Physica A 444 (2016) 713–721

Contents lists available at ScienceDirect

Physica A

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

Personalized recommendation based on heat bidirectional transfer

Wenping Ma *, Xiang Feng, Shanfeng Wang, Maoguo Gong

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China

HIGHLIGHTS

- We proposed a multi-objective population based incremental learning algorithm (MOPBIL).
- MOPBIL combines multi-objective evolutionary algorithms with competitive learning.
- The initialization of learning probability and the probability update mechanism.
- A random perturbation operator is introduced for local optimum problem.
- Simulation results reflect the effectiveness of MOPBIL.

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form 9 September 2015
Available online 30 October 2015

Keywords:
Recommender systems
Information filtering
Bipartite network
Long tail
Heat transfer

ABSTRACT

Personalized recommendation has become an increasing popular research topic, which aims to find future likes and interests based on users' past preferences. Traditional recommendation algorithms pay more attention to forecast accuracy by calculating first-order relevance, while ignore the importance of diversity and novelty that provide comfortable experiences for customers. There are some levels of contradictions between these three metrics, so an algorithm based on bidirectional transfer is proposed in this paper to solve this dilemma. In this paper, we agree that an object that is associated with history records or has been purchased by similar users should be introduced to the specified user and recommendation approach based on heat bidirectional transfer is proposed. Compared with the state-of-the-art approaches based on bipartite network, experiments on two benchmark data sets, *Movielens* and *Netflix*, demonstrate that our algorithm has better performance on accuracy, diversity and novelty. Moreover, this method does better in exploiting long-tail commodities and cold-start problem.

1. Introduction

With the rapid development of Internet and World Wide Web, people who want to find information they need are often confused by abundant online contents. Obviously, effective information provides us convenience, while information overload is likely to destroy our life, because we have to spend a lot of time and energy in finding useful contents carefully in the ocean of big data. Therefore, on one hand we need an effective approach to distinguish useful information and junk; on the other hand, merchants want to get quick turnover by analyzing a mass of previous data. The fact that human needs machine to filter out worthless information leads to the birth of recommender systems. In short, it is difficult to cater to all

* Corresponding author.
E-mail address: wpma@mail.xidian.edu.cn (W. Ma).

http://dx.doi.org/10.1016/j.physa.2015.10.068

tastes, so a list with only popular items cannot satisfy the needs of the multitude. Personalized recommendation is intended to provide accurately reasonable objects for customers with various tastes. Personalized recommender systems have been used to many fields widely, especially in e-commerce [1,2], such as *Amazon.com*, *Netflix.com* and *Youtube.com* and so on. According to statistics, 2/3 of the movies watched are recommended in *Netflix.com*, and recommendation generates 38% more clickthrough for *google* news and 35% sales for *Amazon*. Furthermore, developer recommendation has also become a hot area of research [3–5].

To take advantage of information and offer personalized recommendation to users, various personalized recommendation algorithms have been developed, including collaborative filtering (CF) approaches [6–8], content-based methods [9,10], latent semantic models [11,12] and hybrid algorithms [13]. Traditional methods recommend objects to users on the basis of similarity, either of user-based or object-based or both. Forecast based on only first-order relevance is not accurate for most users who lack enough past information, so accuracy of collaborative filtering is limited. Network-based recommender systems with higher-order relevance are proposed to improve the accuracy of recommendation. Network-based inference (NBI) algorithm proposed in Ref. [14] is a three-steps mass diffusion starting from the target user on bipartite network. This method has higher accuracy than the classical CF algorithms through information transfer, with lower complexity.

Providing accurate suggestions to users is the principal task for recommender systems, but it is not enough. In Ref. [15], McNee et al. showed that recommendation algorithms that focus only on accuracy may not be the best, because of exceptionally similar recommendations provided, which is detrimental to recommender systems. Customers are often fed up with recommended list that lacks surprise, and merchants are also glad to sale inventories that are relatively unpopular. A good personalized recommender systems also need to suggest diverse items not easily discovered by users. The ability of systems to recommend long-tail goods will count for much. In order to measure this ability, evaluation metrics diversity and novelty were introduced [2,16–20]. Recommender systems that are simultaneously accurate, novel and diverse may lead to a conflicting problem. Most traditional approaches that excessively emphasize accuracy and ignore diversity and novelty cannot satisfy personalized requirement, while algorithms based diffusion can generate more diverse and novel suggestions.

In view of advantages on accuracy, diversity and novelty, diffusion-based algorithms have good prospects for development. The heat conduction (Heats) algorithm was proposed in Refs. [21,22]. This method based on the process of heat conduction tries to recommend users some less popular objects, which results in high novelty at the cost of accuracy. In fact, an algorithm with poor accuracy is meaningless, since it will not meet the users needs anymore [23]. Considering that the network-based inference and the heat conduction complement each other, Zhou et al. proposed a hybrid recommendation algorithm (HPH) in Ref. [22] to find the trade-off between forecasting accuracy and user experience. Combining the NBI and Heats methods, the HPH algorithm outperforms both of the separate methods. Some subsequent extensions have been developed to further improve HPH successively by changing fusion strategy, for example, the preferential diffusion [24] and network manipulation [25]. Besides, on the basis of the Heats algorithm, Liu et al. explained theory of Heats in detail and put up an improved method named the biased heat conduction [26] (BHC) which gets apparent improvement on accuracy by taking into account the object degree effect in the last diffusion step. Very recently, a novel algorithm named consistence-based inference (CBI) was proposed to recommend objects on the basis of consistent interests rather than only causality relationship [27]. In the view of the CBI, the object B will be introduced to the target user who has collected the object A only if recommendation strengths from A to B and from B to A are both high.

We get many illumination from The K-Nearest Neighbor Mass Diffuse (KNNMD) method [28] and try to find out an algorithm without setting the predetermined value for K. It is just based on this consideration that we propose an algorithm based on heat bidirectional transfer (HBT). it is easy to get better recommended results by selecting relevant objects among a group of customers who have similar tastes, so we agree that the object B will be introduced to the target user who has collected the object A only if A is associated with B and the target users share common interests with the other users who have purchased B. Four metrics are used to evaluate algorithmic quality and experiments on two famous data sets show the outperformance of this algorithm. Compared with some mainstream approaches, our algorithm makes significant gains in recommending diverse and novel objects, meanwhile accuracy gets improvement.

2. Bipartite network-based algorithms

In commercial applications, merchants need to offer the most appropriate recommendation lists containing a few items to users, and the problem is called Top-N recommendation [29]. So in this paper, what we focus on is not ratings ranged from 1 to 5, but whether one object is collected or not. We suppose the total number of users is m and that of objects is n. Links between users and objects can be represented by a (user–object) binary matrix $A_{m\times n}$, where $a_{i\alpha} = 1$ if object α is collected by user i and $a_{i\alpha} = 0$ otherwise. We assume the length of recommendation list is N. The task of Top-N recommendation is to generate a list with few uncollected objects for the target user.

In the Heats algorithm [19], transfer of information is equivalent to resource allocation, which is a random walk based process. Heats works by assigning objects an initial level of resource denoted by the vector f (where f_α is the resource possessed by object α), and then redistributing it via the transformation $\widetilde{f} = W^H f$, where

$$W^H_{\alpha\beta} = \frac{1}{k(o_\alpha)} \sum_{l=1}^{m} \frac{a_{l\alpha} a_{l\beta}}{k(u_l)} \tag{1}$$