Expert Systems with Applications 42 (2015) 774–788

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Conditional preference in recommender systems

Wenyu Liu [a], Caihua Wu [c], Bin Feng [a], Juntao Liu [a,b,*]

[a] Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[b] 709th Research Institute, China Shipbuilding Industry Corporation, Wuhan 430074, China
[c] Section of Automatic Command, Huang Pi NCO School, Air Force Early-Warning Academy, Wuhan 430345, China

ARTICLE INFO

Article history:
Available online 9 September 2014

Keywords:
Conditional preference
Recommender system
Collaborative filtering
Matrix factorization
List-wise

ABSTRACT

By investigating the users' preferences in MovieLen dataset, we find that the conditional preference exists more widely in rating based recommender systems than imagined. Due to the high space complexity of the existing conditional preference representing models and the high computational complexity of the corresponding learning methods, conditional preference is seldom taken into consideration in the recommender systems. In this paper, we prove that the expressive ability of quadratic polynomial is stronger than that of linear function for conditional preference and propose to use quadratic polynomial to approximate conditional preference. Compared with the existing conditional preference model, the proposed model can save storage space and reduce learning complexity, and can be used in rating based recommender systems to efficiently process large amount of data. We integrate the proposed approximate conditional preference model into the framework of list-wise probabilistic matrix factorization (ListPMF), and verify this recommendation method on two real world datasets. The experimental results show that the proposed method outperforms other matrix factorization based methods.

1. Introduction

Recommender systems are deployed in online business systems widely (Koren & Bell, 2011; Park, Kim, Choi, & Kim, 2012), by which music (Lu & Tseng, 2009), movies (Choi, Ko, & Han, 2012), products (Hong, Li, & Li, 2012) and so on are recommended. One of the fundamental problems in these recommender systems is how to represent the user's preference. In daily life, people's preference may be different under different conditions. For example, we like hot drinks in winter, but we prefer cold drinks in summer. This kind of preference relations is called conditional preference. Conditional preference exists widely in real life. We investigate the user's preferences in MovieLen 1M dataset (Herlocker, Konstan, Borchers, & Riedl, 1999), and find that all of the users' preferences in this dataset are conditional (please see Section 3 for more details).

But most of the existing recommendation methods do not take the conditional preference into consideration. For example, the matrix factorization based methods, such as SVD model (Funk, 2006), Variational Bayesian based Matrix Factorization (VBMF) (Lim & Teh, 2007), Probabilistic Matrix Factorization (PMF) (Salakhutdinov & Mnih, 2008a), Bayesian Probabilistic Matrix Factorization (BPMF) (Salakhutdinov & Mnih, 2008b) and General PMF (GPMF) (Shan & Banerjee, 2010), use linear function to represent the user preference. In these methods, the rating value of user i for item j is estimated by the inner product of the user latent feature vector u_i and item latent feature vector v_j, that is $\hat{R}_{i,j} = u_i v_j^T$. The rating value reflects the favorite degree of user i for item j. That is, the user i's preference is represented by a linear function. However, in this paper, we prove that linear function cannot represent conditional preference. Please see Example 1 and Theorem 1 for more details.

The existing conditional preference representing models, including CP-nets (Boutilier, Brafman, Hoos, & Poole, 1999) and its variations (Brafman, Domshlak, & Shimony, 2006; Boutilier, Bacchus, & Brafman, 2001; Gonzales & Perny, 2005; Lang & Mengin, 2009), have two shortcomings. First, the space complexity is high. In these models, the number of conditional preference rules is exponential to the number of the variables. Second, the corresponding learning problem is intractable even under some simplified conditions (Chevaleyre, Koriche, Lang, Mengin, & Zanuttini, 2011; Dimopoulos, Michael, & Athienitou, 2009). Although some approximate learning methods are proposed (Liu, Xiong, Wu, Yao, & Liu, 2014a; Liu, Yao, Xiong, Liu, & Wu, 2013a), the computational complexity is still high. These two significant shortcomings make the existing conditional preference representing models

* Corresponding author at: 709th Research Institute, China Shipbuilding Industry Corporation, Wuhan 430074, China.

E-mail addresses: liuwy@hust.edu.cn (W. Liu), wucaihua_1999@yahoo.com.cn (C. Wu), fengbin@mail.hust.edu.cn (B. Feng), prolay@163.com (J. Liu).

http://dx.doi.org/10.1016/j.eswa.2014.08.044
0957-4174/

not suitable for recommender systems. The huge amount of data in recommender systems need to be processed efficiently. So the space complexity and the time complexity of the preference representing model are the critical factors to be considered. To integrate the conditional preference into recommender systems, an approximate model should be used. In this paper, we prove that quadratic polynomial can approximate conditional preference better than linear function. For more details, please see Example 1 and Theorem 1. For this reason, we propose to use quadratic polynomial to approximate conditional preference in recommender systems, and integrate it into the framework of ListPMF (Liu, Wu, Xiong, & Liu, 2014b). The proposed method is evaluated on Epinions dataset (Massa & Avesani, 2006) and MovieLen dataset (Herlocker et al., 1999). The recommendation performance is measured by the Normalized Discounted Cumulative Gain (NDCG) (Järvelin & Kekäläinen, 2000) and the Expected Reciprocal Rank (ERR) (Chapelle, Metlzer, Zhang, & Grinspan, 2009), which are the popular measures in information retrieval. The experimental results show our method gets more satisfying recommendation results than the other methods based on matrix factorization.

The main contributions of our work are threefold. (1) We find that most of the users' preferences in rating based recommender systems are conditional. (2) We prove that quadratic polynomial can approximate conditional preference. (3) We integrate the proposed approximate conditional preference model into ListPMF for recommendation and get more satisfying recommendation results.

The rest of this paper is organized as follows. In Section 2, a brief survey of conditional preference and recommendation methods based on matrix factorization is provided. In Section 3, conditional preference in MovieLen dataset is investigated. Section 4 introduces the quadratic polynomial approximation of conditional preference. The ListPMF recommender method integrating approximate conditional preference model is described in Section 5. The experimental results are presented and analyzed in Section 6 followed by the conclusions and further work in Section 7.

2. Background and related work

In this section, we briefly introduce the background and review the related research fields of conditional preference and matrix factorization based recommendation.

2.1. Conditional preference

Conditional preference arises when a user's preferences are different under different conditions. CP-net (Boutilier et al., 1999) is proposed to represent conditional preference. A CP-net is a digraph, whose nodes correspond to variables. Each node is annotated with a conditional preference table. The preferential dependency is represented by the graph structure. CP-nets is extended to improve its expressive ability. For example, Brafman et al. (2006) propose TCP-nets (Trade-off Conditional Preference Networks). Wilson (2004) extends CP-nets with stronger conditional preference statement. In these models, conditional preference is represented by the conditional preference rules. The number of conditional preference rules is exponential to the number of the feature variables. This makes the space complexity of these models too high. To simplify the conditional preference representing model, some approximate models are proposed. For example, Domshlak (2003) adds soft constraint in CP-nets to represent conditional preference approximately. Lang and Mengin (2009) propose separable ceteris paribus structures (SCP-structures), which is a CP-nets without conditional preferential dependency. Dominance testing and consistency testing are still difficult in CP-nets. Here, dominance testing means deciding the preference relation between two given outcomes. Consistency testing means deciding whether a given CP-net is consistent. Dominance testing and consistency testing are the most commonly used operations in recommender systems. The computational complexity of these two testings is exponential to the number of feature variables. So CP-nets and its extensions are seldom used in recommender systems. To facilitate the dominant testing and consistency testing, Boutilier et al. (2001) replace the conditional preference rule with utility function under the assumption of additive independence, and propose UCP-nets. Chatel, Truck, and Malenfant (2010) propose LCP-nets by integrating TCP-nets (Brafman et al., 2006) and UCP-nets (Boutilier et al., 2001). Engel and Wellman (2008) propose UCI-nets under the assumption of conditional independence. Gonzales and Perny (2005) propose GAI-networks under the assumption of additive independence. Computational efficiency of these models is improved, but the assumptions of these models are too strong.

In terms of model learning, Boutilier et al. (2001) propose an algorithm for learning CP-nets over a fixed acyclic digraph. Koriche and Zanuttini (2010) propose an active learning algorithm. Lang and Mengin (2009) address the problem of learning SCP-structures. Dimopoulos et al. (2009) propose an algorithm for generating an acyclic CP-net entailing all examples. The algorithm is proved to be a PAC-learner if the learning examples are transparently entailed by CP-nets (see Definition 5 in Dimopoulos et al., 2009). Liu et al. (2014a) propose a CP-net learning method from inconsistent learning samples. A CP-net learning method based on hypothesis testing is proposed (Liu et al., 2013a), which can get satisfying results with lower computational complexity.

Only a few work using conditional preference in recommender systems. Yu, Yu, Zhou, and Nakamura (2009) propose to model conditional preference in recommender system using domain knowledge. Zhang, Dong, and Zhang (2014) propose to consider conditional preference based on domain knowledge in a simple travel recommender system.

2.2. Matrix factorization

Matrix factorization based recommendation methods are one kind of collaborative filtering methods (Koren & Bell, 2011). Although conventional matrix factorization based recommendation methods (Funk, 2006; Lim & Teh, 2007; Salakhutdinov & Mnih, 2008a, 2008b; Shan & Banerjee, 2010) gain great success, they face the challenges of the data sparsity problem and the cold-start problem same as other collaborative filtering methods (Koren & Bell, 2011). To alleviate these problems, some methods have been proposed to use side information, such as social relation and tag information. These methods can be divided into two categories: regularization-based methods (Jamali & Ester, 2011; Li & Yeung, 2009; Ma, Zhou, Liu, Lyu, & King, 2011a; Zhu, Ma, Chen, & Bu, 2011; Yuan, Chen, & Zhao, 2011; Wu et al., 2012) and factorization-based methods (Ma, Yang, Lyu, & King, 2008; Ma, Zhou, Lyu, & King, 2011b; Singh & Gordon, 2010; Yuan et al., 2011; Zeng & Chen, 2013). In Regularization-based methods, regularization terms, which measure the difference of the related users or items, are added to the loss function. In factorization-based methods, relation matrix representing social relation or other relation is factored as well as the rating matrix. The weighted sum of the relation matrix factorization error and the rating matrix factorization error is minimized. Beside the two kinds of methods mentioned above, there are some recommendation methods proposed to use side information by other ways (Chen, Li, Yang, & Yu, 2013; Liu, Wu, & Liu, 2013b; Zhou, Shan, Banerjee, & Sapiro, 2012).

Traditional matrix factorization based recommendation methods take the sum-of-squares of factorization errors as the objective function. Recently, some researchers propose to optimize