Asymmetric Bayesian Personalized Ranking for One-Class Collaborative Filtering

Shan Ouyang
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, P.R. China
ouyangshan@email.szu.edu.cn

Lin Li
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, P.R. China
lilin20171@email.szu.edu.cn

Weike Pan*
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, P.R. China
panweike@szu.edu.cn

Zhong Ming
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, P.R. China
mingz@szu.edu.cn

ABSTRACT

In this paper, we propose a novel preference assumption for modeling users' one-class feedback such as "thumb up" in an important recommendation problem called one-class collaborative filtering (OCCF). Specifically, we address a fundamental limitation of a recent symmetric pairwise preference assumption and propose a novel and first asymmetric one, which is able to make the preferences of different users more comparable. With the proposed asymmetric pairwise preference assumption, we further design a novel recommendation algorithm called asymmetric Bayesian personalized ranking (ABPR). Extensive empirical studies on two large and public datasets show that our ABPR performs significantly better than several state-of-the-art recommendation methods with either pointwise preference assumption or pairwise preference assumption.

CCS CONCEPTS

• **Information systems** Personalization; • **Human-centered computing** Collaborative filtering.

KEYWORDS

Asymmetric Pairwise Preference Assumption; Bayesian Personalized Ranking; One-Class Collaborative Filtering

ACM Reference Format:
Shan Ouyang, Lin Li, Weike Pan, and Zhong Ming. 2019. Asymmetric Bayesian Personalized Ranking for One-Class Collaborative Filtering. In *Thirteenth ACM Conference on Recommender Systems (RecSys '19), September 16–20, 2019, Copenhagen, Denmark.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3298689.3347051

*Weike Pan and Zhong Ming are corresponding authors for this work.

RecSys '19, September 16–20, 2019, Copenhagen, Denmark

ACM ISBN 978-1-4503-6243-6/19/09…$15.00
https://doi.org/10.1145/3298689.3347051

1 INTRODUCTION

Recommender systems are designed to help users discover items of interest with high efficiency, which play a key role in various web-enabled applications such as online shops[1], music/video streaming platforms[2] and news portals[3]. For now, there have been a variety of methods proposed with different focuses to improve the recommendation quality. The basic idea is to learn some specific preference patterns for each user on the basis of his or her feedback, in which most of them are one-class data such as "thumb up" in social media, "like" in entertainment platforms, and "add-to-cart" in online shopping malls, all indicating relatively positive attitude. The richness of one-class feedback brings both challenges and opportunities for recommender systems. Hence, how to exploit the one-class data to provide better personalized services has become particularly important.

For modeling users' one-class feedback, previous methods assume that a user likes an interacted item and dislikes an un-interacted item to some extent [4, 5, 10], or a user prefers an interacted item to an un-interacted item [14], which are the so-called pointwise preference assumption and pairwise preference assumption, respectively. Recently, a recommendation method extends the horizontal pairwise preference assumption [14] and assumes that the preference of an interacted (user, item) pair is larger than that of an un-interacted (user, item) pair both horizontally and vertically [19]. Specifically, there are two pairwise relationships, including an aforementioned horizontal one and a new vertical one (i.e., an item is preferred by an interacted user to an un-interacted user). However, such an extended assumption may not hold, in particular of the vertical one, because different users may have different evaluation standards, which will then make the preferences of different users uncomparable.

In order to address the fundamental limitation of the symmetric pairwise preference assumption [19], we propose a novel and improved preference assumption, i.e., asymmetric pairwise preference assumption, where we keep the horizontal one and assume that an item is preferred by a group of interacted users to a group of un-interacted users in order to make the vertical one more reasonable

[1]https://www.amazon.com/
[2]https://www.youtube.com/
[3]https://news.yahoo.com/

and comparable. With the proposed first asymmetric assumption for OCCF, we then design a novel recommendation algorithm called asymmetric Bayesian personalized ranking (ABPR). In order to study the effectiveness of different preference assumptions and the corresponding recommendation methods, we conduct extensive empirical studies on two large and public datasets, and find that our ABPR performs significantly better in terms of several commonly used ranking-oriented evaluation metrics.

2 RELATED WORK

In this section, we discuss some related works on modeling one-class feedback. In particular, we describe the existing works from the perspective of pointwise preference assumption and pairwise preference assumption.

2.1 Pointwise Preference Assumption

The basic assumption underlying pointwise preference learning methods is that an interacted (user, item) pair $(u, i) \in \mathcal{R}$ is regarded as a positive instance while an un-interacted (user, item) pair $(u, j) \notin \mathcal{R}$ is taken as a negative instance [10]. Without loss of generality, the pointwise preference assumption can be represented mathematically as follows,

$$\hat{r}_{ui} = 1, \hat{r}_{uj} = 0, \tag{1}$$

where $i \in \mathcal{I}_u$ and $j \in \mathcal{I} \backslash \mathcal{I}_u$ denote an interacted item and an un-interacted item by user u, respectively.

One of the most intuitive ways to model the one-class feedback based on the pointwise preference assumption in Eq.(1) is to take it as a regression problem. For example, some works [3, 4, 10] optimize a weighted square loss function together with some weighting or sampling strategies. Recently, [5] turns to model the one-class feedback from a different perspective of binary classification, instead of the traditional perspective of preference recovery. Specifically, it adopts a logistic loss that is usually used for supervised classification. Very recently, some deep learning based methods [2, 18] are also proposed for the one-class collaborative filtering problem, which adapt classical matrix factorization methods to non-linear multi-layer neural networks.

2.2 Pairwise Preference Assumption

Different from the absolute preference assumption made in pointwise preference learning methods, pairwise preference assumption is a relaxed one, which is usually recognized as a more natural choice for the one-class feedback. In particular, it assumes that compared with the preference of an un-interacted (user, item) pair, the preference of an interacted pair shall be larger [14],

$$\hat{r}_{ui} > \hat{r}_{uj}, \tag{2}$$

where $i \in \mathcal{I}_u$ and $j \in \mathcal{I} \backslash \mathcal{I}_u$.

On the basis of the fundamental pairwise assumption in Eq.(2), there have been numerous methods developed from various points of views. Some works propose to take more pairwise comparisons into consideration, attempting to model the preference difference in a fine-grained manner. For example, [20] divides the un-interacted (user, item) pairs into "unclear" feedback and negative feedback, which results in more pairwise relationships. Instead of using the typical preference assumption with comparison between two (user, item) pairs, [11, 13] introduce pariwise relationships between two (user, item-set) pairs to further relax the assumption in Eq.(2). Correspondingly, some deep learning based methods also successfully apply the pairwise assumption to neural network scheme [15].

All of the aforementioned methods are based on a single-way assumption, which are all modeled from the horizontal dimension of users. With argument that the horizontal pairwise preference assumption is not always sufficient to express users' true tastes, a recent work [19] constructs a mutual pairwise loss to model the relative preference both from the traditional horizontal dimension and a new vertical dimension,

$$\hat{r}_{ui} > \hat{r}_{uj}, \ \hat{r}_{ui} > \hat{r}_{wi}, \tag{3}$$

where $i \in \mathcal{I}_u$, $j \in \mathcal{I} \backslash \mathcal{I}_u$ and $w \in \mathcal{U} \backslash \mathcal{U}_i$. However, because of the probable difference in evaluation standards between two different users, the vertical relationship $\hat{r}_{ui} > \hat{r}_{wi}$ may not be always true.

3 OUR SOLUTION

3.1 Problem Definition

In our studied one-class collaborative filtering (OCCF) problem, we have some observed (user, item) pairs from n users and m items recorded in $\mathcal{R} = \{(u, i)\} \subseteq \mathcal{U} \times \mathcal{I}$, where a pair (u, i) means that a user u has interacted with an item i. For each user $u \in \mathcal{U}$, we exploit his or her interacted items $\mathcal{I}_u$ in order to recommend a personalized ranked list of items from the set of un-interacted items $\mathcal{I} \backslash \mathcal{I}_u$, where $\mathcal{I}$ denotes the whole set of items.

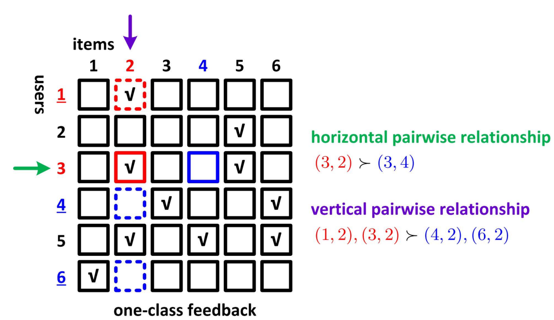

Figure 1: Illustration of our proposed asymmetric pairwise preference assumption. Notice that we have $(3, 2) \succ (3, 4)$ in BPR [14], and $(3, 2) \succ (3, 4)$, $(3, 2) \succ (4, 2)$ in MBPR [19].

3.2 Asymmetric Pairwise Preference Assumption

The mutual preference assumption in Eq.(3) is actually a symmetric pairwise preference assumption that combines a horizontal preference relationship in BPR [14], i.e., $(u, i) \succ (u, j)$ for $(u, i) \in \mathcal{R}$ and $(u, j) \notin \mathcal{R}$, and a vertical preference relationship, i.e., $(u, i) \succ (w, i)$ for $(u, i) \in \mathcal{R}$ and $(w, i) \notin \mathcal{R}$, where "$\succ$" denotes a relative preference relationship. We denote the latter as BPR$^{\mathrm{T}}$ because it adopts a transposed preference assumption of BPR, and also include it in our empirical studies.

However, the one-class feedback $\mathcal{R}$ is essentially not symmetric, because an existence of a (u, i) interaction pair relies more on the horizontal dimension of users, instead of on the vertical dimension of items. More specifically, an item i chosen by a user u is more based on its higher utility in comparison with other un-interacted items according to user u's own evaluation, i.e., $\hat{r}_{ui} > \hat{r}_{uj}$ for $j \in \mathcal{I} \backslash \mathcal{I}_u$, rather than the relative preference between user u and another user $w \in \mathcal{U} \backslash \mathcal{U}_i$ who has not interacted with item i, i.e., $\hat{r}_{ui} > \hat{r}_{wi}$. We have also observed this phenomenon in our empirical studies.

In order to address the fundamental limitation of the symmetric pairwise preference assumption and fully exploit the one-class feedback from both the horizontal dimension and the vertical dimension, we propose an asymmetric pairwise preference assumption. In particular, we relax the vertical pairwise preference assumption in Eq.(3), and assume that the preference of a group of users $\mathcal{P} \subseteq \mathcal{U}_i$ who have interacted with item i is larger than that of a group of users $\mathcal{N} \subseteq \mathcal{U} \backslash \mathcal{U}_i$ who have not interacted with item i, i.e., $\hat{r}_{\mathcal{P}i} > \hat{r}_{\mathcal{N}i}$. Such a vertical relationship is more likely to hold because the comparability between two groups of users is usually higher than that between two single users in Eq.(3). Finally, we reach our asymmetric pairwise preference assumption illustrated in Figure 1,

$$\hat{r}_{ui} > \hat{r}_{uj}, \hat{r}_{\mathcal{P}i} > \hat{r}_{\mathcal{N}i}, \tag{4}$$

where $i \in \mathcal{I}_u$, $j \in \mathcal{I} \backslash \mathcal{I}_u$, $\mathcal{P} \subseteq \mathcal{U}_i$, $\mathcal{N} \subseteq \mathcal{U} \backslash \mathcal{U}_i$ and $u \in \mathcal{P}$. We can see that the difference between Eq.(4) and Eq.(3) is the vertical relationship, and they become exactly the same when $\mathcal{P} = \{u\}$ and $\mathcal{N} = \{w\}$. Hence, the symmetric pairwise preference assumption is a special case of ours.

For instantiation of the relationship between the group preferences $\hat{r}_{\mathcal{P}i}$ and $\hat{r}_{\mathcal{N}i}$, we propose "Many 'Group vs. One' (MGO)" inspired by "Many 'Set vs. One' (MSO)" [13],

$$\hat{r}_{\mathcal{P}i} > \hat{r}_{wi}, w \in \mathcal{N}, \tag{5}$$

where $\hat{r}_{\mathcal{P}i} = \frac{1}{|\mathcal{P}|} \sum_{u' \in \mathcal{P}} \hat{r}_{u'i}$ is the overall preference of user-group $\mathcal{P}$ to item i. Notice that $\hat{r}_{u'i} = U_{u'\cdot} V_{i\cdot}^T + b_{u'} + b_i$ is the prediction rule for the preference of user u' to item i, where $U_{u'\cdot} \in \mathbb{R}^{1 \times d}$ and $V_{i\cdot} \in \mathbb{R}^{1 \times d}$ are latent feature vectors of user u' and item i, respectively, and $b_{u'} \in \mathbb{R}$ and $b_i \in \mathbb{R}$ are the bias of user u' and item i, respectively. As another notice, our asymmetric pairwise preference assumption is defined on both the horizontal dimension and the vertical dimension while those in [12, 13] are solely defined on the horizontal dimension.

3.3 Objective Function and Learning Algorithm

Based on the asymmetric pairwise preference assumption in Eqs.(4-5), we reach an objective function in our asymmetric Bayesian personalized ranking (ABPR) for each quintuple $(u, i, j, \mathcal{P}, \mathcal{N})$,

$$\min_{\Theta} - \ln \sigma(\hat{r}_{uij}) - \frac{1}{|\mathcal{N}|} \sum_{w \in \mathcal{N}} \ln \sigma(\hat{r}_{i\mathcal{P}w}) + \text{reg}(u, i, j, \mathcal{P}, \mathcal{N}), \tag{6}$$

where $\Theta = \{U_{u\cdot}, V_{i\cdot}, b_u, b_i, u \in \mathcal{U}, i \in \mathcal{I}\}$ are the model parameters to be learned, $\hat{r}_{uij} = \hat{r}_{ui} - \hat{r}_{uj}$ and $\hat{r}_{i\mathcal{P}w} = \hat{r}_{\mathcal{P}i} - \hat{r}_{wi}$ denote the corresponding preference differences, and $\text{reg}(u, i, j, \mathcal{P}, \mathcal{N}) = \frac{\alpha}{2}||V_{i\cdot}||^2 + \frac{\alpha}{2}||V_{j\cdot}||^2 + \frac{\alpha}{2}||b_i||^2 + \frac{\alpha}{2}||b_j||^2 + \sum_{u' \in \mathcal{P}}[\frac{\alpha}{2}||U_{u'\cdot}||^2 + \frac{\alpha}{2}||b_{u'}||^2] + \sum_{w \in \mathcal{N}}[\frac{\alpha}{2}||U_{w\cdot}||^2 + \frac{\alpha}{2}||b_w||^2]$ is the regularization term used to avoid overfitting.

We can then have the gradients of the model parameters w.r.t. the tentative objective function in Eq.(6), which will be used in the stochastic gradient descent (SGD) based algorithm.

Algorithm 1 The algorithm of asymmetric Bayesian personalized ranking (ABPR).

for $t = 1, 2, .., T$ **do**
  for $t_2 = 1, 2, .., |\mathcal{R}|$ **do**
    Randomly pick a (user, item) pair (u, i) from $\mathcal{R}$.
    Randomly pick an item j from $\mathcal{I} \backslash \mathcal{I}_u$.
    Randomly pick $|\mathcal{P}| - 1$ users from $\mathcal{U}_i \backslash \{u\}$.
    Randomly pick $|\mathcal{N}|$ users from $\mathcal{U} \backslash \mathcal{U}_i$.
    Calculate the gradients w.r.t. the tentative objection function in Eq.(6).
    Update the corresponding model parameters, i.e., $U_{u\cdot}$, $U_{w\cdot}$, $V_{i\cdot}$, $V_{j\cdot}$, b_i, b_j, b_u, and b_w, where $u \in \mathcal{P}$ and $w \in \mathcal{N}$.
  end for
end for

We describe our ABPR algorithm in Algorithm 1. Notice that the complexity of our ABPR is similar to that of MBPR [19] and BPR [14], because the sizes of the user-groups $\mathcal{P}$ and $\mathcal{N}$ are usually very small such as $|\mathcal{P}| = |\mathcal{N}| = 3$ in our empirical studies.

4 EXPERIMENTAL RESULTS

4.1 Datasets and Evaluation Metrics

We employ two large real-world datasets in our empirical studies, including MovieLens 20M (ML20M) and Netflix, where the former contains approximately 20 million ratings assigned by 138,493 users to 26,744 movies, and the latter contains about 100 million records made by 480,189 users to 17,770 movies. Each record of these two datasets consists of a user ID, an item ID and a corresponding rating score. For Netflix, we randomly sample 50,000 users, and then take all of the associated records, which is thus denoted as NF50KU.

For both ML20M and NF50KU, we follow [13] and keep the records with rating scores larger than 3 as positive interactions; then we randomly pick 60% of records as training data, 20% as test data and the remaining 20% as validation data. We repeat the above procedure for three times and obtain three copies of training data, test data and validation data for each dataset.

In order to make a direct comparison with the closely related works on item recommendation instead of rating prediction as commonly evaluated via MAE (mean absolute error) and RMSE (root mean square error), we adopt five ranking-oriented metrics [1, 9, 16] to evaluate the performance, including Precision@5, Recall@5, F1@5, NDCG@5 and 1-call@5.

4.2 Baselines and Parameter Configurations

In order to study the effectiveness of our proposed asymmetric pairwise preference assumption and the corresponding recommendation algorithm ABPR directly, we include the following closely related baseline methods, including (i) basic matrix factorization with square loss (MF), (ii) matrix factorization with logistic loss (LogMF) [5], (iii) factored item similarity model (FISM) [6], (iv) Bayesian personalized ranking (BPR) [14], (v) BPR with transposed pairwise preference assumption (BPR^{T}), and (vi) mutual BPR (MBPR) [19]. Notice that MF, LogMF and FISM are based on the pointwise preference assumption shown in Eq.(1), and BPR, BPR^{T} and MBPR are based on the pairwise preference assumptions shown in Eqs.(2-3).

Table 1: Recommendation performance of our proposed asymmetric Bayesian personalized ranking (ABPR), three pointwise preference learning methods, i.e., basic matrix factorization with square loss (MF), matrix factorization with logistic loss (LogMF) and factored item similarity model (FISM), and three pairwise preference learning methods, i.e., Bayesian personalized ranking (BPR), BPR with transposed preference assumption (BPR^T), and mutual BPR (MBPR), on ML20M and NF50KU w.r.t. five commonly used ranking-oriented evaluation metrics. The significantly best results are marked in bold (p-value < 0.015). We also include the searched best iteration number T and regularization parameter α for reproducibility.

Dataset	Method				Precision@5	Recall@5	F1@5	NDCG@5	1-call@5
ML20M	Pointwise	MF	$T = 720$	$\alpha = 0.01$	$0.1249_{\pm 0.0014}$	$0.0755_{\pm 0.0015}$	$0.0773_{\pm 0.0012}$	$0.1378_{\pm 0.0014}$	$0.4479_{\pm 0.0028}$
		LogMF	$T = 150$	$\alpha = 0.01$	$0.1622_{\pm 0.0002}$	$0.0918_{\pm 0.0002}$	$0.0951_{\pm 0.0001}$	$0.1805_{\pm 0.0001}$	$0.5269_{\pm 0.0007}$
		FISM	$T = 250$	$\alpha = 0.001$	$0.1351_{\pm 0.0014}$	$0.0821_{\pm 0.0010}$	$0.0836_{\pm 0.0009}$	$0.1505_{\pm 0.0019}$	$0.4755_{\pm 0.0028}$
	Pairwise	BPR	$T = 1000$	$\alpha = 0.001$	$0.1645_{\pm 0.0009}$	$0.0862_{\pm 0.0007}$	$0.0921_{\pm 0.0007}$	$0.1810_{\pm 0.0013}$	$0.5228_{\pm 0.0016}$
		BPR^T	$T = 370$	$\alpha = 0.001$	$0.0632_{\pm 0.0024}$	$0.0364_{\pm 0.0015}$	$0.0373_{\pm 0.0015}$	$0.0693_{\pm 0.0025}$	$0.2536_{\pm 0.0078}$
		MBPR	$T = 990$	$\alpha = 0.01$	$0.1609_{\pm 0.0004}$	$0.0893_{\pm 0.0007}$	$0.0931_{\pm 0.0006}$	$0.1797_{\pm 0.0004}$	$0.5262_{\pm 0.0025}$
		ABPR	$T = 730$	$\alpha = 0.001$	$\mathbf{0.1709}_{\pm 0.0005}$	$\mathbf{0.0940}_{\pm 0.0005}$	$\mathbf{0.0985}_{\pm 0.0004}$	$\mathbf{0.1907}_{\pm 0.0007}$	$\mathbf{0.5482}_{\pm 0.0016}$
NF50KU	Pointwise	MF	$T = 880$	$\alpha = 0.01$	$0.1385_{\pm 0.0002}$	$0.0497_{\pm 0.0009}$	$0.0581_{\pm 0.0005}$	$0.1464_{\pm 0.0008}$	$0.4685_{\pm 0.0019}$
		LogMF	$T = 110$	$\alpha = 0.01$	$0.1665_{\pm 0.0012}$	$0.0580_{\pm 0.0005}$	$0.0671_{\pm 0.0004}$	$0.1773_{\pm 0.0007}$	$0.5185_{\pm 0.0016}$
		FISM	$T = 120$	$\alpha = 0.001$	$0.1447_{\pm 0.0017}$	$0.0520_{\pm 0.0010}$	$0.0606_{\pm 0.0008}$	$0.1536_{\pm 0.0024}$	$0.4833_{\pm 0.0028}$
	Pairwise	BPR	$T = 870$	$\alpha = 0.001$	$0.1664_{\pm 0.0015}$	$0.0560_{\pm 0.0006}$	$0.0654_{\pm 0.0006}$	$0.1765_{\pm 0.0011}$	$0.5112_{\pm 0.0025}$
		BPR^T	$T = 110$	$\alpha = 0.001$	$0.0912_{\pm 0.0048}$	$0.0302_{\pm 0.0014}$	$0.0349_{\pm 0.0017}$	$0.0971_{\pm 0.0057}$	$0.3275_{\pm 0.0139}$
		MBPR	$T = 970$	$\alpha = 0.01$	$0.1794_{\pm 0.0012}$	$0.0659_{\pm 0.0009}$	$0.0748_{\pm 0.0008}$	$0.1928_{\pm 0.0013}$	$0.5489_{\pm 0.0033}$
		ABPR	$T = 610$	$\alpha = 0.01$	$\mathbf{0.1854}_{\pm 0.0008}$	$\mathbf{0.0685}_{\pm 0.0004}$	$\mathbf{0.0777}_{\pm 0.0004}$	$\mathbf{0.1995}_{\pm 0.0004}$	$\mathbf{0.5601}_{\pm 0.0010}$

For all the baseline methods and our ABPR, we implement them in the same SGD-based algorithmic framework written in Java for fair comparison[4]. In particular, we fix the number of latent dimensions $d = 20$, the learning rate $\gamma = 0.01$, and search the best value of the iteration number $T \in \{10, 20, 30, \ldots, 990, 1000\}$ and the best value of the tradeoff parameter on the regularization terms $\alpha \in \{0.001, 0.01, 0.1\}$ for each method on each dataset via the performance of NDCG@5 on the validation data. For MF, LogMF and FISM, we randomly sample three times of un-interacted (user, item) pairs as negative one-class feedback to augment the interacted (user, item) pairs for preference learning [6]. For our ABPR, we fix the number of user-group as $|\mathcal{P}| = |\mathcal{N}| = 3$ [13].

4.3 Results

We report the results in Table 1, from which we can have the following observations: (i) our ABPR performs significantly better (p-value[5] is smaller than 0.015) than all the baseline methods in all cases, which clearly shows the effectiveness of our proposed asymmetric pairwise preference assumption in modeling one-class feedback; (ii) MBPR performs similar to BPR on ML20M and better than BPR on NF50KU, which shows the sensitivity of the mutual pairwise preference assumption in MBPR w.r.t. different datasets (notice that our ABPR performs better than BPR on both datasets showcasing the superiority of our relaxed vertical pairwise relationship in our asymmetric assumption); (iii) the performance of BPR is much better than that of BPR^T, which is expected as the horizontal preference relationship, i.e., $(u, i) \succ (u, j)$, is more reasonable than the transposed one, i.e., $(u, i) \succ (w, i)$, considering the probable incomparability of preferences between different users (notice that our ABPR with asymmetric assumption, i.e., $(u, i) \succ (u, j)$ and $(\mathcal{P}, i) \succ (\mathcal{N}, i)$, performs the best); and (iv) for the methods with pointwise preference assumption, LogMF performs similar to BPR and better than MF and FISM, which shows the importance of an appropriate loss function in modeling one-class feedback.

[4]The source code is available at http://csse.szu.edu.cn/staff/panwk/publications/ABPR/.
[5]We conduct statistical significance test by using the MATLAB two-sample t-test function ttest2.m shown at https://ww2.mathworks.cn/help/stats/ttest2.html.

5 CONCLUSIONS AND FUTURE WORK

In this paper, we study an important recommendation problem called one-class collaborative filtering (OCCF) with users' one-class feedback such as "likes" in many online and mobile applications. In particular, we propose a novel preference assumption called asymmetric pairwise preference assumption, where we assume that a user prefers an interacted item to an un-interacted one as well as an interacted item is preferred by a group of interacted users to a group of un-interacted users. As far as we know, it is the first asymmetric assumption for OCCF. We then design a novel recommendation algorithm called asymmetric Bayesian personalized ranking (ABPR), and find it performs significantly better than several pointwise preference learning methods and pairwise preference learning methods on two large and public datasets.

For future works, we are interested in studying the proposed asymmetric preference assumption in more learning paradigms and problem settings such as listwise preference learning [17], deep learning [2, 8], and sparsity reduction and cold-start recommendation [7].

ACKNOWLEDGEMENT

We thank the support of National Natural Science Foundation of China Nos. 61872249, 61836005 and 61672358. Weike Pan and Zhong Ming are corresponding authors for this work.

REFERENCES

[1] Harr Chen and David R. Karger. 2006. Less is More: Probabilistic Models for Retrieving Fewer Relevant Documents. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'06)*. 429–436.

[2] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *Proceedings of the 26th International Conference on World Wide Web (WWW'17)*. 173–182.

[3] Xiangnan He, Hanwang Zhang, Min-Yen Kan, and Tat-Seng Chua. 2016. Fast Matrix Factorization for Online Recommendation with Implicit Feedback. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'16)*. 549–558.

[4] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative Filtering for Implicit Feedback Datasets. In *Proceedings of the 8th IEEE International Conference on Data Mining (ICDM'08)*. 263–272.

[5] Christopher C. Johnson. 2014. Logistic Matrix Factorization for Implicit Feedback Data. In *Proceedings of the NeurIPS 2014 Workshop on Distributed Machine Learning and Matrix Computations*.

[6] Santosh Kabbur, Xia Ning, and George Karypis. 2013. FISM: Factored Item Similarity Models for Top-N Recommender Systems. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13)*. 659–667.

[7] Yeon-Chang Lee, Sang-Wook Kim, and Dongwon Lee. 2018. gOCCF: Graph-Theoretic One-Class Collaborative Filtering Based on Uninteresting Items. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI'18)*. 3448–3456.

[8] Jianxun Lian, Xiaohuan Zhou, Fuzheng Zhang, Zhongxia Chen, Xing Xie, and Guangzhong Sun. 2018. xDeepFM: Combining Explicit and Implicit Feature Interactions for Recommender Systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'18)*. 1754–1763.

[9] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press, New York, NY, USA.

[10] Rong Pan, Yunhong Zhou, Bin Cao, Nathan Nan Liu, Rajan M. Lukose, Martin Scholz, and Qiang Yang. 2008. One-Class Collaborative Filtering. In *Proceedings of the 8th IEEE International Conference on Data Mining (ICDM'08)*. 502–511.

[11] Weike Pan and Li Chen. 2013. CoFiSet: Collaborative Filtering via Learning Pairwise Preferences over Item-sets. In *Proceedings of SIAM International Conference on Data Mining (SDM'13)*. 180–188.

[12] Weike Pan and Li Chen. 2013. GBPR: Group Preference Based Bayesian Personalized Ranking for One-class Collaborative Filtering. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*. 2691–2697.

[13] Weike Pan, Li Chen, and Zhong Ming. 2019. Personalized Recommendation with Implicit Feedback via Learning Pairwise Preferences over Item-sets. *Knowledge and Information Systems* 58, 2 (2019), 295–318.

[14] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence (UAI'09)*. 452–461.

[15] Bo Song, Xin Yang, Yi Cao, and Congfu Xu. 2018. Neural Collaborative Ranking. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM'18)*. 1353–1362.

[16] Daniel Valcarce, Alejandro Bellogín, Javier Parapar, and Pablo Castells. 2018. On the Robustness and Discriminative Power of Information Retrieval Metrics for top-N Recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys'18)*. 260–268.

[17] Liwei Wu, Cho-Jui Hsieh, and James Sharpnack. 2018. SQL-Rank: A Listwise Approach to Collaborative Ranking. In *Proceedings of the 35th International Conference on Machine Learning (ICML'18)*. 5311–5320.

[18] Yao Wu, Christopher DuBois, Alice X. Zheng, and Martin Ester. 2016. Collaborative Denoising Auto-Encoders for Top-N Recommender Systems. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining (WSDM'16)*. 153–162.

[19] Lu Yu, Ge Zhou, Chu-Xu Zhang, Junming Huang, Chuang Liu, and Zi-Ke Zhang. 2016. RankMBPR: Rank-Aware Mutual Bayesian Personalized Ranking for Item Recommendation. In *Proceedings of the 17th International Conference Web-Age Information Management (WAIM'16)*. 244–256.

[20] Runlong Yu, Yunzhou Zhang, Yuyang Ye, Le Wu, Chao Wang, Qi Liu, and Enhong Chen. 2018. Multiple Pairwise Ranking with Implicit Feedback. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM'18)*. 1727–1730.