Knowledge-Based Systems 22 (2009) 261–265

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Collaborative filtering adapted to recommender systems of e-learning

J. Bobadilla *, F. Serradilla, A. Hernando, MovieLens

Universidad Politecnica de Madrid, Computer Science, Crta. De Valencia, Km 7, 28031 Madrid, Spain

ARTICLE INFO

Article history:
Available online 3 February 2009

Keywords:
Recommender systems
Collaborative filtering
e-Learning

ABSTRACT

In the context of e-learning recommender systems, we propose that the users with greater knowledge (for example, those who have obtained better results in various tests) have greater weight in the calculation of the recommendations than the users with less knowledge. To achieve this objective, we have designed some new equations in the nucleus of the memory-based collaborative filtering, in such a way that the existent equations are extended to collect and process the information relative to the scores obtained by each user in a variable number of level tests.

1. Introduction

Recommender systems (RS) cover an important field within collaborative services that are developed in the Web 2.0 environment [1–3] and enable user-generated opinions to be exploited in a sophisticated and powerful way. RS can be considered as social networking tools that provide dynamic and collaborative communication, interaction and knowledge.

RS cover a wide variety of applications [4–6], although, those related to movie recommendations are the most well-known and most widely-used in the research field [7–9]. Nevertheless, the collaborative e-learning field is strongly growing [10,11], converting this area in an important receiver of applications and generating numerous research papers [12,13] into the computer science field [14,15] and into different areas [16,17]. The endeavor to create distributed, federation [18] and grid [19] collaborative e-learning services are particularly interesting.

The RS stage that normally has the greatest influence on the quality of the results obtained is the collaborative filtering (CF) phase [20,21]. CF is based on making predictions about a user's preferences or tastes based on the preferences of a group of users that are considered similar to this user. A substantial part of the research in the area of CF centers on how to determine which users are similar to the given one; in order to tackle this task, there are fundamentally 3 approaches: memory-based methods, model-based methods and hybrid approaches.

Memory-based methods [22,23] use similarity metrics [21] and act directly on the ratio matrix that contains the ratings of all users who have expressed their preferences on the collaborative service; these metrics mathematically express a distance between two users based on each of their ratios. Model-based methods [22] use the ratio matrix to create a model from which the sets of similar users will be established. Among the most widely-used models we have: Bayesian classifiers [24], neural networks [25] and fuzzy systems [26]. Generally, commercial RS use memory-based methods [27], while model-based methods are usually associated with research RS.

Regardless of the method used in the CF stage, the technical objective generally pursued is to minimize the prediction errors, by making the accuracy [28–31] of the RS as high as possible; nevertheless, there are other objectives that need to be taken into account: avoid overspecialization phenomena, find good items, credibility of recommendations, precision, recall measures, etc.

Memory-based methods work on a table of U users who have rated I items. The prediction of a non-rated item i for a user u is computed as an aggregate of the ratings r of the K most similar users (k-neighborhoods) for the same item i. The most common aggregation approaches are the average and the weighted sum; the similarity approaches usually compute the similarity between two users x and y: sim(x,y) based on their ratings of items that both users have rated. The most popular similarity metrics are Pearson correlation and cosine.

2. e-Learning memory based filtering

One of the ideas underlined in the philosophy of the actuation of the RS is based on the equality between its users, not only on their possibilities of access to the service, but also above all with regards to the contribution by each one of them to the recommendations that the rest could receive. The usual RS generate the recommendations for each user based on the ratios supplied by the users with contributions most similar to them.

The equal treatment between users is adequate and convenient in the majority of the RS, for example, there is no reason, in advance, to believe that one user is more qualified than another to offer recommendations about movies, journeys, blogs, etc. However,

* Corresponding author. Tel.: +34 670711147; fax: +34 913367522.
E-mail address: jesus.bobadilla@upm.es (J. Bobadilla).

0950-7051/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2009.01.008

there exists a group of RS in which this situation does not make so much sense. The RS of e-learning are the most paradigmatic in this asymmetric situation; in these RS one can easily make a distinction between advanced and novice users, for example between the ratios generated by teachers and those generated by students, or those supplied by advanced students (for example final years) and those who are starting their studies.

One approach with which to confront this new characteristic could be to divide all the users into a list of distinct classes, each one of which could contain one of the reference groups (teachers, advanced students, novice students, etc.), and each one of the reference groups a weighting value for their importance. In our case, we have opted for a more generic and progressive solution, in which we have avoided the establishment of distinct classes establishing a weighting value for each one of the users of the system.

In our model of CF adapted to the RS of e-learning we start from the usual two-dimensional matrix of ratios R of U users for I items and we add to it another two-dimensional matrix C for scores of U users and T level tests or exams. In this way, each user of the educational system (fundamentally students) will be characterized by I possible ratings of items (in our case valuations of documentation, teachers, subjects, etc.) and by T scores from academic level tests, which demonstrate their knowledge of the educational material. The number T of level tests can be as large as desired, and can include scores of automatically corrected tests, scores of subjects that have been passed, scores from exercises, practicals carried out, etc.

The fundamental idea that we would like to formalize is the weighting of the recommendations which stem from the user details, not only like the traditional similarity between their ratios and those of the rest of the users, but also to take into account that the recommendations of the users with better scores have a greater weight than the recommendations of the users with lower scores.

In order to evaluate the importance of the knowledge of a user x (Cx) over the recommendations that will be received from him a user y with knowledge Cy, a large number of metrics can be established. In this paper it has been decided to use a simple and asymmetric metric that can be established by means of the function f (1), although, using other metrics such as those shown in (2) would also be feasible. The choice of one metric or another comes from the manner in which it is desired to weight the relationship of knowledge demonstrated by each pair of users and by the nature of the RS details themselves

$$f = \begin{cases} Cx - Cy, & Cx > Cy \\ 0, & Cx \leqslant Cy \end{cases} \tag{1}$$

Thus, in the metric (1), if the knowledge of user x is 0.7 (on a scale of 0–1) and that of user y is 0.2 (on the same scale), the weighting of the knowledge of user x to user y would be 0.5, while the weighting of knowledge of user y to user x would be zero

$$f = Cx - Cy, \quad f = (Cx - Cy)^2, \quad f = e^{Cx - Cy} \tag{2}$$

The metric that contains the exponential function in (2) is that which offers the smoothest and most progressive results, resulting in it being the most suitable metric when there are no important reasons for designing a new one that would be adapted to the specific needs of the RS.

The new measurement for similarity between the users x and y, which we call *importance*, can be established as defined in (3). The first term of the equation refers to the importance of the scores, while the second term refers to the similarity of the users according to their ratios and applying some of the traditional metrics (Pearson, Correlation, MSD...).

The sum serves to discover the arithmetic mean of the T scores that evaluate the knowledge of the user; a score that is not evaluated must be initialized with the minimum score (0 on the scale of 0–1). Cxt represents the knowledge of the user x on the t subject, test, etc.

$$\mathrm{imp}(x,y) = \left[\frac{1}{T}\sum_{t=1}^{T} f(Cxt, Cyt)\right] \times \mathrm{sim}(x,y) \tag{3}$$

The values of importance obtained between the pairs of users serve to obtain the desired k-neighborhoods of each user, just as is done with the traditional metrics of CF, and in this way, recommendations can be made based on the evaluations given to the k users most similar to each other.

3. Testing of the accuracy of the recommendations

In the proposed CF adapted to the RS of e-learning, the value of the item i (5) is computed according to the values of the ratios of all the users that have rated this item weighting each one of their mean scores (4). That is to say, the predicted value is compared not only with the value rated by the user, but with that expected according to the values rated by the users with better scores of the system (6), where $r_{u,i}$ denotes the rating of user U over the item I, whereas $\tilde{K}$ denotes the set of k-neighborhoods. The set of equations that express these ideas mathematically are:

$$\bar{c}_u = \frac{1}{T}\sum_{t=1}^{T} c_{u,t}, \quad c_{u,t} \in [0..1] \tag{4}$$

$$\begin{cases} e_i = \frac{1}{\mu}\sum_{u=1}^{|\tilde{U}|} \bar{c}_u r_{u,i} \quad \forall i | \exists r_{u,i} \neq \phi, \quad \tilde{U} \subset U | \exists r_{u,i} \neq \phi, \quad \mu = \sum_{u=1}^{|\tilde{U}|} \bar{c}_u \\ e_i = 0 \quad \forall i | \neg \exists r_{u,i} \neq \phi \end{cases} \tag{5}$$

$$\mathrm{error}_{u,i} = \frac{1}{|\tilde{K}|}\sum_{k=1}^{|\tilde{K}|} r_{k,i} - [(1-\alpha)e_i + \alpha r_{u,i}] \quad |r_{\mathrm{ui}} \neq \phi \wedge \exists r_{k,i} \neq \phi \tag{6}$$

The parameter α (with the range of 0–1) serves to adjust the weight that we give to the values rated by each user with regards to the weight that we give the evaluations by the users with better scores. Expressed in another way, the parameter $(1 - \alpha)$ allows variation of the importance given to the evaluations by people with greater knowledge (scores).

Fig. 1 shows the meaning of the presented equations.

4. Empirical tests performed: experiment design

Due to the lack of any well-known data base for e-learning, publicly accessible for research and which contains information about the scores of the users, we used a known RS database from a field that is different from e-learning; in order to test our approach of CF adapted to e-learning we took the first five items of the MovieLens database [32] as five scores which have been evaluated by each user, in such a way that in Eq. (4) T has the value 5 and we are able to obtain the mean score for each user. Previously a 0 is inserted for those items that have not been rated, therefore indicating that the knowledge of a user in a test not performed is nil. The remainder of the items is used to discover the similarity between pairs of users.

In all the experiments carried out, for each item that each user has rated, the average value of the ratios given by their k-neighborhoods for that item has been calculated and the prediction has been compared with the value rated by the user (6) weighted with its estimated value (5), thus obtaining the calculation of the mean absolute error (MAE).

The previous process was carried out for each of the following k-neighborhoods values: 15, 30, 60, 90, covering from 1.6% to 5.3% of the total number of users. Additionally, the values of the parameter α: 0.3–0.8 with steps of 0.1 have been used. The first

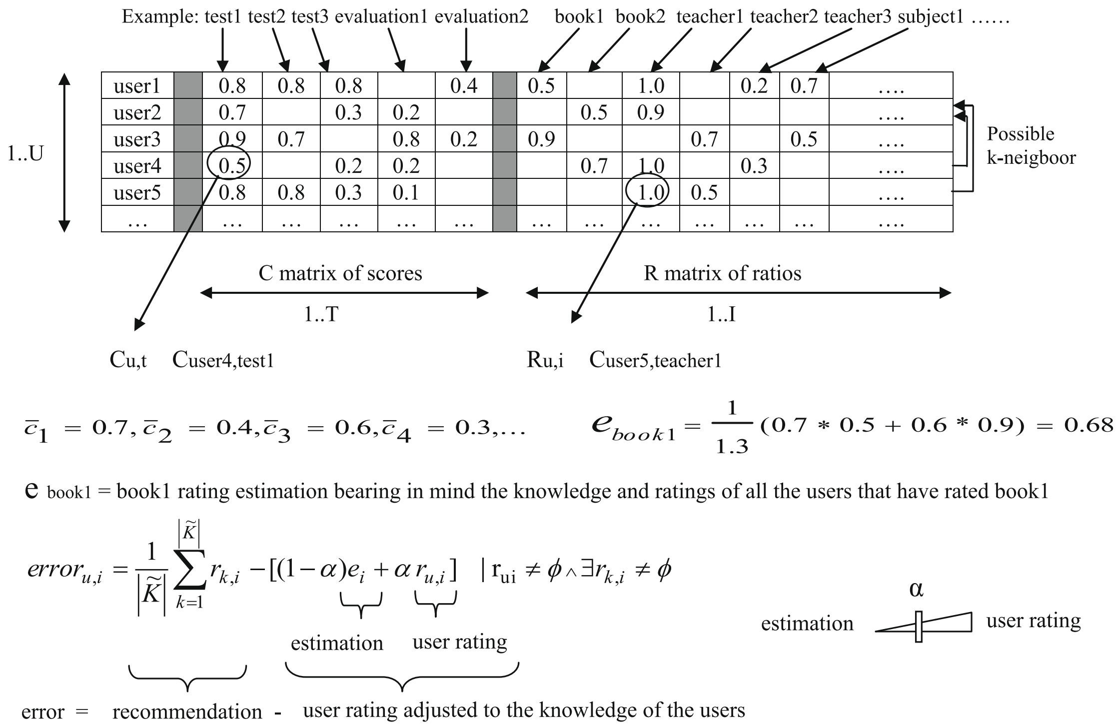

Fig. 1. Meaning of the equations.

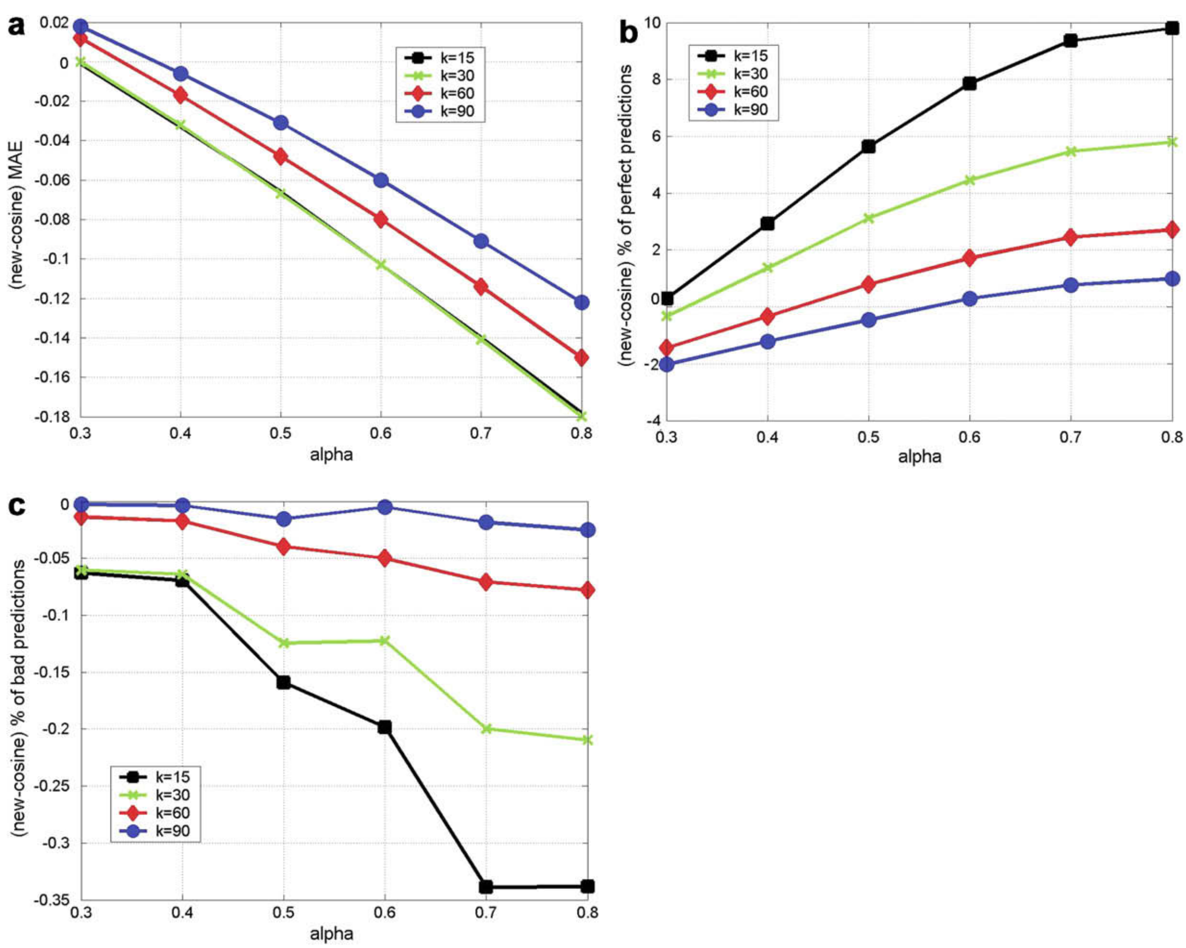

Fig. 2. Evolution of the cosine metric compared with the proposed metric in different values of the α and k-neighborhoods parameters: (a) Mean absolute error (MAE), (b) percentage of perfect predictions and (c) percentage of bad predictions.

values of α (0–0.2) were not included due to the fact that they give a disproportionate weight to the knowledge of the users, while the ultimate values of α (0.9 and 1) have not been used because they take us back to the traditional RS in which the knowledge of the users is not taken into account.

In order to evaluate the importance of the knowledge of a user x over the recommendations received from a user y, the cosine metric equation and the function f defined in (1) have been used. The results obtained were compared, in all cases, with those obtained using only the cosine metric, as is done in the traditional RS.

The expected results should show an improvement using the metric adapted to e-learning CF (3) as opposed to the use of the traditional metric, due to the fact that the error value (6) with which the accuracy of the system was calculated has been adapted to include the importance of the knowledge of the users. The total number of experiments carried out is 144 (4 k-neighborhoods $*$ 6 levels of $\alpha * 2$ metrics $*$ 3 types of results); the experiments have been grouped in such a way that the following can be determined: accuracy, number of perfect predictions and number of bad predictions.

We consider a perfect prediction to be each situation in which the prediction of the number of stars recommended for one user in one film matches the value rated by that user for that film. We consider a bad prediction to be each situation in which the prediction of the number of stars recommended for one user in one film is different by more than 2 stars from the value rated by that user for that film.

5. Results

The results of the experiments are presented in Fig. 2. This figure shows the evolution of the cosine metric related to the proposed metric in the different values of the k-neighborhoods studied and in the different values of α (from 0.3 to 0.8).

Into the figure, the MAE, percentage of perfect predictions and percentage of bad predictions are presented. The two latter measurements give an idea of the variation of the MAE and give some very important information about the virtues and defects of the results. The percentages are calculated based on the total number of predictions obtained in each case.

In Fig. 2a it can be seen how the MAE of the cosine metric is greater than the MAE of the proposed metric in almost all the cases, since the subtraction of both produces negative values on the y axis of the graph; therefore we know that the accuracy levels of the RS are better when the new metric is applied, and this is due to the favorable weighting of the knowledge of the users. We can also see how the proposed metric offers its best results (as opposed to the cosine metric) when the values of α increase and those of k decrease; this same circumstance is repeated in Fig. 2c, where the percentage of incorrect predictions is particularly minimal having used the new metric (when compared to the cosine metric) with high values of α and low values of k (15 and 30).

In Fig. 2b an improvement in the percentage of the correct predictions can be seen when the proposed metric is used with low k values and high α values. The cosine metric obtains better results of correct predictions only when k is high and α is low simultaneously.

6. Conclusions

The recommender systems of e-learning allow the possibility of weighting the importance of the recommendations that each user generates, depending on their level of knowledge.

In order to include the knowledge level of the users in the collaborative filtering step, it is necessary to design new metrics which, being based on the current ones, incorporate the additional information with regards to the scores obtained by each user.

The validation of the new metrics requires a modification of the traditional equations used in order to measure the total error of the system when these metrics are used; the mean absolute error, mean squared error, or whatever other measurement of the total accuracy of the system must also reflect the knowledge level of the users.

The new metric proposed in the paper has obtained better results than the traditional equivalent when both have been subjected to the processing of the total accuracy of the system using a MAE measure adapted to include the knowledge of the users. The remaining indicators studied (percentage of correct predictions and percentage of particularly erroneous predictions) has also produced better results using the proposed metric.

Although the experiments have not been carried out with an e-learning database, both the equations designed and the methodology used could be used in the same way in different recommender systems of e-learning. The solidity of the database used, the large number of experiments carried out and the quality of the results obtained permit us to face developments in the distinct recommender systems in the sense of differentiating the users by some characteristic, such as by knowledge in the sphere of collaborative e-learning.

A well defined field of research exists in collaborative filtering when the nature of the recommender systems allows the incorporation of weighting in the importance of each one of the users, and the collaborative systems of e-learning are named to lead the developments in this new field of investigation.

Acknowledgement

Our acknowledgements to the GroupLens Research Group.

References

[1] M. Knights, Web 2.0, IET Communications Engineer, February–March 2007, pp. 30–35.
[2] K.J. Lin, Building Web 2.0, Computer, May 2007, pp. 101–102.
[3] T. Janner, C. Schroth, Web 2.0 and SOA: converging concepts enabling the internet of services, IT Pro, May–June 2007, pp. 36–41.
[4] H. Jinghua, W. Kangning, F. Shaohong, A survey of e-commerce recommender systems, in: Proceedings of the International Conference on Service Systems and Service Management, 2007, pp. 1–5, doi:10.1109/ICSSSM.2007.4280214.
[5] R. Baraglia, F. Silvestri, An online recommender system for large web sites, in: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, 2004, pp. 199–205, doi:10.1109/WI.2004.10158.
[6] D.R. Fesenmaier, U. Gretzel, C. Knoblock, C. Paris, F. Ricci, S. Stabb, H. Werther, A. Zipf, Intelligent systems for tourism, Intelligent Systems 17 (6) (2002) 53–66.
[7] J.A. Konstan, B.N. Miller, J. Riedl, PocketLens: toward a personal recommender system, ACM Transactions on Information Systems 22 (3) (2004) 437–476.
[8] N. Antonopoulus, J. Salter, Cinema screen recommender agent: combining collaborative and content-based filtering, IEEE Intelligent Systems 2006 35–41.
[9] P. Li, S. Yamada, A movie recommender system based on inductive learning, in: Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems, vol. 1, December 2004, pp. 318–323.
[10] H. Denis, Managing collaborative learning processes in e-learning applications, in: 29th International Conference on Information Technology Interfaces, June 2007, pp. 345–350.
[11] H.M. Huang, S.S. Liaw, Investigating learner attitudes toward a collaborative e-learning system, in: 10th International Conference on Computer Supported Cooperative Work in Design, May 2006, pp. 1–5.
[12] G. Costagliola, F. Ferrucci, G. Polese, G. Scanniello, An approach for the creation of collaborative environments, in: 2nd International Conference on Information Technology: Research and Education, June 2004, pp. 184–188.
[13] Ch. Gang, Ch. Zeyu, S. Ruimin, W. Jiajun, collaborative education model and its application in e-learning, in: Sixth IEEE/ACIS International Conference on Computer and Information Science, July 2007, pp. 856–860.
[14] C. Bouras, G. Hornig, V. Triantafillou, T. Tsiatsos, Architectures supporting e-learning through collaborative virtual environments: the case of INVITE, in: IEEE International Conference on Advanced Learning Technologies, August 2001, pp. 13–16.
[15] S.R. Balasundaram, B. Ramadoss, Object oriented analysis learning tool using collaborative learning, in: Seventh International Conference on Information Technology Based Higher Education and Training, July 2006, pp. 811–816.

[16] T. Okamoto, E-collaborative learning technologies and e-pedagogy, in: Third IEEE International Conference on Advanced Learning Technologies, July 2003, pp. 516–517.
[17] H.J. Wu, S.Ch. Huang, A dynamic e-learning system for the collaborative business environment, in: Seventh IEEE International Conference on Advanced Learning Technologies, July 2007, pp. 46–50.
[18] S. Aguirre, A. Fumero, J. Quemada, J. Salvachua, A. Tapiador, Joint degrees in e-learning systems: a web services approach, in: International Conference on Collaborative Computing: Networking, Applications and Worksharing, November 2006, pp. 1–5.
[19] GuiLing Wang, YuShun Li, ShengWen Yang, ChunYu Miao, Jun Xu, MeiLin Shi, Service-oriented grid architecture and middleware technologies for collaborative e-learning, in: IEEE International Conference on Services Computing, vol. 2, July 2005, pp. 67–74.
[20] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, IEEE Transactions on Knowledge and Data Engineering 17 (6) (2005) 734–749.
[21] J.L. Herlocker, J.A. Konstan, J.T. Riedl, L.G. Terveen, Evaluating collaborative filtering recommender systems, ACM Transactions on Information Systems 22 (1) (2004) 5–53.
[22] J.S. Breess, D. Heckerman, C. Kadie, Empirical analysis of predictive algorithms for collaborative filtering, in: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, 1998, pp. 43–52.
[23] F. Kong, X. Sun, S. Ye, A comparison of several algorithms for collaborative filtering in startup stage, in: Proceedings of the IEEE Networking, Sensing and Control, March 2005, pp. 25–28.
[24] S.B. Cho, J.H. Hong, M.H. Park, Location-based recommendation system using bayesian user's preference model in mobile devices, Lecture Notes on Computer Science 4611 (2007) 1130–1139.
[25] H. Ingoo, J.O. Kyong, H.R. Tae, The collaborative filtering recommendation based on SOM cluster-indexing CBR, Expert Systems with Applications 25 (2003) 413–423.
[26] R.R. Yager, Fuzzy logic methods in recommender systems, Fuzzy Sets and Systems 136 (2) (2003) 133–149.
[27] G. Linden, B. Smith, J. York, Amazon.com recommendations, IEEE Internet Computing 7 (2003) 76–80.
[28] G.M. Giaglis, G. Lekakos, Improving the prediction accuracy of recommendation algorithms: approaches anchored on human factors, Interacting with Computers 18 (3) (2006) 410–431.
[29] Y. Li, R. Nayak, L.T. Weng, Y. Xu, An improvement to collaborative filtering for recommender systems, in: Proceedings of the IEEE International Conference on Computational Intelligence for Modelling, Control and Automatitation, 2005.
[30] I. Fuyuki, T.K. Quan, H. Shinichi, Improving accuracy of recommender systems by clustering items based on stability of user similarity, in: Proceedings of the IEEE International Conference on Intelligent Agents, Web Technologies and Internet Commerce, 2006.
[31] Y. Manolopoulus, A. Nanopoulus, A.N. Papadopoulus, P. Symeonidis, Collaborative recommender systems: combining effectiveness and efficiency, Expert Systems with Applications 34 (2007), doi:10.1016/j.eswa.2007.05.013.
[32] http://www.movielens.org.