HYBRID RECOMMENDER SYSTEM: COLLABORATIVE FILTERING AND DEMOGRAPHIC INFORMATION FOR SPARSITY PROBLEM

Kouadria ABDERRAHMANE[1], Mohammad Yahya H. AL-SHAMRI[2], Nouali OMAR[3]

Recommender systems (RSs) have attracted the attention of many researchers by their applications in various interdisciplinary fields. They help users to overcome the information overload problem and help businessmen to make more profits. Many recommendation algorithms have been discussed and proposed in the literature. Nevertheless, the performance of every single algorithm is limited, and each has its strengths and weaknesses. This paper tries to exploit the power of four RSs, namely, user demographic RS, item demographic RS, user-based and item-based collaborative RSs. Actually, a hybrid recommendation system consisting of two-level hybridization between four individual systems that vary according to the filtering method and the target is proposed. This allows each system to express itself before obtaining the result of the final prediction. Hence, the prediction value will reflect the individual system point of view and consequently different points of view will be available to aggregate the result. For this regard, many schemes for weighing the importance of the prediction result of each system are used. The proposed framework reduces the problem of sparsity and improves the accuracy. Experimental results confirm this finding and show that the proposed framework reduces error and improves coverage compared to traditional ones.

Keywords: Collaborative filtering; demographic filtering; hybrid recommender system; weighting schemes

1. Introduction

Web and its enormous amount of information make it increasingly difficult to find the necessary and useful information quickly and efficiently. Recently, RSs have become the key tool for helping users to filter out this information based on the user's preferences. Today, RSs help a large number of people to find their preferred friends, interesting music, favorite videos, books, CDs, clothing and so on. Some applications for that are Amazon, eBay, Taobao, Facebook, Ringo and YouTube [1]. RSs can be Content-based RS (CBRS),

[1] Ecole nationale Supérieure d'Informatique (ESI), BP 68M, 16309, Oued-Smar, algiers, and Ibn Khaldoun University Tiaret, Tiaret, Algeria, e-mail: Ab_kouadria@esi.dz

[2] Computer Engineering Department, King Khalid University, Abha, Saudi Arabia; Faculty of Engineering and Architecture, Ibb University, Ibb, Yemen, e-mail: mohamad.alshamri@gmail.com

[3] Research Centre in Scientific and Technical Information (CERIST), Algeria, e-mail: onouali@cerist.dz

Collaborative RS (CRS), Demographic RS (DRS), or Hybrid RS (HRS) [2]. The user profile of CBRS reflects the user needs in terms of the content specifications of the items the user has seen before. However, many items cannot be represented by a set of keywords that should be selected from the item contents.

CRS considers the interaction between users. It is the most common one because of its simplicity to obtain a set of ratings implicitly or explicitly from users. CRS can deal with a large number of users and items and has the ability to filter items whose content is not easily analyzed by automated processes. Moreover, CRS can support social factors by taking into account the interests of like-minded users. This is used by Amazon library "users who liked this book also liked another book". Despite the overall success of CRS, it suffers from the problems of missing data (sparsity) and cold start. Usually, each user evaluates only a very limited number of items compared to the total number that can be recommended. The relevance of the recommendations for a given user depends on the number of evaluations performed. If the user profile is empty or contains very few ratings, it will be difficult for CRS to associate this user to any group of users and hence recommendations will be impossible or unsatisfactory. CRS can be implemented as user-based [3,4] or item-based [5] according to the way of comparing the ratings. If the comparison is done between users in terms of their ratings, then it is a user-based CRS. Otherwise, if the comparison is done between the ratings given for an item and the other items, then it is an item-based CRS.

The demographic attributes of users or items are used by DRS for profiling users and hence producing recommendations. The user profile consists of a list of demographic data such as age, gender, occupation, and country [6]. DRS does not provide individualized recommendations, as users of the same type will have the same recommendations. For this reason, DRS are not regarded as personal RSs but as group referral systems. The main advantage of DRS is that it may not require a history of user ratings as CRS and CBRS need. Grundy system [7] is an example of DRS which recommends books based on the personal information collected through interactive dialogue. HRS exploits the complementarity of two or more individual RSs by combining them according to many ways [8].

The most well-known dataset for recommendation applications is the MovieLens dataset which has information about users rating some movies. The dataset gives many demographic data for its users like age, gender, occupation and zip code. The resulting user-item matrix is usually sparse which degrades the performance. For solving this, we propose a HRS consisting of two-level hybridization between four individual systems, namely user-based CRS (UCRS), item-based CRS (ICRS), user DRS (UDRS) and items DRS (IDRS) as shown in Fig. 1. These systems vary on either the filtering method or the target (users or items). Accordingly, we proposed one-level hybridization and two-level hybridization at the predication phase of the normal RS. This allows individual system to express itself before we get the final prediction result. By this way, the prediction value will express the individual system point of view and hence different views will be there for aggregating the result. Many schemes can be used for weighing the importance of the prediction result of each system.

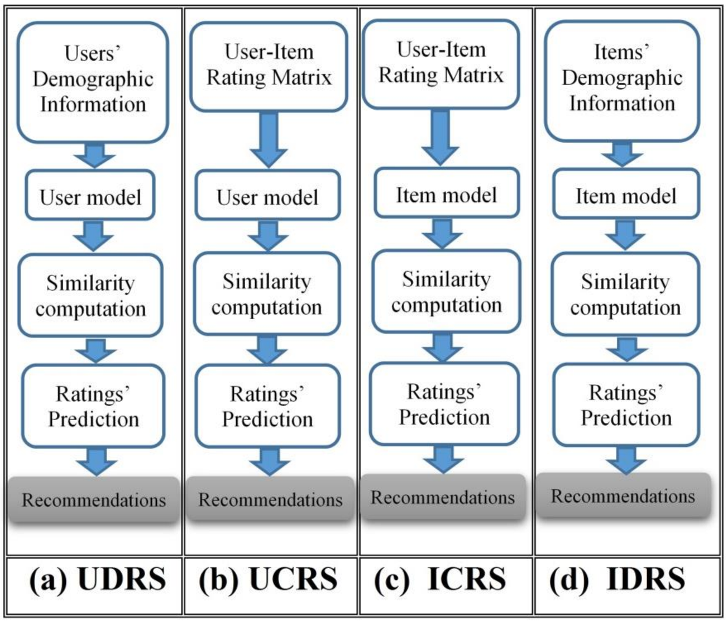

Fig. 1. Individual recommender systems

This paper is organized as follows; the second section outlines the related research while section 3 details the proposed hybrid system. Section 4 gives the experiments and their analysis while the last section concludes the paper and gives some future research directions.

2. Related work

Memory-based CRSs use all the information available to generate the predictions to share opinion between users and can be implemented based on users or items [9]. UCRS looks for users who share similar preference patterns (in terms of ratings for items) with the active user and recommend items that are rated high by those similar users. In contrast, ICRS considers the similarity between items instead of users. Some researchers use the hybridization between UCRS and ICRS to improve the predictive accuracy of the system. For example, the authors of [10] proposed hybrid predictive UCRS and ICRS with smoothing sparse data HSPA algorithm. In fact, this framework smoothed data using ICRS then predicts the model based on both users' aspects and items' aspects. Wang and Yin [11] proposed a synthetically CF model by combining popular items of both ICRS and UCRS to suggest a more diverse recommendation list of items for the target user.

On the other hand, DRSs rely on demographic information which is different from ratings and hence can be used to alleviate CRS problems [12-14]. In this direction, Al-Shamri [15] discussed many profiling approaches for DRSs. Two major weaknesses of CRSs are the cold start problem and the sparsity. Chikhaoui et al. [12] proposed an approach that combines CRS, CBRS and DRS to solve the cold start problem. Jyoti and Jayant [16] proposed a weighted scheme between the predictions of ICRS and DRS. Saurabh and Shailendra [17] combined nearest neighbor ratings prediction with rating computed from UDRS and IDRS. The new user cold start problem is resolved by generating immediate ratings based on user demographics while item cold start problem is addressed by using items' cluster. Xie et al. [18] proposed adaptive weighted prediction by employing final ratings from UCRS and ICRS. Fan et al [19] proposed hybrid similarity measures based on the user ratings and user attributes. He and Jin [20] proposed a similarity measure that considers both co-rated items and the users' preferences for item attributes.

3. Two-level hybrid recommender system

HRS consisting of four parallel systems with two cascaded steps is proposed here to alleviate data sparsity and cold start problem. The following subsections will discuss individual systems in details.

3.1. User-based collaborative recommender system

Usually, UCRS uses Pearson correlation coefficient (PCC) to compare users based on their rating history (co-rated items). However, the number of co-rated items can have an influence when items receive fewer ratings [2]. For example, a similarity value of 0.91 between two users having only 7 co-ratings is not as reliable as a similarity value of 0.81 obtained from 115 co-rated items. As a solution for this issue, Jaccard similarity [21,22] is used to adjust PCC.

$$sim_{UCRS}(a,u) = \frac{U_a \cap U_u}{U_a \cup U_u} * \frac{\sum_{i \in U_{au}} (r_{a,i} - m_a)(r_{u,i} - m_u)}{\sqrt{\sum_{i \in U_a} (r_{a,i} - m_a)^2} \sqrt{\sum_{i \in U_u} (r_{u,i} - m_u)^2}} \qquad (1)$$

The Jaccard similarity can be considered as a scaling factor for the PCC similarity value. The set of symbols used in this paper are listed below:

$r_{a,i}$	Rating of user *a* on item *i*.
U_a	Set of items rated by user *a*.
U_{au}	$U_a \cap U_u$: Set of items rated by both users *a* and *u*.
I_i	Set of users rating item *i*.
I_{ij}	$I_i \cap I_j$: Set of users rating both items *i* and *j*.
m_a	Average rating of user *a*.
N	Neighborhood set size for a user (most similar users to that user)
$\bar{R}_i$	Average rating of item *i*.

K	Neighborhood set size for an item (most similar items to that item)
$pr_{u,i}$	Predicted rating value of item *i* for user *u*.
$\lvert TestSet \rvert$	Number of all pairs (u,i) in the test set.
N_i^P	Total number of predicted items for user *u*
M_A	Total number of the active users.

After calculating the similarity values, RS finds estimated values for all items not seen by the active user. For recommendations, the top-N most similar items are displayed to the user. The most popular prediction formula for UCRS and UDRS is given below [9], where *x*=C for UCRS and *x*=D for UDRS.

$$pr_{UxRS}(a,i) = m_a + \frac{\sum_{u=1}^{N} sim_{UxRS}(u,i) \times (r_{u,i} - m_u)}{\sum_{k=1}^{K}(\lvert sim_{UxRS}(u,i)\rvert)} \tag{2}$$

3.2. Item-based collaborative recommender system

This system uses PCC and Jaccard similarity also for calculating the similarity but between items instead of users.

$$sim_{ICRS}(i,j) = \frac{I_i \cap I_j}{I_i \cup I_j} * \frac{\sum_{u \in I_{ij}}(r_{u,i} - \bar{R}_i)(r_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in I_i}(r_{u,i} - \bar{R}_i)^2}\sqrt{\sum_{u \in I_j}(r_{u,j} - \bar{R}_j)^2}} \tag{3}$$

The predicted value for both ICRS and IDRS is calculated using the following formula [5], where *x*=C for ICRS and *x*=D for IDRS.

$$pr_{IxRS}(u,i) = \frac{\sum_{k=1}^{K} sim_{IxRS}(i,k) \times r_{u,k}}{\sum_{k=1}^{K}(\lvert sim_{IxRS}(i,k)\rvert)} \tag{4}$$

3.3. User-based demographic recommender system

UDRS represents each user by a demographic binary vector of size 27 attribute (Table 1) covering three categories of information (age-4, gender-2 and occupation-21) [6].

Table 1

Structure of User Demographic Vector [6]

	Feature	**Feature contents**	**Comments**
Age (categorical)	1	age ≤18	True slot takes value of 1 while the rest remain 0 (False)
	2	18 <age ≤ 29	
	3	29 <age ≤ 49	
	4	age >49	
Gender (categorical)	5	Male	True gender is 1 while the other slot is 0
	6	Female	
Occupation (categorical)	7-27	7 | … | 27	True occupation is 1 while the rest remain 0

A binary 1 indicates that the user has that demographic information. To encode the gender, we used 1 for male and 0 for female while age attribute is quantized into 4 groups (non-adults (age ≤18), young adults (18 <age ≤ 29), prime-age adults (29 <age ≤ 49) and mature adults (age >49)). Every user belongs to a single age group which takes value 1.

With regard to occupation, a single slot describes the user occupation and it is 1 for that occupation. The demographic correlation between the active user and each member of its neighborhood is calculated using cosine vector similarity measure [9].

$$sim_{UDRS}(a,u) = \frac{\sum_{i\in U_{au}}(r_{a,i})(r_{u,i})}{\sqrt{\sum_{i\in U_a}(r_{a,i})^2}\sqrt{\sum_{i\in U_u}(r_{u,i})^2}} \tag{5}$$

3.4. Item-based demographic recommender system

A demographic vector for each item is constructed using the demographic information available for movies. There are 18 distinct movie genres, ranging from Children's to Horror. An additional slot called "unknown" is utilized for films that cannot be categorized under any one of the existing genres; hence we will have a vector of 19 features [6]. It is important to point out that a film can belong to more than one genre at the same time. In that case, the slots which correspond to each of these categories should take a value of 1 (True), while the rest will stay fixed at 0 (False). Based on the items' demographic vectors, we calculate the items demographic similarity by adopting Cosine similarity [9].

$$sim_{IDRS}(i,j) = \frac{\sum_{u\in I_{ij}}(r_{u,i})(r_{u,j})}{\sqrt{\sum_{u\in I_i}(r_{u,i})^2}\sqrt{\sum_{u\in I_j}(r_{u,j})^2}} \tag{6}$$

3.5. Hybrid demographic/collaborative recommender system

DRS and CRS require no content description and they rely only on a set of neighbors for their recommendations. Many approaches try to hybridize DRS and CRS to solve the cold start problem or to reduce sparsity. However, they try only a limited number of the available alternatives and at the early stages of the recommendation process. For example, Al-Shamri and Bharadwaj [23] built a hybrid user model from demographic and rating information of the user. This paper uses the hybridization at the prediction level between the four individual systems. This direction keeps each system as it is until the prediction stage where we have to merge prediction values based on some criteria. This allows each individual system to give a prediction value or generates a zero value and hence even for zero value from some systems, we still have the opportunity to find a prediction value based on the opinion of the other parallel systems. Moreover, we can weight each prediction value according to its importance for the user. Three weighting schemes are used in this paper, namely, *Fixed weighting*, *Weighted average*, and *Gradient-descent weighting*.

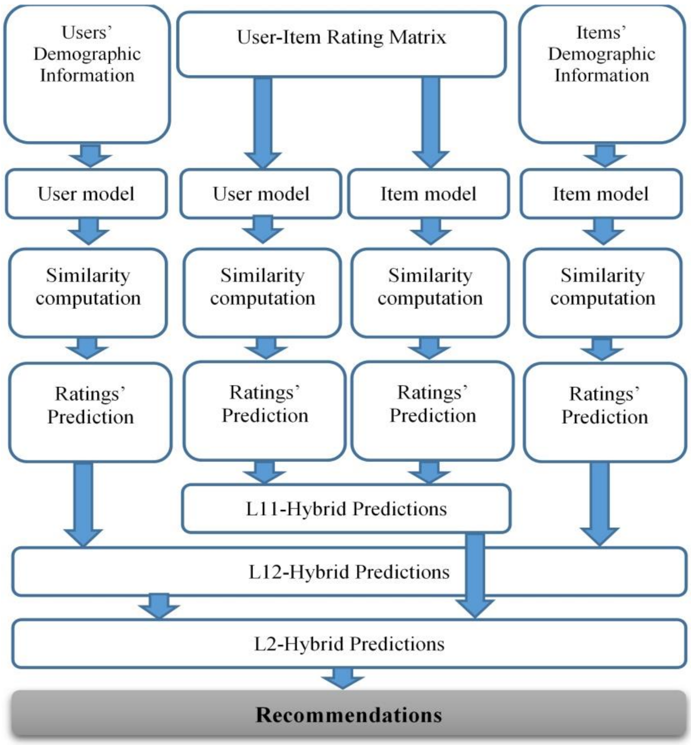

Fig. 2. The proposed hybrid recommender system

Fig. 2 gives a block diagram for the proposed hybrid system in which UDRS and IDRS are grouped together and called L12. Similarly, UCRS and ICRS are grouped together and called L11. The final result is obtained from the results of these homogenous groups and called L2. This forms a two-level hybrid system which will reduce the effect of sparsity and alleviate cold start problem and also generate more accurate recommendations. The first level (L11) prediction values of CRS and DRS will be computed as follows:

$$pr_{HxF}(u,i) = \begin{cases} 0 & pr_{UxRS}(u,i) = pr_{IxRS}(u,i) = 0 \\ pr_{UxRS}(u,i) & pr_{IxRS}(u,i) = 0 \\ pr_{IxRS}(u,i) & pr_{UxRS}(u,i) = 0 \\ w_{UxRS} * pr_{UxRS}(u,i) + w_{IxRS} * pr_{IxRS}(u,i) & \text{Otherwise} \end{cases} \tag{7}$$

where *x*=C for CRS and *x*=D for DRS. The weighting values w_{UxRS} and w_{IxRS} can be fixed or variable based on a weighting scheme as below for the weighted average scheme.

$$w_{UxRS} = \frac{pr_{UxRS}(u,i)}{pr_{UxRS}(u,i)+pr_{IxRS}(u,i)} \quad (8a)$$

$$w_{IxRS} = \frac{pr_{IxRS}(u,i)}{pr_{UxRS}(u,i)+pr_{IxRS}(u,i)} \quad (8b)$$

Moreover, the weights can be calculated using a fast heuristic gradient descent scheme which is adopted by Koren [24] and Gedikli et al. [25]. To obtain the final prediction value, we combine the level one prediction values using [26]:

$$pr_{HRS}(u,i) = w_{HCF} * pr_{HCF}(u,i) + w_{HDF} * pr_{HDF}(u,i) \quad (9)$$

where w_{HCF} and w_{HDF} are the weights to control the influence of each hybrid system. The aim is to minimize the prediction error for each single user and item.

4. Experiments and analysis

As discussed before, we conduct our experiments using the well-known 100K MovieLens dataset that has five sets called, user-1 to user-5 datasets. Each one is divided into 80% for training and 20% for testing. We calculated similarity values for all individual systems, UCRS ICRS, UDRS and IDRS in the offline phase then we apply KNN algorithm to select the top-N most similar items and users for the target item or user to produce individual predictions. Afterwards, we combine predictions to produce level one and level two predictions. The weights are determined using many schemes as discussed before.

4.1. Evaluation measures

This paper evaluates the system performance using coverage and two error measures, namely MAE (Absolute Mean Error) and RMSE (Root Mean Square Error) [21,27]. MAE is defined as the average absolute difference between the ground truth and the predicted rating value.

$$MAE = \frac{\sum_{(u,i)\in TestSet}|r_{u,i}-pr_{u,i}|}{|TestSet|} \quad (10)$$

The RMSE is given by [12]:

$$RMSE = \sqrt{\frac{\sum_{(u,i)\in TestSet}(r_{u,i}-pr_{u,i})^2}{|TestSet|}} \quad (11)$$

The coverage measures the percentage of items that the system can predict them. This measure over all the active users is [15]:

$$\text{Coverage} = \frac{\sum_{i=1}^{M_A} N_i^P}{|TestSet|} \quad (12)$$

4.2. Experiments analysis

The results of the individual and hybrid systems are listed in Tables 2 and 3, respectively. The fixed weighting used 0.6 for user related systems and 0.4 for items related systems. In the second level we used 0.6 for CRSs and 0.4 for demographic systems. The results of individual systems give advantage for collaborative filtering approaches. However, UCRS is the best among all in terms of MAE and RMSE. ICRS has the best performance in terms of coverage.

Table 2

MAE, RMSE and Coverage of individual systems

Metric	UDRS	IDRS	UCRS	ICRS
MAE	0.8269	0.9301	0.7552	0.7587
RMSE	1.0623	1.1562	0.9658	0.9721
Coverage	0.9235	0.9260	0.9685	0.9729

Table 3

MAE, RMSE and coverage of the hybrid system

		Gradient-descent	Fixed-weighting	Weighted average
L11	MAE	0.7229	0.7260	0.7247
	RMSE	0.9253	0.9297	0.9290
	Coverage	0.9859	0.9859	0.9859
L12	MAE	0.8004	0.8013	0.8007
	RMSE	1.0177	1.0209	1.0184
	Coverage	0.9938	0.9938	0.9938
L2	MAE	0.7418	0.7436	0.7423
	RMSE	0.9502	0.9515	0.9509
	Coverage	0.9971	0.9971	0.9971

We can see clearly that hybrid systems are much better than individual ones in all aspects. In terms of weighting schemes, the results of Table 3 are very close to each other with some advantage to gradient descent weighting. Again the results of the CRSs are better than that of DRSs in terms of error. However, coverage of DRSs is better than that of collaborative ones. This is logical as the CRSs usually suffer from the cold start problem and rely on the common ratings between users or items which may be not available for some items. Hence, hybridization between CRSs and DRSs will come up with a solution for those items. In terms of weighting, again gradient-descent is better than the others with very close results for the favor of second level hybridization.

Individually, ICRS has the best coverage value because it compares items in terms of users and the number of items is much more than the number of users. The DRSs at both user and item levels show low coverage values compared to collaborative ones. This is logical as we have different descriptions for the items and different demographic information for users. However, these results changed with L12 hybridization between DRSs of items and users where the coverage values of L21 jumped to 99.34% and becomes better than that of L12. The improvement percentage between individual DRSs and L12 is 7.32%. The results show that L11 coverage values are less than that of L12 because of usual sparsity problem of CRSs. Fig. 3 illustrates the coverage values of L11, L12 and L2 hybridization with the weighting schemes.

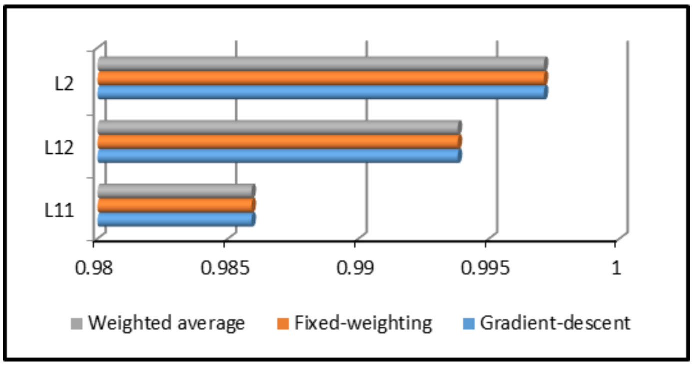

Fig. 3. Coverage values of the hybridization at different levels with weighting schemes

All weighting schemes have the same coverage value but differ in the error values. The results of L2 hybridization show that the coverage is improved while the errors become little more compared to the results of L11 and L12 hybridization. This is an expected result because more items now are involved in the error calculations compared to the previous case.

5. Conclusions

The proposed hybrid recommender addresses the issue of sparsity and cold start problems by utilizing the users' and the items' ratings and demographic information. Actually, the hybridization is done between two different systems with two different types of information. The first level of the hybridization goes through two branches; the first one hybridizes CRSs of user-based and item-based systems. This hybridization exploits the comparison between users in terms of their common ratings and between items in terms of their common users. The second branch hybridizes the DRSs of users and items which have two different types of underlying demographic information. This reduces the effect of the sparsity and alleviates the cold start problem. The second level hybridizes the two resulting systems coming from the two branches and hence does the hybridization at the system level.

The results show that the proposed hybrid system outperforms the individual systems and improves the coverage by covering more items. In terms of weighting schemes, the gradient-decent weighting performs better than the other weighting schemes with very close results. This issue can be explored more using other heuristic approaches to guide the weighting process to the different choices of the users.

R E F E R E N C E S

[1]. *F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors,* Recommender Systems Handbook, Springer, 2011.

[2]. *J. Bobadilla, F. Ortega, A. Hernando and A. Gutiérrez,* "Recommender systems survey," in Knowledge-Based Systems, vol. 46, 2013, pp. 109–132.

[3]. *J. Herlocker, J. Konstan, A. Borchers, and J. Riedl,* "An algorithmic framework for performing collaborative filtering", in Proceedings of ACM SIGIR conference on Research and development in information retrieval, 1999, pp. 230-237.

[4]. *W. Qin, L. Xin and H. Liang,* "Unifying user-based and item-based algorithm to improve collaborative filtering accuracy," in Energy Procedia, vol. 13, 2011, pp. 8231-8239.

[5]. *G. Linden, B. Smith and J. York,* "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," in IEEE Internet Computing, vol. 7, no. 1, 2003, pp. 76-80.

[6]. *M.G. Vozalis, and K.G. Margaritis,* "Collaborative filtering enhanced by demographic correlation," in Proceedings of AIAI Symposium on Professional Practice in AI, 2004.

[7]. *E. Rich*, "User Modeling via Stereotypes," Cognitive Science, vol. 3, 1979, pp. 329-354.

[8]. 7*E. Çano and M. Morisio*, ''Hybrid recommender systems: A systematic literature review,'' in Intelligent Data Analysis, vol. 21, no. 6, 2017, pp. 1487–1524.

[9]. *J. S. Breese, D. Heckerman, and C. Kadie,* "Empirical analysis of predictive algorithms for collaborative filtering", in Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98), July 1998, pp 43–52.

[10]. *R. Hu, Y. Lu,* "A hybrid user and item-based collaborative filtering with smoothing on sparse data," in Proceedings of IEEE Artificial Reality and Telexistence-Workshops, ICAT'06, November 2006, pp. 184-189.

[11]. *J. Wang, J. Yin, "*Combining user-based and item-based collaborative filtering techniques to improve recommendation diversity," in Proceedings of Biomedical Engineering and Informatics (BMEI), December 2013, pp. 661–665.

[12]. *B. Chikhaoui, M. Chiazzaro and S. Wang, "*An improved hybrid recommender system by combining predictions*,"* in Proceedings of IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA), 2011, pp. 644-649.

[13]. *S. Laila and S. Akram,* "Exploiting user demographic attributes for solving cold-start problem in recommender system", in Lecture Notes on Software Engineering, vol. 1, no. 3, 2013, pp. 303-307.

[14]. *W. Zhao, S. Li, Y. He, L. Wang, J.-R. Wen, and X. Li* "Exploring demographic information in social media for product recommendation," in Knowledge Information Systems, vol. 49, no. 1, 2015, pp. 61–89.

[15]. *M. H. Y. Al-Shamri,* "User profiling approaches for demographic recommender systems", in Knowledge-Based Systems, vol. 100, 2016, pp. 175–187.

[16]. *G. Jyoti and G. Jayant,* "Performance Analysis of Recommendation System Based On Collaborative Filtering and Demographics," in Proceedings of Communication, Information & Computing Technology (ICCICT), January 2015, pp. 1-6.

[17]. *K. T. Saurabh, and K. S. Shailendra,* "An Approach for Recommender System by Combining Collaborative Filtering with User Demographics and Items Genres," in International Journal of Computer Applications, vol. 128, no.13, October 2015, pp. 16-24.

[18]. *F. Xie, M. Xu, and Z. Chen,* “RBRA: A simple and efficient rating-based recommender algorithm to cope with sparsity in recommender systems,” in Proceedings of 26th IEEE International Conference on Advanced Information Networking and Applications Workshops, March 2012, pp. 306-311.

[19]. *X. Fan, Z. Chen, L. Zhu, Z. Liao, and B. Fu*, “A novel hybrid similarity calculation model,” Scientific Programming, vol. 2017, Article ID 4379141, 9 pages, 2017.

[20]. *X. He and X. Jin*, “Collaborative filtering recommendation algorithm considering users’ preferences for item attributes”, in the proceedings of IEEE International Conference on Big Data and Computational Intelligence (ICBDCI), Feb. 2019.

[21]. *L. Candillier, F. Meyer, and F. Fessant,* “Designing specific weighted similarity measures to improve collaborative filtering systems,” in: Perner P. (eds) Advances in Data Mining. Medical Applications, E-Commerce, Marketing, and Theoretical Aspects, vol. 5077, 2008, pp. 242-255.

[22]. *J. Feng, X. Feng, N. Zhang, J. Peng*, “An improved collaborative filtering method based on similarity” in PLoS ONE, vol. 13, no. 9, 2018, pp. 1-18.

[23]. *M. Y. H. Al-Shamri, and K.K. Bharadwaj,* “A Compact User Model for Hybrid Movie Recommender System,” in Proceedings of 7th IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA’07), 2007, pp. 519-524.

[24]. *Y. Koren,* “Factor in the neighbors: Scalable and accurate collaborative filtering,” in ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 4, no. 1, 2010, pp. 1-24.

[25]. *F. Gedikli, F. Bagdat, M. Ge and D. Jannach*, “RF-REC: Fast and accurate computation of recommendations based on rating frequencies,” in Proceedings of IEEE 13th Conference on Commerce and Enterprise Computing (CEC), September 2011, pp. 50-57.

[26]. *D. Jannach, Z. Karakaya, and F. Gedikli,* “Accuracy improvements for multi-criteria recommender systems”, in Proceedings of the 13th ACM conference on electronic commerce, June 2012, pp. 674-689.

[27]. *J. Herlocker, J. Konstan, L. Terveen, J. Riedl,* “Evaluating collaborative filtering recommender systems,” in ACM Transactions on Information Systems, vol. 22, no. 1, 2004, pp. 5–53.