Physica A 456 (2016) 271–280

Contents lists available at ScienceDirect

Physica A

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

Improved personalized recommendation based on a similarity network

Ximeng Wang, Yun Liu *, Fei Xiong

Key Laboratory of Communication & Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, Beijing 100044, PR China

HIGHLIGHTS

- We improve network-based inference model based on a similarity network.
- User influence is introduced into network-based recommender systems and make recommendations more personalized.
- We improve resource-allocation process and make it more reasonable.
- Our proposed approaches obtain better recommendations on accuracy and diversity than some other network-based recommendation models.

ARTICLE INFO

Article history:
Received 13 October 2015
Received in revised form 7 February 2016
Available online 2 April 2016

Keywords:
Similarity network
Resource-allocation
Network-based inference
Bipartite graphs
Personalized recommendation

ABSTRACT

A recommender system helps individual users find the preferred items rapidly and has attracted extensive attention in recent years. Many successful recommendation algorithms are designed on bipartite networks, such as network-based inference or heat conduction. However, most of these algorithms define the resource-allocation methods for an average allocation. That is not reasonable because average allocation cannot indicate the user choice preference and the influence between users which leads to a series of non-personalized recommendation results. We propose a personalized recommendation approach that combines the similarity function and bipartite network to generate a similarity network that improves the resource-allocation process. Our model introduces user influence into the recommender system and states that the user influence can make the resource-allocation process more reasonable. We use four different metrics to evaluate our algorithms for three benchmark data sets. Experimental results show that the improved recommendation on a similarity network can obtain better accuracy and diversity than some competing approaches.

1. Introduction

In recent years, with the explosive development of the Internet, the amount of information worldwide is increasing far more quickly than our ability to process it. This information overload problem becomes more and more serious and increases the significance of having effective information filtering methods. To solve this problem, research into different areas invented various tools, such as search engines and recommender systems. However, these search engines do not consider user preference and return the same results to people who may have different habits. To the contrary, recommender systems

* Corresponding author.
E-mail addresses: wangxm@bjtu.edu.cn (X. Wang), liuyun@bjtu.edu.cn (Y. Liu), xiongf@bjtu.edu.cn (F. Xiong).

http://dx.doi.org/10.1016/j.physa.2016.03.070

are a personalized information filtering technology that when used to predict an active user will like a particular item based on the particular history of user choices. Using the contributions of the recommender system, researchers in physics and computer science have paid huge attention to them and recently proposed several advanced recommendation algorithms.

Collaborative filtering (CF) [1–3], content-based analysis [4,5] and matrix factorization [6,7] have been proposed by computer scientists, while physicists introduced diffusion-based approaches, such as network-based inference (NBI) [8,9] and heat conduction (HC) [10,11]. Collaborative filtering is one of the successful and fashionable algorithms that includes two different kinds of methods called user-based CF and item-based CF. More specifically, the user-based CF mainly evaluates the tastes that differ between users, while the item-based CF uses historical data to calculate the similarities between items. Generally, item-based CF has higher accuracy in practice [12]. Some researchers have further improved collaborative filtering by unifying the user-based and item-based approaches to obtain good performance [13,14], while simultaneously, others have enhanced effects by combining collaborative filtering with clustering models [15,16]. Nevertheless, the performance of collaborative filtering is strongly limited by the sparseness of data, because users do not willingly rate purchased items and the already rated items are only a small fraction of the huge number of total items [17,18]. As a means of improving algorithms to resolve this problem, some accessorial information can be regarded as prior information to raise that accuracy [19,20].

Personalized recommendation based on physical processes has attracted the attention of many researchers [21]. Some researchers have analyzed the memory effect of online ratings and online user preferences to find the user behavior pattern, indeed regarded as a significant feature and then added into recommender algorithms [22–24]. A lot of effort has been devoted to designing better recommendation algorithms based on bipartite graphs [9,25]. Network-based inference (NBI) is a typical recommendation approach that is based on bipartite network [8], which successfully increases recommendation effectiveness and reduces the dimensions of user–item vectors for computation [26]. Then, many recommendation approaches have been conducted later on NBI related models, such as combining ProbS with HeatS and providing an integrated recommendation [27]. In addition, NBI related models are widely used in a variety of domains, including user behavior analysis, railway transportation systems, user link predictions, and others [28–31].

NBI related models allocate all resources equally between neighbors in a user–item bipartite network without considering the user's special preference, which will cause the recommendation results not to be personalized. Furthermore, NBI cannot consider the influence between users, so similar users always have the same choice preference in general, which is a very significant characteristic of recommender systems. In this paper, we improve the NBI model through a similarity network that combines user similarity and item similarity in a bipartite network. We propose an information filtering approach to improve certain aspects of the recommender system, such as precision, recall and others. The cold-start problem is used to describe the dilemma of sparse data [32], and we enhance the diversity of recommendation lists to help new items be pushed out, thereby relieving the dilemma of the cold-start problem. Users always pay more attention to the items at the front of recommendation lists, so our approach provides better ranking of recommendation lists to users, which means that users will have a greater probability of choosing the items in our recommendation lists.

The remainder of this paper is organized as follows: Section 2 presents the original NBI model and our improved recommendation approach. Section 3 presents our experiments, and Section 4 concludes this paper with discussion and findings.

2. Model and methods

2.1. The network-based inference model

The NBI model is widely used for object-oriented recommendations in traditional bipartite networks. It takes advantage of resource allocation process to recommend non-collected items to users on the bipartite network, and these consist of a set of user nodes and item nodes. The relationship between the user and the item can be presented by an adjacent matrix, where every element a_{ij} is defined in Eqs. (1):

$$a_{ij} = \begin{cases} 1, & \text{user } j \text{ collects item } i \\ 0, & \text{else.} \end{cases} \tag{1}$$

The original NBI model is proposed in Ref. [8]. It first distributes the resource from item to its neighboring users on average and then redistributes the total resources of every user to its neighboring items with an equal probability. The recommendation list can then be generated according to the final resource of items, so that the item that has a relatively large degree can get more resources and is placed at the front of the recommendation list. The model is presented in Eq. (2):

$$f'(I_i) = \sum_{j=1}^{n} \frac{1}{k(I_j)} \sum_{\beta=1}^{m} \frac{a_{i\beta} a_{j\beta}}{k(U_\beta)} f(I_j) \tag{2}$$

where $f(I_j)$ and $f'(I_i)$ are the original value of item j and the final value item i; $k(I_j)$ and $k(U_\beta)$ mean the degree of item j and user β, respectively.

Nevertheless, distributing the resources on a bipartite network with equal probability may cause some problems. First, the results of recommendations are not personalized because the average distribution is not aimed at the target user,