Journal of Engineering and Applied Sciences 12 (16): 4108-4112, 2017
ISSN: 1816-949X

Collaborative Filtering Recommendation using Personalized Page Rank Algorithm with New Personalized Parameters

[1]Hayder M. Naji and [2]Ghaidaa A. Al-Sultany
[1]Department of Software,
[2]Department of Information Network, Babylon University, Babil, Iraq

Abstract: Collaborative filtering recommendation system shares the user's interests and recommends items to a user based on the interests of the other users whom are similar to his/her owntendencies. Basically, the Personalized Page Rank Algorithm (PPR) suggests items with respect to the target user by personalizing him/her only. In this study, Iteratively with each target user, the remaining users are personalized according to their rating patterns by supporting them withnew Personalized Parameters (PP). The personalized parameters have a role of personalized measure from which each user's rank will affect and be affected on the other user's ranks depending on the PP values. The achievement of more accurate recommender system needsmore personalization to satisfy user's tendencies so we Present a Personalized Recommendation system using PPR algorithm with more personalization method. Finally, classification accuracy measures have been used to evaluate the outcome top-N recommendation list on a MovieLens dataset in comparison with the outcome of traditional PPR.

Key words: Recommendation system, personalized page rank, personalized parameters, collaborative filtering, dataset, accurate

INTRODUCTION

Recommender system is a most powerful and popular software tool to extract and predict the relevant information and provide suggestions of items that are extremely likely of concern to a particular user (Burke, 2007). Recommender system is subclass of information filtering to filter a huge of the information and to offer a personal service to users.

Electronic retailer offers a huge selection of products, modern consumers are flooded with choices and so happy to meet a variety of special content that is appropriated to them tastes and this is key to increase user satisfaction and loyalty.

Recommender systems mainly can be classified into three broad approaches based on different numbers of technologies classified as (Ricci *et al*., 2015): (collaborative filtering approach, content-based approach, hybrid-based approach), collaborative filtering (social filtering) is one of common widely techniques of the recommender system. Personalized recommendation in this technique based on the ratings of the others users that have similar tastes and interesting items and the history of the user's rating. Collaborative filtering can often be grouped as being either: memory-based or model-based (Breese *et al*., 1998). Computing the similarity measures consider the core of memory-based on contrary, Personalize Parameters (PP) have a role of voting and force up the collaborative filtering meaning and personalizing each user with different weights according to his/her watching. Consequently, the retrieved movies for each user would be relevant to the other similar users of watching history.

Literatures review: Page Rank algorithms gives an equal weights for each entity in the process of gathering the page rank value for it as (Goel, 2009a) also in the bipartite graph based recommendation algorithms they are gives the same equal chance for the users to force up their items, Personalized Page Rank algorithm gives the target user a high boost of his page rank according to the Eq. 1 (Goel, 2009b). To a certain page i, we can calculate the pagerank value $\pi(i)$ for it as follow:

$$\pi(i) = \begin{cases} (1\text{-}\varepsilon) \sum\limits_{j:j\,\text{links to}\,i} \dfrac{\pi(i)}{|\text{out}(j)|} & \text{if } i \neq x \\ \varepsilon + (1\text{-}\varepsilon) \sum\limits_{j:j\,\text{links to}\,i} \dfrac{\pi(i)}{|\text{out}(j)|} & \text{if } i = x \end{cases} \quad (1)$$

Where:

x = The target node with probability
ε = Jump to x, probability
$(1\text{-}\varepsilon)$ = Click on a random hyperlink within page i

Corresponding Author: Hayder M. Naji, Department of Software, Babylon University, Babil, Iraq

If we wanted to personalize page rank to a certainsubset of the nodes in the network (e.g., personalize to all Californians men women 18-27 year olds, etc.) the Page Rank algorithm personalized to a subset A would be:

$$\pi(i) = \begin{cases} (1\text{-}\varepsilon) \sum_{j:j \text{ links to } i} \frac{\pi(i)}{|\text{out}(j)|} & \text{if } i \notin A \\ \frac{\varepsilon}{|A|} + (1\text{-}\varepsilon) \sum_{j:j \text{ links to } i} \frac{\pi(i)}{|\text{out}(j)|} & \text{if } i \in A \end{cases} \quad (2)$$

Also, the resulted movies by this way are recommended to the subset A, i.e., movies are recommended to public users are belongs to a community which represent the subset A. But, we want to recommend a movies to a specific user with movies are relevant to a subset of users and personalized with target user so we added the personalize parameters to personalize users are relevant and their movies are candidate.

Normalizing the page rank values for each type entities are very important, according to the empirical result by Bahmani *et al*. (2010) suggest that personalized page rank with normalized terms over-performs other methods while personalized page rank without normalizing terms performs rather poorly.

MATERIALS AND METHODS

The achievement of a more accurate recommender needs more personalizing to comprehend the user taste hence we support our system with parameters (p, n) for each useras shown in the Fig. 1. The parameters values in the user layer indicate the similarity among theusers through analyzingthe rating/tagging assignment of the allconnected to increase their ranks. The counter p is increased by 1 when two users give a positive rate or a negative rate and we increase the counter n when they give a dissimilar rate. Likewise, the user tags were utilized in the same way in which if there were a common tag on a movie, the parameter p would be increased by 2 which it is the duplicate of the rate as presented in the Algorithm 1.

Algorithm 1; calculating personalize parameters:

```
Input
    User_current, users rating on movies
Output
    p, n for each user
Begin
    Initial each movie and user with the following values
        p←1
        n←1
    for each movie rated by the current user called m
        for each user rated the same movie called mm
            if (m.rate>2.5 &mm.rate>2.5)|(m.rate<2.5 &mm.rate<2.5)
        then mm.user.p ++
            else     mm.user.n++
End
```

After calculation the values of (p, n) parameters they had used to update the ranks of the users by adding the value of local epsilon to the value of each user's rank. Local epsilon reflectsthe personalization of each user with different weight according to his PP's (Personalized Parameters) values as shown in the Eq. 3.

$$\text{local\_}\varepsilon = (p\text{-}n)/\min \quad (3)$$

where, min is the minimum number between the current user's movies and the other user's movies who wants to update his rank. The value of local_ε updates the rank of each user according to the Eq. 4.

$$\pi_{user} = \pi_{user} + \left(\text{local\_}\varepsilon^{\lambda} \times 10^{\lambda}\right) \quad (4)$$

where, π_{user} is the page rank for the user we power eps to λ to get on a high contrast of personalization weight among the users and multiply it with the value 10^{λ} to get on update of >1 for the user rank for all (local_ε>0.1) if we choose (λ = 6) which will be discussed more in the experimental results section. The Algorithm 2 presents how to rank the movies to be recommended if it is among the top-n movies with respect to the personalized parameters modification.

Algorithm 2; calculating movies ranks:

Input

User$_{current}$, users rating on movies

Output

Movies ranks

Begin

Calculate users personalization parameters (p, n), call algorithm 1.

Initial π_m for all movies,

π_m=1/M where π_m is the pagerank of movie m, M is the number of movies for each iteration:

1. Calculate π_u for all users where π_u is a pagerank for the user u

   $\pi_u = \sum_{\substack{m \in \text{Movies} \\ u \text{ rate } m}} \frac{\pi_m}{d_m}$ where is the number of users rated m.

2. Update π_u for all users, personalize them with a different local eps.

   Apply equation 1.

   Apply equation 2.

3. Normalize user's pagerank to be sum to 1.
4. Personalize the User$_{current}$'s pagerank, for all users u,

   $\pi_u = \pi_u\,(1\text{-}\varepsilon) + \begin{cases} \varepsilon & \text{if } u = \text{User}_{current} \\ 0 & \text{otherwise} \end{cases}$

5. Calculate for all movies,

   $\pi_m = \sum_{\substack{u \in \text{Users} \\ u \text{ rate } m}} \frac{\pi_u}{d_u}$ where is the number of movies rated by u.

Sort the movies descending according to the movies ranks

End

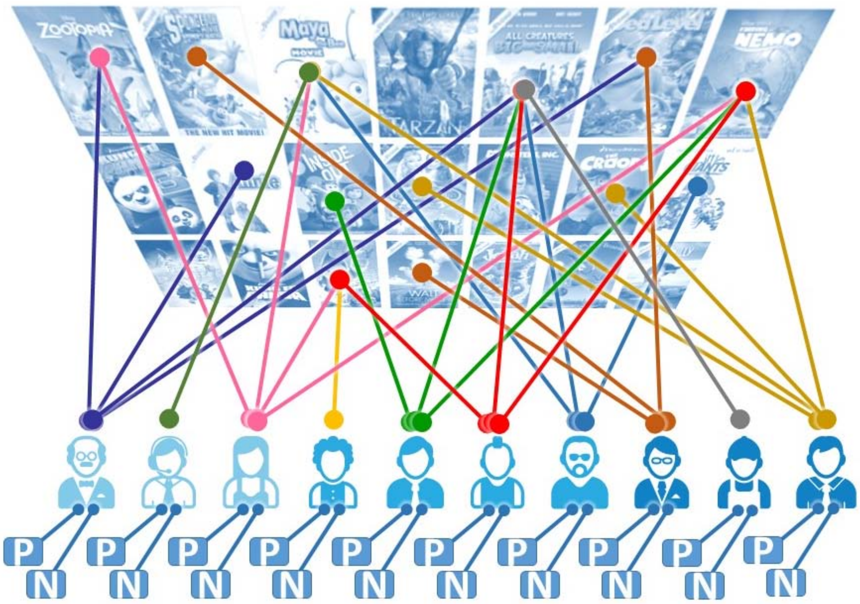

Fig. 1: Users with personalized parameters

RESULTS AND DISCUSSION

Evaluation and experimental results: Based on the classification methods, the recommendations made can be divided into four kinds. If the user is interested in what the system has suggested to him/her, the system has a True Positive (TP) otherwise, if the item is uninteresting a False Positive (FP) suggestion has been made. If the system cannot predict an interesting item we have a False Negative (FN). If the system does not suggest an item not interesting for the user then we have a True Negative (TN).

The classification measures that measure the suggested item is correct or incorrect and it is include three measures: precision, recall and F-measure (Uluyagmur *et al*., 2012). The values of these measures depend on the confusion matrix as a Table 1.

To evaluate our recommendation system's performance and compare it with the others we used precision, recall and F-measure metrics. Precision is a ratio of the relevant suggested items to total number of items suggested shown in Eq. 5.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{5}$$

Recall is a ratio of the relevant suggested items to total number of relevant items available shown in Eq. 6.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{6}$$

Table 1: Confusion matrix

	Predictive model	
Recommendation	Yes	No
Actual recommended		
Yes	True Positive (TP)	False Negative (FN)
No	False Positive (FP)	True Negative (TN)

F-measure is combine the precision and recall shown in Eq. 7.

$$\text{F-measure} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{7}$$

For example, given N = 10 and the user have 20 movie in the test data, so if the recommendation system results 5 movies from the user test data among the top 10 then precision will be 0.5, recall will be 0.25 and F-measure will be 0.33.

We used the MovieLens 2K dataset to evaluate the recommendation system, it has 10, 197 movies and 2, 113 users in average of 85 movie rating per user we evaluate the performance of our system on 500 user and splitting the dataset into 3:1, i.e., 75% for training and 25% for testing.

The system has been evaluated with different values of N to calculate thesystem accuracy thus we test it 10 times with N = {5, 10, ..., 50} each time the system recommend top-N movies for each user. The total accuracy measure were computed from the average of user's accuracy values and compared against the baseline recommender system.

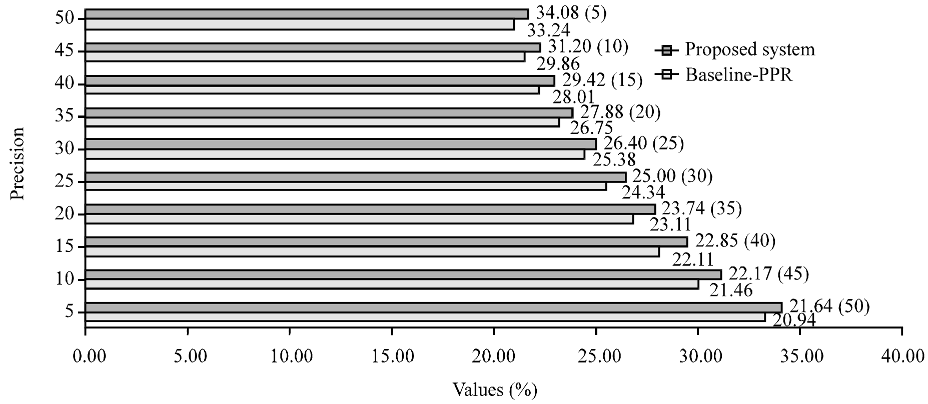

Fig. 2: Comparison of systems using precision accuracy measure

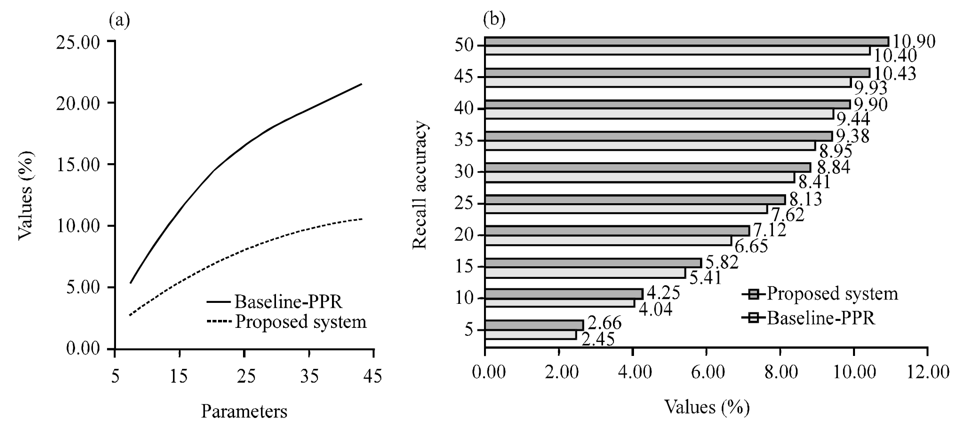

Fig. 3: Comparison of systems using recall accuracy measure

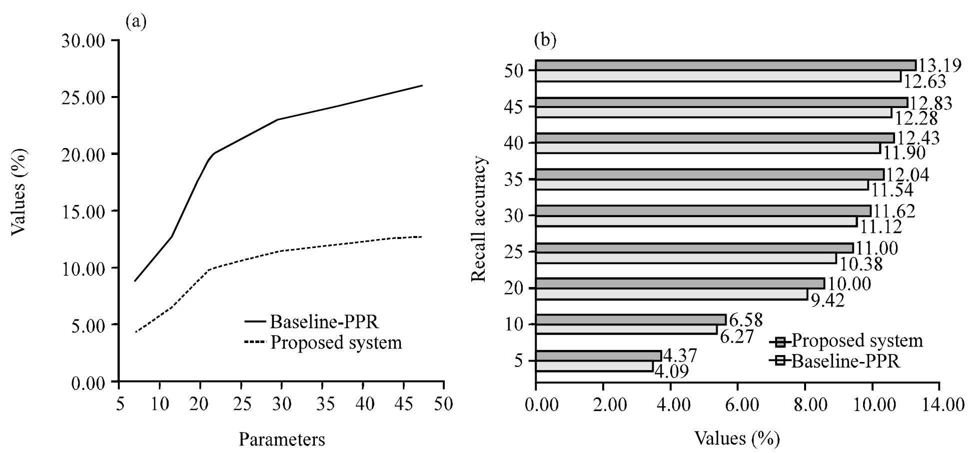

Fig. 4: Comparison of systems using F-measure

Figure 2-4 show the performance of the system of personalized Page rank as a base line and the proposed system supported with a personalized parameters and local epsilon for personalizing each user with respect to the other the user's ratingtendency. After we calculate ε by the Eq. 3, we update page rank for the user according to the Eq. 4, we power ε by the value of λ to get on high contrast between the different values of ε. This power makes the values of ε verysmall so we multiply it with 10^{λ} to return back its effect, Fig. 5 shows the effect of using Eq. 4 with λ = 6.

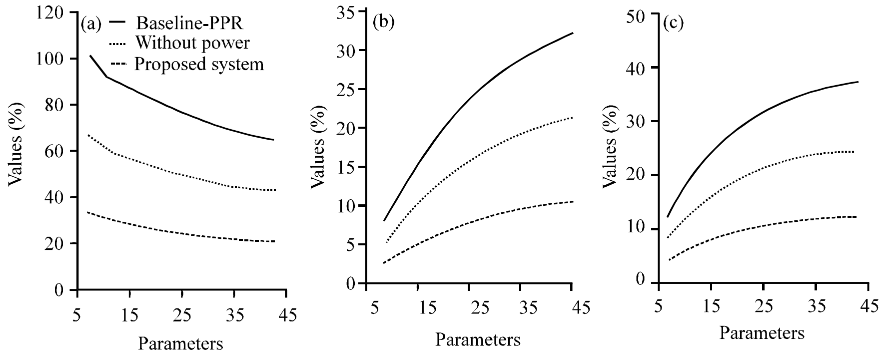

Fig. 5: The effect of power ε: a) Precision; b) Recall; c) F-measure

CONCLUSION

The recommendation system is supported with new personalized parameters (p, n) for each user that are adapted with respect to the other the user's ratingtendency was proposed on this research. The v recommendation system results outperform the baseline traditional system in terms of the evaluation measures with significant find outcomes. The next stage of this work focuses on using the personalized parameters to explicit user own community by a threshold on the value of ε and deal with it as a similarity scale for community detection and then we can recommend movies to a set of users.

REFERENCES

Bahmani, B., A. Chowdhury and A. Goel, 2010. Fast incremental and personalized Page Rank. Proc. VLDB Endowment, 4: 173-184.

Breese, J.S., D. Heckerman and C. Kadie, 1998. Empirical analysis of predictive algorithms for collaborative filtering. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, Jul 24-26, 1998, Madison, WI., pp: 43-52.

Burke, R., 2007. Hybrid Web Recommender Systems. In: The Adaptive Web, Brusilovsky, P., K. Alfred and N. Wolfgang (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-72079-9, pp: 377-408.

Goel, A., 2009a. More recommendation systems (Lectures). Stanford University Computer Science, California, USA.

Goel, A., 2009b. Reputation systems and page rank (Lectures). Stanford University Computer Science, California, USA.

Ricci, F., L. Rokach and B. Shapira, 2015. Recommender System Handbook. 2nd Edn., Springer, Heidelberg, Germany, ISBN: 978-1-4899- 7636-9, Pages: 1003.

Uluyagmur, M., Z. Cataltepe and E. Tayfur, 2012. Content-based movie recommendation using different feature sets. Proceedings of the World Congress on Engineering and Computer Vol. 1, October 24-26, 2012, WCECS, San Francisco, USA., ISBN: 978-988-19251-6-9, pp: 17-24.