様式８

論　文　内　容　要　旨

| 報告番号 | 甲　先　第　432　号 | 氏　名 | 张 茜 |
|---|---|---|---|
| 学位論文題目 | Research on Recommendation Models for One-class Collaborative Filtering<br>(1クラス協調フィルタリングの推奨モデルに関する研究) | | |

内容要旨


Recommender systems have become an indispensable tool for real-world applications. One-class collaborative filtering has attracted much attention in recommendation communities because the "one-class" is more suitable to describe data of many applications. Many recommendation methods have been proposed for realizing personalized ranking with one-class feedback (implicit feedback). Pairwise ranking methods with relative preference assumptions are widely used for dealing with the one-class problem due to their high performance. Bayesian Personalized Ranking (BPR) is one of the most popular pairwise methods, assuming users prefer the observed item to the unobserved item.

BPR assumes the equal importance of each user's unobserved items. However, existing some items that users have not seen yet. It is not appropriate to treat each user's all unobserved items equally. Additionally, the parameters in BPR are learned by the stochastic gradient descent (SGD) optimization algorithm. The previous work has shown that the vanishing gradient problem exists in the learning process when the user's preference difference between the observed item and the unobserved item is very large.

In order to alleviate the problems of the previous model, three recommendation models are studied for one-class collaborative filtering in this thesis, including PBPR (Prior-based Bayesian Pairwise Ranking), PBPR* (Improving PBPR) and DBPL (Double Bayesian Pairwise Learning). All three recommendation models consider users' preference differences between their unobserved items and can be realized without any additional social information. In addition, the users' potential preference scores on their unobserved items are calculated based on users' historical interactions for further distinguishing the relative preference of each user's any two unobserved items. The key contributions in this thesis are summarized as below:

(1) Motivated by the discovery that the user may be interested in items that their like-minded users have observed, users' potential preference scores on their unobserved items could be calculated by the similarities between users and the similarities between items. The similarities between users and the similarities between items are measured at the item level and the entity level considering that the user might like the item or entity. Experiments on the real-world dataset demonstrate the results of the *UIIU* (user-based item similarity and item-based user similarity) method are the best in most cases. The potential preference scores calculated by the *UIIU* method are used for further studying.

(2) With the observation that each user has his/her own chosen intention on different service systems and most people's chosen intention about items have continuity and do not change suddenly, the Latent Dirichlet Allocation (LDA) model is used to realize this observation. The user's chosen intention is considered as the hidden variable, and two distributions (user-chosen intention distribution and item-chosen intention distribution) are updated during the learning process of the model. The users' potential preference scores can be obtained by the inner product of two distributions. Experimental results of the LDA-based method are better than BPR across all evaluation metrics on three datasets.

(3) For alleviating the assumption in BPR that equal importance of the huge unobserved items, the novel model PBPR is proposed. It relaxes the simple pairwise preference assumption in BPR by further considering the pairwise preference between any two unobserved items. PBPR considers the situation of existing fine-grained preference difference between any two unobserved items of a user. It assumes the user prefers an unobserved item with a higher potential preference score over another unobserved item. PBPR* is proposed to enhance the performance of PBPR by conducting several strategies to overcome shortcomings in PBPR, for more accurate recommendation results.

(4) With the consideration that the user's preference difference between the observed item and the unobserved item can be reduced by fusing a relatively smaller preference difference between another pair of items, DBPL is proposed by taking two pairwise preferences into the previous pairwise learning model. DBPL also takes into account each user's fine-grained preference differences between unobserved items. For each user, the unobserved item, which has a higher potential preference score, is assumed to have a smaller preference difference with the observed item of the user. Theoretically, DBPL could alleviate the vanishing gradient problem in the previous algorithm's learning procedure and obtain more accurate recommendations.

(5) A series of experiments over three real-world datasets are conducted to validate three recommendation models. Experimental results show the effectiveness of recommendation models for solving the one-class collaborative filtering problem. The experimental results of PBPR, PBPR* and DBPL are better than BPR, showing the effectiveness of assumptions proposed for recommendation models. Experimental results of the PBPR*-based method are better than the PBPR-based method in most cases. Experimental results of the DBPL-based recommendation method outperform other recommendation methods across all evaluation metrics on all datasets.