A Collaborative Recommender System based on Asymmetric User Similarity

Marta Millan[1], Maria Trujillo[1] and Edward Ortiz[1]
[1] School of Systems and Computer Engineering of Universidad del Valle, Ciudad Universitaria Melendez, Cali, Colombia
{millan, mtrujillo}@univalle.edu.co

Abstract. Recommender systems could be seen as an application of a data mining process in which data collection, pre-processing, building user profiles and evaluation phases are performed in order to deliver personalised recommendations. Collaborative filtering systems rely on user-to-user similarities using standard similarity measures. The symmetry of most standard similarity measures makes it difficult to differentiate users' patterns based on their historical behaviour. That means, they are not able to distinguish between two users when one user' behaviour is quite similar to the other but not *vice versa*. We have found that the *k*-nearest neighbour algorithm may generate groups which are not necessarily homogenous. In this paper, we use an asymmetric similarity measure in order to distinguish users' patterns. Recommendations are delivered based on the users' historical behaviour closest to a target user. Preliminary experimental results have shown that the similarity measure used is a powerful tool for differentiating users' patterns.

Keywords: Recommender systems, collaborative filtering, asymmetric similarity measure, *k*-nearest neighbours.

1 Introduction

Recommender systems could be seen as an application of a data mining process [21] in which data collection, pre-processing, building user profiles and evaluation phases are performed in order to deliver personalised recommendations. The goal of recommender systems is to provide a user with personalized recommendations based on either his/her tastes and preferences or based on a group of people with similar tastes and preferences [1]. Five classes of recommendation techniques are proposed in terms of the background data, input data and the algorithm to generate recommendations: collaborative, content-based, demographic, utility-based and knowledge-based [5].

Collaborative filtering techniques, in particular, rely on user-to-user similarities but have three major limitations: sparsity, scalability, and cold-star. Several methods have been proposed for solving these limitations based on clustering and machine learning techniques [3][4][14][16][19]. ClustKNN [23] addresses the scalability problem applying the *k*-means algorithm for building a user model and the *k*-nearest neighbour (KNN) clustering algorithm for calculating predictions. Kim *et al.* [16] propose a probabilistic model generated under the assumption that items are not related to each other in any way, *i.e.* they are independent. A smoothing-based method is introduced under a hybrid collaborative filtering approach in [34]. From training data, initial clusters are calculated using the *k*-means algorithm. The Pearson correlation coefficient is used as a similarity measure function. Smoothing strategies are applied to unrated items.

In Godoy *et al.* [11] user-profiling approaches to develop agents that help users in finding, filtering and getting information tasks are reviewed. These approaches need data and information about users in order to capture user's profiles. To construct user's profiles it is necessary to infer information based on the user-system interaction. User profiles are based on knowledge acquired implicitly or explicitly from this interaction. Some sources of this information, proposed in Godoy *et al*. [11], include historical user navigation, access logs and relevance feedback – explicit or implicit. Explicit feedback requires that users assign values to an examinee item. Instead, implicit feedback is inferred based on implicit interest indicators.

In McLaughlin and Herlocker [19], a user nearest-neighbour (NN) algorithm is analysed and a belief distribution algorithm is introduced in order to improve user modelling. In this proposal, a predictive algorithm, solves two main drawbacks of NN approaches: few neighbours who have rated an item are available for a target user; neighbours with a very low correlation score to target user rated an item. The Pearson correlation is used to calculate the most similar N-users for a target user from account historical item rating data. Taking into account that user's rating are subjective a belief difference distribution is introduced from calculating correlations.

In [31], an approach of collaborative filtering was introduced in which user neighbourhood is calculated based on demographic data, psychographic data and users' historical behaviour. A weight similarity measure is proposed for clustering users in order to take into account dynamic human being behaviour. This measure is characterized by providing a way to define which characteristics are more important at a specific moment. Moreover, characteristics are used in a nominal scale of measurement since users' behaviour has no order when time is not taken into account.

We have found that the *k*-nearest neighbour (KNN) algorithm may generate groups which are not necessarily homogenous. This is probably due to the size of the available data. We have detected that these differences may underestimate the similarity between users.

In this paper, a collaborative recommender system based on an asymmetric measure is introduced. It is derived from the need to distinguish between two users when one user' behaviour is quite similar to other but not *vice versa*. We use an asymmetric similarity measure for distinguishing users' patterns [6][7]. In this approach, a user-to-user similarity matrix is built and clusters are extracted through thresholding. Recommendations are delivered based on the users' historical behaviour closest to a target user. Preliminary experimental results have shown that the similarity measure is a powerful tool for differentiating users' patterns.

The paper is organized as follows. Collaborative recommender systems basic concepts are described in Section 2. In Section 3, we introduce a recommender system based on asymmetric users' patterns. A digital library experimental framework where our approach has been implemented is described in Section 4. Preliminary evaluation and final remarks are presented in Sections 5 and 6, respectively.

2 Collaborative Recommender Systems

In a collaborative filtering process, there is a set of *m* users, U= $\{u_1, u_2, \ldots, u_m\}$ and a set of *n* items, I=$\{i_1, i_2, \ldots, i_n\}$. To each user u_j, a list of items $L_j = \{i_k\text{: } 1 \leq k \leq n_j.\}$ is associated. L_j contains items on which a user has explicitly shown interest by assigning a rating score to them or implicitly, based on user behaviour. Collaborative filtering algorithms are applied in order to find interesting items for a target user. Interesting items can be obtained in two ways: prediction and recommendation [31]. In the prediction way, calculating a predictive score $P_{score}(i_q,u_j)$ that represents the level of interest that an item $i_q \notin L_j$ may have for user u_j. In the recommendation way, a list L of N-top items, such that $L \subset I$ and $L \cap L_j = \varnothing$, is delivered to user u_j. Thus, L contains items that could be interesting for user u_j.

Collaborative filtering user-based recommendation approaches try to identify neighbourhoods for each user based on similar features – *e.g.* demographics, psychographics, behavioural – [13][17][22][24][27]. Most collaborative techniques work based on ratings about items provided for users.

Collaborative filtering techniques can be classified as either model-based or memory-based [1]. The former builds a model from historical data to recommend other items [3] [4] [14] [25]. The latter uses a utility function in order to calculate similarity between users to build neighbourhoods [9][12] [15] [17][24][27] [32][33]. If a user is included in a neighbourhood – similar tastes are shared – it is possible to predict the utility of an item to him/her based on others items rated by users in the same neighbourhood.

Similarity measures are commonly evaluated using ratings as though they were quantitative values regardless of the fact that these measures are not defined for attributes or subjective user evaluations. The Pearson correlation and the cosine distance are often used to assess similarity between users [12] [20] [29] [32]. However, – according to Choi *et al.* [9] – these similarity measures have some weaknesses: scalability problems, applying limitations depending on the domain, and assuming that attributes are mutually independent. Garden *et al.* [10] and Herlocker *et al.* [12] have considered the use of the Spearman correlation, which is just the Pearson correlation coefficient computed for the ranks of the ratings.

3 A Recommender System based on Asymmetric User Similarity

The proposed approach is integrated into an experimental computer sciences digital library. In this context, users are students and items are digital documents, such as papers, books, research reports and theses. This approach uses an unobtrusive method for recommendation calculations that takes into account information related to the documents he/she has previously downloaded.

The approach combines advantages of memory-based and model-based collaborative recommendation systems. Memory-based – in an off-line phase – deals with user neighbourhood calculations in order to reduce the computational complexity – scalability problem. Model-based – in an online phase – uses a probability model for the preferences prediction based on user neighbourhood information. Thus, a user preferences pattern is calculated in an off-line phase. His/her preferences prediction is calculated in an on-line phase.

3.1 Asymmetric Users' Similarity

Collaborative recommender systems are based on an important feature of human behaviour which is the tendency to consume a limited set of items. Thus, a set of items becomes fashionable for a group of people with similar preferences. Collaborative filtering systems are based on similarity of users in which a neighbourhood to each user is generated. A neighbourhood is built taking into account user's similarity.

Although users hardly ever give explicit feedback and user interests change over time [26], most collaborative techniques work based on ratings about items provided for users. Ratings can be obtained explicitly or implicitly. Explicit ratings are subjective user evaluations or voting. The similarity between two users is evaluated using ratings as numeric values. Similarity measures – *e.g.* Pearson correlation coefficient and cosine distance – are applied regardless of the fact that these functions are not defined for subjective user evaluations. That is, the average of "very satisfied", "somewhat dissatisfied" and "very dissatisfied" does not exist [8] [28].

The normalised rank transformation, the Spearman correlation, the Kendall correlation, the Pearson correlation and the Footrule, the Cayley, the Hamming, the Ulam, the Chebyshev/Maximum value and the Minkowski distances are commonly used to calculate similarity or distance between two ranks or two rating vectors [30]. These measures are not asymmetric, that is, they can not capture differences between two users when one user has a lot more historical information than the other. The users' similarity measure has to distinguish when a user u_b has a lot more historical information than user u_a whilst user u_a has a quite similar historical information as user u_b. This could be seen as quantifying the similarity between u_a and u_b when $L_a \subseteq L_b$. In this case, the similarity score between user u_a and user u_b has to be large whilst the similarity score between user u_b and user u_a has to be small.

Similarity between users is defined as either exact or approximate historical information matches or coincidences of patterns. Thus, the similarity between u_a and u_b could be measured as follows [6] [7]:

$$S(u_a, u_b) = \frac{|L_a \cap L_b|}{|L_a|}, \tag{1}$$

where $|\cdot|$ is the cardinality of a set or a list of items.

By asymmetry, the similarity between u_b and u_a is given by:

$$S(u_b, u_a) = \frac{|L_a \cap L_b|}{|L_b|}. \tag{2}$$

A user-to-user similarity matrix is built using the similarity measure. In this matrix, the *j-th* row represents the similarity between user *j* and the set of users.

For instance, Figure 1 represents a set of users and their lists of downloaded data.

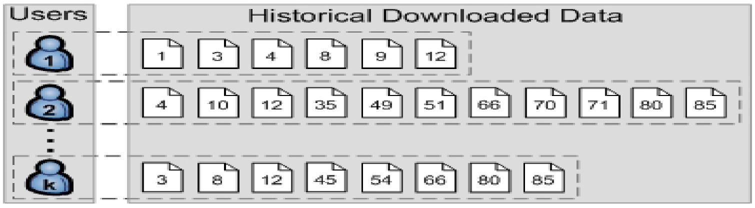

Fig. 1. Illustration of downloaded data

Thus, the user-to-user similarity matrix has to be fully calculated. A threshold is used for selecting the most similar user to the target user in order to obtain the target user profile. For instance, setting the threshold equal to 0.7, the nearest neighbours of u_j are selected as all users who have a similarity score larger than 0.7 in the *j-th* row in the user-to-user similarity matrix.

3.2 Predicted Scores

Recommendations are based on the probability that a user has a preference for a particular document. If this probability is high it is more likely that a recommendation on a document will be useful to the user. For the sake of completeness, we outline the notation used in this section.

Let u_j be the target user. Let $\boldsymbol{C}_{u_j}$ be the cluster of users in the nearest neighbours of u_j. Let i_q be the target document. Let $\boldsymbol{C}_{i_q,u_j}$ be the cluster of users in the nearest neighbourhood of u_j who downloaded i_q.

The predicted score that u_j has preferences for i_q is proportional to the probability that users in $\boldsymbol{C}_{u_j}$ have shown preferences for i_q.

$$P_{score}\left(i_q, u_j\right) = \frac{\left|\boldsymbol{C}_{i_q,u_j}\right|}{\left|\boldsymbol{C}_{u_j}\right|}, \tag{3}$$

where |.| is the cardinality of a set of users.

The predicted score is calculated based on the target user neighbourhood preferences on the target document.

4 Deployment

Our approach is implemented into a digital library experimental framework. The architecture of the integrated system is shown in Fig. 2. The digital library is supported on a data base that contains information on user registrations and documents. The recommender system is supported on a data mart. Information such as document metadata, documents downloaded by user, and characteristic vectors is stored in the data mart. Recommendations are generated based on historical downloaded user's data. This is an unobtrusive approach and takes a downloaded document action as an implicit feedback.

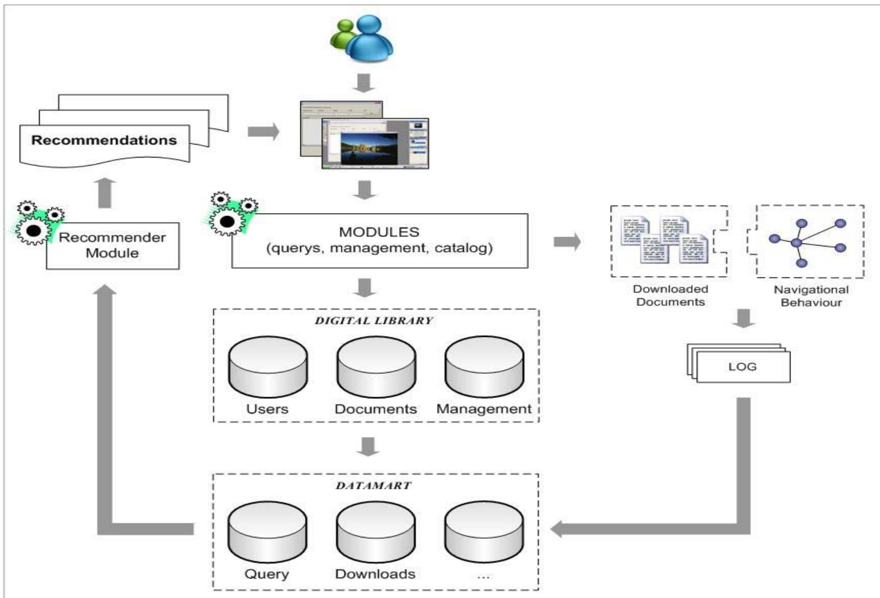

Fig. 2. System Architecture

Recommendations are twofold: a list of recommendations based on collaborative filtering and a second list based on downloaded document frequencies. Documents in the former, as in other recommender systems, are ranked according to their predicted score – in Equation 3. A verification process of previous recommendations is carried out to avoid making the same recommendations. However, collaborative filtering techniques provide recommendations regardless of users current interests. As for the second list, a user current interest is shown as his/her navigational behaviour. After his/her first search using key words, the main area of knowledge associated to those key words is used in a documents query and documents are sorted out by download frequency. A list of *n*-top documents is generated and a recommendation window is shown. Moreover, users have the choice of entering the window to check the recommendations included in the list or closing that window for the time being. Whatever courses of action a user takes are registered in the data mart.

5 Preliminary Evaluation

We have presented an approach to personalised information retrieval tasks in a digital library environment. According to Adomavicius *et al*. [1] the personalisation process is integrated by three stages: understanding customers, delivering personalised offerings, measuring personalisation impact. In this paper, we focused on understanding users and delivering personalised offering phases. Moreover, once the digital library data mart contains enough information – in this specific domain – we will be able to evaluate our approach.

The performance of the asymmetric similarity measure is evaluated using the MovieLens data set, developed at the University of Minnesota [Available at http:// www.grouplens.org]. The dataset contains 1,000,209 anonymous ratings of approximately 3,900 movies made by 6,040 MovieLens users who joined MovieLens in 2000. In a pre-processing phase, the most frequent rated genre, which is drama, was selected for the experimental validation.

We take as implicit rating the action of rating a movie. Information from u_1 to u_{100} was used in this evaluation. The asymmetric similarity measure was evaluated with the selected data set. The maximum similarity value reached between users was twelve times, as shown in Table 1. Users u_{53}, u_{70} and u_{77} rated the same movies that u_{21} rated. However, u_{53}, u_{70} and u_{77} rated more movies than u_{21} as the similarity score between u_{53} and u_{21} shows. In a similar manner, u_{33}, u_{58} and u_{88} rated the same movies that u_{94} did.

Table 1. MovieLens data set: List of users with the asymmetric similarity measure equal to one

Asymmetric Similarity Score	Asymmetric Similarity Score	Asymmetric Similarity Score	Asymmetric Similarity Score
$S(u_3,u_{62}) = 1$	$S(u_{62},u_3) = 0.3055560$	$S(u_{61},u_{48}) = 1$	$S(u_{48},u_{61}) = 0.0328947$
$S(u_7,u_{48}) = 1$	$S(u_{48},u_7) = 0.0526316$	$S(u_{61},u_{92}) = 1$	$S(u_{92},u_{61}) = 0.0515464$
$S(u_{20},u_{48}) = 1$	$S(u_{48},u_{20}) = 0.0394737$	$S(u_{94},u_{33}) = 1$	$S(u_{33},u_{94}) = 0.0122699$
$S(u_{21},u_{53}) = 1$	$S(u_{53},u_{21}) = 0.00340136$	$S(u_{94},u_{58}) = 1$	$S(u_{58},u_{94}) = 0.0150376$
$S(u_{21},u_{70}) = 1$	$S(u_{70},u_{21}) = 0.0588235$	$S(u_{94},u_{88}) = 1$	$S(u_{88},u_{94}) = 0.0416667$
$S(u_{21},u_{77}) = 1$	$S(u_{77},u_{21}) = 0.0588235$	$S(u_{98},u_{58}) = 1$	$S(u_{58},u_{98}) = 0.0601504$

Based on a user-to-user similarity matrix, the u_{100} neighbourhood was determined setting the threshold equal to 0.6; u_{100} rated 24 movies. The list of users belonging to u_{100} neighbourhood is shown in Table 2.

Table 2. MovieLens data set: List of users belonging to the u_{100} neighbourhood using the asymmetric similarity measure and a threshold equal to 0.6

The u_{100} Neighbourhood	Rated Movies	Asymmetric Similarity Score
u_{58}	113	0.791667
u_{48}	152	0.791667
u_{53}	294	0.666667
u_{62}	242	0.625000
u_{36}	111	0.625000

The list of users belonging to the u_{100} neighbourhood using 5-NN, based on a user-to-user similarity matrix calculated with the Jaccard similarity coefficient [8] is shown in Table 3. However, the u_{100} neighbourhood shown in Table 3 changes when a

Jaccard is used due to the fact that the measure is affected by the large quantities of movies rated by users in Table 2.

Table 3. MovieLens data set: List of users belonging to the u_{100} neighbourhood using 5-NN

The u_{100} Neighbourhood	Rated Movies	Jaccard Score
u_{93}	23	0.270270
u_{13}	22	0.243243
u_{47}	20	0.222222
u_{95}	40	0.185185
u_7	8	0.185185

A prediction accuracy metric, the relative frequency with which the system makes correct decisions about whether a movie is of interest to a user, was used in the preliminary evaluation. Available information was divided into 90% training set and 10% cross-validation set. That is, when a user has 10 ratings, 9 ratings are used for building the model and 1 rating is used for validating the model.

When a list with 8 recommendations was generated for each user using the proposed approach, 78% of the users had rated at least one of the recommended movies in his/her cross-validation set. When a list with 8 recommendations was generated for each user using 5-NN, 56% of the users had rated at least one of the recommended movies in his/her cross-validation set. When a list with 8 recommendations was generated for each user using 10-NN, 47% of the users had rated at least one of the recommended movies in his/her cross-validation set.

6 Final Remarks

We had detected that the difference between document downloaded quantities may underestimate the similarity between users. We have proposed the use of an asymmetric similarity measure for reducing the impact of comparing users on the basis of the number of downloaded documents (large versus small quantities). The measure is used to identify a neighbourhood whose traits are strongly similar to those of an active user's behaviour thus reducing the possibility of generating irrelevant recommendations.

This approach has two characteristics: users' neighbourhood is dependent of a similarity score value rather than of a predefined number and a user does not always belong to the neighbourhood of the users who belong to his/her own neighbourhood. For generating recommendations, a prediction score is calculated based on the target user neighbourhood preferences on the target document.

Our next step, with data generated by users of the digital library, is to build a lifetime model for evaluating recommendations and to use a Bayesian approach for taking into account information no longer used in recommendations calculation, as *a priori* knowledge.

Acknowledgments. This work was partially supported by Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología “Francisco José de Caldas” (Colciencias) under Grant No. 1406-14-17902.

References

1. Adomavicius, G., Tuzhilin, A. Toward a Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Transaction on Knowledge and Data Engineering 17(6) (2005) 734-749
2. Anderberg, M. Cluster Analysis for Applications, New York, Academic Press. (1973).
3. Billsus, D., Pazzani, M. Learning Collaborative Information Filters. In Proceedings of the Fifteenth International Conference on Machine Learning (1998) 46-54
4. Breese, J., Heckerman, D. Kadie, C. Empirical Analysis of predictive algorithms for collaborative filtering. In Proceedings of 14th Annual Conference on Uncertainty in Artificial Intelligence, Morgan Kaufman (1998) 43-52
5. Burke, R. Hybrid Recommender Systems: Survey and Experiments. User Modelling and User-Adapted Interaction 12(4) (2002) 331-370
6. Chen, H., Lynch, K. Automatic Construction of Networks of Concepts Characterizing Document Databases. Journal of IEEE Transactions on Systems, Man and Cybernetics 22(5) (1992) 885-902
7. Chen, H., Martínez, J., Ng, T., Schatz B. A Concept Space Approach to Addressing the Vocabulary Problem in Scientific Information Retrieval. An Experiment on the Worm Community System. Journal of the American Society for Information Science 48(1) (1998) 17-31
8. Chen, P. Y., Popovich, P. M. Correlation: Parametric and Nonparametric Measures. Sage University Papers Series: quantitative application in the social science, 2002
9. Choi, S.H., Kang, S., Jeon, J.Y. Personalized recommendation system based on product specification values. Expert Systems with Applications 31 (2000) 607-616
10. Garden, M. and Dudek, G. Mixed Collaborative and Content-Based Filtering with User-Contributed Semantic Features. In Proceedings of the Twenty-First National Conference on Artificial Intelligence (2006) Menlo Park, CA: AAAI Press
11. Godoy, D., Amadi, A. User profiling in personal information agents: a survey. The Knowledge Engineering Review Vol. 00:0 1-33 (2005) Cambridge University Press
12. Herlocker, J., Konstan, J., Borchers, A., Riedl, J. An Algorithm Framework for Performing Collaborative Filtering. In Proceedings of ACM SIGIR’99, Berkeley, California (1999) 230-237
13. Jin, R., Chai, J., Si L. An Automatic Weighting Scheme for Collaborative Filtering. In Proceedings of SIGIR’04, Sheffield,UK (2004) 337-344
14. Kamahara, J., Asawaka, T., Shimojo, S., Miyahara, H. A Community-based Recommendation System to Reveal Unexpected Interest. In Proceedings of the 11th International Multimedia Modelling Conference (MMM’05) IEEE Computer Society (2005) 433-438
15. Karpis, G. Evaluation of Item-Based Top-N Recommendation Algorithms. In Proceedings of the Tenth International Conference of Information and Knowledge Management, Atlanta, USA (2001) 247-254
16. Kim, H., Kim, J. Herlocker, J. Feature-Based Prediction of Unknown Preferences for Nearest-Neighbour Collaborative Filtering. Data Mining, 2004. ICDM 2004. In Proceedings of Fourth IEEE International Conference (2004) 435 – 438

17. Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., Riedl, J. GroupLens: applying collaborative filtering to Usenet news. Communications of the ACM 40(3) (1997) 77-87
18. Linden, G., Smith, B., York, J. Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 7(1) (2003) 76-80
19. McLaughlin, M., Herlocker, J. A Collaborative Filtering Algorithm and Evaluation Metric that Accurately Model the User Experience. In Proceedings of SIGIR'04, Sheffield, UK (2004) 329-336
20. Montaner, M., Rosa, J.L. A taxonomy of recommender agents on the Internet. Artificial Intelligence Review, 19 (2003) 285-330
21. Papagelis, M., Plexousakis, D., Rousidis, I., Theoharopoulos, E. Qualitative Analysis of User-based and Item-based Prediction Algorithms for Recommendation Systems. Journal of Engineering Applications of Artificial Intelligence, 18 (2005), 781-789
22. Pennock, D.M., Horvitz, E., Lawrence, S, Giles, C. L. Collaborative filtering by personality diagnosis: A hybrid memory and model-based approach. In Proceedings of the 16th Conference on Uncertainty and Artificial Intelligence, Stanford California. (2000) 473-480
23. Rashid, A.M., Lam, H., Karypis, G., Riedl, J. ClustKNN A highly Scalable Hybrid Model-& Memory-Based CF Algorithm. In Proceedings of WEBKDD´06, Philadelphia, Pennsylvania USA (2006)
24. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J. Grouplens: an open architecture for collaborative filtering of netnews. In Proceedings of the ACM 1994 Conference on Computer Supported Cooperative Work, Chapel Hill, NC USA (1994) 175-186
25. Sarwar, B., Karypis, G., Konstan, J., Riedl, J. Analysis of recommendation algorithms for E-commerce. In EC'00 Proceedings of the 2nd ACM Conference on Electronic Commerce. Minneapolis, Minnesota, United States (2000) 158-167
26. Schwab, I., Kobsa, A., Koychev, I. Learning User Interests through Positive Examples Using Content Analysis and Collaborative Filtering. Draft from Fraunhofer Institute for Applied Information Technology, Germany, 2001. Available at http://citeseer.ist.psu.edu
27. Shardanand, U., Maes, P. Social Information Filtering: Algorithms for Automating "Word of Mouth". Conference on Human Factors in Computing Systems. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Denver Colorado, USA (1995) 210-217
28. Shoukri, M. M. Measures of Interobserver Agreement, Chapman Hall/CRC, 2004.
29. Sugiyama K., Hatano K., Yoshikawa M. Adaptive Web Search based on user profile constructed without any effort from users. In Proceedings of WWW 2004, New York USA 675-684
30. Teknomo, Kardi. Similarity Measurement. Available at http://people.revoledu.com/ kardi/ tutorial/Similarity/
31. Trujillo, M., Millan, M., Ortiz, E. A Recommender System Based on Multi-features. International Conference on Computational Science and Its Applications – ICCSA, LNCS 4706, Part II, (2007) 370-382
32. Wang, J., de Vries, A., Reinders, M. Unifying User-based and Item-based Collaborative Filtering Approaches by Similarity Fusion. In SIGIR '06 Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval Seattle, Washington, USA (2006) 501-508
33. Xi, W., Fox E., Fan, W., Zhang, B., Chen, Z., Yan J., Zhuang, D. SimFusion: Measuring Similarity using Unified Relationship Matrix. In Proceedings of SIGIR'05 Salvador, Brazil (2005) 130-137
34. Xue G., Lin C., Yang Q., Xi W., Zeng H., Yu Y. Scalable Collaborative Filtering Using Cluster-based Smoothing. In Proceedings of SIGIR'05 Salvador, Brazil (2005) 114-121