Engineering Applications of Artificial Intelligence 67 (2018) 283–295

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Social personalized ranking with both the explicit and implicit influence of user trust and of item ratings

Gai Li [a,*], Youfen Chen [a], Zhiqiang Zhang [a], Jianghong Zhong [a], Qiang Chen [b]

[a] School of Electronics and Information Engineering, Shunde Polytechnic, Guangdong Shunde 528333, China
[b] Department of Computer Science, Guangdong University of Education, Guangdong Guangzhou 510303, China

A R T I C L E I N F O

Keywords:
Recommended systems
Collaborative filtering
Social personalized ranking
Implicit feedback
Explicit feedback

A B S T R A C T

Due to the inherent deficiency of social collaborative filtering algorithms based on rating prediction, social personalized ranking algorithms based on ranking prediction have recently received much more attention in recommendation communities due to their close relationship with real industry problem settings. However, most existing social personalized ranking algorithms focus on either explicit feedback data or implicit feedback data rather than making full use of the information in the dataset. Until now, no studies have been done on social personalized ranking algorithms by exploiting both the explicit and implicit influence of user trust and of item ratings. In order to overcome the defects of prior researches and to further solve the problems of data sparsity and cold start of collaborative filtering, a new social personalized ranking model (SPR_SVD++) based on the newest xCLiMF model and TrustSVD model was proposed, which exploited both the explicit and implicit influence of user trust and of item ratings simultaneously and optimized the well-known evaluation metric Expected Reciprocal Rank (ERR) Experimental results on practical datasets showed that our proposed model outperformed existing state-of-the-art collaborative filtering approaches over two different evaluation metrics $NDCG$ and ERR, and that the running time of SPR_SVD++ showed a linear correlation with the number of users in the data collection and the number of observations in the rating and trust matrices. Due to its high precision and good expansibility, SPR_SVD++ is suitable for processing big data and has wide application prospects in the field of internet information recommendation.

1. Introduction

With the exponential growth of information generated on the World Wide Web, recommender systems, as one of the more efficient information filtering techniques, have attracted a lot of attention in the last decade. Recommender systems focus on solving the information overload problem by suggesting items that are of potential interest to users. Recommender systems have been applied to many areas on the Internet, such as the e-commerce system Amazon, the DVD rental system Netflix, and Google News (Adomavicius and Tuzhilin, 2005; Liu et al., 2013; Yang et al., 2013). Recommender systems are usually classified into three categories based on how recommendations are made: Content-based recommendations, Collaborative Filtering (CF), and Hybrid approaches (Adomavicius and Tuzhilin, 2005). The core idea of the content-based recommendations is to recommend an item to a user based upon a description of the item and a profile of the user' s interests. Collaborative Filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). Hybrid approaches are based on content-based recommendations and collaborative filtering, which combine both techniques to improve the quality of the recommendation. Compared to content-based approaches, Collaborative Filtering enjoys the advantage of being content-agnostic. In other words, Collaborative Filtering can recommend items without the additional computational expense or copyright issues involved with processing items directly. Collaborative filtering is widely acknowledged as one of the most successful recommender techniques (Adomavicius and Tuzhilin, 2005; Liu et al., 2013; Yang et al., 2013; Li and Chen, 2016; Li and Ou, 2016; Li et al., 2016).

Collaborative filtering algorithms can be divided into two categories: Collaborative Filtering (CF) algorithms based on rating prediction and Personalized Ranking (PR) algorithms based on ranking prediction (Li and Ou, 2016; Li et al., 2016; Pessiot et al., 2007; Shi et al., 2013c).

* Corresponding author.
E-mail address: ligai999@126.com (G. Li).

https://doi.org/10.1016/j.engappai.2017.10.006
Received 6 April 2017; Received in revised form 22 August 2017; Accepted 2 October 2017
0952-1976/

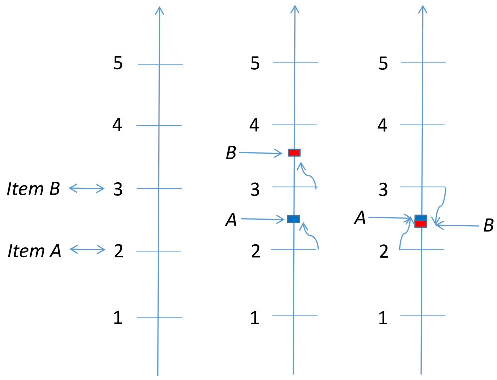

Fig. 1. The inherent deficiency of collaborative filtering algorithms based on rating prediction.

	Item			
User				2
		5		
		4		
	4		3	

	Item			
User	1		1	
		1		1
	1		1	
		1	1	

Fig. 2. The Explicit (Right) and implicit (Left) feedback data processed by collaborative filtering algorithm.

In collaborative filtering algorithms based on rating prediction, one predicts the actual rating for an item that a customer has not yet rated, and then ranks the items according to the predicted ratings. On the other hand, for Personalized Ranking algorithms based on ranking prediction, one predicts a preference ordering over the yet unrated items without going through the intermediate step of rating prediction.

However, Collaborative Filtering (CF) algorithms based on rating prediction have an inherent deficiency, which causes two equally good methods of predicting the ratings to possibly perform differently at predicting the rankings. A simple example can be found in Fig. 1: letting vector [2, 3] be the true ratings of items A and B respectively, r_1 =[2.5, 3.6] and r_2 =[2.5, 2.4] be two prediction vectors obtained from two different methods. r_1 and r_2 are equivalent with respect to the squared error (both errors equal $0.5^2 + 0.6^2$), while only r_1 predicts the correct rankings, as it scores B higher than A. Social collaborative filtering algorithms based on rating prediction also use the squared error to evaluate the performance of the models, so they have the same inherent deficiency as described above. A more detailed study of the performance evaluation problem in collaborative filtering can be found in Pessiot et al. (2007). Considering the inherent deficiency of collaborative filtering algorithms based on rating prediction, and from the recommendation perspective, the order over the items is more important than their rating in a real application. Therefore, in this paper, we focus on Personalized Ranking algorithms based on ranking prediction.

Traditional personalized ranking algorithms only utilize the user-item rating matrix for recommendation. The data in the user-item rating matrix processed by personalized ranking algorithm are divided into two categories: explicit feedback data (e.g.: ratings, votes) and implicit feedback data (e.g.: clicks, purchases) (Li and Chen, 2016; Li and Ou, 2016; Li et al., 2016). With the advent of online social networks, social trust aware personalized rankings have drawn lots of attention (Krohn-Grimberghe et al., 2012; Guo et al., 2015; Zhao et al., 2014; Yao et al., 2014). However, most existing social personalized ranking methods focus on either explicit feedback data or implicit feedback data rather than make full use of the information in the dataset. In fact, in most real-world recommender systems containing social networks both explicit and implicit user feedback are abundant and could potentially complement each other. It is desirable to unify these two heterogeneous forms of user feedback and social networks of users in order to generate more accurate recommendations. The idea of unifying these two heterogeneous forms of user feedback and social networks of users was first used in the TrustSVD model (Guo et al., 2016), where the three forms of feedback were combined via a factorized neighborhood model (called Singular Value Decomposition++, SVD++) (Koren, 2010), an extension of a traditional nearest item-based model in which the item–item similarity matrix was approximated via low rank factorization. However, TrustSVD is based on rating prediction and also has the same inherent deficiency as described in Fig. 1. Until now, nobody has studied social personalized ranking algorithms by exploiting both the explicit and implicit influence of user trust and of item ratings simultaneously.

In order to overcome the defects of prior researches and to further solve the problems of data sparsity and cold start of collaborative filtering, a new social personalized ranking model (SPR_SVD++) based on the newest xCLiMF model (Shi et al., 2013b) and TrustSVD model was proposed, which exploited both the explicit and implicit influence of user trust and of item ratings simultaneously and optimized the well-known evaluation metric Expected Reciprocal Rank (ERR). Experimental results on practical datasets showed that our proposed algorithm outperformed existing state-of-the-art collaborative filtering approaches over two different evaluation metrics $NDCG$ and ERR, and that the running time of SPR_SVD++ showed a linear correlation with the number of users in the data collection and the number of observations in the rating and trust matrices. With its high precision and good expansibility, SPR_SVD++ is suitable for processing big data, and has wide application prospects in the field of internet information recommendation.

The rest of this paper is organized as follows: Section 2 introduces previous related work; Section 3 demonstrates the problem formalization and TrustSVD model; a new social personalized ranking model (SPR_SVD++) is proposed in Section 4; the experimental results and discussion are presented in Section 5, followed by the conclusion and future work in Section 6.

2. Related Work

2.1. Collaborative Filtering

Collaborative filtering algorithms have been widely studied in both academic and industrial fields. Collaborative filtering algorithms can be divided into two categories: Collaborative filtering algorithm based on explicit feedback data (e.g.: ratings, votes) and Collaborative filtering algorithm based on implicit feedback data (e.g.: clicks, purchases) (Li and Ou, 2016; Li et al., 2016; Guo et al., 2016; Koren, 2010; Liu et al., 2010). Explicit feedback data are more widely used in the research fields of recommender systems. They are often in the form of numeric ratings from users expressing their preferences regarding specific items (Shi et al., 2013c; Yao et al., 2014; Srebro et al.). Implicit feedback data are easier to collect. The research on implicit feedback about CF is also called One-Class Collaborative Filtering (OCCF), in which only positive implicit feedback or only positive examples can be observed (Li and Chen, 2016; Li and Ou, 2016; Li et al., 2016; Krohn-Grimberghe et al., 2012; Guo et al., 2015; Zhao et al., 2014; Pan et al., 2015; Shi et al., 2013a; Pan and Chen, 2013; Rendle et al., 2009; Pan et al., 2008). The explicit and implicit feedback data can be expressed in matrix form as shown in Fig. 2. In the explicit feedback matrix, an element can be any real number, but often ratings are integers in the range (1∼5), such as the ratings on Netflix, where a missing element represents a missing example. In the implicit feedback matrix, the positive-only user preferences data can be represented as a single-valued matrix.