Incorporating Singular Ratings into Collaborative Filtering

Ruzhi Xu, Shuaiqiang Wang, Xuwei Zheng
School of Computer Science and Technology
Shandong University of Finance and Economics
7366 East 2nd Ring Road, Jinan, 250014 China
{xrz,swang,zxw}@sdufe.edu.cn

Abstract—Collaborative filtering (CF) is an effective technique addressing the information overload problem, where each user is represented as a set of rating scores on items. Given a target user, conventional CF algorithms measure similarity between two users by utilizing each pair of rating scores on common rated items but discarding scores rated by either of them. In this paper, we called the former as dual ratings while the latter as singular ratings. Our experiments show that only about 10% ratings are dual ones and can be used for similarity evaluation while the left 90% are singular ones and discarded. For making full use of the limited data resource, in this paper, we present SingCF, which attempts to incorporate singular ratings for accuracy improvement of CF algorithms. In particular, we first estimate the unrated scores for singular ratings and transform them into dual ones. Then we perform a CF process to discover neighborhood users and make predictions for each user. Experiments in comparison with the state-of-the-art methods demonstrate the promise of our approach.

Keywords: Collaborative Filtering, Ranking-Oriented CF, Recommender Systems

1. Introduction

Ever since the thriving of the Web, the world has been flooded with an overwhelming amount of information, which represents one of today’s major challenges on the Web. As an effective technique addressing the problem, recommender systems attempt to make predictions and recommendations based on a large collection of users’ historical ratings, and become a de facto standard and must-own tool for e-commerce to promote business and help customers find products [1]. Prominent examples include eBay, Amazon, Last.fm, Netflix, Facebook, and LinkedIn.

Collaborative filtering (CF) is one of the most successful approaches to build recommender systems. CF algorithms are based on the assumption that users will rate or act on other items similarly if they have rated items similarly or had similar behaviors [2], [3]. CF utilizes the user-item rating matrix to make predictions and recommendations, avoiding the need of collecting extensive information about items and users. In addition, CF can be easily adopted in different recommender systems without requiring any domain knowledge [4].

Given the effectiveness and convenience, many CF methods have been proposed, which fall into two categories: memory-based [3], [5], [4], [6] and model-based [7], [8], [9], [10]. Memory-based methods make predictions based on similarities between users or items, while model-based methods estimate or learn a model to make predictions.

In this study, we focus on memory-based CF. In comparison with model-based CF, memory-based algorithms are relatively easy to implement with strong robustness and comparable effectiveness [11]. Besides, many commercial systems such as Amazon.com are memory-based.

Memory-based CF algorithms can be rating-oriented or ranking-oriented [6]. Rating-oriented methods predict rating scores for items based on users’ historical rating scores on items, while ranking-oriented techniques predict a ranking of items based on users’ preferences on items.

Generally, memory-based CF, either rating-oriented or ranking-oriented, works in the following two phases: (I) discovery of neighborhood users and (II) prediction for recommendation. For each user, Phase I discovers a set of most similar users as the neighborhood users, based on which Phase II predicts rating scores or preferences on items for recommendation purpose.

In existing memory-based CF algorithms, each user is represented as a set of scores on rated objects, either items or preferences. Phase I measures similarity between users by utilizing each pair of rating scores on common rated objects but discarding scores rated by either of them. In this paper, we called the former as dual ratings while the latter as singular ratings.

For example, Pearson correlation coefficient [3], [12], a widely used similarity measure in rating-oriented CF, is based on two users’ rating scores on the set of common items [3], [5]. Kendall tau correlation coefficient [13], a popular similarity measure in ranking-oriented CF, is based on two users’ preferences on the same set of items [4], [6].

Our experimental results on two movie rating datasets show that only about 10% ratings are dual ones and can be used for similarity evaluation in CF algorithms, while the left 90% are singular ones and discarded. Since data sparsity is one of the most acute challenges in CF [14], it
merged to be an important issue to explore a practical way of making full use of the limited data resource, especially the singular ratings.

In light of this, in this study, we present SingCF, which attempts to incorporate singular ratings for accuracy improvement of CF algorithms. In particular, we first estimate the unrated scores for singular ratings and transform them into dual ones. Then we perform a CF process to discover neighborhood users and make predictions for each user. Besides, we prove the equivalence of the similarity measure and the prediction formula between rating-oriented and ranking-oriented CF algorithms, with which ranking-oriented algorithms can be directly built based on rating-oriented techniques. Then we implement two versions of our SingCF algorithms for validation, a rating-oriented and a ranking-oriented. Experiments in comparison with the state-of-the-art methods demonstrate the promise of our approach.

Why can SingCF achieve improvement in recommendation accuracy? Among the additional data introduced by SingCF, half of them are the ground truth scores rated by users, i.e., the singular ratings. Besides, the mean error of the estimated scores in the other half group are quite low, which is only less than 20% higher than that of the final predictions. Resulting from the additional high-quality data, SingCF can achieve more accurate performance.

The effectiveness of SingCF can be understood from another perspective. The unrated ratings of the singular ones are the missing values. Data cleaning is an essential technique in data mining, and one possible step is to attempt to fill in the missing values automatically with a measure of central tendency [15]. SingCF proposes an effective method to fill in the missing values for singular ratings and effects on predicting accuracy, as it would for recommendation accuracy.

We make the following contributions. (1) We propose SingCF, a collaborative filtering algorithm which incorporates singular ratings for making full use of the limited data resource and improving recommendation accuracy. (2) We prove the equivalence of the similarity measure and the prediction formula between ranking-oriented and rating-oriented CF, with which ranking-oriented algorithms can be directly built based on rating-oriented techniques.

The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 presents the preliminaries on memory-based collaborative filtering. Section 4 proposed the SingCF algorithm. Section 5 reports the experimental results. Section 6 concludes the paper.

2. Related Work

Given the effectiveness and convenience, many collaborative filtering (CF) algorithms have been proposed, which fall into two categories: memory-based or model-based. Memory-based methods make predictions based on similarities between users or items, while Model-based methods estimate or learn a model to make predictions.

2.1 Memory-based CF

Memory-based CF algorithms can be rating-oriented or ranking-oriented. Rating-oriented methods recommend items for users based on their historical rating scores on items. The user-based paradigm [3], [5] is more common, which estimates the unknown ratings of a target user based on the ratings by a set of neighboring users that tend to rate similarly to the target user. In the item-based paradigm [16], [17], item-item similarity is used to select a set of neighboring items that have been rated by the target user and the ratings on the unrated items are predicted based on his ratings on the neighboring items. Since the number of items is usually much less than the number of users in most applications, item-item similarities are less sensitive to the data sparsity problem.

Ranking-oriented methods are able to capture the preference similarity between users even if their rating scores differ significantly. Recently, the formulation of recommendation problem is shifting away from rating-oriented to ranking-oriented [18]. EigenRank [4] measured the similarity between users with Kendall tau rank correlation coefficient for neighborhood selection, predicted the relative preferences of items with the preference function, and aggregated these preferences into a total ranking. VSRank [6] introduced a novel degree-specialty weighting scheme to ranking-oriented CF based on vector space model.

Many commercial systems such as Amazon.com are memory-based since they are relatively easy to implement with strong robustness and comparable effectiveness [11]. In this study, we focus on memory-based CF, incorporating singular ratings to seek accuracy improvement of CF algorithms.

2.2 Model-based CF

Model-based CF algorithms can also be classified into rating-oriented and ranking-oriented. As a conventional CF paradigm, many rating-oriented algorithms have been proposed. For example, Shani et al. [7] used a Markov decision processes (MDPs) model for recommender systems, which viewed the recommendation process as a sequential optimization problem. Si and Jin [8] presented a flexible mixture model (FMM) for collaborative filtering. FMM is an extension of partitioning/clustering algorithms, which clusters both users and items together simultaneously without assuming that each user and item should only belong to a single cluster. Comprehensive surveys of rating-oriented CF can be found in [19], [20], [14].

As a new CF paradigm, Ranking-oriented techniques also received attentions recently. For example, Weimer et al. [9] used maximum margin matrix factorization to optimize ranking of items for collaborative filtering. Liu et al. [10]
adopted a probabilistic latent preference analysis model that made ranking predictions by directly modeling user preferences with respect to a set of items rather than the rating scores on individual items. Rendle et al. [21] proposed a Bayesian probabilistic model for personalized ranking from implicit feedback. Sun et al. [22] defined novel content-based and rating-oriented meta-level features, and adapted learning to rank to hybrid recommender systems.

3. Memory-based Collaborative Filtering

Memory-based collaborative filtering (CF) algorithms can be classified into two categories: rating-oriented and ranking-oriented.

3.1 Rating-Oriented CF

Let $U$ be a set of users and $I$ be a set of items. In a recommender system, each user $u \in U$ rates scores on a set of items $I_u \subseteq I$, and each item $i \in I$ is rated by a set of users $U_i \subseteq U$. Let $R_{m \times n}$ be a user-item rating matrix with $m$ users and $n$ items, where each element $r_{u,i} \in \mathbb{N}$ is the rating score of the item $i$ with respect to $u$, and $\mathbb{N}$ is the natural number set indicating different relevance scores. $\overline{r}_u$ is the mean score of user $u$ over all the items rated by $u$. For a target user $u$, a set of neighborhood users $N_u \subset U$ are selected according to their similarity to $u$, based on which the rating scores on the unrated items are predicted.

Rating-oriented CF recommends items based on historical rating scores of items. User-user paradigm is a widely model for rating-oriented CF, where the Pearson correlation coefficient is used to evaluate the similarity $s_{u,v}$ between two users $u$ and $v$ with their normalized ratings on the set of common items $I_{u,v} = I_u \cap I_v$: \[
s_{u,v} = \sum_{i \in I_{u,v}} \frac{(r_{u,i} - \overline{r}_u)(r_{v,i} - \overline{r}_v)}{\sqrt{\sum_{i \in I_{u,v}} (r_{u,i} - \overline{r}_u)^2 \sum_{i \in I_{u,v}} (r_{v,i} - \overline{r}_v)^2}} = \frac{\langle v, u \rangle}{\sqrt{\langle v, v \rangle \langle v, u \rangle}} \tag{1}\]

The rating of item $i$ for user $u$ can be predicted by the scores of $i$ rated by a set of neighborhood users $N_u$ of $u$: \[
\hat{r}_{u,i}^{(P)} = \overline{r}_u + \frac{\sum_{v \in N_u} s_{u,v}(r_{v,i} - \overline{r}_v)}{\sum_{v \in N_u} s_{u,v}} \tag{2}\]

3.2 Ranking-Oriented CF

Ranking-oriented CF predicts users’ preference ranking on items to make recommendations based on their relative preferences on common pairs of items [4, 6].

In ranking-oriented CF, Kendall tau correlation coefficient [13] is a common measure for evaluating similarity between two users based on users’ preference scores on the same set of the pairwise preferences on items: \[
s_{u,v} = \frac{N_c - N_d}{\frac{1}{2}k(k-1)} \tag{4}\]

where $N_c$ and $N_d$ are the numbers of the concordant pairs and discordant pairs respectively, and $N_c + N_d = \frac{1}{2}k(k-1)$.

Let $p_{u,(i,j)} \in \{+1, -1\}$ be the preference function, where $p_{u,(i,j)} = +1$ indicates user $u$ prefers item $i$ to $j$ while $p_{u,(i,j)} = -1$ indicates $u$ prefers $j$ to $i$, formally: \[
p_{u,(i,j)} = \begin{cases} +1, & \text{if } r_{u,i} > r_{u,j} \\ -1, & \text{if } r_{u,i} < r_{u,j} \end{cases} \tag{5}\]

Let $N_u(i,j)$ is the set of neighborhood users of $u$ who hold certain preferences on the pair of items $i$ and $j$. For user $u$, the preference on a pair of items $(i, j)$ can be predicted with a preference function $\Psi_u(i, j)$ as follows. \[
\Psi_u(i, j) = \sum_{v \in N_u(i,j)} s_{u,v}p_{u,(i,j)} \tag{6}\]

Based on the predicted pairwise preferences, a total ranking of items for user $u$ can be obtained by applying a preference aggregation algorithm.

4. The SingCF Algorithm

In this section, we first investigate the equivalence between two paradigms of CF algorithms, rating-oriented and ranking-oriented. Then we mainly introduce our algorithm SingCF based on rating-oriented CF. The ranking-oriented version of SingCF can be easily implemented resulting from the equivalence between two CF paradigms.

4.1 Ranking-Oriented CF vs. Rating-Oriented CF

In ranking-oriented CF, each user $u$ is represented as a set of relative preferences on pairs of items. In this study we propose another representation based on user-(item-item) preference matrix, which is defined as follows:
Definition 1: Let $R_{m \times n}$ be the user-item score matrix with $m$ users and $n$ items. The matrix $P_{m \times n(n-1)}$ is called a user-(item-item) preference matrix with $m$ users and $n(n-1)$ preferences on all possible pairs of items, where each element $p_{u,(i,j)} = \{+1, -1\}$ indicates the preference of the target user $u$ on the pair of items $i$ and $j$, as shown in Equation (5).

According the definition of the user-(item-item) preference matrix $P_{m \times n(n-1)}$, the following corollary is satisfied:

Corollary 1: The average rating scores of each user is 0, formally $\forall u \in U : \overline{r}_u = 0$.

Since for each user $u \in U$ who has rated a pair of items $i$ and $j$, $p_{u,(i,j)} + p_{u,(j,i)} = 0$, and thus Corollary 1 holds.

Theorem 1: The ranking-oriented similarity measure of the Kendall tau correlation coefficient based on the common preference sets is equivalent to the Pearson correlation coefficient based on the representation of the user-(item-item) preference matrix.

Proof: According to the definition of the user-(item-item) preference matrix $P_{m \times n(n-1)}$, $p_{u,(i,j)} = 1$.

Let $I_C$ be the set of items where both $u$ and $v$ hold certain preference on each pair of items, formally, $I_C \subseteq I_{u,v} \land \forall i,j \in I_C \rightarrow p_{u,(i,j)}p_{v,(i,j)} \neq 0$. Based on Corollary 1, the Pearson correlation coefficient (Equation 1) can be rewritten as follows:

$$s_{u,v} = \frac{\sum_{i,j \in I_C} (p_{u,(i,j)} - \overline{r}_u)(p_{v,(i,j)} - \overline{r}_v)}{\sqrt{\sum_{i,j \in I_C} (p_{u,(i,j)} - \overline{r}_u)^2 \sum_{i,j \in I_C} (p_{v,(i,j)} - \overline{r}_v)^2}} = \frac{\sum_{i,j \in I_C} p_{u,(i,j)}p_{v,(i,j)}}{\sqrt{\sum_{i,j \in I_C} p_{u,(i,j)}^2 \sum_{i,j \in I_C} p_{v,(i,j)}^2}} = \frac{\sum_{i,j \in I_C} p_{u,(i,j)}p_{v,(i,j)}}{k(k-1)}$$

For a target pair of items $(i, v)$ and two users $u$ and $v$, $p_{u,(i,j)}p_{v,(i,j)}$ can be considered as an indicator function, such that it is equal to $+1$ if the preference on items $i$ and $j$ is concordant in users $u$ and $v$ while $-1$ if the preference is discordant. Formally,

$$p_{u,(i,j)}p_{v,(i,j)} = \begin{cases} +1, & \text{if } (r_{u,i} - r_{u,j})(r_{v,i} - r_{v,j}) > 0 \\ -1, & \text{if } (r_{u,i} - r_{u,j})(r_{v,i} - r_{v,j}) < 0 \end{cases}$$

Thus the sum of $p_{u,(i,j)}p_{v,(i,j)}$ for all possible item pairs is $2(N_c - N_d)$, and $s_{u,v}$ can be rewritten as follows:

$$s_{u,v} = \frac{2(N_c - N_d)}{k(k-1)},$$

which is equivalent to the Kendall tau correlation coefficient based on the common preference sets (Equation 4).

Theorem 2: The ranking-oriented prediction formula based on the common preference sets is equivalent to the rating-oriented prediction formula based on the representation of the user-(item-item) preference matrix.

Proof: Based on Corollary 1, the ranking-oriented prediction formula based on $P_{m \times n(n-1)}$ (Equation 2) can be rewritten as follows:

$$r^{(P)}_{u,(i,j)} = \overline{r}_u + \frac{\sum_{v \in N_u} s_{u,v}(r_{v,(i,j)} - \overline{r}_v)}{\sum_{v \in N_u} s_{u,v}}$$

which is the very formula of the rating-oriented prediction formula based on the common preference sets (Equation 6).

Besides, for memory-based algorithms, ranking-oriented CF performs same as rating-oriented CF, where a set of most similar users are discovered as the neighborhood users, based on which the scores/relationships of the unrated items/preferences are predicted for recommendation.

Thus, a ranking-oriented CF is equivalent to a rating-oriented CF based on user-(item-item) preference matrix.

In this section, we mainly introduce our algorithm SingCF based on the rating-oriented CF. The ranking-oriented version can be easily implemented resulting from the equivalence of the similarity measure and the prediction formula between ranking-oriented and rating-oriented CF algorithms.

4.2 Estimation for Singular Rating

In existing memory-based CF algorithms, the similarity between users is evaluated by with dual ratings without considering any singular one. The dual ratings are pairs of scores on certain items rated by a pair of users, and the singular ratings are scores rated by either of two users.

Definition 2: Let $I_u$ and $I_v$ be two sets of items rated by user $u$ and $v$ respectively. Let $I_{u,v} = I_u \cap I_v = \{i_1, i_2, \ldots, i_k\}$ be the common set of items
rated by users \( u \) and \( v \). All pairwise rating scores rated by \( u \) and \( v \) on \( I_{u,v} \) are called dual ratings, i.e., \( \{(r_{u,i_1}, r_{v,i_1}), (r_{u,i_2}, r_{v,i_2}), \ldots, (r_{u,i_k}, r_{v,i_k})\} \). The rating scores rated by \( u \) on items \( I_u \setminus I_{u,v} \) and those by \( v \) on \( I_v \setminus I_{u,v} \) are called singular ratings.

Example 1: Let \( \{i_1, i_2, i_3, i_4\} \) be four items, Let \( \{u, v\} \) be two users, where their rating scores are listed as follows:

\[
\begin{align*}
    r_{u,i_1} &= 0, r_{u,i_2} = 1, r_{u,i_3} = 2, \\
    r_{v,i_2} &= 3, r_{v,i_3} = 4, r_{v,i_4} = 5.
\end{align*}
\]

According to our definitions, \( (r_{u,i_2}, r_{v,i_2}) = (1, 3) \) and \( (r_{u,i_3}, r_{v,i_3}) = (2, 4) \) are dual ratings. \( r_{u,i_1} = 0 \) and \( r_{v,i_4} = 5 \) are singular ones.

For incorporating singular ratings, a most straightforward way is to estimate their corresponding unrated scores and transform the singular ratings into dual ones. In Example 1, \( r_{u,i_1} \) and \( r_{u,i_3} \) are corresponding unrated scores of \( r_{u,i_1} \) and \( r_{v,i_4} \) respectively.

In doing this, accurate estimation is very important in SingCF, which can avoid introducing plenty of noises into the rating matrix \( R \).

In this study, we fully consider the historical rating scores assigned by the target user \( u \) to items \( I_u \), and the scores assigned by users \( U_i \) to the target item \( i \):

\[
\tilde{r}_{u,i} = \overline{r}_u + \frac{\sum_{v \in U_i} (r_{v,i} - \overline{r}_v)}{|U_i|} \quad (7)
\]

\[
\tilde{u}_{i} = \frac{\sum_{i \in I_u} r_{u,i}}{|I_u|} = \overline{r}_u \quad (8)
\]

In particular, Equations (7) and (8) are special cases of Equations (2) and (3) respectively, where all users \( U_i \) who have rated \( i \) are considered as the neighborhood users of \( u \), all items \( I_u \) rated by \( u \) as the neighborhood items of \( i \), and all similarities are 1.

In SingCF, we linearly combine \( \tilde{r}_{u,c} \) and \( \tilde{r}_{u,c} \) to estimate the unrated ratings for singular data:

\[
\begin{align*}
    r_{u,i}^{(E)} &= \alpha \cdot \tilde{r}_{u,i} + (1 - \alpha) \cdot \tilde{u}_i \\
    &= \alpha \cdot \left( \overline{r}_u + \frac{\sum_{v \in U_i} (r_{v,i} - \overline{r}_v)}{|U_i|} \right) + (1 - \alpha) \cdot \overline{r}_u \\
    &= \overline{r}_u + \alpha \cdot \frac{\sum_{v \in U_i} (r_{v,i} - \overline{r}_v)}{|U_i|}
\end{align*}
\]

where \( U_i \) is the set of users who has rated \( i \), \( \alpha \) is a real number and \( 0 \leq \alpha \leq 1 \).

A very interesting observation is that Equation (9) is very similar to Equation (2), where the first part measures the average scores assigned by \( u \), and the second part evaluates the quality of \( i \), whether its rating is higher than the average score or lower. These two equations are the same when the following conditions hold:

- The similarity between \( u \) and each user who has rated \( i \) is “1”, formally, \( \forall v \in U_i : s_{u,v} = 1 \).
- Each user who has rated \( i \) is considered as a neighborhood user of \( u \), formally, \( \forall v \in U_i : v \in N_u \), and \( N_u \equiv U_i \).
- The parameter \( \alpha \) in Equation (9) is “1”.

The first two conditions are easily understood. We have no prior knowledge on similarity between users, and have to assign each to “1”. Thus each user who has rated \( i \) is considered as a neighbor of \( u \).

Since not all of users in \( U_i \) are similar to the target user \( u \), we lack full confidence on the quality estimation on \( i \), and \( \alpha \) can be regarded as a confidence rate.

Thus if the first two conditions are satisfied, Equation (9) generalizes (2), and they are the same if the confidence rate \( \alpha = 1 \).

Similarly, for ranking-oriented CF, let \( U_{i,j} \) be the set of users who hold certain preferences on items \( i \) and \( j \), i.e., \( U_{i,j} \subseteq U_i \cap U_j \land \forall v \in U_{i,j} \rightarrow p_{u,v(i,j)} \neq 0 \). According to Corollary 1, for each user, \( \overline{r}_u = 0 \). The estimating formula of Equation (9) can be rewritten as follows:

\[
f(U_{i,j}) = \alpha \cdot \frac{\sum_{v \in U_{i,j}} p_{v(i,j)}}{|U_{i,j}|} \quad (10)
\]

\( f(U_{i,j}) \) returns a real number while the preference on \( i, j \) should be +1 or -1. Let \( \theta \) be a threshold where \( 0 \leq \theta \leq 1 \). The preference on \( (i, j) \) can be estimated as follows:

\[
p_{u(i,j)} = \begin{cases} 
    +1, & \text{if } f(U_{i,j}) > \theta \\
    -1, & \text{if } f(U_{i,j}) < -\theta 
\end{cases} \quad (11)
\]

4.3 Discussion

The effectiveness of SingCF can be understood from the following two perspectives.

- **Utilizing additional high-quality data.** Among the additional data introduced by SingCF, half of them are the ground truth scores rated by users, i.e., the singular ratings. Besides, the mean error of the estimated scores in the other half group are quite low, which is only less than 20% higher than that of the final predictions. Resulting from the additional high-quality data, SingCF can achieve more accurate performance.

- **Data cleaning.** The unrated ratings of the singular ones are the missing values. Data cleaning is an essential technique in data mining, and one possible step is to attempt to fill in the missing values automatically with
a measure of central tendency [15]. SingCF proposes an effective method to fill in the missing values for singular ratings and effects on predicting accuracy, as it would for recommendation accuracy.

Thus SingCF works in three phases: (I) data cleaning, (II) neighbor discovery and (III) rating prediction. First of all, Phase I estimates the unrated rating scores and transforms singular ratings into dual ones. Then for each user, Phase II discovers a set of most similar users as the neighborhood users. Then Phase III makes predictions and recommendations based on the rating scores of the neighborhood users.

5. Experiments

We used MovieLens\textsuperscript{1}, a real movie rating datasets in our experiments. In our experiments, we randomly selected 80\% rated items for training and used the remaining 20\% for testing. In order to guarantee that there are adequate number of common rating items between each neighborhood user and the target user, we filtered those users who have rated less than 50 items. We ran each algorithm 5 times and reported the average performance.

5.1 Rating-Oriented CF

For rating-oriented collaborative filtering (CF), we used two standard evaluation criterions, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), to evaluate our algorithm.

MAE is the most widely used evaluation metric in CF research literature. Let $R_T$ be the test data, where each element $r_{u,i}^{(T)} \in R_T$ is the score assigned by user $u$ to item $i$. MAE computes the average of the absolute difference between the predictions and true ratings:

$$\text{MAE} = \frac{\sum_{r_{u,i}^{(T)} \in R_T} |r_{u,i}^{(T)} - r_{u,i}^{(P)}|}{|R_T|}$$

RMSE is another popular metric, which was used in the Netflix prize\textsuperscript{2} for movie recommendation performance:

$$\text{RMSE} = \sqrt{\frac{\sum_{r_{u,i}^{(T)} \in R_T} (r_{u,i}^{(T)} - r_{u,i}^{(P)})^2}{|R_T|}}$$

We chose a conventional user-based CF method as our comparison partner, which measured similarity between users using the Pearson correlation coefficient, predicted ratings using Equation (2). Our implementation of SingCF was based on the conventional CF. A direct comparison of the two will provide valuable and irreplaceable insights.

Table 1 shows the performance comparisons under the MAE and RMSE measures. From the table we can see that our proposed SingCF algorithm can achieve significant accuracy improvement, resulting from effectively incorporate singular ratings.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>0.8199</td>
<td>1.0546</td>
</tr>
<tr>
<td>SingCF</td>
<td>0.8798</td>
<td>3.8339</td>
</tr>
</tbody>
</table>

5.2 Ranking-Oriented CF

For ranking-oriented CF, the wildly used measure is the Normalized Discounted Cumulative Gain (NDCG) [23] metric, which is popular in information retrieval for evaluating ranked document results.

In the context of collaborative filtering, item ratings assigned by users can naturally serve as graded relevance judgements. Specifically, the NDCG metric is evaluated over some number $n$ of the top items on the ranked item list. Let $U$ be the set of users and $r_{u,p}$ be the rating score assigned by user $u$ to the item at the $p$th position of the ranked list from $u$. The average NDCG at the $n$th position with respect to all users $U$ is defined as follows.

$$\text{NDCG}_{avg} = \frac{1}{|U|} \sum_{u \in U} \sum_{p=1}^{n} \frac{2^{r_{u,p}} - 1}{\log(1 + p)}$$

The value of NDCG ranges from 0 to 1. A higher value indicates better ranking effectiveness. NDCG is very sensitive to the ratings of the highest ranked items. This is modeled by the discounting factor $\log(1 + p)$ that increases with the position in the ranking. This is a highly desirable characteristic for evaluating ranking quality in recommender systems. This is because, just as in web search, most users only examine the first few items from the recommended list. The relevance of top-ranked items are far more important than other items [4].

We used EigenRank [4], a state-of-the-art ranking-based CF algorithm, as our main comparison partner. Our implementation of SingCF was based EigenRank. A direct comparison of the two will provide valuable and irreplaceable insights. Besides, we also used CoFiRank, another state-of-the-art ranking-based collaborative filtering algorithms as our comparison partner. In particular, EigenRank measured similarity between users with Kendall tau rank correlation coefficient for neighborhood selection, predicted pairwise preferences of items with a preference function, and aggregated the predicted preferences into a total ranking with a greedy algorithm. CoFiRank used Maximum Margin Matrix Factorization and employed structured output prediction to directly optimize ranking scores instead of ratings. In addition, we also included comparisons with UPR, a conventional user-based CF method, which measured similarity between users using the Pearson correlation coefficient, predicted ratings using Equation (2), and then ranked the items for each user according to their predicted rating scores for the purpose of obtaining a ranking of items.

Figure 1 shows the performance comparisons under the NDCG measure. From the figure we can see that

---

\textsuperscript{1}http://www.grouplens.org/
\textsuperscript{2}http://www.netflixprize.com/
our proposed SingCF outperformed all comparison partners. In particular, SingCF achieved the best performances on NDCG@3–5, and the second best performances on NDCG@1–2, only slightly lower than CoFiRank but significantly higher than EigenRank and UPR.

6. Conclusion

In this paper, we have proposed SingCF, a collaborative filtering (CF) algorithm which incorporates singular ratings for making full use of the limited data resource and improving recommendation accuracy. In particular, we first estimate the unrated ratings and transform the singular ratings into dual ones. Then we perform the CF process to discover neighborhood users and make predictions for recommendations. There are two paradigms for CF algorithms: rating-oriented and ranking-oriented. In this paper we prove the equivalence between the two paradigms, based on which we also provide two versions of SingCF for rating-oriented and ranking-oriented CF. Experiments have validated the effectiveness of our algorithm.

There are several interesting directions for future work. Firstly, we plan to perform a systematic study on SingCF, investigating the various factors that may affect its performance. Secondly, it is very interesting to study other possible formulae for estimating the unrated ratings for singular data. Last but not least, the proposed algorithm is not limited to ranking-based CF, and we plan to adapt it to model-based CF and examine its effectiveness.

References


