Personalized Recommendation Considering Secondary Implicit Feedback

Siyuan Liu[†], Qiong Wu[†], and Chunyan Miao[†‡]
[†]*Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly (LILY)*
[‡]*School of Computer Science and Engineering*
Nanyang Technological University
Singapore
{syliu, wu.qiong, ascymiao}@ntu.edu.sg

***Abstract*—In e-commerce, recommendation is an essential feature to provide users with potentially interesting items to purchase. However, people are often faced with an unpleasant situation, where the recommended items are simply the ones similar to what they have purchased previously. One of the main reasons is that existing recommender systems in e-commerce mainly utilize primary implicit feedback (i.e., purchase history) for recommendation. Little attention has been paid to secondary implicit feedback (e.g., viewing items, adding items to shopping cart, adding items to favorite list, etc), which captures users' potential interests that may not be reflected in their purchase history. We therefore propose a personalized recommendation approach to combine the primary and secondary implicit feedback to generate the recommendation list, which is optimized towards a Bayesian objective criterion for personalized ranking. Experiments with a large-scale real-world e-commerce dataset show that the proposed approach presents a superior performance in comparison with the state-of-the-art baselines.**

***Index Terms*—recommendation, personalized ranking, implicit feedback**

I. Introduction

In current age, online shopping has become an inseparable part of people's life with the prosperity of e-commerce systems such as Amazon and Taobao. Recommendation is an essential feature in these systems to help users to find their preferred items among tons of millions of products based on the interaction information (i.e., feedback) between users and items. Feedback is usually explicit or implicit. Explicit feedback directly reflects users' preferences towards items in terms of ratings, while implicit feedback indirectly suggest users' potential interest on items.

In real life, users are often faced with an unpleasant situation, where the recommended items are simply the ones similar to what they have purchased previously but not those of particular interest for them. For example, after a user bought an iPhone online, he may receive recommendations of other brands of smartphones. However, as he just purchased iPhone, he may not have interest to buy another smartphone. One of the main reasons leading to such a situation is that most recommendations made in e-commerce are based on users' explicit feedback (i.e., ratings) or users' purchase history, or what we call as *primary implicit feedback* [1]. Little attention has been paid to users' other interactions with e-commerce systems, such as viewing items, adding items to shopping cart, and adding items to favorite list, or what we call as *secondary implicit feedback*. However, in the e-commerce context, these secondary implicit feedback is in fact of great value for generating recommendations as it contains rich information about users' potential interests. For example, a user may explore a lot of products before finally making a purchase. While viewing items, the user may also add those of high interest into the shopping cart or the favorite list. These user-system interaction data carry important hints on the user's latest preferences, which cannot be captured by existing recommendation models based on purchased history alone.

To date, many recommendation methods have been proposed based on implicit feedback, such as WRMF [2], EALS [3], Hu et al. [4], BPR [5], CLiMF [6], MRLR [7], EFM [8], Costa Fortes et al. [9] and Liu et.al. [10]. Most of these approaches mainly make use of implicit feedback to provide a personalized ranking of items to the user. However, these methods generally predict users' preferences based on primary implicit feedback, e.g., users' purchase history [5] or item properties (e.g., category, brand, etc.) [7], which ignore the abundant secondary implicit feedback.

In this work, we propose *PS*Rank, a personalized ranking based recommendation approach that integrates both primary and secondary implicit feedback. To capture users' latest preferences to make recommendations, we propose to jointly factorize the primary and secondary implicit feedback with shared user and item latent factors. These latent factors are learned towards a Bayesian objective criterion using a Stochastic Gradient Descent (SGD) based approach. We further conduct experiments over a large scale real-world dataset to study the performance of *PS*Rank in providing personalized recommendations. The experimental results demonstrate that *PS*Rank consistently outperforms the state-of-art methods in terms of Area Under ROC Curve (AUC) [5].

II. Related Works

In the literature for recommender systems, it is typical to formulate recommendation as either a rating prediction problem or a personalized ranking one. Rating prediction is to make use of explicit feedback in terms of ratings to predict a user's preferences towards items. The preferences are reflected as numerical scores, based on which items are ordered and recommended to the user. Many approaches have been developed for the purpose of improving the accuracy in predicting scores. Typical methods include neighbourhood methods [11], matrix factorization [12], and probabilistic matrix factorization [13].

Due to the simplicity on model optimization and evaluation, rating prediction based approaches have become dominant in the research community of recommender systems in the early stage. However, it has been shown that recommendation is better modelled as a personalized ranking problem [14] [5]. In addition, in real-world scenarios, user and item interactions are usually in the form of implicit feedback, such as purchasing products and watching movies, which are easily tracked in an automatic means. As rating prediction based approaches are not suitable for implicit feedback due to its one-class nature, a number of personalized ranking methods based on implicit feedback have been proposed, such as WRMF [2], EALS [3], Hu et al. [4], BPR [5], CLiMF [6], MRLR [7], and EFM [8].

Among these works, WRMF proposes to use weighted low rank approximation and negative example sampling to address the scenarios where only positive implicit feedback is given. Rating information is also combined for a more accurate recommendation. To speed up the computation of WRMF, EALS assigns weights to missing data based on the popularity of items instead of imposing a uniform-weight restriction on them. Hu et. al. propose a factor model which couples an estimate on users' preferences with a confidence level. BPR optimizes ranking performance through a general pair-wise ranking function and infers users' preferences over items by utilizing a negative sampling strategy. CLiMF models the binary relevance data by means of directly optimizing the Mean Reciprocal Rank. MRLR proposes a multi-level representation learning model for personalized ranking based recommendation by introducing item categories as the intermediate level of item organization. To jointly recommend items and lists, EFM makes use of user-item interactions and user-generated list (e.g., playlist and songlist) to discover the relationship between the list and its items with embedding-based algorithms.

Sharing some similarities with the proposed work, Costa Fortes et al. [9] and Liu et.al. [10] propose to exploit multiple feedback to make recommendations. In particular, Costa Fortes et al. [9] propose a framework to combine the recommendation predictions achieved based on individual types of feedback. The combination is through a liner regression algorithm based on a Bayesion optimization criterion. Liu et al. [10] propose a BPR based model to integrate multiple feedback through extending WRMF by optimizing towards a Bayesian objective criterion.

III. The Proposed Approach

In this section, we first introduce the Bayesian-based personalized ranking model (BPR). We then present how to integrate primary and secondary implicit feedback together through extending the BPR optimization criterion to jointly factorize the primary and secondary implicit feedback, followed by the learning and optimization of latent factors using multi-relational stochastic gradient descent.

A. Bayesian-based Personalized Ranking

In an e-commerce system, let U be the set of all users and I be the set of all items. For each user $u \in U$, he has a preference order over I, which is defined by a pair-wise preference order $>_u \subset I^2$, where $i >_u j$ indicates that u prefers $i \in I$ to $j \in I$. The goal of BPR is to learn the preference order for each user. To do so, BPR formulates its optimization criterion of learning based on Bayesian analysis. Specifically, suppose $\mathbf{\Theta}$ represents the parameters of Matrix Factorization (MF)[1], the Bayesian formulation of finding the optimal personalized ranking for all items $i \in I$ is to maximize the following posterior probability:

$$p(\mathbf{\Theta}| >_u) \propto p(>_u |\mathbf{\Theta})p(\mathbf{\Theta}). \tag{1}$$

Assuming the ordering of each pair of items i, j for user u is independent of the ordering of any other pair, $p(>_u |\mathbf{\Theta})$ in Eq. 1 is rewritten as:

$$p(>_u |\mathbf{\Theta}) = \prod_{u \in U} p(>_u |\mathbf{\Theta}) \quad = \prod_{(u,i,j) \in \mathcal{D}} p(i >_u j|\mathbf{\Theta}), \tag{2}$$

where $\mathcal{D}$ is the observed set of $>_u$. Suppose the individual probability that u prefers item i to item j is:

$$p(i >_u j|\mathbf{\Theta}) := \sigma(\hat{x}_{uij}(\mathbf{\Theta})), \tag{3}$$

where σ is the logistic sigmoid:

$$\sigma(x) = \frac{1}{1+e^{-x}}, \tag{4}$$

and $\hat{x}_{uij} = \hat{x}_{ui} - \hat{x}_{uj}$. $\hat{x}_{ui}$ is the predicted preference of u for i and is estimated by MF. Suppose $p(\mathbf{\Theta})$ is a normal distribution with zero mean and variance-covariance matrix $\Sigma_{\mathbf{\Theta}}$. The optimization criterion for BPR is:

$$\begin{aligned} \mathcal{L} &= \ln p(\mathbf{\Theta}| >_u) \\ &\propto \ln p(>_u |\mathbf{\Theta})p(\mathbf{\Theta}) \\ &\propto \ln \prod_{(u,i,j)\in\mathcal{D}} \sigma(\hat{x}_{uij})p(\mathbf{\Theta}) \\ &\propto \sum_{(u,i,j)\in\mathcal{D}} \ln \sigma(\hat{x}_{uij}) + p(\mathbf{\Theta}) \\ &\propto \sum_{(u,i,j)\in\mathcal{D}} \ln \sigma(\hat{x}_{uij}) - \lambda_{\mathbf{\Theta}}||\mathbf{\Theta}||, \end{aligned} \tag{5}$$

where $\lambda_{\mathbf{\Theta}}$ is the model specific regularization parameter.

In MF, the problem of predicting $\hat{x}_{ui}$ can be considered as a task of estimating a matrix $\mathbf{X} : |U| \times |I|$, which is approximately estimated by the matrix product of two low-rank matrices $\mathbf{W} : |U| \times k$ and $\mathbf{H} : |I| \times k$,

$$\hat{\mathbf{X}} := \mathbf{W}\mathbf{H}^T, \tag{6}$$

[1]BPR is suitable for MF and kNN. In this paper, we use MF as an example to illustrate how BPR works.

where k is the dimensionality of the approximation. The prediction formula for $\hat{x}_{ui}$ can then be written as:

$$\hat{x}_{ui} =< \mathbf{w}_u, \mathbf{h}_i >= \sum_{d=1}^{k} w_{ud} \cdot h_{id}. \tag{7}$$

The optimization of BPR criterion is done through Stochastic Gradient Descent (SGD) with bootstrap sampling of training triples [5]. The model parameter for MF is $\mathbf{\Theta} = (\mathbf{W}, \mathbf{H})$. Therefore, the optimization of BPR criterion only lies on $\frac{\partial \hat{x}_{uij}}{\partial \mathbf{\Theta}}$.

B. Incorporating Secondary Implicit Feedback

In this part, we introduce how to extend the BPR optimization criterion to incorporate secondary implicit feedback. Let $F_u^P \subseteq U \times I$ be the observed primary implicit feedback (e.g., users' purchase history for items) and $F_u^S \subseteq U \times I$ be the observed secondary implicit feedback (e.g., user-item interactions except purchase, such as viewing items, adding items to favorite, and adding items to cart). We further define the set of items interacted by u in primary and secondary implicit feedback as $I_u^P = \{i|(u,i) \in F_u^P\}$ and $I_u^S = \{i|(u,i) \in F_u^S\}$, respectively. We then revise the definition of the pair-wise preference order $>_u$ in BPR as:

$$i >_u j : \forall i \in I_u^P, \forall j \notin I_u^P; \tag{8}$$
$$i >_u j : \forall i \in I_u^S, \forall j \in I \setminus (I_u^P \cup I_u^S). \tag{9}$$

The meaning of extended definition of $>_u$ can be further illustrated using Figure 1.

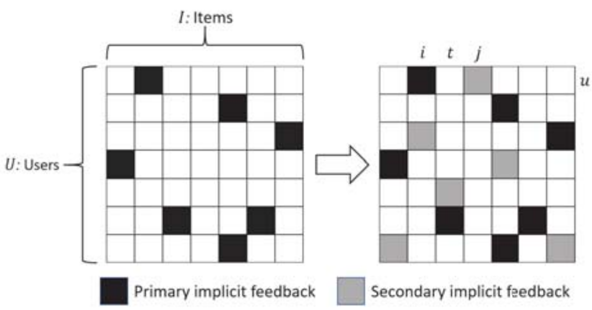


Figure 1. The illustration of the extended pair-wise preference order.

In Figure 1, there are two matrices. The left matrix represents the primary implicit feedback and shows the original definition of $>_u$ in BPR. The cell (u,i) shaded with black means that the corresponding user u purchased the item i, while the cell (u,j) without shading suggests that there is no purchase observed from u for item j. In other words, it indicates that u prefers i to j. The right matrix represents the integration of primary and secondary implicit feedback. In this matrix, we have a new shading color, i.e., gray. A cell (u,j) shaded in gray means that there is secondary feedback from u towards j. For example, if u clicks j but have not purchased it, it will be shaded in gray to suggest that u prefers j to another item t if there is no observed interactions of purchase or clicking from u towards t (i.e, (u,t) is not shaded in any color). Therefore, we will have the following preference order $i >_u j$, $i >_u t$, and $j >_u t$, where $j >_u t$ is not captured in the original definition, which actually provides a hint on u's potential preference.

Following the BPR optimization criterion, the optimization criterion for learning the preference order defined in Eq. 8 is given by:

$$\mathcal{L}_P \propto \sum_{(u,i,j) \in \mathcal{D}_P} \ln \sigma(\hat{x}_{uij}^P) - \lambda_{\mathbf{\Theta}} ||\mathbf{\Theta}||, \tag{10}$$

where $\hat{x}_{uij}^P = \hat{x}_{ui}^P - \hat{x}_{uj}^P$ and $\mathcal{D}_P$ is the observed set of preference order satisfying Eq. 8. Similarly, the optimization criterion for learning the preference order defined in Eq. 9 is given by:

$$\mathcal{L}_S \propto \sum_{(u,i,j) \in \mathcal{D}_S} \ln \sigma(\hat{x}_{uij}^S) - \lambda_{\mathbf{\Theta}} ||\mathbf{\Theta}||, \tag{11}$$

where $\hat{x}_{uij}^S = \hat{x}_{ui}^S - \hat{x}_{uj}^S$ and $\mathcal{D}_S$ is the observed set of preference order satisfying Eq. 9. Here, $\hat{x}_{ui}^P$ and $\hat{x}_{ui}^S$ are estimated through MF. The main difference between Eq. 10 and Eq. 11 is that the former captures users' preferences over items based on primary implicit feedback, while the latter obtains user's preference over items based on secondary implicit feedback. To capture both preferences in the latent factors of users and items, we propose to sew $\mathcal{L}_P$ and $\mathcal{L}_S$ together by using the following joint optimization criterion:

$$\begin{aligned} \mathcal{L} \propto \mathcal{L}_P + \alpha \mathcal{L}_S \propto & \sum_{(u,i,j) \in \mathcal{D}_P} \ln \sigma(\hat{x}_{uij}^P) \\ & + \alpha \times \sum_{(u,i,j) \in \mathcal{D}_S} \ln \sigma(\hat{x}_{uij}^S) - \lambda_{\mathbf{\Theta}} ||\mathbf{\Theta}||, \end{aligned} \tag{12}$$

where α is the weight for the preference order reflected by secondary implicit feedback.

C. Optimization

The optimization procedure for learning the model parameters (i.e., user and item latent factors) can be realized via the SGD strategy. As the first step, we calculate the gradient of Eq. 12 with respect to the model parameter, as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{\Theta}} &= \frac{\partial \mathcal{L}_P}{\partial \mathbf{\Theta}} + \alpha \times \frac{\partial \mathcal{L}_S}{\partial \mathbf{\Theta}} \\ &= \sum_{(u,i,j) \in \mathcal{D}_P} \frac{\partial \ln \sigma(\hat{x}_{uij}^P)}{\partial \mathbf{\Theta}} \\ &+ \alpha \times \sum_{(u,i,j) \in D_S} \frac{\partial \ln \sigma(\hat{x}_{uij}^S)}{\partial \mathbf{\Theta}} - \lambda_{\mathbf{\Theta}} \frac{\partial ||\mathbf{\Theta}||^2}{\partial \mathbf{\Theta}}, \end{aligned} \tag{13}$$

where

$$\frac{\partial \mathcal{L}(\hat{x}_{uij}^P)}{\partial \mathbf{\Theta}} = \frac{-e^{-\hat{x}_{uij}^P}}{1 + -e^{-\hat{x}_{uij}^P}} \cdot \frac{\partial \hat{x}_{uij}^P}{\partial \mathbf{\Theta}} - \lambda_{\mathbf{\Theta}} \cdot \mathbf{\Theta}; \tag{14}$$

$$\frac{\partial \mathcal{L}(\hat{x}_{uij}^S)}{\partial \mathbf{\Theta}} = \frac{-e^{-\hat{x}_{uij}^S}}{1 + -e^{-\hat{x}_{uij}^S}} \cdot \frac{\partial \hat{x}_{uij}^S}{\partial \mathbf{\Theta}} - \lambda_{\mathbf{\Theta}} \cdot \mathbf{\Theta}. \tag{15}$$

The derivative of $\hat{x}^P_{uij}$ is:

$$\frac{\partial \hat{x}^P_{uij}}{\partial \boldsymbol{\Theta}} = \begin{cases} h_{id} - h_{jd} & \text{if } \theta = w_{ud}, \\ w_{ud} & \text{if } \theta = h_{id}, \\ -w_{ud} & \text{if } \theta = h_{jd}, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

where w_{ud}, h_{id}, and h_{jd} are the dth latent feature of the corresponding latent factor $\mathbf{w}_u$, $\mathbf{h}_i$, and $\mathbf{h}_j$ in Eq 7, respectively. As $\hat{x}^P_{uij}$ and $\hat{x}^S_{uij}$ share the same latent factors for users and items, the derivative form of $\hat{x}^S_{uij}$ is exactly the same as $\hat{x}^P_{uij}$ as shown in Eq. 16.

Next, we adopt the multi-relational SGD strategy [15] to optimize the joint factorization procedure. Specifically, the optimization procedure is conducted alternatively with respect to $\mathcal{D}_P$ and $\mathcal{D}_S$. Firstly, some training instances are randomly sampled from $\mathcal{D}_P$ and $\mathcal{D}_S$, respectively. Then at each iteration, a gradient descent step is performed for all related parameters according to the loss of the training instance. The details of the procedure are shown in Algorithm 1.

Algorithm 1: The Optimization for Joint Factorizing the Primary and Secondary Implicit Feedback

1 Randomly initialize $\boldsymbol{\Theta}$ with small values;
2 Randomly sample for $\mathcal{D}_P$ and $\mathcal{D}_S$;
3 Draw N instances (u, i, j) from $\mathcal{D}_P$;
4 Draw N instances (u, m, n) from $\mathcal{D}_S$;
5 **for** $t = 1$ **to** $maxiters$ **do**
6   **foreach** (u, i, j) **do**
7         $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta} + \mu(\frac{e^{-\hat{x}^P_{uij}}}{1+-e^{-\hat{x}^P_{uij}}} \cdot \frac{\partial \hat{x}^P_{uij}}{\partial \boldsymbol{\Theta}} + \lambda_{\boldsymbol{\Theta}} \cdot \boldsymbol{\Theta})$;
8   **end**
9   **foreach** (u, m, n) **do**
10         $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta} + \mu(\alpha \frac{e^{-\hat{x}^S_{umn}}}{1+-e^{-\hat{x}^S_{umn}}} \cdot \frac{\partial \hat{x}^S_{umn}}{\partial \boldsymbol{\Theta}} + \lambda_{\boldsymbol{\Theta}} \cdot \boldsymbol{\Theta})$;
11   **end**
12   **if** $\mathcal{L}$ *has converged* **then**
13   break
14   **end**
15 **end**

In Algorithm 1, $\mathcal{D}_P$ and $\mathcal{D}_S$ are the primary and secondary implicit feedback, respectively. μ is the learning rate, α is the weight for the preference order reflected by secondary implicit feedback, $\lambda_{\boldsymbol{\Theta}}$ is the regularization coefficient, and $maxiters$ is the maximum number of iterations. Lines 3-4 are to sample user-specific ranking tuples for training. To reduce computation complexity, we randomly sample the training tuples instead of using all observed ones. Lines 6-8 and lines 9-11 are to update parameters using the samples of primary and secondary implicit feedback, respectively. The whole procedure will stop until the maximum iteration number is reached or $\mathcal{L}$ converges.

IV. EXPERIMENTS

To study the performance of the proposed approach in making personalized recommendations, we conduct extensive experiments on a large-scale real-life dataset and compare the proposed approach with BPR in terms of the Area Under the Roc Curve (AUC).

A. Experimental Setup

1) Dataset: We use a dataset[2] published in AlibabaCloud Tianchi Data platform[3] to conduct evaluation. The dataset is provided by TMall, one of the most popular e-commerce websites in China. Besides the primary implicit feedback (i.e. the purchase history for users towards items), the dataset also includes various types of secondary implicit feedback, such as clicking an item to view the details, adding an item to a user's favorite list, and adding an item to a user's shopping cart. The statistics of the dataset is summarized in Table I.

Table I
STATISTICS OF THE DATASET.

# users: 424,170; # items:1,090,390			per-user
Primary	# purchase	3,292,144	7.76
Secondary	# click	48,550,713	114.46
	# add-to-favorite	3,005,723	7.09
	# add-to-cart	76,750	0.18
Total		54,925,330	129.49

2) Comparing Methods: In the experiments, we compare the proposed approach with the BPR approach [5]. BRP is a sampling-based latent factor method for personalized ranking, which optimizes the pair-wise ranking between observed instances and sampled negative ones. This method uses a Bayesian objective and only considers primary implicit feedback for parameter learning.

Based on different types of incorporated secondary implicit feedback, we report three sets of results for the proposed approach, which are *PS*Rank with click, *PS*Rank with favorite, and *PS*Rank with cart. They make use of purchase history together with the data of click, favorite list or shopping cart, respectively, as secondary implicit feedback to predict whether a user will purchase an item in the future. For simplicity, in the following parts, we will call them as click, favorite, and cart, respectively.

We empirically find the best performing parameters for all methods. Specifically, we set the learning rate μ to 0.01 and the regularization parameter $\lambda_{\boldsymbol{\Theta}}$ to 0.0025 for all methods. For the proposed approach, we apply the grid search strategy to find the best performing α, which is the weight for secondary implicit feedback. All experiments are run 100 times to achieve a statistical accuracy and the average results are reported.

3) Training Dataset and Evaluation Metric: To conduct evaluation, we firstly sort the full dataset based on the timestamp of each interaction and then we divide the full dataset into training and testing data according to the timestamp order. It happened to be that the interactions before November 11

[2]https://tianchi.aliyun.com/datalab/dataSet.htm?spm=5176.100073.888.13.2c2795e4khBUN9&id=1

[3]https://tianchi.aliyun.com/

compose about 80% (i.e., 80.73%) of the full datasest, which are used for training, and the purchase interactions in the rest data are used for testing. In the experiment, the average Area Under the ROC curve (AUC) is employed to evaluate the performance of *PS*Rank (i.e., click, favorite, and cart) and BPR. AUC is a commonly used metric for evaluating the quality of personalized ranking [5], and the average AUC is computed as:

$$\mathrm{AUC} = \frac{1}{|U|}\sum_{u\in U}\frac{1}{|E(u)|}\sum_{(i,j)\in E(u)}\delta(\hat{x}_{ui} > \hat{x}_{uj}), \quad (17)$$

where $\delta(\hat{x}_{ui} > \hat{x}_{uj})$ is an indicator function, which returns 1 if $\hat{x}_{ui} > \hat{x}_{uj}$ is true, and 0 otherwise. $\hat{x}_{ui}$ is the predicted preference of u over i. $E(u)$ is the set of evaluation pairs for u, which is given by:

$$E(u) := \{(i,j)|(u,i) \in \mathcal{D}_{test} \wedge (u,j) \notin (\mathcal{D}_{test} \cup \mathcal{D}_{train})\}, \quad (18)$$

where $(u,i) \in \mathcal{D}_{test}$ (or $\mathcal{D}_{train}$) means u is observed to purchase i in the testing dataset $\mathcal{D}_{test}$ (or the training dataset $\mathcal{D}_{train}$). A larger AUC value indicates a better performance, and the AUC value of a random guess is 0.5.

B. Experimental Results

In this part, we will first introduce the overall performance of PSRank, and then we will conduct analysis on the weight parameter α for secondary implicit feedback, followed by the analysis with respect to the problem of data sparsity.

1) Overall Performance: The overall performance of *PS*Rank in terms of AUC comparing with BPR is summarized in Table II.

Table II
OVERALL PERFORMANCE OF ALL METHODS IN TERMS OF AUC (@$\alpha = 1$), WHERE THE BEST PERFORMANCE IS HIGHLIGHTED IN BOLD.

d	BPR	*PS*Rank		
		Click	Cart	Favorite
$d = 10$	0.6024	**0.7866**	0.7602	0.7834
$d = 20$	0.6180	**0.8017**	0.7768	0.7939
$d = 50$	0.6333	**0.8257**	0.7748	0.8139

It can be seen that *PS*Rank consistently outperforms BPR when $d = 10$, $d = 20$, and $d = 50$, no matter which of the three types of secondary implicit feedback is incorporated. It is interesting to notice that incorporating the click information contributes to improving the performance most remarkably, while cart information achieves a relatively less improvement on the performance. A possible reason is that comparing to the items clicked or added to favorite list by a user, the items in the user's shopping cart may share more similarities with the items that the user purchased previously. For example, we may encounter the situation where there are a number of similar items in an e-commerce system, and we may simply add all the similar items into shopping carts at the beginning and then choose one to purchase in the end. Therefore, the items left in the cart actually present some similarities as the the purchased item, and may not reflect a user's other potential preferences as much as the click or favorite information.

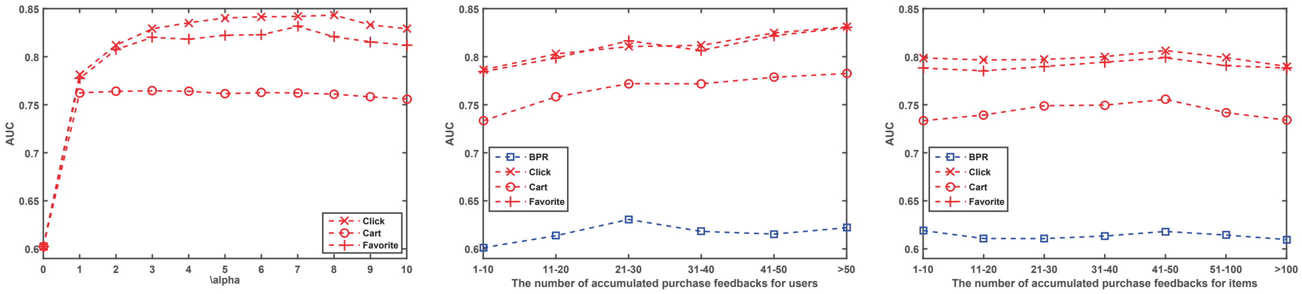
2) Impact of Parameter α: In *PS*Rank, α is the parameter for the consideration of the preference order learned from secondary implicit feedback as shown in Eq. 12. Intuitively, a larger α imposes more weights on secondary implicit feedback during parameter learning. To study the impact of α on the performance of *PS*Rank, we conducted a grid search of the best performing α in the range of [0, 10] with the step size of 1. Fig. 2(a) shows the effect of α on *PS*Rank for $d = 10$ when different types of implicit feedback are employed.

In Fig. 2(a), the x-axis is the α value, and the y-axis is the AUC value. When $\alpha = 0$, it means that no secondary implicit feedback is employed. Therefore, the three lines all start from the same point, which is the AUC value when BPR is adopted. For cart, α has a relatively small impact on the performance, and a small descending trend can be noticed after $\alpha = 4$. For favorite or click, the best performance is achieved when $\alpha = 7$ and $\alpha = 8$, respectively. A decreasing trend is noticed after the two values. It also can be noticed that the AUC values for cart are smaller than the AUC values for click or favorite, which is consistent with the results shown in Table II. In the following experiments, we will use the α values that present the best performance for the proposed approach.

3) Performance Analysis w.r.t Data Sparsity: The problem of data sparsity is a serious one that may impact the performance of personalized ranking approaches. In this part, we specially study the performance of *PS*Rank with respect to accumulated implicit feedback for users or items to explore how *PS*Rank performs when faced with the data sparsity problem. Firstly, we dispose the results obtained in Table II by selecting users who have purchased a number of items and the number is located in a specific range, i.e., 1-10, 11-20, 21-30, 31-40, 41-50, >50. Here, 1-10, 11-20, 21-30, 31-40, 41-50, and >50 mean that users purchased 1-10, 11-20, 21-30, 31-40, 41-50, and more than 50 items, respectively. The performance of various approaches is shown in Fig. 2(b).

It can be seen that the performance of *PS*Rank employing any of the three types of secondary implicit feedback outperforms BPR across all ranges. In particular, click and favorite present almost the same performance for all the ranges of the number of purchased items. Although the performance of cart is not as good as click or favorite, the AUC value for cart is still remarkably higher than the AUC value presented by BPR. More specially, for the users with accumulated number of purchase feedback in all ranges, *PS*Rank can achieve an AUC value of more than 0.7 or even 0.8, which is greater than BPR by about 30%. Even for the users with accumulated number of purchase feedback in the range of 1-10, click or favorite achieves an AUC value of approximately 0.8, and the AUC value for the cart approximates to 0.75. This suggests that *PS*Rank can effectively alleviate the data sparsity problem even when users purchases only a few items.

Secondly, we study the performance of *PS*Rank with respect to accumulated implicit feedback for items. Following the same way for studying the performance of *PS*Rank with respect to accumulated implicit feedback for users, we select the items which have been purchased by a specific number



(a) (b) (c)

Figure 2. Experimental results: (a)The impact of parameter α on *PS*Rank (@d=10); (b) Micro-analysis of *PS*Rank (@d=10) with respect to users with different scale of accumulated purchase feedback; (c) Micro-analysis of *PS*Rank (@d=10) with respect to items with different scale of accumulated purchase feedback.

of users and the number is located in the following ranges: 1-10, 11-20, 21-30, 31-40, 41-50, 51-100, and >100. Here, 1-10, 11-20, 21-30, 31-40, 41-50, 51-100, and >100 mean that items are purchased by 1-10, 11-20, 21-30, 31-40, 41-50, 51-100, and more than 100 users, respectively. The performance of various approaches is shown in Fig. 2(c).

It can be seen that the performance of *PS*Rank outperforms BPR across all ranges when any of the three types of secondary implicit feedback is employed. In particular, click presents the best performance, which is slightly better than favorite. Similar to the results presented in Fig. 2(a), though cart does not present a good performance as click or favorite, it still achieves an AUC value of above 0.7 in all ranges of accumulated implicit feedback for items, which is greater than BPR by about 15%.

V. Conclusion

In this paper, we propose a personalized ranking based recommendation approach. The proposed approach incorporates primary (e.g., purchase history) and secondary (e.g., clicking, adding items to favorite and adding items to shopping cart) implicit feedback together to make recommendations to users. In particular, we first extend the preference order defined in BPR model to incorporate the preference order inferred by secondary implicit feedback. Then we propose an optimization criterion to sew up the learning model for the first and secondary implicit feedback by jointly sharing latent factors for users and items. We further propose how to optimize the latent factors based on SGD. We conducted experiments on a real-world dataset to study the performance of the proposed approach in terms of AUC by comparing with other baseline methods. The experimental results show that the proposed approach consistently outperforms the comparing methods when different types of secondary implicit feedback are adopted. We also find that the proposed approach presents a superior performance even when there is serious data sparsity.

Acknowledgment

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative.

References

[1] L. Tang, B. Long, B.-C. Chen, and D. Agarwal, "An empirical study on recommendation with multiple types of feedback," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 283–292.

[2] R. Pan, Y. Zhou, L. Cao, B., N. N., R. Lukose, M. Scholz, and Q. Yang, "One-class collaborative filtering," in *Data Mining*, 2008, pp. 502–511.

[3] X. He, H. Zhang, M. Y. Kan, and T. S. Chua, "Fast matrix factorization for online recommendation with implicit feedback," in *The 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2016, pp. 549–558.

[4] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, 2008, pp. 263–272.

[5] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *The Conference on Uncertainty in Artificial Intelligence*, 2009, pp. 452–461.

[6] Y. Shi, A. Karatzoglou, L. Baltrunas, M. Larson, N. Oliver, and A. Hanjalic, "Climf: learning to maximize reciprocal rank with collaborative less-is-more filtering," in *The 6th ACM Recommender Systems Conference*. ACM, 2012, pp. 139–146.

[7] Z. Sun, J. Yang, J. Zhang, A. Bozzon, Y. Chen, C. Xu, and C. Sierra, "Mrlr: Multi-level representation learning for personalized ranking in recommendation," in *The 26th International Joint Conference on Artificial Intelligence*, 2017, pp. 2807–2813.

[8] D. Cao, L. Nie, X. He, S. Zhu, and T. S. Chua, "Embedding factorization models for jointly recommending items and user generated lists," in *The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2017, pp. 585–594.

[9] A. da Costa Fortes and M. G. Manzaoto, "Ensemble learning in recommender systems: Combining multiple user interactions for ranking personalization," in *Proceedings of the 20th Brazilian Symposium on Multimedia and the Web*, 2014, pp. 47–54.

[10] J. Liu, C. Shi, B. Hu, S. Liu, and S. Y. Philip, "Personalized ranking recommendation via integrating multiple feedbacks," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2017.

[11] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," in *The 22nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1999, pp. 230–237.

[12] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[13] A. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," in *The 20th Annual Conference on Neural Information Processing Systems (NIPS)*, 2008, pp. 1257–1264.

[14] M. Weimer, A. Karatzoglou, Q. V. Le, and A. J. Smola, "Cofi rank-maximum margin matrix factorization for collaborative ranking," in *The 20th Annual Conference on Neural Information Processing Systems (NIPS)*, 2008, pp. 1593–1600.

[15] A. Krohn-Grimberghe, L. Drumond, C. Freudenthaler, and L. Schmidt-Thieme, "Multi-relational matrix factorization using bayesian personalized ranking for social network data," in *The 5th ACM International Conference on Web Search and Data Mining*, 2012, pp. 173–182.