected and interrelated. The SpiderWeb axons are also interdependent, allowing the sharing of same, similar, or different characteristics among each other.

The analysis of the axon components of the SpiderWeb model presented in the previous part aimed primarily at providing the basic key concepts for collecting proper system requirements. These concepts are to be used as guidelines for gathering critical information that may affect the functional and non-functional behavior of the RS under development. A form of small-scale ethnography analysis is conducted for collecting and analyzing information for the three axons described before.

Our method includes focus questions produced in the form of questionnaires. These questions are distributed among the targeted group or are used as part of the interviewing process, and the answers are recorded, analyzed and evaluated.

The SpiderWeb methodology is integrated with the WebE process (Pressman, 2000), the latter being used for the development of Web applications.

As previously mentioned, the SpiderWeb model is utilized to guide the creation of the recommendation engine so as to provide the right recommendations. The engine employs both offline and online data gathering and processing procedures as follows:

Offline Operation Procedures

During the Analysis Phase the following offline operation procedures are used:

Step 1 – Administer focus questions to groups of users and collect HSCO information on country characteristics, user requirements and the RS itself.

Step 2 – Estimate the preferences' priorities (e.g. purchasing decisions or searching patterns) according to the user groups and the type of the application

Step 3 – Compute similarity measures between user groups

Step 4 – Derive the possibility measures

Online Operation Procedures

Step 1 – Set up basic parameters by establishing an initial dialogue with users to collect HSCO information (e.g. age, gender)

Step 2 – Record on-click and identify users' interests and preferences patterns

Step 3 – Analyze search and browse patterns, build categories

Step 4 – Match related content with categories

Step 5 – Provide recommendations

First, historical data are selected and added into datasets. For every search, if frequently occurring patterns (classifiers) are found then those of good quality are used for recommendations. When new information is requested, the system identifies the corresponding class labels using multiple classifiers. Finally, the performance of the RS is evaluated by investigating the accuracy of the recommendations offered.

Typically, a client is accessing the RS through her/his Web browser where she/he can search and retrieve information (fig. 2). The Web server will receive her/his request and information and then process the data. The requirements pre-processing subsystem will receive the request and search the semantic rule database and if there is a relationship then there is a match to the original requirements. Offline classifiers are built and stored in the classifier rule database for new classifications. For every new customer requirements, the server will find all the classifiers that meet the conditions of these requirements by searching the classifier rule database and the new recommendations are provided.

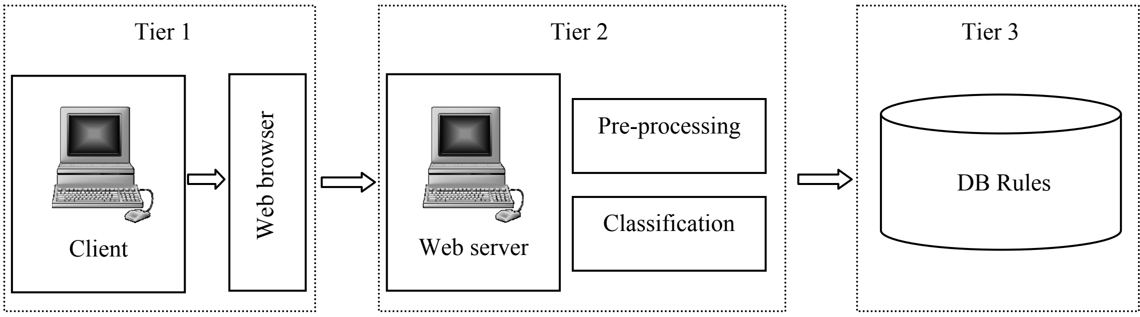

Figure 2: The 3-Tier Architecture of our RS.

3 RS CASE STUDY

We have developed an application to provide information on treatments and health care providers around the world based on patients' requests for medical tourism. Finally, we enhanced the system by adding a RS to provide more options to patients on health care providers as well as treatments.

RS are typically either homogeneous (i.e., content-based filtering) or heterogeneous (i.e., collaborative filtering) for product recommendations (Yuan and Cheng, 2004; Schafer, Konstan and Riedl, 1999). The Content-Based Filtering (CBF) approach recommends products to target customers according to the preferences of their neighbors while the Collaborative Filtering recommends products to target customers based on their past preferences. We have developed a hybrid RS based on CBF and CF approaches. Therefore, we combined both collaborative and content association rules to form hybrid association rules.

Focusing on the HSCO factors we have isolated three such factors for demonstration purposes, namely Gender, Age and Country. A user is searching the site on the treatment and destination he/she is interested e.g. Cosmetic Surgery in Cyprus or browsing the site for the treatment he/she is interested (fig. 3). The user may select one of the available providers to check his profile information such as, description, facilities, medical team, etc. The RS system provides recommendations for more available related treatments (on cosmetic surgery) offered by the provider. In addition, he/she will receive recommendations on more available providers that offer the treatment requested, in this case cheek implants.

If we now take into account the aforementioned three HSCO factors then recommendations become more accurate and closer to the user (hidden) preferences. According to user's profile the system provides some recommendations that are directed towards that specific gender like *Masks*, *Lipsticks*, *Makeup* etc (fig. 4a). According to Age, a simple rule inserted into the system as a result of the pre-processing stage the system suggests certain types of cosmetic surgery or/and treatments that are highly popular among a certain range of age, like face lifting, botox, breast augmentation or lift, etc. (fig. 4b).

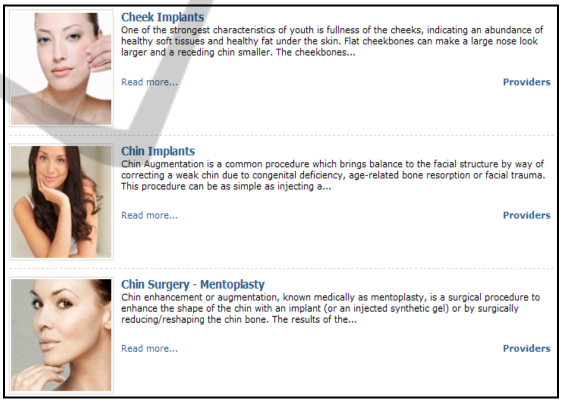

Figure 3: Browsing Cosmetic Surgery Treatments.

Country characteristics may also lead to suggesting other sources of information, like site seeing in the specific country, monuments etc (fig. 4c). The recommendations are also prioritized based on the city of his/her choice and neighboring cities within the country of origin based on distance proximity.

The system was validated using a population of 75 people both male and female persons, with average age of 43 years old.

A questionnaire was distributed to the population to evaluate the effectiveness and efficiency of the RS. The responses of the users revealed the correct nature of the recommendations which were made based on the HSCO factors gathered via the SpiderWeb. Female subjects of our population were

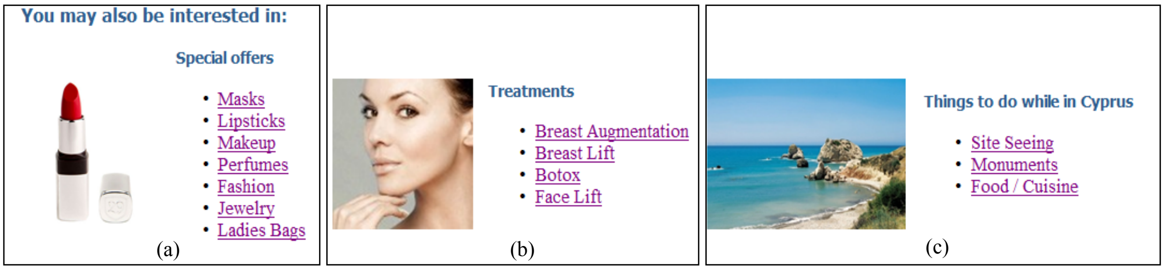

Figure 4: Recommendations of information based on (a) Gender, (b) Age, (c) Country.

excited to read the recommendations on cosmetics, fashion and the alternative treatments (surgical) suggested by the system. The male subjects did not receive any such recommendations and were satisfied they did not have to go through information of no practical benefit to them. In addition, ratings of the recommendations ranged between good and excellent, with people finding very helpful the suggestions of the system, understandable and clear. It is worth noting that 5 out of our 35 female persons were reluctant to give their age (and did not actually provide this data till the end of the session) as they felt it was a sensitive personal information. Fortunately, 4 of our remaining female persons that actually inserted their age to the system were over 40 and thus they received the information regarding additional treatments based on their age range, something which was found at first amusing but also quite interesting as these specific persons had limited knowledge on the suggested therapies. Overall, 72 people replied that they prefer to use a Web application with recommendations.

4 CONCLUSIONS

In this paper we have proposed a method based on the SpiderWeb model for identifying significant human, social, cultural, and organizational (HSCO) requirements for enhancing the development of RS.

We have developed a small pilot application with which the efficacy and efficiency of a RS that takes into consideration HSCO factors was demonstrated. A short scale user validation suggested that indeed when taking such factors into consideration for producing recommendations satisfaction between users is greatly enhanced as the time to search for and locate useful information is minimized and often the suggested information broadens the views of users on certain subjects.

Future research will concentrate on the use of Fuzzy Cognitive Maps (FCM) so as to approach user behaviour from a quantitative point of view.

REFERENCES

Andreou, A. S., Mavromoustakos, S. M., Schizas, C. N. 2002. Towards the Identification of Human, Social and Organizational Requirements for Successful E-Commerce Systems Development. In *Proceedings of the International Conference on Financial Engineering, E-Commerce and Supply Chain, Athens, Greece, pp. 24-30.*

Baraglia, R., Silvestri, F., 2007. Dynamic Personalization of Web sites Without User Intervention. *Communications of the ACM, Vol.50, No.2, pp.63-67.*

Jian Chen, J., Shtykh, R., Jin, O. 2009. A Web Recommender System Based on Dynamic Sampling of User Information Access Behaviors. *9th IEEE International Conference on Computer and Information, Technology, Xiamen, China,, pp. 172-177.*

Pressman, R. S., 2000. *Software Engineering: A Practitioner's Approach.* London: Mcgraw-Hill.

Yuan, S. T., Cheng, C., 2004. Ontology-Based Personalized Couple Clustering For Heterogeneous Product Recommendation. In *Mobile Marketing. Expert Systems* With Applications, 26, *pp.461–476.*

Schafer, J. B., Konstan, J. A., Riedl, J., 1999. Recommender Systems in E-Commerce. In *Proceedings of the 1st ACM EC, pp.158–166.*