IJICTR
International Journal of Information & Communication Technology Research

Volume 14- Number 1 – 2022 (48 -56)

Personality-Based Matrix Factorization for Personalization in Recommender Systems

Mazyar Ghezelji
Computer Engineering Faculty
K. N. Toosi University of Technology
Tehran, Iran
maz.ghezelji@email.kntu.ac.ir

Chitra Dadkhah*
Computer Engineering Faculty
K. N. Toosi University of Technology
Tehran, Iran
dadkhah@kntu.ac.ir

Nasim Tohidi
Computer Engineering Faculty
K. N. Toosi University of Technology
Tehran, Iran
n.tohidi@email.kntu.ac.ir

Alexander Gelbukh
Centro de Investigación en Computación
Instituto Politécnico Nacional
Mexico City, Mexico
gelbukh@cic.ipn.mx

Received: 8 February 2022 – Revised: 1 March 2022 - Accepted: 12 March 2022

***Abstract*—Recommender systems are one of the most used tools for knowledge discovery in databases, and they have become extremely popular in recent years. These systems have been applied in many internet-based communities and businesses to make personalized recommendations and acquire higher profits. Core entities in recommender systems are ratings given by users to items. However, there is much additional information which using it can result in better performance. The personality of each user is one of the most useful data that can help the system produce more accurate and suitable recommendations for active users. It is noteworthy that the characteristics of a person can directly affect his/her behavior. Therefore, in this paper, the personality of users is identified, and a novel mathematical and algorithmic approach is proposed in order to utilize this information for making suitable recommendations. The base model in our proposed approach is matrix factorization, which is one of the most powerful methods in model-based recommender systems. Experimental results on MovieLens dataset demonstrate the positive impact of using personality information in the matrix factorization technique, and also reveal better performance by comparing them with the state-of-the-art algorithms.**

Keywords: recommender system; matrix factorization; knowledge discovery; personality.

Article type: Research Article

[DOI: 10.52547/itrc.14.1.48]

I. INTRODUCTION

The rapid growth of e-commerce and social media has led to a great source of information on the web, which causes information overload [1]. This information overload creates some difficulties for users to find their desired items or information. Therefore, this specific problem will decrease user's satisfaction and loyalty to the system [2]. In order to handle this problem, recommender systems have been introduced.

A recommender system is a special type of Knowledge Discovery in Databases (KDD) techniques that generates personalized recommendations by analyzing the patterns of user's behavior and interests [3, 4]. The main idea of a recommendation system is to find a set of items that a user will be interested in. Various types of recommender systems have been developed by utilizing three main entities in the system: users, items, and user-item interactions [5].

* Corresponding Author

There are different categories of recommender systems: content-based, collaborative filtering, knowledge-based and hybrid recommender systems [6]. In the content-based approach, recommendations are generated based on similarities in active user preferences and item's attributes. The collaborative filtering approach does not use the item's attributes, instead, it calculates the ratings of items based on the opinions of other users [7]. Knowledge-based recommender systems suggest products (items) based on inferences about a users' requirements and preferences [6]. Finally, hybrid recommender systems are a combination of different approaches, which give better results and increase the accuracy of recommendations. Collaborative filtering has gained many researchers and e-retailers attention due to its special attributes [8]. This approach does not require domain knowledge, which means it can be applied in systems where obtaining attributes of items is difficult or in domains in which this process cannot be automated [8, 9]. In addition, collaborative filtering can provide a serendipitous recommendation as it does not only recommend items similar to the target user history or with specific attributes, but it takes advantage of a group of users, which helps in discovering new items [9]. Two major approaches in collaborative filtering are neighborhood-based and model-based approaches in which matrix factorization is the most common method [9-11].

Neighborhood-based approach uses similarity measures between users and items to predict the preferences of the target user, but matrix factorization approach represents users and items with a set of latent factors to make it directly comparable. The neighborhood-based approach can find local relations among users and can give a relatively good result when the number of users is not large and the rating matrix is not so sparse. However, as the size of online communities' increase and the number of users grows, the precision of the collaborative filtering method's recommendations decreases, and it gets harder to implement it at scale. Matrix factorization method can find overall structure, unlike the neighborhood-based approach, and has good performance on large amounts of data [6, 12, 13]. Besides, it has integrated with auxiliary information to mitigate cold start problem and has improved the precision of recommendations.

Most of the basic recommender systems use item ratings from users in order to predict their preferences, but there are auxiliary information like the context of ratings and personality information that could improve the performance of this method [10, 14, 15]. Human's nature and personality affects all aspects of their life including behavior in social networks, online shops, etc. The correlation of personality and rating behavior has been widely studied before and the results show that leveraging personality information can help collaborative filtering methods to get a better result and to solve cold start problem [4, 16, 17].

Previous methods which used personality information, mostly extracted personality information from other data sources or by mining the contents of items and profiles of users. However, these data sources are not always available. In this paper we show that just having access to the rating matrix is not a limitation, and personality information can still be extracted.

As it has been mentioned before, the personality of users affects their rating behavior and could obtained from it too. It has been shown that the personality of a user, in terms of optimism and pessimism, can be extracted from his/her ratings to items and other users. It can also have effects on his/ her relations with other users [18]. In this regard, we propose a novel matrix factorization method using the personality of users that is extracted from their rating behavior.

The structure of this paper is organized as follows: in section 2, some of the main previous studies related to matrix factorization and personality detection are reviewed. In section 3, the proposed method is explained in detail and a hypothesis which is used in matrix factorization is introduced. Then in section 4, the proposed method is evaluated and compared with other related methods. Finally, in section 5, conclusion and some of the possible future works are mentioned.

II. Related Works

Many studies have shown the success and the positive effect of collaborative filtering method in recommender systems. Each of the two important types of collaborative filtering, neighborhood-based and model-based, has its advantages and disadvantages. Hence, both of these approaches are being used in either research or industry.

Koren [2] represented the importance of explainability in neighborhood-based methods and proposed a neighborhood-based recommender system, which worked based on optimizing a global cost function. It maintained the explainability of the neighborhood-based method, which is the key to this method's vast usage, and used implicit ratings, which decreased the errors conceivably. Besides, Koren presented the top-k recommender evaluation technique to distinguish the quality of different recommender systems properly.

George and Merugu [19] proposed a novel collaborative filtering approach based on the weighted Bregman co-clustering algorithm. The main idea of this work was to find neighborhoods faster than matrix factorization approaches, and to be able to use average ratings of co-clusters and user biases in order to generate predictions. Their experiments showed that this model can be trained

[DOI: 10.52547/itrc.14.1.48]

much faster than formal matrix factorization or SVD models.

Koren et al. in [4] reviewed the theory and applications of the very popular matrix factorization method, in recommender systems. They boosted the plain matrix factorization method with biases, implicit feedback, and temporal dynamics. As a result, they got better results with comparing to other state-of-the-art methods through their implementations.

Luo et al. [5] presented a non-negative matrix factorization algorithm for recommender systems. They used a single-element-based approach, which resulted in computational efficiency and ease of use for industrial applications. The result of their algorithm showed that it can outperform classic and weighted non-negative matrix factorization algorithms in terms of efficiency and accuracy.

One of the most important capabilities of matrix factorization is the ability to include different parameters in the learning process with the goal of decreasing error. Hu and Pu [9] addressed cold start problem by utilizing personality information of users in the matrix factorization collaborative filtering approach. They defined personality as it was introduced in psychology previously; the personality of a person can be defined by five different bipolar dimensions: Openness to Experience, Conscientiousness, Extroversion, Agreeableness, and Neuroticism. Experimental results illustrated that their proposed cascade method performed better than the classic rating-based collaborative filtering systems.

As it has been mentioned before, personality of users can play an important role in some areas. Khwaja et al. [10] presented an approach for developing an activity recommender system for improving subjective wellbeing. They demonstrated that the link between subjective wellbeing and alignment between activities with personality can considerably improve the accuracy of the recommendations. Also, Gupta et al. [12] showed the correlations between personality traits of a person (age, gender, lifestyle, etc.) and his or her musical choice. They believed that personalities can vary from person to person and over time, besides, it can be extracted and exploited for more accurate recommendations.

Khan et al. [16] investigated the interaction of users in social networks like Twitter and IMDB in order to extract psychological information about them. Their proposed model used the extracted information from social networks about users for recommending movies. Their experimental results demonstrated the effectiveness of their model in the movie recommendation task in comparison with other models.

There are two kinds of trust in trust-aware recommender systems: implicit and explicit trust [14]. Explicit trust considers all social links among different users in social networks, which are established directly by users. Using this concept, Yakhchi et al. [20] introduced a new two-level model called TAP. In their research, they designed a mathematical model based on the matrix factorization method to consider both personality and trust information at the same time. In other words, TAP analyzed behavior of different users with two goals: 1) Detecting the personality type of each user, 2) Combining trust information in order to give more personalized recommendations to each user.

Zhang et al. [21] suggested a single-step matrix factorization process, called FeatureMF, that used features of various items for recommendation. Their proposed model planned all available attribute data in each of the item features into the same latent factor space with items and users. Hence, they could design a representation for items by matrix factorization. Experimental results showed that FeatureMF, as a scalable model, outperformed other related models and it could solve the cold start and the data sparsity problems.

III. THE PROPOSED METHOD

In the previous section it was mentioned that side information, like links between users, can be used for increasing the accuracy of matrix factorization method. However, there are a few studies on finding user's attributes in rating matrix and using them for matrix factorization. To get the maximum outcome from the limited data source, we define personality attributes for each user. These attributes will be used in the optimization process. The proposed method consists of two main parts:

- Personality information extraction
- Personality-based matrix factorization in terms of optimism and pessimism

In the following paragraphs, first the problem is defined and then the solution, which is the proposed method, is explained.

Assume $U = \{u_1 . u_2 . \dots . u_n\}$ is the set of users and $I = \{i_1 . i_2 . \dots . i_m\}$ is the set of items, n is the size of users and m is the size of items in the system. Then, $G \in R^{n \times m}$, which is the rating matrix would be used. Each entry $G_{u.i}$ of G is a rating from user u for item i. The ratings are in the range of 1 to 5, and 0 means the user has not seen or rated the item. Regarding the available data, the defined problem is:

Considering the rating matrix G and the personality vector $\vec{P}$, the goal is to fill the 0 entries in G by making accurate predictions.

[DOI: 10.52547/itrc.14.1.48]

A. *Personality*

Optimism and pessimism are two important characteristics of users that have a huge impact on their ratings. Obviously, if a user's ratings are mostly higher than the ratings' average, he or she is more optimistic and if a user's ratings are mostly lower than the ratings' average, he or she is more pessimistic [3, 20]. In this regard, the personality of a user is defined as Eq. (1).

$$Personality(u) = O_u - P_u \qquad (1)$$

Where O_u is the optimism degree of user u, P_u is the pessimism degree of user u, and *Personality* indicates the total propensity of the user to optimism and pessimism, which is a real number in the range of -1 and 1. *Personality* of 1 means totally optimistic and -1 means totally pessimistic.

The value of O_u and P_u calculate as Eqs. (2) and (3), respectively.

$$O_u = \frac{|\{r_{ui}|r_{ui}>3.\ \bar{r}_i\leq 3\}|}{|\{r_{ui}|r_{ui}\neq 0,\ \bar{r}_i\leq 3\}|} \qquad (2)$$

Where r_{ui} is the rating that user u have given to item i and $\bar{r}_i$ is the average of all ratings given to item i. This equation indicates that the optimism degree of a user depends on the number of high ratings that user u had given to the less popular items.

$$P_u = \frac{|\{r_{ui}|r_{ui}<3.\ \bar{r}_i\geq 3\}|}{|\{r_{ui}|r_{ui}\neq 0.\ \bar{r}_i\geq 3\}|} \qquad (3)$$

Eq. (3) indicates that the pessimism degree of a user depends on the number of low ratings that user u had given to the more popular items. Less popular items are items with the average rating of less than 3 and more popular items are items with the average rating of greater than 3.

In the numerator of both Eq. (2) and Eq. (3), r_{ui} = 3 is not counted, because in the proposed model the rating 3 given by a user to an item indicates he/she being neutral, neither optimistic nor pessimistic.

B. *Hypothesis*

There are several relations between user characteristics and the decisions that he or she tends to make. In the case of social networks and recommender systems, the influence of their characteristics on their actions and ratings is undeniable.

In this paper, after analyzing the MovieLens 100k user-item matrix [22], user ratings, and their personalities, a hypothesis about the relation between personality and ratings is constructed. An important point is that this hypothesis is about all users who have rated more items than a threshold. The MovieLens dataset has been already pre-processed, so the minimum threshold is considered as 20. Besides, the variable T is defined as the set of all users, who have rated more items than the threshold. The hypothesis is stated in Eq. (4).

$$\forall_{u\in T}\ P_u > \bar{P}_T \rightarrow \bar{r}_u > \bar{r}_T \qquad (4)$$

Where u indicates user, P_u indicates the personality of user u, $\bar{P}_T$ is the average personality of all users in the T set, $\bar{r}_u$ is the average of ratings given by user u and $\bar{r}_T$ is the average of ratings given by all users in the T set. In other words, the hypothesis is as follows:

Among all users in the set T, for each user u, if the personality of user u is higher than the average personality of all users in T, then the average ratings of user u is higher than the average ratings of all users in T.

In order to test this hypothesis, a two-sample C-test on the MovieLens 100k dataset was applied [22]. In this regards, two arrays $\bar{tp}$ and $\bar{fp}$ were used indicating true positive and false positive, respectively. Besides, every user throughout the dataset was checked. If the personality of user u is greater/less than the average personalities of all users and the average of all ratings given by user u is greater/less than the average of all ratings in rating matrix, then the value 1 was added to $\bar{tp}$ and the value 0 was added to $\bar{fp}$. Otherwise, the value 0 and 1 was added to $\bar{tp}$ and $\bar{fp}$, respectively.

According to $\bar{tp}$ and $\bar{fp}$ vectors, the null hypothesis H_0 and the alternative hypothesis H_1 are defined as Eq. (5).

$$H_0: \bar{tp} \leq \bar{fp} \qquad H_1: \bar{tp} > \bar{fp} \qquad (5)$$

The result of *C*-test [23] shows that the null hypothesis is rejected with *t*-statistics [23] equal to 18.5745 and *p*-value [23] equal to 3.724477401546685e-71. Therefore, it is inferred that the proposed hypothesis is true. It means that if a user's personality is more than the average, then his or her average rating is also more than the total average.

C. *Personality-Based Matrix Factorization (PBMF)*

Matrix factorization (MF) is a very popular model-based collaborative filtering technique. Its scalability, accuracy, ability to integrate regularizations, and ability to provide prediction when there is lack of data (cold start problem), has been proven in literature [1, 2]. Therefore, it has been used as the basic model here.

The main idea of MF is to decompose the rating matrix into two smaller matrices, of which the product of them will be the actual rating matrix. In this regard, R is the rating matrix with $m \times n$ dimensions, U is the latent users with $m \times k$ dimensions and V is latent items with $n \times k$ dimensions, m is the number of users, n is the number of items and k is the number of latent factors [4]. MF process is as Eq. (6).

$$R = U \cdot V^T \qquad (6)$$

[DOI: 10.52547/itrc.14.1.48]

In order to find the U latent and the V latent, the optimization problem in Eq. (7) should be solved.

$$\min_{U.V} \parallel W \odot (R - U.V^T) \parallel_F^2 + \lambda_1 \cdot \parallel U \parallel_F^2 + \lambda_2 \cdot \parallel V \parallel_F^2 \quad (7)$$

Where W is the weight matrix. Usually, $W_{ui} = 1$ if there is a rating from user u for item i, otherwise $W_{ui} = 0$. The term $\parallel\cdot\parallel$ is Frobenius norm [24] and $\odot$ is an element-wise product of two matrixes. $\parallel U \parallel_F^2$ and $\parallel V \parallel_F^2$ are regularization terms, which prevent overfitting. The most well-known method for this optimization problem is stochastic gradient descent, which will be discussed later in this paper.

With the aim of taking advantage of the mentioned hypothesis, some regularizations were applied. Suppose that the personality of a user is higher than the average personality, but his or her average rating is less than average rate of all users. In this case, the hypothesis would be violated. Hence, the difference of his or her average ratings and the total average ratings were added as a penalty to the optimization function. In this regard, the term that was added to the optimization problem to be minimized is stated in Eq. (8).

$$\min \sum_{u|P_u > \bar{P}} \max(0, \bar{r} - \bar{r}_i)^2 + \sum_{u|\bar{P} > P_u} max(0, \bar{r}_i - \bar{r})^2 \quad (8)$$

Where u indicates every user in the system. The first part of Eq. (8) is for the case that the personality of the user is higher than the average personality and the second part is related to the case that the personality of user is lower than the average personality.

Considering all of the parts mentioned above, the optimization problem can be written as Eq. (9).

$$\min_{U.V} \parallel W \odot (R - U.V^T) \parallel_F^2 + \lambda_1 \cdot \parallel U \parallel_F^2 + \lambda_2 \cdot \parallel V \parallel_F^2 + \sum_{u|P_u > \bar{P}} max(0, \bar{r} - \bar{r}_i)^2 + \sum_{u|\bar{P} > P_u} max(0, \bar{r}_i - \bar{r})^2 \quad (9)$$

It should be noted that, as the max function is used here, there is no closed-form solution for this problem. Therefore, the gradient descent [25] is applied to get an acceptable local minimum.

Gradient descent method uses derivative of the optimization function in order to shift the solution towards a better one. Using gradient descent requires a matrix form equation. Therefore, the formulation should be rewritten. First, some new terms should be defined for the new equation, as follows:

- Vector A with m elements, where its elements are 1 if $P_i \leq \bar{P}$ and $r_i \geq \bar{r}$, -1 if $P_i > \bar{P}$ and $r_i < \bar{r}$ and 0 otherwise.
- Vector X with n elements, where its elements are equal to $\frac{1}{n}$.
- Vector Y with m elements, where its elements are equal to $\bar{r}$.

Using new terms, the formulation is as Eq. (10).

$$\min_{U.V} \frac{1}{2} \cdot \parallel W \odot (R - U.V^\intercal) \parallel_F^2 + \frac{\lambda_1}{2} \cdot \parallel U \parallel_F^2 + \frac{\lambda_2}{2} \cdot \parallel V \parallel_F^2 + \frac{\lambda_3}{2} \cdot (A^\intercal \cdot U \cdot V^\intercal \cdot X - Y))^2 \quad (10)$$

Now for the updating step in gradient descent process, the derivative of the optimization function is taken with respect to U and V. Considering the optimization function F, the derivative of F with respect to U is stated in Eq. (11).

$$\frac{\partial F}{\partial U} = \lambda_1 \cdot U - (W \odot (R - U.V^\intercal) \odot W) \cdot V + \lambda_3 \cdot (X^\intercal \cdot V \cdot U^\intercal + (-Y)^\intercal) \cdot A \cdot A \cdot (X^\intercal \cdot V) \quad (11)$$

And the derivative of F with respect to V is stated in Eq. (12).

$$\frac{\partial F}{\partial V} = \lambda_2 \cdot V - (W^\intercal \odot (R^\intercal - V.U^\intercal) \odot W^\intercal) \cdot U + \lambda_3 \cdot A^\intercal \cdot (U \cdot V^\intercal \cdot X - Y) \cdot X \cdot (A^\intercal \cdot U) \quad (12)$$

For generalization of the solution, the element-wise product of W and W was not written with, as it is a matrix of 1s and 0s.

Algorithm 1 presents the proposed method, PBMF, in the pseudo-code form. After running the algorithm, the predicted rating matrix will be $U \cdot V^\intercal$.

Algorithm 1: The proposed algorithm (PBMF)

Input: the rating matrix **G**, regularization coefficients λ_1, λ_2, λ_3 number of latent factors **K**, number of iterations **i**, learning rate α.
Output: **U** and **V**

```
Calculate W, X, Y;
Calculate Personalities;
Initialize U, V with random elements;
Define l = 0;
while l != i do
        Calculate A;
        Calculate ∂F/∂U regarding calculated parameters;
        Calculate ∂F/∂V regarding calculated parameters;
        Update U ⟵ U − α ∂F/∂U;
        Update V ⟵ V − α ∂F/∂V;
        Update l ⟵ l + 1;
End
Return U, V;
```

IV. EVALUATION

For evaluation the MovieLens 100k dataset [26] was used. MovieLens[1] is a dataset derived from a non-commercial web-based recommender system for movies. Each user can rate each movie from 1 to 5. Furthermore, users can attach tags based on the content of the movie, so that the accuracy of the recommender will be increased. This dataset consists of 100,000 ratings from 943 users on 1682 movies. The dataset has been pre-processed so that each user has rated at least 20 movies. Demographic information like age, gender and occupation is also stated for each user. The complementary information about the personality of users in the MovieLens dataset is mentioned in Table 1.

TABLE I. STATISTICS OF MOVIELENS 100K

-	Rate	#Rates of each user	Personality	Optimism	Pessimism
Min	1	20	-0.79	0	0
Max	5	737	1	1	0.79

[1] The datasets applied in the current study is available here.

[DOI: 10.52547/itrc.14.1.48]

Mean	3.52	106.04	0.10	0.23	0.12

To compare the different methods and represent the error of their predictions, two evaluation metrics have been used in this paper: MAE and RMSE.
MAE stands for the phrase Mean Absolute Error and is defined as Eq. (13).

$$MAE = \frac{\sum_{i \in R} |r_i - \hat{r}_i|}{N} \qquad (13)$$

MSE stands for Mean Squared Error and is defined as Eq. (14).

$$MSE = \frac{\sum_{i \in R} (r_i - \hat{r}_i)^2}{N} \qquad (14)$$

RMSE is the Root of MSE as Eq. (15).

$$RMSE = \sqrt{\frac{\sum_{i \in R} (r_i - \hat{r}_i)^2}{N}} \qquad (15)$$

In Eqs. (13), (14) and (15), r_i is the actual rate, $\hat{r}_i$ is the predicted rate and N is the total number of ratings.

The algorithms selected for comparison with our proposed methods are as follows:

- Matrix Factorization (MF) [2, 6]: It is a well-known model-based collaborative filtering algorithm. It is the base algorithm for many powerful methods, because of its flexibility and scalability.
- K-Nearest Neighbor (KNN) [2]: This lazy clustering algorithm is one of the primary clustering methods in data mining. It has been used extensively in model-based collaborative filtering approaches.
- Co-Clustering (CC) [19]: It is a data mining method that relates to a simultaneous clustering of the rows and columns of a matrix. It has been used as a dynamic collaborative filtering approach for many recommender systems.
- Non-negative Matrix Factorization (NMF) [5]: It is a general method for dimensional reduction and feature extraction on non-negative data and has been used in collaborative filtering approaches. It works based on matrix factorization techniques and considers just positive data about users and items.

These algorithms have been selected because they have proven their acceptable performances in related works [2, 7, 10, 21, 27-29].

V. EXPERIMENTAL RESULTS

In this paper, the cross-validation procedure is used for evaluation. In order to use cross-validation, first the dataset should be split into equal parts, then each time, one part should be set as test data and the remaining parts as train data. Therefore, here the MovieLens dataset is split into 5 parts. The training part is 80% and the test part is 20% of the whole dataset. These parts are taken randomly, and they are used for 5-fold cross-validation. This procedure is repeated for all parts and final results are the average result of all repeats.

In the following some parameter settings related to the selected algorithms are defined.

The common similarity measure in KNN method is cosine similarity. Here, the maximum number of k is set as 40 and the minimum number of k as 1. The reason for using maximum and minimum values for k is that some users have fewer neighbors than maximum k, also users with negative similarity cannot be considered as friends.

Number of factors used in the NMF algorithm is 15. Regularization terms for users and items are both 0.06 and it has been optimized in this paper using stochastic gradient descent for 50 epochs.

For the CC algorithm, the number of user and item clusters are both 3 and the number of iterations for the optimization loop is 20. As the result of CC and NMF algorithms are dependent on the initial state, the evaluation is repeated with 5 different random initial states and the average of them is used for the result. Besides, for the MF and PBMF the initial states are matrices that all of their elements have the value 1.

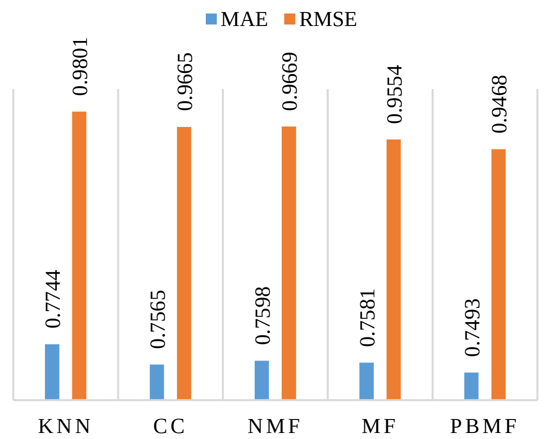
The number of iterations of optimization loops in both MF and PBMF is 80. In addition, λ_1 and λ_2 in both methods are 0.05. Number of latent factors in MF is 2 and in PBMF is 3. λ_3 in PBMF is 4. The appropriate λ_1, λ_2, λ_3 and k would be discussed later in this section.

The detailed comparison of these four algorithms is shown in Fig. 1.

The selected methods can be divided into two groups:

1. Matrix factorization-based algorithms (MF-based algorithms)
2. Non-matrix factorization-based algorithms (Non-MF-based algorithms)

Considering these two groups, in Table 2 the performance of PBMF is compared with other approaches. It shows that the proposed method could outperform other algorithms, regardless of whether they have used MF or not.



[DOI: 10.52547/itrc.14.1.48]

Fig. 1. Comparing MAE and RMSE of the proposed method and the selected algorithms.

TABLE II. MAE AND RMSE OF DIFFERENT APPROACHES

Algorithm	RMSE	MAE
MF-based algorithms	0.9611	0.7589
Non-MF-based algorithms	0.9733	0.7654
PBMF	**0.9468**	**0.7493**

In order to get the best result using PBMF, the optimum regularization terms and number of latent factors should be found. In this regard, the grid search process [27] is used for tuning the hyper-parameters. To do so, the whole dataset is used as the train data and MSE is calculated for each combination of hyper-parameters in a 20 iteration learning loop. The results are shown in Fig. 2.

Fig. 2 demonstrates that the MSE decreases as the number of factors get higher and then increases when the number of latent factors is more than 3. The MSE also decreases as regularization term increases and then increases when the regularization term is more than 4. Therefore, the optimal regularization term and number of latent factors that are used in this paper are 4 and 3, respectively.

It can be observed that MF-based algorithms have a relatively better performance compared to other types of collaborative filtering methods. Among the Non-MF-based algorithms, the CC algorithm gives 0.0136 reduction over KNN in terms of RMSE and 0.0179 reduction in terms of MAE.

Among MF-based algorithms, PBMF has the best performance with RMSE of 0.9468, which is 0.0086 less than MF and 0.0201 less than NMF. Thus, Fig. 2 clearly illustrates the positive impact of personality information on the performance of MF.

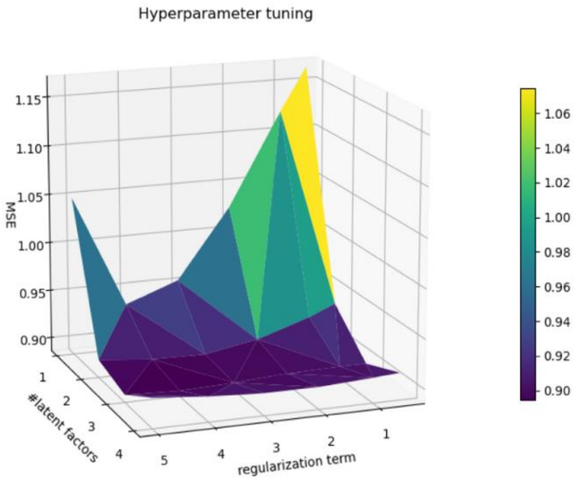


Fig. 2. Hyper-parameter tuning.

According to Fig. 2, the difference between RMSE and MAE of the CC algorithm is more than other algorithms. Considering this difference, it can be inferred that the variance of individual errors in the sample is greater than others. Comparison of RMSE and MAE for MF and PBMF algorithms through different iterations can be seen in Fig. 3.

In Fig. 3, the huge difference between two methods can be seen, in which MF starts with the large initial steps and steps get smaller as it converges. Therefore, it can be said that for any number of iterations, PBMF gives better results than MF.

VI. CONCLUSION AMD FUTURE WORKS

Matrix factorization is one of the most widely used algorithms of collaborative filtering approach in recommender systems. This algorithm breaks down the user-item matrix and generates two rectangular matrices with lower dimensionality. Furthermore, there are some recommender systems that use personality information to increase the accuracy of their recommendations. These recommender systems have better performance in comparison with the conventional recommendation methods, particularly ones that should directly handle data sparsity and cold start problems.

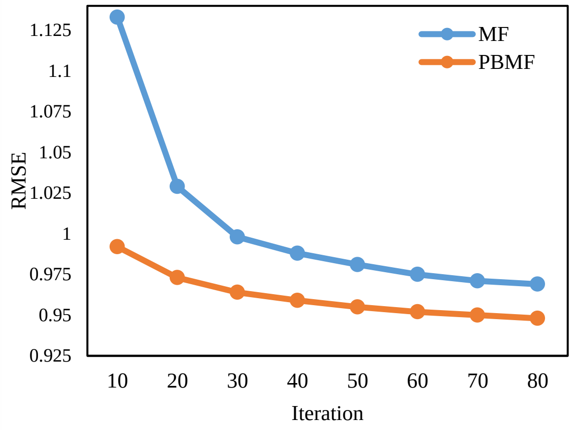


Fig. 3. Comparison of MF and PBMF in different iterations.

In this paper, we proposed a new matrix factorization method for recommender systems named Personality-Based Matrix Factorization (PBMF). Experiments on one of the most well-known standard dataset, called MovieLens 100k, showed the ability of personality information in terms of optimism and pessimism in empowering the matrix factorization algorithm. The main point of identifying the optimism and pessimism of user's personality is to use it as a penalty function in the matrix factorization process, so that the predictions for a user would be closer to their personality. As a result, predicted ratings are more similar to reality and recommendations are more accurate. In this regard, MAE and RMSE on MovieLens dataset were calculated for PBMF and results indicate that their values were smaller than the other related algorithms in this area. Our proposed method improved the system's performance and generated accurate recommendations.

[DOI: 10.52547/itrc.14.1.48]

Our proposed method, Personality-Based Matrix Factorization, can be highly beneficial in scenarios where the only available data source is the rating matrix, including ratings given by users for items, such as movie or music review websites. It can reduce the error rate of predictions and recommendations by extracting human's behavioral factors, which is referred to as personality, without having access to additional information. This approach of extracting personality information and using them to enhance the usability of the matrix factorization technique can be developed further with more diverse data sources and new perspectives of analyzing human's decision making.

For further studies, the minimum ratings that a user should give can be considered, besides, personality information can be applied in other types of recommender systems too.

As another future work, some sources of item information can be exploited. For example, one probable resource is item reviews. Accordingly, different kinds of Natural Language Processing (NLP) methods could be applied to extract useful information related to various items.

VII. Acknowledgement

The authors would like to thank Dr. Belmond Yoberd, Divisional Director, Mott MacDonald Limited, United Kingdom for this literature review contribution.

VIII. References

[1] Mehta, Rachana; Rana, Keyur, "A review on matrix factorization techniques in recommender systems," in *2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA)*, Mumbai, India, 2017.

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

[3] Sarwar, Badrul; Karypis, George; Konstan, Joseph; Riedl, John, "Application of Dimensionality Reduction in Recommender System - A Case Study," Minnesota Univ Minneapolis Dept of Computer, 2000.

[4] Koren, Yehuda; Bell, Robert; Volinsky, Chris, "Matrix factorization techniques for recommender," *Computer*, vol. 42, no. 8, pp. 30-37, 2009.

[5] Luo, Xin; Zhou, Mengchu; Xia, Yunni; Zhu, Qingsheng, "An Efficient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1273-1284, 2014.

[6] A. Charu C., Recommender Systems, Springer International Publishing, 2016.

[7] Schafer, J. Ben; Frankowski, Dan; Herlocker, Jon; Sen, Shilad, "Collaborative Filtering Recommender Systems," in *The Adaptive Web. Lecture Notes in Computer Science*, Berlin, Heidelberg, Springer, 2007, pp. 291-324.

[8] Tohidi, Nasim; Dadkhah, Chitra, "Improving the performance of video Collaborative Filtering Recommender Systems using Optimization Algorithm," *International Journal of Nonlinear Analysis and Applications*, vol. 11, no. 1, pp. 283-295, 2020.

[9] Hu, Rong; Pu, Pearl, "Enhancing collaborative filtering systems with personality information," in *Association for Computing Machinery*, New York, United States, 2011.

[10] Khwaja, Mohammed; Ferrer, Miquel; Iglesias, Jesus Omana; Faisal, A Aldo; Matic, Aleksandar, "Aligning daily activities with personality: towards a recommender system for improving wellbeing," in *13th ACM Conference on Recommender Systems*, 2019.

[11] Abdi, Mohamed Hussein; Okeyo, George Onyango; Mwangi, Ronald Waweru, "Matrix Factorization Techniques for Context-Aware Collaborative Filtering Recommender Systems: A Survey," *Computer and Information Science*, vol. 11, no. 2, pp. 1-10, 2018.

[12] Gupta, Sonali; Gulati, Payal; Bhatia, Surbhi; Madaan, Rosy, "An automatic approach to music recommendations based on individual personality traits," in *the International Conference on Innovative Computing & Communications (ICICC)*, Delhi, India, 2020.

[13] Cheng, Chen; Xia, Fen; Zhang, Tong; King, Irwin; Lyu, Michael R., "Gradient boosting factorization machines," in *8th ACM Conference on Recommender systems*, 2014.

[14] Son, Juyeon; Choi, Wonyoung; Choi, Sang-Min, "Trust information network in social Internet of things using trust-aware recommender systems," *International Journal of Distributed Sensor Networks*, vol. 16, no. 4, 2020.

[15] Feng, He; Qian, Xueming, "Recommendation via user's personality and social contextual," in *22nd ACM international conference on Information & Knowledge Management*, 2013.

[16] Khan, Euna Mehnaz; Mukta, Md Saddam Hossain; Ali, Mohammed Eunus; Mahmud, Jalal, "Predicting Users' Movie Preference and Rating Behavior from Personality and Values," *ACM Transactions on Interactive Intelligent Systems*, vol. 10, no. 3, pp. 1-25, 2020.

[17] Gao, Rui; Hao, Bibo; Bai, Shuotian; Li, Lin; Li, Ang; Zhu, Tingshao, "Improving user profile with personality traits predicted from social media content," in *7th ACM conference on recommender systems*, 2013.

[18] Beigi, Ghazaleh; Tang, Jiliang; Liu, Huan, "Signed link analysis in social media networks," in *10th International Conference on Web and Social Media*, Cologne, Germany, 2016.

[19] George, Thomas; Merugu, Srujana, "Ascalable collaborative filtering framework based on co-clustering," in *Fifth IEEE International Conference on Data Mining (ICDM'05)*, Houston, TX, USA, 2005.

[20] Yakhchi, Shahpar; Ghafari, Seyed Mohssen; Orgun, Mehmet, "TAP: A Two-Level Trust and Personality-Aware Recommender System," in *International Conference on Service-Oriented Computing*, Dubai, United Arab Emirates, 2020.

[21] Zhang, Haiyang; Ganchev, Ivan; Nikolov, Nikola S.; Ji, Zhanlin; O'Droma, Máirtín, "FeatureMF: An Item Feature Enriched Matrix Factorization Model for Item Recommendation," *IEEE Access*, vol. 9, pp. 65266-65276, 2021.

[DOI: 10.52547/itrc.14.1.48]

[22] Tang, Jiliang; Aggarwal, Charu; Liu, Huan, "Recommendations in signed social networks," in *The 25th International Conference on World Wide Web*, 2016.

[23] C. H. Goulden, Methods of Statistical Analysis, vol. 31, Wiley, 1956, pp. 989-990.

[24] W. Ford, "Algorithms," in *Numerical Linear Algebra with Applications*, Elsevier, 2015, pp. 163-179.

[25] Christensen, Jesper; Bastien, Christophe, "Introduction to General Optimization Principles and Methods," in *Nonlinear Optimization of Vehicle Safety Structures*, Elsevier, 2016, pp. 107-168.

[26] Harper, F. Maxwell; Konstan, Joseph A., "The movielens datasets: History and context," *ACM Transactions on Interactive Intelligent Systems*, vol. 5, no. 4, pp. 1-19, 2015.

[27] Chan, Simon; Treleaven, Philip, "Continuous Model Selection for Large-Scale Recommender Systems," in *Handbook of Statistics*, Elsevier, 2015, pp. 107-124.

[28] Beigi, Ghazaleh; Ranganath, Suhas; Liu, Huan, "Signed Link Prediction with Sparse Data: The Role of Personality Information," in *The 2019 World Wide Web Conference*, 2019.

[29] Sagdic, A.; Tekinbas, C.; Arslan, E.; Kucukyilmaz, Tayfun, "A Scalable K-Nearest Neighbor Algorithm for Recommendation System Problems," in *43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2020.

Mazyar Ghezelji obtained his B.Sc. in Software Engineering from Urmia University, Urmia, Iran in 2019, and then spent one year at K. N. Toosi University of Technology studying Artificial Intelligence under the supervision of Dr. Dadkhah. He is currently continuing his graduate studies at the University of Victoria, Victoria, British Columbia, Canada. His key research interests are Recommender Systems, Machine Learning, Software Engineering, and Bioinformatics.

Chitra Dadkhah received the B.Sc. degree in Software Engineering from Shahid Beheshti University, Tehran, Iran, in 1990, the M. Sc. degree in Computer Engineering, Artificial Intelligence, from the IASI Department, University of Paris 11, Orsay, France, 1993. She also received the Ph.D. degree in Computer Engineering, Artificial Intelligence, from the Department of Computer & IT, Amirkabir University (Polytechnic), Tehran, Iran, 2005. She is Assistant Professor in Computer Engineering Faculty of K. N. Toosi University of Technology, Tehran, Iran. Her research interests include Evolutionary & Swarm Algorithms, Natural Language Processing, Recommender Systems, Robotic (simulation).

Nasim Tohidi received the B.Sc. and M.Sc. degrees in Computer Engineering (software) from Alzahra University, Tehran, Iran, in 2015 and 2018, respectively. Now, she is a Ph.D. candidate in Computer Engineering (Artificial Intelligence) at K. N. Toosi University of Technology, Tehran, Iran. She has been awarded honorary admission to the graduate program and also received INEF Scholarships for having high academic records and achievements. Her research interests include Natural Language Processing and Evolutionary Algorithms.

Alexander Gelbukh is currently a Research Professor and the Head of the Natural Language Processing Laboratory, Center for Computing Research, Instituto Politécnico Nacional, Mexico, and an Honorary Professor of Amity University, India. He has authored or coauthored more than 500 publications in computational linguistics, Natural Language Processing, and Artificial Intelligence, recently with a focus on sentiment analysis and opinion mining. He is a member of the Mexican Academy of Sciences, a Founding Member of the Mexican Academy of Computing, and a National Researcher of Mexico (SNI) at excellence level 3 (highest). He is the Editor-in-Chief, an associate editor, or an editorial board member of more than 20 international journals, and has been the chair or the program committee chair of over 50 international conferences.

[DOI: 10.52547/itrc.14.1.48]