Personalized Recommendation via Cross-Domain Triadic Factorization

Liang Hu
Department of CSE
Shanghai Jiaotong University
lianghu@sjtu.edu.cn

Jian Cao*
Department of CSE
Shanghai Jiaotong University
cao-jian@sjtu.edu.cn

Guandong Xu
Advanced Analytics Institute
University Technology Sydney
guandong.xu@uts.edu.au

Longbing Cao
Advanced Analytics Institute
University Technology Sydney
longbing.cao@uts.edu.au

Zhiping Gu
Department of Electrical Engineering
Shanghai Technical Institute of Electronics & Information
guzhiping@stiei.edu.cn

Can Zhu
Department of CSE
Shanghai Jiaotong University
0627wshhg@sjtu.edu.cn

ABSTRACT
Collaborative filtering (CF) is a major technique in recommender systems to help users find their potentially desired items. Since the data sparsity problem is quite commonly encountered in real-world scenarios, Cross-Domain Collaborative Filtering (CDCF) hence is becoming an emerging research topic in recent years. However, due to the lack of sufficient dense explicit feedbacks and even no feedback available in users' uninvolved domains, current CDCF approaches may not perform satisfactorily in user preference prediction. In this paper, we propose a generalized Cross Domain Triadic Factorization (CDTF) model over the triadic relation *user-item-domain*, which can better capture the interactions between domain-specific user factors and item factors. In particular, we devise two CDTF algorithms to leverage user explicit and implicit feedbacks respectively, along with a genetic algorithm based weight parameters tuning algorithm to trade off influence among domains optimally. Finally, we conduct experiments to evaluate our models and compare with other state-of-the-art models by using two real world datasets. The results show the superiority of our models against other comparative models.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: Information Filtering; I.2.6 [**Artificial Intelligence**]: Parameter Learning

General Terms
Algorithms, Performance, Experimentation, Human Factors

Keywords
Recommender System, Cross-Domain Collaborative Filtering, Triadic Factorization

1. INTRODUCTION
ACM 978-1-4503-2035-1 /13/05.The huge and ever fast increasing amount of information on the Internet has penetrated every corner of our life. However, we become more easily overwhelmed by so much information and unable to find what we really desire. When we follow events on Facebook, buy books on Amazon or add apps into a smartphone, systems may record our feedbacks, e.g., a rating assigned to a book. Based on such observed feedbacks (or ratings) collected from like-minded users, collaborative filtering (CF) in recommender systems can predict personalized preferences to unconsumed items. In general, CF methods can be sub-divided into neighborhood-based and model-based approaches [5; 22; 26]. Therein, latent factor model based on matrix factorization (MF) [6; 9] has gained the dominance in recent years.

The essence to success in CF is highly dependent on the feedback data. However, users are not always willing to provide feedbacks due to various personal reasons. Even some applications possess the data sparsity problem in nature, for instance, users who has bought a new car recently may not have a new car purchase plan in next five years. Thus most CF methods, including MF, suffer from the data sparsity [26] and cold-start [9; 23] issues. The lack of reliable feedback data has become a major barrier for CF methods.

To deal with the sparsity issue, Cross-Domain Collaborative Filtering (CDCF), which leverages the information from multiple related domains, is an emerging research topic in recent years. Some CDCF algorithms have been proposed in literatures [11], where the basic idea is based on the assumption of the existence of multiple related domains and the user preference learned from one dense domain, e.g. movies, can be re-used to make prediction in a sparse domain, e.g. books (i.e. cross domain learning) [12; 18]. An early neighborhood based CDCF (N-CDCF) was mentioned in [1], but it can only provide a very local optimum solution as done by neighborhood based CF models (further analysis provided in the next section). Recently, some cross-domain matrix factorization (CDMF) models [18; 24] have been proposed to overcome the local optimum problem of N-CDCF. The underlying idea of CDMF can be illustrated using Figure 1 (b), where user factor matrix **U** serves as the bridge to transfer knowledge from auxiliary domain (*A*) to target domain (*T*).

Most CDMF models assume the auxiliary data is relative dense for all users or items [18]. However, we argue that this assumption is not always true in real world. In general, our argument is based on the well-known power law, as illustrated in Figure 1(a), only the minority of users are rating frequently while the majority of users are quite inactive in providing feedback. This observation might impact the hypothesis of traditional CDCF approaches, therefore resulting in the deterioration of recommendation performance.

*Jian Cao is the corresponding author

WWW 2013, May 13--17, 2013, Rio de Janiero, Brazil.
ACM 978-1-4503-2035-1 /13/05.