Mining Mood-specific Movie Similarity with Matrix Factorization for Context-aware Recommendation

Yue Shi
Multimedia Information Retrieval Lab
Delft University of Technology
Delft, The Netherlands
y.shi@tudelft.nl

Martha Larson
Multimedia Information Retrieval Lab
Delft University of Technology
Delft, The Netherlands
m.a.larson@tudelft.nl

Alan Hanjalic
Multimedia Information Retrieval Lab
Delft University of Technology
Delft, The Netherlands
a.hanjalic@tudelft.nl

ABSTRACT

Context-aware recommendation seeks to improve recommendation performance by exploiting various information sources in addition to the conventional user-item matrix used by recommender systems. Recommendations should also usually strive to satisfy a specific purpose. Within the Moviepilot mood track of the context-aware movie recommendation challenge, we propose a novel movie similarity measure that is specific to the movie property demanded by the challenge, i.e., movie mood. Our measure is further exploited by a joint matrix factorization model for recommendation. We experimentally validate the effectiveness of the proposed algorithm in exploiting mood-specific movie similarity for the recommendation with respect to several evaluation metrics, demonstrating that it could outperform several state-of-the-art approaches. In particular, mood-specific movie similarity is demonstrated to be more beneficial than general mood-based movie similarity, for the purpose of mood-specific recommendation.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *Information Filtering*

General Terms

Algorithms, Performance, Experimentation

Keywords

Recommender systems, collaborative filtering, context-aware recommendation, matrix factorization, mood-specific movie similarity

1. INTRODUCTION

Recommender systems can be improved by enrichment with various sorts of contextual information, e.g., relationship among users in social media sites, tags of products, background message of products, timestamp of user actions, e.g., [1][3]. Recently, exploiting contextual information for improved recommendation has experienced an upsurge of interest in recommender systems community, e.g., [4][5][19]. Context-aware recommendation differs from traditional recommendation not only in the sense that it makes use of contextual information that is available to be exploited, but also in the sense that the purpose of recommendation is more specific, e.g., recommending movies specifically for Christmas week, or recommending movies specifically featured with an emotional property [15]. [1] Accordingly, the research challenge in the area of context-aware recommendation mainly concerns two aspects: First, it is expected that the new recommendation technique/model could retain the benefits of conventional recommender approaches, e.g., collaborative filtering (CF), which mines recommendation from user-item matrix [2], while also allowing the integration of additional contextual information to exert impact on recommendation. Second, it is expected that the new recommendation technique could be "aware" of the specific purpose for which the recommendation is being generated. In other words, it is especially desired that the specific purpose could be taken into account in the recommendation model.

In this paper, we propose a new recommendation model for context-aware recommendation that specifically addresses these two aspects of the research problem. First, we propose a joint matrix factorization model that not only factorizes the user-item rating matrix but also exploits contextual information (i.e., movie mood in this paper) as a regularization term. In this way, we incorporate into the proposed model the ability to learn from user-item matrix (essentially similar to the conventional CF) but at the same time also allow contextual information to be fused into the recommendation process. Second, we propose a novel similarity measure that could reflect the relationship between items with respect to a specific property (i.e., a specific movie mood in this paper). In this way, the proposed model could direct the recommendation towards the specific purpose (i.e., recommending movies with a specific mood in the paper). Throughout this paper, we extend our study based on the Moviepilot challenge, which will be introduced in detail in next section. Although we focus on mining mood-specific movie similarity with a joint matrix factorization model for context-aware recommendation in this paper, the proposed model could be easily applied to integrate other contextual information for other specific recommendation purposes.

CAMRa2010, September 30, 2010, Barcelona, Spain.

[1] http://www.dai-labor.de/camra2010/challenge

The key contribution of our work lies in proposing a novel mood-specific movie similarity, which is exploited in a joint matrix factorization model for improved context-aware (mood-specific) movie recommendation.

The remainder of the paper is structured as follows. In the next section, we present an overview on the Moviepilot mood track of context-aware movie recommendation challenge (denoted as Moviepilot challenge in the following). Then, in section 3, we summarize related work and position our approach with respect to it. The proposed mood-specific movie similarity and the joint matrix factorization model are described in detail in section 4, after which we present experimental evaluation on the Moviepilot challenge dataset. The last section sums up the key aspects of the proposed algorithm and briefly addresses the direction for future work.

2. Overview of Moviepilot Challenge

2.1 Problem statement

The problem of Moviepilot challenge can be described as: Based on both user-item rating matrix and other provided contextual information, e.g., movie-mood tag matrix, to recommend a list of movies that have specific mood property (indicated by mood tags) to each target/test user [15]. The recommendation list should contain relevant[2] movies as many possible, and as highly ranked as possible. In other words, it is expected that recommended movies need to be interesting to the user but also of the specific mood.

2.2 Characteristics

Here, we summarize several characteristics of this challenge, which are necessarily to mention in order to distinguish it from the traditional collaborative filtering problem.

First and foremost, the evaluation of the challenge is based on recommendation list, e.g., by precision, mean average precision (MAP), and area under curve (AUC) and not based on the rating prediction error rate, e.g., root mean square error used in Netflix contest[3].

Second, the evaluation is based on the recommended movies with the specified mood. Two conceptualizations of the problem are possible. Under the first, one could use known recommendation approaches to generate the initial recommendation, and then remove those movies without the specified mood. However, under this approach the mood-specific demand of users would not actually be exploited during the recommendation process. This situation could lead to recommendation performance far short of what is targeted. Under the second, one could conceptualize the recommendation as not only involving the movies that the user would be interested in, but also emphasizing movies with the specified mood. Our work in this paper indeed follows the latter conceptualization.

Third, apart from a large scale user-item rating matrix, various sorts of other contextual information are provided in this challenge, e.g., mood tags of movies, movie location, intended audience of movies. For this reason, various information sources can contribute to generating recommendations. In this paper, only the mood tags are used.

Fourth, there are quite a few movies in the provided evaluation/validation set that actually have only very low ratings, as can be seen (in Fig. 1) from the distribution of ratings in the pre-defined evaluation set. The implied goal is that these low rated movies should also be recommended to the target user when appropriate. All of the evaluation metrics consider movies of different ratings in the evaluation set as equally relevant. On one hand, this state of affairs suggests that relying solely on rating prediction to generate recommendation might not be sufficient. On the other hand, it is also difficult to apply ranking-oriented approaches [13][14][18], since either a pair-wise ranking or list-wise ranking approach requires training examples annotated with item ratings.

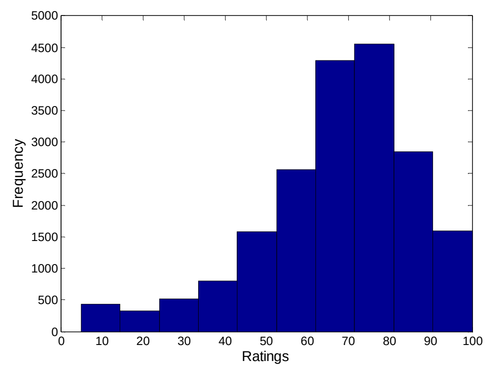

Fig. 1 The distribution of ratings in the evaluation set of Moviepilot challenge

3. RELATED WORK

This section briefly summarizes the existing related research in CF and context-aware recommendation.

3.1 Collaborative Filtering

A comprehensive survey of collaborative filtering approaches can be found in [1][10]. Collaborative filtering is based on the concept that users with similar preference would like the same items in future (i.e., user-based perspective), or items with similar purchase/ consumption history would be liked by users who already like one of those items (i.e., item-based perspective). Various algorithms have been devoted to improving either user-based approaches [9][17][21], or item-based approaches [7][16][20]. However, measuring similarities among users or items on a large scale is computationally expensive, which also degrades the usefulness of these similarities for context-aware recommendation scenario.

Matrix factorization (MF) techniques have attracted much research attention, due to the advantages of scalability and accuracy [12], especially for large-scale data, as exemplified by the Netflix contest. Generally, MF techniques learn latent features of users and items from the observed ratings in the user-item rating matrix. These latent features are further used to predict unobserved ratings. However, MF is specifically rating prediction oriented. Additional consideration needs to be made in order to apply it to recommendation task. The joint matrix factorization framework is an extension of MF. It is widely proposed for different purposes, e.g., fusing document content and graph link

[2] A movie is considered "relevant" for a user if it has been rated by that user and also has the specified mood used for evaluation.

[3] http://www.netflixprize.com

information for document retrieval or web page classification [6][23], fusing geographical location features and people activity correlation for location-based recommendation [22]. In this paper, we also exploit joint matrix factorization to fuse mood-specific movie similarity with factorization of user-movie rating matrix.

Many ranking-oriented approaches have been investigated in the literature as well, including exploitation of pair-wise preference between items for users, e.g., probabilistic latent preference analysis [13] and Bayesian probabilistic ranking [14], as well as exploitation of list-wise preference [18], recently. However, the basis for ranking in those approaches, either pair-wise or list-wise preference, is derived from item ratings, which effectively imposes the assumption that low rated items should be ranked lower than high rated items. This property does not satisfy the requirement in the Moviepilot challenge as mentioned in section 2.2, since there are quite a few relevant movies actually having very low ratings. The difference can also be reflected by the evaluation metrics, i.e., the suggested metrics for Moviepilot challenge, e.g., MAP and AUC, are rating insensitive, while the Normalized Discounted Cumulative Gain (NDCG) used in most of the aforementioned ranking-oriented approaches is rating sensitive. For this reason, we did not extend our study following those approaches, but rather leave them for future work.

3.2 Context-aware Recommendation

Recently, much research effort has been devoted to context-aware recommendation.[4] Basically, contextual information in recommender systems can be regarded as any available information beyond the user-item rating matrix. Contextual information is regarded as additional dimensions to the user and item rating vectors in [1], that are used to obtain item recommendation for a given context. More recently, context-aware recommendation has become more application-specific, e.g., exploiting contextual information for travel recommendation [5], for news recommendation [4], and for music recommendation [19]. Compared to this work, the proposed algorithm in this paper enjoys the advantages of being adaptive to any new application domain and being able to handle large-scale data.

Recommendation models for fusing contextual information, especially for social relationship, have been studied, recently. A random walk with restart model [11] is exploited to generate item recommendation from a hyper-graph, which can include multiple types of nodes and links including user-user friendship, item-tag indication, etc. However, all the links in the hyper-graph are generally considered as binary, which could hardly be common in the general case. For example, in the Moviepilot challenge, the user-movie relationship is reflected by scaled ratings, while the movie-mood relationship is reflected by binary indication. Fusing the two into a hyper-graph is difficult, since simply thresholding the ratings to be binary ones could result in a critical information loss that would negatively influence the recommendation performance.

4. THE ALGORITHMS

In this section, we introduce our proposed algorithm for the context-aware movie recommendation specified in Moviepilot challenge, i.e., recommending movies with specific mood property to the target users. We first present a novel movie (item) similarity, which encodes the connection between two movies w.r.t. the specified mood. Then, we present a joint matrix factorization framework, which integrates mood-specific movie similarity as a regularization modality into conventional user-item rating matrix factorization. A locally optimal solution can be obtained by applying an efficient learning process.

4.1 General Mood-based Movie Similarity

According to traditional item-based CF, item-to-item similarity can be defined as the cosine similarity between two item rating vectors [7]. Similarly, given the movie-mood (binary) matrix M (consisting of N movies and E mood tags), in which M_{jk}=1 indicates that the movie j has the mood tag k, otherwise 0, we can define the general mood-based movie-to-movie similarity between movie j and n, as shown below:

$$S_{jn}^{(Movie)} = \frac{\sum_{k=1}^{E} M_{jk} M_{nk}}{\sqrt{\sum_{k=1}^{E} M_{jk}^2}\sqrt{\sum_{k=1}^{E} M_{nk}^2}} \qquad (1)$$

However, the mood-based similarity only gives general closeness measurement of two movies in terms of all their mood properties. For example as shown in Fig.2, two movies (**A** and **B**) sharing a different mood property could be equally similar to another movie (**D**). If the required mood of movie is specified, this similarity fails to differentiate the two movies. For this reason, we expect that a fruitful approach would involve adjusting the movie-to-movie similarity when the mood is specific. In the following section, we propose a novel mood-specific movie similarity to address this issue.

4.2 Mood-specific Movie Similarity

Rather than the general mood-based movie-to-movie similarity, the mood-to-mood similarity between mood i and k can be also defined as:

$$S_{ik}^{(Mood)} = \frac{\sum_{j=1}^{N} M_{ji} M_{jk}}{\sqrt{\sum_{j=1}^{N} M_{ji}^2}\sqrt{\sum_{j=1}^{N} M_{jk}^2}} \qquad (2)$$

Once the mood-to-mood similarity matrix is obtained, for a given mood m, we can generate a mood-specific (to m) movie-mood matrix, as expressed in Eq. (3):

$$M_{jk}^{(m\text{-}spec)} = \begin{cases} S_{mk}^{(Mood)}, & if\ k \neq m \\ M_{jk}, & if\ k=m \end{cases} \qquad (3)$$

Note that the mood-specific movie-mood matrix $M^{(m\text{-}spec)}$ is not binary. Then, the mood-specific movie-to-movie similarity can be expressed as:

$$S_{jn}^{(m\text{-}spec)} = \frac{\sum_{k=1}^{E} M_{jk}^{(m\text{-}spec)} M_{nk}^{(m\text{-}spec)}}{\sqrt{\sum_{k=1}^{E} M_{jk}^{(m\text{-}spec)2}}\sqrt{\sum_{k=1}^{E} M_{nk}^{(m\text{-}spec)2}}} \qquad (4)$$

For the example in Fig.2, suppose that we are given the specific demand of mood “anxious”. We can then derive mood-specific movie similarity particularly for the mood “anxious”. In contrast with the case of general mood-based similarity, in the case of mood-specific movie similarity the movie **D** is shown to be more similar to the movie **A** than to the movie **B**. It can be observed that the mood-specific movie similarity could help to distinguish

[4] http://ids.csom.umn.edu/faculty/gedas/cars2009

movies according to the specification of the demand of the mood property. Fore this reason, the mood-specific movie similarity could have the potential of further contributing to the context-aware recommendation.

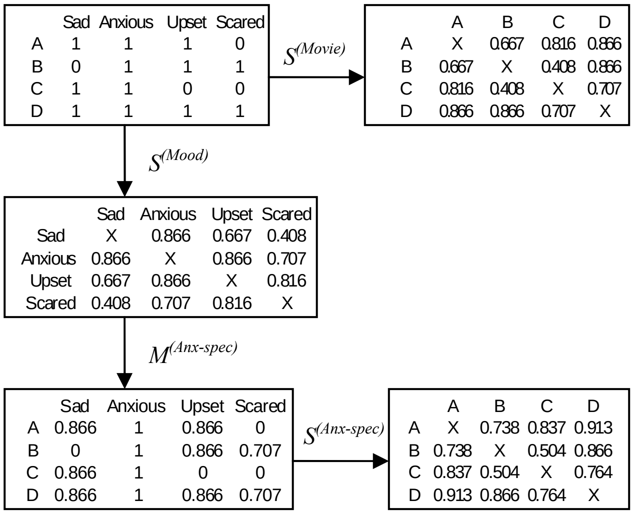

Fig. 2 An illustrative example of mood-specific movie similarity

4.3 Joint Matrix Factorization

The basic matrix factorization [12] can be formulated as in Eq. (5):

$$U,V = \arg\min_{U,V} \frac{1}{2}\sum_{u=1}^{K}\sum_{j=1}^{N} I_{uj}\left(R_{uj} - U_u^T V_j\right)^2 + \frac{\lambda_U}{2}\|U\|_F^2 + \frac{\lambda_V}{2}\|V\|_F^2 \quad (5)$$

Given the user-movie rating matrix R consists of K users and N items, MF represents the user-movie rating matrix R with two low-rank matrices, U and V. A d-dimensional set of latent features is used to represent both users (in U) and items (in V). Note that we use U_u to denote a column d-dimensional feature vector of user u and V_j a column d-dimensional feature vector of movie j, and R_{uj} denotes the user u's rating on movie j. I_{uj} denotes an indicator function that is equal to 1 when $R_{uj} > 0$, and 0 otherwise. Finally, λ_U, λ_V are regularization parameters for which we set $\lambda_U = \lambda_V = \lambda$ for simplification in this paper.

A joint matrix factorization model can be formulated by taking into account the mood-specific movie similarity as an additional regularization modality (denoted as JMF-MS), which is expressed as:

$$L(U,V) = \frac{1}{2}\sum_{u=1}^{K}\sum_{j=1}^{N} I_{uj}\left(R_{uj} - U_u^T V_j\right)^2 + \frac{\alpha}{2}\sum_{j=1}^{N}\sum_{n=1}^{N} J_{jn}\left(S_{jn}^{(m\text{-}spec)} - V_j^T V_n\right)^2 + \frac{\lambda}{2}\left(\|U\|_F^2 + \|V\|_F^2\right) \quad (6)$$

where J_{jn} is also an indicator function that is equal to 1 if $S_{jn}^{(m\text{-}spec)} > 0$, and 0 otherwise. α is a tradeoff parameter for weighting the contribution of regularization by mood-specific movie similarity. The objective function consists of two parts, i.e., factorizing latent user features and movie features in terms of user-movie rating matrix and simultaneously factorizing latent movie features in terms of mood-specific movie similarity matrix, which is independent of rating matrix. The advantages of the joint matrix factorization model are twofold: First, it could alleviate data sparseness problem in the rating matrix where movies with few ratings could benefit from additional modality, i.e., mood-specific movie similarity, for learning latent features. Second, as our purpose is to recommend movies with specified mood tag, the mood-specific movie similarity particularly helps to learn latent movie features with respect to the specified mood. Note that the first advantage could also be realized by using another movie modality, e.g., the general mood-based movie similarity. However, we emphasize that the second advantage is unique in this model. In the experimental section, we compare the benefits from general mood-based movie similarity and mood-specific movie similarity.

Minimization of the objective function can be solved by gradient descent with alternatively fixed U and V. This process results in a local minimum solution. The gradients of L(U,V) with respect to U and V can be computed as:

$$\frac{\partial L}{\partial U_u} = \sum_{j=1}^{N} I_{uj}\left(U_u^T V_j - R_{uj}\right)V_j + \lambda U_u \quad (7)$$

$$\frac{\partial L}{\partial V_j} = \sum_{u=1}^{K} I_{uj}\left(U_u^T V_j - R_{uj}\right)U_u + 2\alpha\sum_{n=1}^{N} J_{jn}\left(V_j^T V_n - S_{jn}^{(m\text{-}spec)}\right)V_n + \lambda V_j \quad (8)$$

Note that in Eq. (8) we exploit the symmetry of $S^{(m\text{-}spec)}$. The details of the algorithm JMF-MS are described in Fig. 3.

Algorithm JMF-MS

Input: user-movie rating matrix R,
mood-specific movie-to-movie similarity $S^{(m\text{-}spec)}$,
tradeoff parameter α,
regularization parameter λ,
maximal number of iterations *iterMax*,
stop condition ε.

Output: Complete user-movie relevance matrix $\hat{R}$.

```
Initialize U^(0), V^(0) with random values;
t=0;
f=0;                          // indicator of convergence
Compute L^(t) as in Eq. (6);
while (t < iterMax and f = 0) do
    η=1;
    Compute ∂L/∂U^(t), ∂L/∂V^(t) as in Eq. (7) and (8)
    while ( L(U^(t) − η ∂L/∂U^(t), V^(t) − η ∂L/∂V^(t)) ≥ L^(t) ) do
        η=η/2;     // maximize learning step size
    U^(t+1) = U^(t) − η ∂L/∂U^(t), V^(t+1) = V^(t) − η ∂L/∂V^(t);
    Compute L^(t+1) as in Eq. (6)
    if (L^(t) − L^(t+1) ≤ ε) do
        f = 1;
    t = t+1;
Return R̂ = U^(t)T V^(t);
```

Fig. 3 Algorithm description for the proposed JMF-MS

4.4 Complexity Analysis

By exploiting the data sparseness, the computation of the objective function in Eq. (6) is of complexity $O(d|R|+d|S^{(m\text{-}spec)}|+d(K+N))$, where $|R|$ denotes the number of observed ratings in a given user-movie matrix, and $|S^{(m\text{-}spec)}|$ denotes the number of non-zero similarities in the mood-specific similarity matrix. The complexity of the gradients in Eq. (7) and Eq. (8) is $O(d|R|+dK)$ and $O(d|R|+d|S^{(m\text{-}spec)}|+dN)$, respectively. Considering the fact that we often have $|R| >> K, N$, the total complexity in one iteration is $O(d|R|+ d|S^{(m\text{-}spec)}|)$. For further simplification, one could only concern the movies with the specified mood tag, since movies without the specific mood are not expected to be recommended. Therefore, the number of movies involved could be much less than the total number of movies, which could lead to $|S^{(m\text{-}spec)}|<< |R|$. In practice, the complexity could be equivalent to $O(d|R|)$, which is linear with the number of observed ratings in the user-movie matrix. This analysis indicates that the proposed algorithm is computationally efficient and can be applied to large-scale cases.

5. EXPERIMENTS AND EVALUATION

In this section, we present the experiments to evaluate the proposed algorithm. The research questions that need to be answered through the experiments can be formulated as follows:

(1) Does minimizing the objective function indeed contribute to improving recommendation performance?

(2) Can the proposed algorithm JMF-MS outperform other state-of-the-art approaches?

(3) How helpful is the mood-specific movie similarity for the purpose of mood-specific recommendation in the Moviepilot challenge?

5.1 Experimental Setup

5.1.1 Dataset

Our experiments are conducted on the dataset of the "Moviepilot mood track", which consists of around 4.5M ratings (scale 0–100) assigned by around 105K users to a collection of around 25K movies. Apart from user-movie rating matrix, various contextual information is provided, e.g., gender and age of users, production year of movies, intended audience of movies, etc. Among all of the contextual information, we only exploit the mood tags of movies in this work, since the recommendation purpose is to recommend movies with a specific mood tag. The movie-mood tag (binary) matrix consists of around 25K movies and 16 mood tags, which in total involves 6,712 mood tag indications on movies.

5.1.2 Evaluation metrics

Four widely used evaluation metrics for measuring the quality recommendation list are adopted, i.e., Precision at top-5 recommended movies (P@5), P@10, mean average precision (MAP) and the area under the (receiver operating) curve (AUC) [8][10]. P@5 reflects the average ratio of the number of relevant movies over the top-5 recommended movies for all users. P@10 is defined similarly as P@5. The definition of MAP is expressed as below:

$$MAP = \frac{1}{K_{ts}} \sum_{u=1}^{K_{ts}} \frac{\sum_{j=1}^{N_u} \left(rel_u(j) \times P_u@j \right)}{\sum_{j=1}^{N_u} rel_u(j)} \tag{9}$$

where K_{ts} is the number of users for testing, and N_u denotes the number of recommended movies for the user u. $rel_u(j)$ is a binary indicator, which equals to 1 if the movie of rank j is relevant to user u, otherwise 0. $P_u@j$ is the precision of the top j recommended movies for the user u, i.e., the ratio of movies in the top j recommendation are relevant to the user u.

The definition of AUC [14] for user u is expressed as:

$$AUC_u = \frac{1}{Pair(u)} \sum_{(j,n)\in Pair(u)} h\left(rank(j) < rank(n)\right) \tag{10}$$

in which $h(x)$ is an indicator function that equals to 1 if x>0 (or logically true), and is 0 otherwise. $Pair(u)$ is a set of movie pairs to be evaluated for user u, i.e.,

$$Pair(u) = \{(j,n) \mid j \in Ts(u), n \notin Tr(u), n \notin Ts(u)\} \tag{11}$$

where $Tr(u)$ is the training set of movies for user u that are already in the user's profile, and $Ts(u)$ is the test set of movies for user u that are expected to be recommended. The best achievable AUC is of value 1. In our experiments, we report the AUC averaged across all test users.

Among all of the metrics used in this paper, we emphasize that the AUC is less indicative than other precision-related metrics. The reason is the AUC involves all pairs containing any relevant movie in the test set as well as all irrelevant movies (which are not in either training set or test set). However, the number of irrelevant movies is much larger than that of the relevant ones in the dataset, which means that the AUC could be less sensitive to the variation in the ordering of movies.

5.1.3 Experimental protocol

We use the predefined training set and evaluation (validation) set from the Moviepilot challenge to evaluate the performance of our proposed algorithm. The ratings in the training set are used to generate recommendations for the test users in the evaluation set. Moreover, the evaluation only concerns the movies with a specific mood tag (i.e., the one with identifier 16) to be the true positives for the test users. Recommended movies that are in the evaluation set but are without the specified mood tag are counted as false positives for the test users, i.e., they do not contribute to improvement as reflected by the evaluation metrics.

5.2 Impact of tradeoff parameter α

The tradeoff parameter α in the proposed algorithm influences the relative contribution from the mood-specific movie similarity. Fig. 4 illustrates the impact of α on the evaluation metrics, in which we can see that for most of the metrics, i.e., P@5, P@10 and MAP, the optimal α is nearly 0.1. For AUC the impact of α is unclear. However, as mentioned before, we consider AUC not to be a meaningful measure for the context-aware recommendation. For this reason, we focus on the precision related metrics and further keep the α set to 0.1. Note that in this paper we fix the dimensionality of latent features to be 10. Although tuning this parameter would influence the performance, we focus in this paper only on the usefulness of the mood-specific movie similarity in the joint matrix factorization model. The regularization parameter λ is set to 1 based on the observation from the performance of the basic matrix factorization, which is also indicated in section 5.4.

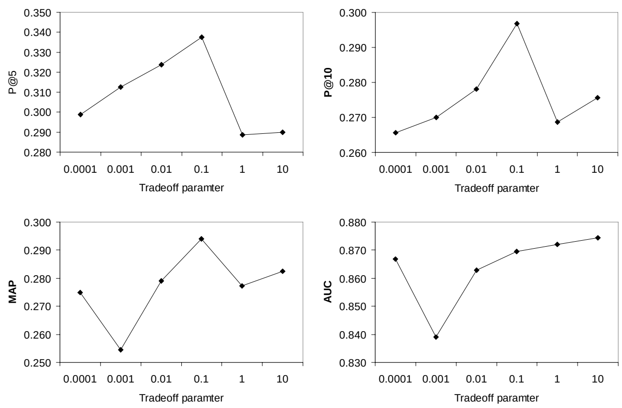

Fig. 4 The impact of tradeoff parameter on the recommendation performance of the proposed algorithm

5.3 Effectiveness

We further investigate the research question (1), namely to which extent the minimization of the objective function contributes to the improvement of the recommendation. For this reason, we demonstrate the variation of the output of the objective function (normalized for demonstration purpose) and evaluation metrics, e.g., P@5 and MAP, simultaneously during the iterations of optimization, as shown in Fig. 5. As observed, the recommendation performance generally increases when the objective function is minimized.

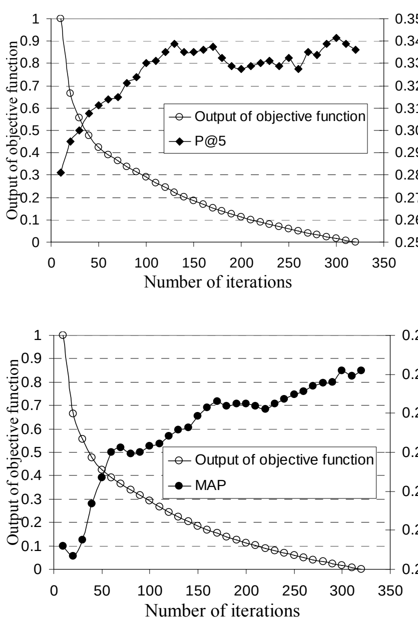

Fig. 5 The effectiveness of the proposed algorithm for recommendation performance

5.4 Performance Comparison

In order to provide answer to the research question (2), we compare in this section the performance of the proposed algorithm JMF-MS with some other state-of-the-art recommendation approaches listed below.

- **PopRec**: Movies are recommended to users based on their popularity, which is defined in terms of the number of users who watched them in the training set. We consider it is a non-personalized and naive baseline since every test user will have same recommendation.
- **RWR**: This algorithm is a state-of-the-art recommendation approach that uses random walk with restart (RWR) [11], which takes the user-item rating matrix as a bi-partite graph. Here, we set the restart probability to 0.8 to give the optimal performance on the evaluation set.
- **MF**: This algorithm represents a basic state-of-the-art matrix factorization approach as in [12], which is also equivalent to Eq. (5). The dimensionality of the latent user and movie features is set to 10. The regularization parameter λ is set to 1, which also gives the optimal performance on the evaluation set. Note that the same corresponding parameters are used in the proposed algorithm.
- **JMF-MB**: This algorithm is a joint matrix factorization approach that incorporates the general mood-based movie similarity rather than mood-specific movie similarity in Eq. (6). It also shares the same corresponding parameters as used in MF. In addition, we set the tradeoff parameter to 0.01 to again give the best performance on the evaluation set. The JMF-MB is used to compare with the proposed algorithm especially for validating the usefulness of the mood-specific movie similarity.

The results are shown in Table.1, from which we can see that the proposed algorithm outperforms other approaches with respect to P@5, P@10 and MAP. Although the AUC performance is worse than the performance of PopRec or RWR, we emphasize again that the precision related metrics are more critical indicators for this recommendation task. This observation holds especially for the early recommendation performance, which is reflected by P@5 and P@10. The proposed algorithm achieves over 10% improvement over MF with respect to P@5, P@10 and MAP, which indicates that the mood-specific movie similarity indeed brings up additional benefits apart from user-movie rating matrix. Regarding the research question (3), the proposed algorithm also achieves over 5% improvement over JMF-MB with respect to P@10 and MAP, which indicates that the mood-specific movie similarity is more beneficial for the mood-specific recommendation purpose compared to mood-based movie similarity.

Table 1. Comparison of recommendation performance between the proposed algorithm and other baseline approaches.

Algorithms	P@5	P@10	MAP	AUC
PopRec	0.239	0.259	0.264	0.874
RWR	0.266	0.273	0.281	**0.883**
MF	0.301	0.264	0.273	0.862
JMF-MB	0.333	0.286	0.281	0.861
Proposed JMF-MS	**0.338**	**0.297**	**0.294**	0.869

6. DISCUSSION AND CONCLUSIONS

In this paper, we proposed a novel mood-specific movie similarity concept, which is suitable for the purpose of mood-specific movie recommendation. A joint matrix factorization model is proposed that factorizes both the user-movie rating matrix and the mood-specific movie similarity matrix. The proposed algorithm is scalable for application to the task of recommendation with specific purpose—the complexity is linear with the number of observed ratings in rating matrix. It is also experimentally demonstrated to be effective for mood-specific movie recommendation and outperforms the state-of-the-art approaches. In addition, we specifically validate the usefulness of the mood-specific movie similarity compared to general mood-based movie similarity, which indeed leads to a substantial performance improvement.

In our future work, we will investigate exploitation of ranking-oriented approaches for context-aware recommendation. Additionally, we will investigate other movie modalities that might be able to influence mood-specific movie recommendation.

7. ACKNOWLEDGEMENTS

The research leading to these results was carried out within the PetaMedia Network of Excellence and has received funding from the European Commission's 7th Framework Program under grant agreement n° 216444.

8. REFERENCES

[1] Adomavicius, G., Sankaranarayanan, R., Sen, S., and Tuzhilin, A., 2005. Incorporating contextual information in recommender systems using a multidimensional approach. ACM TOIS, 23, 1, 103-145.

[2] Adomavicius G., and Tuzhilin, A., 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE TKDE, 17, 6, 734-749.

[3] Burke, R., 2002. Hybrid recommender systems: survey and experiments. User Modeling and User-Adapted Interaction, 12, 4, 331-370.

[4] Cantador, I., and Castells, P., 2009. Semantic contextualisation in a news recommender system. In CARS ′09.

[5] De Carolis, B., Mazzotta, I., Novielli, N., and Silvestri, V., 2009. Using common sense in providing personalized recommendations in the tourism domain. In CARS ′09.

[6] Deng, H., Lyu, M. R., and King, I., 2009. Effective latent space graph-based re-ranking model with global consistency. In WSDM ′09, 212-221.

[7] Deshpande, M., and Karypis, G., 2004. Item-based top-N recommendation algorithms. ACM TOIS, 22, 1, 143-177.

[8] Gunawardana A., and Shani, G., 2009. A survey of accuracy evaluation metrics of recommendation tasks. JMLR, 10, 2935-2962.

[9] Herlocker, J., Konstan, J., Borchers, A., and Riedl, J., 1999. An algorithmic framework for performing collaborative filtering. In SIGIR ′99, 230-237.

[10] Herlocker, J., Konstan, J., Terveen, L. G., and Riedl, J. 2004. Evaluating collaborative filtering recommender systems. ACM TOIS, 22, 1, 5-53.

[11] Konstas, I., Stathopoulos, V., and Jose, J. M., 2009. On social networks and collaborative recommendation. In SIGIR ′09, 195-202.

[12] Koren, Y., Bell, R., and Volinsky, C., 2009. Matrix factorization techniques for recommender systems. IEEE Computer, 42, 8, 30-37.

[13] Liu, N. N., Zhao, M., and Yang, Q., 2009. Probabilistic latent preference analysis for collaborative filtering. In CIKM ′09, 759-766.

[14] Rendle, S., Freudenthaler, C., Gantner, Z., and Schmidt-Thieme L., 2009. BPR: Bayesian personalized ranking from implicit feedback. In UAI ′09, 452-461.

[15] Said, A., Berkovsky, S., and De Luca, E.W., 2010. Putting things in context: Challenge on context-aware movie recommendation. In CAMRa2010: Proceedings of the RecSys '10 Challenge on Context-aware Movie Recommendation, to appear.

[16] Sarwar, B., Karypis, G., Konstan, J., and Reidl, J., 2001. Item-based collaborative filtering recommendation algorithms. In WWW ′01, 285-295.

[17] Shi, Y., Larson, M., and Hanjalic, A., 2009. Exploiting user similarity based on rated-item pools for improved user-based collaborative filtering. In RecSys ′09, 125-132.

[18] Shi, Y., Larson, M., and Hanjalic, A., 2010. List-wise learning to rank with matrix factorization for collaborative filtering. In RecSys ′10, to appear.

[19] Su, J.-H., Yeh, H.-H., Yu, P. S., and Tseng, V., 2010. Music recommendation using content and context information mining. IEEE Intelligent Systems, 25, 1, 16-26,

[20] Yildirim, H., and Krishnamoorthy, M. S., 2008. A random walk method for alleviating the sparsity problem in collaborative filtering. In RecSys ′08, 131-138.

[21] Zhang, J. and Pu, P., 2007. A recursive prediction algorithm for collaborative filtering recommender systems. In RecSys ′07, 57-64.

[22] Zheng, V. W., Zheng, Y., Xie, X., and Yang, Q., 2010. Collaborative location and activity recommendations with GPS history data. In WWW ′10, 1029-1038.

[23] Zhu, S., Yu, K., Chi, Y., and Gong, Y., 2007. Combining content and link for classification using matrix factorization. In SIGIR ′07, 487-494.