Similarity Measures for Collaborative Filtering to Alleviate the New User Cold Start Problem

Hemlata Katpara, Prof. V.B. Vaghela

Computer Engineering Department, L. D. College of Engineering, Ahmedabad, Gujarat

***Abstract*--Collaborative filtering is one of the most useful methods of product recommendation to users of online store. The most critical component of this method is finding similarities among users using user-item rating matrix so that the product can be recommended to the user based on similarities. The varieties of measures used for finding similarities are Cosine, Pearson correlation coefficient, mean squared difference, etc. An important issue of recommendation system that is viewed by researchers is the new user cold start problem, which occurs when a new user is coming to the system and no rating is available in user-item rating matrix. The mentioned measures are not suitable for new user cold start problem. In this paper, we discussed about various measures like PIP, PSS and NHSM for the new user cold start problem and we also show that how the NHSM measure is improving the recommendation performance under cold start condition. We also mentioned the advantage and disadvantage of each of these measures. The possible experiments can be done on three different datasets and the result show the superiority of NHSM similarity measure in recommendation system.**

***Keywords:* Collaborative filtering, Similarity measures, Cold-start, Recommendation system**

I. INTRODUCTION

The Internet, with hundreds of millions of pages worldwide has become the greatest source of information. In this context, information retrieval tools are helping the users in information seeking. Particularly, the users demand personalized search systems that will provide more adequate information for their particular taste and interest [1].

This is the aim of recommendation systems. They use information about users, user profiles, to predict the utility of a particular item. Thus providing personalized recommendation. Recommender systems have proven to be useful in the system like e-commerce, and they surely have a promising future in many other domains like Web search engines, digital TV program recommenders, etc. The recommender system have been used particularly in two task [1].

First, they have been used to predict the utility of a given item. According to this context, the user first selects the item in which he or she is interested. This is usually done after performing a search, browsing an online catalog. Then the recommender system predicts the rating of an item that the user would give to that item [1].

Second, the recommender system has been used to recommend the list of items to the user. This is often called find good items task. In this case, the system chooses the items that are most relevant according to user's search [1].

Many recommender system algorithms have been developed in various application such as e-commerce, digital library and online advertising [2,3,4]. These algorithms can be categorized into two different classes. Content based filtering, collaborative filtering and combination of these can also be found in the literature. In content based filtering, In content-based filtering [4], an item is recommended to an active user by analyzing profile of the user, profile of the item and profiles of items she preferred in past. For example, keywords of purchased books of a user might be used to find other books that contain the same or similar keyboards for recommendation [5]. The disadvantage of content-based filtering is its inability to evaluate the quality of an item. For example, it cannot distinguish a good article from a bad one if both articles use similar words. In fact, the quality of an item is a highly subjective feature that depends on the tastes, ideas, culture, etc., of each person and that would be difficult for a machine to analyze. Finally, content-based filtering does not have a way of finding serendipitous items that are interesting for the user, that is, really good items that are not apparently related to the user profile [1].

The collaborative filtering is the most successful and widely used recommendation system [1,6]. It makes recommendation according to the similar users with the active user or the similar items with the items which are rated by the active user. Collaborative filtering systems are less sensitive to the problems that are mentioned above since they are not based on the content of items but rather on the opinions of other users. The system will recommend items that have received high ratings by other users with similar taste or interests.

In collaborative filtering, the user profile is the set of rating given to different items. These rating can be captured explicitly that is by asking the user or implicitly by observing his/her interaction with the system. The ratings are stored in the form of user-item matrix representation [1].

The collaborative filtering includes memory-based method and model-based method [1,6]. The memory-based method first calculates the similarities among users and then selects the most similar users as the neighbors of the active user. Finally, it gives the recommendations according to the

neighbors. However, the model-based method first constructs a model to describe the behaviour of users and therefore to predict the ratings of items.

The memory-based method can give considerable recommended accuracy, but the computing time will grow rapidly with the increasing of users and items. The model-based method tends to be faster in prediction time than the memory-based method, because the construction of the model can be finished in a considerable amount of time and this process is executed off-line. The shortcoming of the model-based method is that the recommendation performance is not as good for the memory-based method. The memory based methods are more sensitive than the model based to some common problems of recommender system, among which we highlight the following [1].

A. *Sparsity of the rating matrix*

In most recommender systems, each user rates only a small subset of the available items, so most of the cells in the rating matrix are empty. In such cases, finding similarities among different users or items are challenging [1].

B. *Cold-start*

Related to the previous problem, this one deals with the difficulty in making recommendations for users recently introduced into the system. In such cases, the users have not rated enough items yet , so the recommender system is unable to guess their interests. Some systems overcome this problem by forcing the user first to rate a given set of items. However, these initial ratings could introduce biases into the system. Finally, note that the cold-start problem also affects new items as they will not be recommended until enough users have rated them [1].

C. *Shilling*

Recommender systems could suffer spam attacks, mainly from users interested in misleading the system to recommend a certain product. Several techniques affecting both neighbor-based and model-based algorithms have been studied in the past.

Some authors have combined techniques from both model based and memory based algorithms to take the advantage of best of both worlds. And that third type of recommender system is hybrid recommender system that combines content based methods with collaborative filtering [1].

This paper focuses on various measures that related to new user cold start problem and they are Proximity-Impact-Popularity (PIP), Proximity-Significance-Singularity (PSS) , NHSM (New heuristic similarity measure) . Here, we compare the user-user similarity for all these three measures for the given user-item rating matrix. We also show that the NHSM similarity measure is giving more accuracy in recommendation system. We also discuss the advantage and disadvantage of this NHSM similarity measure with other similarity measures for cold start problem mentioned above.

II. LITERTURE REVIEW

A. *Cold Start Problem*

An important issue for RSs that has greatly captured the attention of researchers is the cold-start problem. This problem has two variants: the new user cold-start problem and the new item cold-start problem. The new item cold-start problem occurs when there is a new item that has been transferred to the system. Because it is a new product, it has no user ratings (or the number of ratings is less than a threshold as defined in some equivalent papers) and is therefore ranked at the bottom of the recommended items list [7,8].

Moreover, this problem can be partially handled by staff members of the system providing prior ratings to the new item. Thus, the concentration of the cold-start problem is dedicated to the new user cold-start problem when no prior rating could be made due to the privacy and security of the system. It is difficult to give the prediction to a specific item for the new user cold-start problem because the basic filtering methods in RSs, such as collaborative filtering and content-based filtering, require the historic rating of this user to calculate the similarities for the determination of the neighborhood. For this reason, the new user cold-start problem can negatively affect the recommender performance due to the inability of the system to produce meaningful recommendations [7,8].

B. *Similarity measures for Cold-Start problem*

1) *Proximity-Impact-Popularity (PIP) measure*

The measure is composed of three factors of similarity, Proximity, Impact, and Popularity, and hence, named PIP. With the PIP measure, the similarity between two users u_i and u_j is calculated as :

$$SIM(u_i, u_j) = \sum_{k \in c_{i,j}} PIP(r_{ik}, r_{jk}) \quad (1)$$

Where r_{ik} and r_{jk} are the ratings of item k by user i and j respectively . $C_{i,j}$ is the set of co-rated items by users u_i and u_j and PIP(r_{ik} , r_{jk}) is the PIP score for the two ratings r_{ik} and r_{jk}. For any two ratings, r_1 and r_2,

$$PIP(r_1, r_2) = Proximity(r_1, r_2) \times Impact(r_1, r_2) \times Popularity(r_1, r_2) \quad (2)$$

Proximity factor is based on the simple arithmetic difference between two ratings [5,9]. It also consider that whether the two ratings are in agreement or not, giving penalty to ratings while in disagreement. Means that if two ratings are on the same side of given rating scale which is divided by its median then they are regarded to be in agreement. In the (a) of fig. 1 , the pair (a_1 , a_2) is given further penalty since the two ratings

are in disagreement, one preferring the item , the other disliking , compared with the (b_1 ,b_2) pair located on the same side .The penalty is given by doubling the distance between the ratings , which is then squared[5,9].

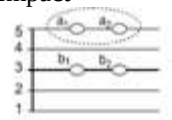
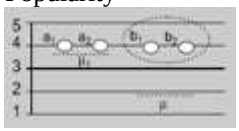
Proximity	Impact	Popularity
 Both pairs(a1,a2) and (b1,b2) have the same distance of 2. A1 has positive value and a2 has negative value . Penalty is given by distance between them. In the pair (b1,b2) , b1 has positive value and b2 is neutral and hence no penalty is given	 Both pairs (a1, a2) and (b1,b2) are at zero distance. The (a1, a2) pair is at stronger preference level then the pair (b1,b2). So that the pair (a1,a2) is having more impact.	 The pairs (a1,a2) and (b1,b2) are at zero distance and same impact factor. Consider μ_2 is the average rating for the commonly rated items. And in such a case similarity between two ratings are showing stronger similarity when the average rating for the commonly rated items is μ_1.

Fig. 1. Description of three factors of PIP measure

Second, the Impact factor considers how strongly an item is preferred or disliked by users. In (b) of fig.1, both pairs (a_1 , a_2) and (b_1 ,b_2) are showing zero distance in agreement and having same proximity factor . The ratings a_1 and a_2 are showing stronger preference than the other ratings [5,9].

Third, the Popularity factor gives the bigger value to a similarity for ratings that are further from the average rating of a co-rated item. In (c) of fig.1 , the two ratings b_1 and b_2 are far from average rating and showing stronger similarity between two users.

The below are the equations for solving PIP measure.

Agreement----

For any two ratings r_1 and r_2 , let R_{max} be the maximum rating and R_{min} the minimum in the rating scale and

$$R_{med} = \frac{R_{max} + R_{min}}{2}$$

A Boolean function $Agreement(r_1, r_2)$ is defined as follows:

$Agreement(r_1, r_2)$ = false $if\ r_1 > R_{med}\ and\ r_2 < R_{med})\ or\ (r_1 < R_{med}\ and\ r_2 > R_{med})$

$Agreement(r_1, r_2)$ = true otherwise

Proximity----

A simple absolute distance between the two ratings is defined

$D(r_1, r_2) = |r_1 - r_2|$, $if\ Agreement(r_1, r_2)$ is true

$D(r_1, r_2) = 2.|r_1 - r_2|$ $if\ Agreement(r_1, r_2)$ is false

Then the $Proximity\ (r_1, r_2)$ is defined as

$$Proximity(r_1, r_2) = \{\{2.(R_{max} - R_{min}) + 1\} - D(r_1, r_2)\}^2$$

Impact---

$Impact(r_1, r_2)$ is defined as :

$Impact(r_1, r_2) = (|r_1 - R_{med}| + 1)(|r_2 - R_{med}| + 1)$ $if\ Agreement(r_1, r_2)$ is true

$Impact(r_1, r_2) = \frac{1}{(|r_1 - R_{med}| + 1)(|r_2 - R_{med}| + 1)}$ $if\ Agreement(r_1, r_2)$ is false

Popularity----

Let μ_k is the average rating of item k by all users

$Popularity(r_1, r_2)$ is defined as :

$$Popularity(r_1, r_2) = 1 + \left(\frac{r_1 + r_2}{2} - \mu_k\right)^2$$

$if\ (r_1 > \mu_k\ and\ r_2 > \mu_k)\ or\ (r_1 < \mu_k\ and\ r_2 < \mu_k)$

$Popularity(r_1, r_2) = 1\ otherwise$

Now, if we calculate the user-user similarity for the given PIP measure and for the given user-item matrix, it is as follows. In the example of the user-item matrix, the missing ratings represented by the symbol - .

	Item1	Item2	Item3	Item4
User1	4	3	5	4
User2	5	3	-	-
User3	4	3	3	4
User4	2	1	-	-
User5	4	2	-	-

Table 1 User-item rating matrix

$$\begin{array}{c} \\ u1 \\ u2 \\ u3 \\ u4 \end{array} \begin{array}{c} \begin{array}{cccc} u2 & u3 & u4 & u5 \end{array} \\ \begin{pmatrix} 0.743 & 1.0 & 0.167 & 0.506 \\ & 0.743 & 0.162 & 0.763 \\ & & 0.167 & 0.506 \\ & & & 0.767 \end{pmatrix} \end{array}$$

(I) PIP

$$\begin{array}{c} \\ u1 \\ u2 \\ u3 \\ u4 \end{array} \begin{array}{c} \begin{array}{cccc} u2 & u3 & u4 & u5 \end{array} \\ \begin{pmatrix} 0.1901 & 0.4247 & 0.06 & 0.1905 \\ & 0.2001 & 0.0531 & 0.1245 \\ & & 0.06 & 0.1905 \\ & & & 0.0639 \end{pmatrix} \end{array}$$

(II) PSS

$$\begin{array}{c} \\ u1 \\ u2 \\ u3 \\ u4 \end{array} \begin{array}{c} \begin{array}{cccc} u2 & u3 & u4 & u5 \end{array} \\ \begin{pmatrix} 0.02089 & 0.05520 & 0.00475 & 0.02440 \\ & 0.0183 & 0.00464 & 0.03561 \\ & & 0.00636 & 0.02500 \\ & & & 0.01531 \end{pmatrix} \end{array}$$

(III) NHSM

Fig.2. User-User Similarity Matrix

The problems with PIP measure are as follows:

1. As shown in fig.2, for PIP measure, the similarity between user3 and user5 is lower than the PIP similarity between user4 and user5 . But from the user-item rating matrix we can see that the former is more similar than the latter [6].
2. The PIP measure only considers the set of common ratings and the absolute value. This will lead to low accuracy. For example, user1 and user2 have four common rated items, where user1 and user2 have rated 6 and 8 items respectively. It will be more similar than user1 and user3 who have four common rated items, where user3 has rated 100 items.
3. PIP measure only considers the local context information of common ratings. So there exists the misleading of similarity and it can be eliminate, if we consider the global information of the user behaviour.
4. The formula for the PIP measure is very complex. It uses different formula at different conditions. That is also not normalized. It is also not possible to combine this similarity measure with the other one [6].

Due to the above mentioned problems, there is a need for a similarity measure that will punish bad similarity and reward the good similarity. So we adopt a non-linear function that is sigmoid function in our improved measure. And that improved PIP measure is called PSS.

2) PSS (Proximity-Significance-Singularity) measure

The user PSS similarity can be calculated as follows:

$$sim(u,v)^{pss} = \sum_{p \in I} PSS(\ r_{u,p}, r_{v,p}\ ) \qquad (3)$$

Where the PSS($r_{u,p}$, $r_{v,p}$) is the PSS value of user u and v , it is defined as follows :

PSS($r_{u,p}$, $r_{v,p}$) =

Proximity($r_{u,p}$, $r_{v,p}$) x Significance ($r_{u,p}$, $r_{v,p}$) x Singularity($r_{u,p}$, $r_{v,p}$) (4)

Here, from the equation, we can see that PSS measure is composed of three factors like Proximity , Significance and Singularity[6]. The Proximity factor is same as PIP and it only consider the distance between two ratings. The second factor is significance and we assume that the ratings are more significant if they are more far from the median rating. Now we take one example , if two users rate two items as (4,4) or (2,2) and the other two users rate the two items as (5,3) or (4,2) . Then we can say that the previous ratings by two users are more significant than the other two users. And the last factor is Singularity. It represents how two ratings are different with other ratings [6].

The Formulas for these three factors are as follows.

$$\text{Proximity}\ (\ r_{u,p}\ , r_{v,p}\ ) \ = 1 - \frac{1}{1+\exp(\ -|r_{u,p} - r_{v,p}|)}$$

$$\text{Significance}\ (r_{u,p}\ , r_{v,p}) = \frac{1}{1+\exp(\ -|\ r_{u,p} - r_{med}\ |\ .|\ r_{v,p} - r_{med}\ |\ )}$$

$$\text{Singularity}\ (r_{u,p}\ , r_{v,p}) \ = 1 - \frac{1}{1+\exp(\ -|\ \frac{r_{u,p} + r_{v,p}}{2} - \mu_p\ |)}$$

where μ_p is the average rating of item p . $r_{u,p}$ is the rating of item p by user u .

Advantages of PSS measure with PIP measure

1. As shown in fig.2, by comparing PSS measure with PIP measure, the similarity between user3 and user5 is higher than the similarity between user4 and user5 and it is according to user-item rating matrix.
2. More accurate result is obtained in user-user similarity matrix that is shown in fig. 2
3. The formula for the PSS is not so complex compare with PIP measure.

3) The New heuristic similarity measure(NHSM)

This measure is composed of same factors like PSS that are Proximity, Significance and Singularity. All these three factors are same as PSS.NHSM integrate the modified Jaccard and User rating preference in the design. The formulas for NHSM are as follows [7].

$$sim(u,v)^{NHSM} \ = \ sim(u,v)^{JPSS}.\ sim(u,v)^{URP} \qquad (5)$$

$$sim(u,v)^{URP} = \ 1 - \frac{1}{1+\exp(\ -|\ \mu_u - \mu_v\ |\ .|\sigma_u - \sigma_v\ |)} \qquad (6)$$

$$sim(u,v)^{JPSS} \ = \ sim(u,v)^{PSS}.\ sim(u,v)^{Jaccard} \qquad (7)$$

$$sim(u,v)^{Jaccard} \ = \ \frac{|\ I_u \cap I_v\ |}{|\ I_u\ | \times |\ I_v\ |} \qquad (8)$$

$$sim(u,v)^{pss} = \sum_{p \in I} PSS(\ r_{u,p}, r_{v,p}\ ) \qquad (9)$$

Advantage of NHSM:

1. From the fig.2 of user-user similarity matrix , we can see that the similarity between user1 and user3 is higher than the similarity between user1 and user2 , which is also correct in PIP and PSS measure[6].
2. The similarity between user3 and user5 is also higher than that of similarity between user4 and user5 , which is not correct in case of PIP measure.
3. If the high difference exist between the two user's rating , then similarity should be very low . And this becomes true for the similarity between user2 and user4 that is shown in user-user similarity matrix[6].
4. From the fig. 2 of user-user similarity matrix , we can say that each user has different similarities and that is why each user becomes comparable . This is not seen in other user-user similarity matrix.

III. EVALUATION METHDOLOGY

A. Data Set

The comparison of recommendation performance for the collaborative filtering method for the measures like PIP , PSS & NHSM can be done by using the following three dataset[9]. For demonstration of performance, each dataset can be divided in two parts - 20% users are selected as testing users and 80% users are selected as training users

Table 2 Summary of dataset

Dataset	Description	Profile	Availability
MovieLens	Ratings of movie s in the scale of 1-5	247753 users 586994 tag application across 34208 movies 22884377 ratings	Available at http://www.grouplens.org/
Jester jokes	Ratings of jokes in the scale of -10 to 10	73496 users 100 jokes 4.1 million continuous rating	Available at http://www.ieor.berkeley.edu/~goldberg/jester-data/
Netflix	Ratings of movie s in the scale of 1-5	480000 customer 17000 movies 100 million movie rating	Available at http://www.netflixprize.com

B. Evaluation metrics

There are several types of evaluation metrics available that compare the quality of recommender system [4]. They can be divided into two types: Predictive accuracy metrics, Classification & Rank accuracy metrics.

1) Predictive Accuracy :

It represents the quantitative accuracy of the predicted value of a target item. And the metrics for that are MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error)[4].

MAE (Mean Absolute Error):

It is the average of absolute errors over all the predictions made by the collaborative filtering algorithms. For showing better accuracy, the value of MAE should be small.

$$MAE = \frac{\sum_{i=1}^{MAX} | r_i - \hat{r}_i |}{MAX} \quad (10)$$

Where r_i and $\hat{r}_i$ are actual and predicted rating of an active user on an item by collaborative filtering algorithm. The variable MAX shows the number of times prediction performed by the Collaborative filtering algorithm.

RMSE (Root Mean Squared Error) :

It is the square root of the mean/average of the square of all of the error. Compared to the similar Mean Absolute Error, RMSE amplifies and severely punishes large errors [4].

$$RMSE = \sqrt{\frac{1}{MAX} \sum_{i=1}^{MAX} (r_i - \hat{r}_i)^2} \quad (11)$$

2) Classification Accuracy :

The qualitative performance of a recommendation system can be measured by classification accuracy. In recommender system, the list of recommendation items is provided to an active user. Based on that, to evaluate the quality of a RS the following metrics are used [4].

Precision: It is the fraction of items in Lr that are relevant.

Recall: It is the fraction of total relevant items that are in the recommended list Lr.

A set of items on which the user has made high ratings is called the list of relevant items Lrev. The formulas for both of these metrics are as follows [4].

$$\text{Precision} = \frac{| L_r \cap L_{rev} |}{| L_r |} \quad (12)$$

$$\text{Recall} = \frac{| L_r \cap L_{rev} |}{| L_{rev} |} \quad (13)$$

It is good for a system to have high precision and recall values. Both of these metrics are inversely related such that when precision increases, recall decreases. The measure F1 combines both the measure precision and recall and it is as follows.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (14)$$

3) Receiver Operating Characteristic (ROC)

It is a graphical technique that uses two metrics, true positive rate (TPR) and false positive rate (FPR). We can use it as an alternative to precision/recall. With the help of this

metric, we can distinguish graphically good items (relevant) from bad (non-relevant) items [1].

Similar to Precision and Recall measures, ROC curves make an assumption of binary relevance. Items recommended are either successful recommendations (relevant) or unsuccessful recommendation (non-relevant). One consequence of this assumption is that the ordering among relevant items has no consequence on the ROC metric—if all relevant items appear before all non-relevant items in the recommendation list, you will have a perfect ROC curve.

The area under the curve, known as Swets' Measure and it is measuring the system's capacity to differentiate the good and bad items[1].

4) Half -life utility metric

The half-life utility metric is proposed by Breese et al. (1998). It postulates that the probability that the user will select a relevant item drops exponentially down the list [1]. It is describing the usefulness of an ordered list of recommendation items. According to theorem, the items in the beginning of the list have higher probability of being seen by the user. Probability decreases as we go down the list. The Formula for this metric is as follows [1].

$$R_{\propto} = \frac{\sum_j \max(v_{aj} - d, 0)}{2^{(j-1)/(\propto-1)}} \quad (15)$$

where d is neutral vote , that is slightly negative and α is the viewing half-life , means the position of the item in the list where there is a possibility of the item is being seen by a user .

IV. CONCLUSION

This paper first discuss about the cold-start problem in recommendation system. It is presenting various measures for cold start problem. First , it analyzes the Proximity-Impact-Popularity(PIP) measure and derive the user-user similarity matrix for the given user-item rating matrix. This paper is presenting various disadvantages of this measure with respect to user-user similarity matrix. Then in order to overcome the problem of PIP measure, the paper presented a new measure called PSS (Proximity-Significance-Singularity) and also compare it with PIP measure. And finally the paper is presented a new heuristic measure called NHSM that will give more accurate similarity between users compare to PIP and PSS. This paper shows the superiority of NHSM measure in recommendation performance.

REFERENCES

[1]. Carneiro, D. Fernández, V. Formoso, Comparison of collaborative filtering algorithms: limitations of current techniques and proposals for scalable, high- performance recommender system, ACM Trans. Web 5 (1) (2011) 1–33.

[2]. P. Resnick, H.R. Varian, Recommender systems, Commun. ACM 40 (3) (1997) 56–58.

[3]. B.N. Miller, I. Albert, S.K. Lam, J.A. Konstan, J. Riedl, MovieLens unplugged: experiences with an occasionally connected recommender system, in: proceedings of the 8th International Conference on Intelligent user interfaces, 2003, pp. 263–266.

[4]. Bidyut Kr. Patra, Raimo Launonen, Ville Ollikainen Sukumar Nandi : A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data Knowl. Based Syst. (2015),

[5]. H.J. Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Inform. Sci.178) (2008) 37–51.

[6]. Haifeng Liu, Zheng Hu, Ahmad Mian, Hui Tian, uzhen Zhu : A new user similarity model to improve he accuracy of collaborative filtering ,Knowledge-Based Systems 56 (2014) 156–166

[7]. Le Hoang Son : Dealing with the new user cold-start problem in recommender systems: A comparative review , Information Systems (2014)

[8]. L. Safoury, A. Salah, Exploiting user demographic attributes for solving cold-start problem in recommender system, Lect. Notes Softw. Eng. 1 (3) (2013) 303–307.

[9]. HYUNG JUN AHN , A Hybrid Collaborative Filtering Recommender System using a new similarity measure, Hangzhou, China , April 15-17 , 2007

[10]. J. Bobadilla, F. Ortega, A. Hernando, J. Bernal, A collaborative filtering approach to mitigate the new user cold start problem, Knowledge-Based Syst. 26 (2011) 225–238.

[11]. J. Bobadilla, F. Ortega, A. Hernando, A collaborative filtering similarity measure based on singularities, Inform. Process. Manage. 48 (2012) 204–217.

[12]. J. Bobadilla, A. Hernando, F. Orteqa, A. Gutirrez, Collaborative filtering based on significances, Inform. Sci. 185(2012)1–17.