Information Processing and Management 000 (2017) 1–19

Contents lists available at ScienceDirect

Information Processing and Management

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

A social recommendation method based on an adaptive neighbor selection mechanism

Sajad Ahmadian, Majid Meghdadi*, Mohsen Afsharchi

Department of Computer Engineering, University of Zanjan, Zanjan, Iran

ARTICLE INFO

Article history:
Received 9 November 2016
Revised 30 January 2017
Accepted 10 March 2017
Available online xxx

Keywords:
Recommender systems
Adaptive neighbor selection
Confidence
Reliability
Trust

ABSTRACT

Recommender systems are techniques to make personalized recommendations of items to users. In e-commerce sites and online sharing communities, providing high quality recommendations is an important issue which can help the users to make effective decisions to select a set of items. Collaborative filtering is an important type of the recommender systems that produces user specific recommendations of the items based on the patterns of ratings or usage (e.g. purchases). However, the quality of predicted ratings and neighbor selection for the users are important problems in the recommender systems. Selecting suitable neighbors set for the users leads to improve the accuracy of ratings prediction in recommendation process. In this paper, a novel social recommendation method is proposed which is based on an adaptive neighbor selection mechanism. In the proposed method first of all, initial neighbors set of the users is calculated using clustering algorithm. In this step, the combination of historical ratings and social information between the users are used to form initial neighbors set for the users. Then, these neighbor sets are used to predict initial ratings of the unseen items. Moreover, the quality of the initial predicted ratings is evaluated using a reliability measure which is based on the historical ratings and social information between the users. Then, a confidence model is proposed to remove useless users from the initial neighbors of the users and form a new adapted neighbors set for the users. Finally, new ratings of the unseen items are predicted using the new adapted neighbors set of the users and the *top_N* interested items are recommended to the active user. Experimental results on three real-world datasets show that the proposed method significantly outperforms several state-of-the-art recommendation methods.

1. Introduction

In recent years, the volume of data on the web has been increasing at an unprecedented rate and the information overload creates difficulties for the Internet users. Therefore, it is difficult for the users to find useful information among the available choices. Recommender systems (RSs) are used to make personalized recommendations for the users to overcome the information overload problem. In other words, the RSs help the users to find items (e.g. books, movies, news, etc.) of interest from a plethora of available choices. In order to provide high quality recommendations by the RSs, these systems

* Corresponding author.
E-mail addresses: s.ahmadian@znu.ac.ir (S. Ahmadian), meghdadi@znu.ac.ir (M. Meghdadi), afsharchim@znu.ac.ir (M. Afsharchi).

http://dx.doi.org/10.1016/j.ipm.2017.03.002
0306-4573/

need to predict and compare the utility of items and then decide what items to recommend based on this comparison (Ricci, Rokach, Shapira, & Kantor, 2011).

Collaborative filtering (CF) is one of the recommender systems methods which is used to exploit information about the past behavior or the opinions of an existing user to predict unrated items and find the items that the current user of the system will most probably like or be interested in. The basic idea of the CF-based methods is that if the users shared the same interests in the past then they will have similar tastes in the future (Cechinel, Sicilia, Alonso, & Barriocanal, 2013; Domingues, Jorge, & Soares, 2013; Park, Park, Jung, & Lee, 2015). In these methods, a matrix of given user-item ratings is used as the input to calculate similarity values between the users. Moreover, these methods use the similarity values to determine the distance between a pair of the users and also form neighbors set of the active user. Therefore, the opinions of the users in the neighbors set of the active user are used to provide suitable recommendations for the active user in these methods. The CF-based methods are often classified into two groups including memory-based and model-based methods. In the memory-based methods, the original user-item ratings matrix is stored in memory and used directly for making the recommendations. On the other hand, in the model-based methods, the raw data are processed offline as item-based filtering or some dimensionality reduction approaches. Then, the pre-computed or learned model is used to make predictions at run time. Some of the model-based methods have been presented in the literature including probabilistic models (Javari & Jalili, 2015; Ma, Zhang, Liu, Li, & Yuan, 2016), clustering models (Bilge & Polat, 2013; Tsai & Hung, 2012), dimensionality reduction methods (Hernando, Bobadilla, & Ortega, 2016; Hong, Zheng, & Chen, 2016), pattern mining techniques (Tsai & Lai, 2015), latent semantic models (Hofmann, 2004), and Markov decision process models (Shani, Brafman, & Heckerman, 2005). Although the memory-based methods are more widely used than the model-based methods, such methods face problems of scalability of the systems with tens of millions of users and millions of items (Ricci et al., 2011).

The CF-based methods often suffer from several shortcomings which lead to reduce of the system performance. These methods use the user-rating matrix to compute the similarity values between the users and form the neighbors set of the active user (Kaššák, Kompan, & Bieliková, 2016; Liu, Hu, Mian, Tian, & Zhu, 2014). Therefore, the CF-based methods need the extensive data which contains exclusively the ratings made by the users over most of the items to make efficient recommendations. However, in real-world applications, the user-rating matrixes tend to be very sparse which means the users typically provide ratings for only a small number of the items. This problem is called data sparsity problem which makes to reduce the performance of the CF-based methods for finding the neighbors set of the active user. Moreover, the cold-start problem can be viewed as a special case of the data sparsity problem. This problem is about the new users that have not rated any item yet or the users who have rated a few items. In addition, recommending the items that have not been rated or bought yet is another challenge of the CF-based methods, as there are not available sufficient feedbacks on these items. Malicious attacks are another problem of the CF-based methods which can reduce the reliability of the systems (O'Mahony, Hurley, & Silvestre, 2005). A malicious attack occurs when a user tries to influence the functioning of the system intentionally. In other words, a malevolent user might try to influence the behavior of the recommender system in such a way that includes a certain item very often in its recommendation list.

To overcome the mentioned problems of the CF-based methods, several approaches have been proposed in the literature. Trust-aware recommender systems aim to exploit the social information from trust networks of the users to improve the system performance of the CF-based methods (Fang, Guo, & Zhang, 2015; Kim & Phalak, 2012; Yan, Zheng, Chen, & Wang, 2013). In these systems, the opinions of the trusted neighbors can be used as a starting point for predicting unrated items and making recommendations to the users. Therefore, these systems can be able to alleviate the cold-start problem and improve on the user coverage measure. Moreover, the trust statements between the users help to make the recommender systems more robust against malicious attacks. It is due to this fact that desired trust relationships to a fake profile cannot easily be injected into a recommender database.

Neighbor selection for the users is one of the most important issues in recommender systems which it has a high impact on the accuracy of the predicted ratings in these systems. Most of the recommendation methods use the identified neighbors set of the users to predict unseen items and produce some recommendations to the users. However, these neighbors of the users may not be useful to predict all of the unseen items and it leads to reduce the accuracy of the predicted ratings in the recommendation process. In this paper, a social recommendation method is proposed which is based on an adaptive neighbor selection mechanism to improve the accuracy of ratings prediction and resolve neighbor selection problem in the recommender systems. To this end, the initial neighbors set of the users is calculated using clustering algorithm which is based on the combination of the similarity values and the trust information between each pair of the users. Then, the initial ratings of the unseen items are predicted using the initial neighbors set of the users and also the quality of the predicted ratings is evaluated using a reliability measure which is based on the similarity values and the trust statements between the users. Then, an adaptive neighbor selection mechanism is proposed to calculate a new neighbors set for the users who their initial neighbors sets have a low reliability measure to predict a target item. For this purpose, a novel confidence model between the users is proposed to identify and remove the useless users from the initial neighbors set of the users. Finally, the final ratings of the unseen items are predicted using the new adapted neighbors set of the users and the *top_N* interested items are recommended to the active user.

The remainder of this paper is organized as follows: Related studies are reviewed in Section 2. Section 3 introduces the proposed method. In Section 4, the proposed method is compared with the state-of-the-art methods by performing several experiments on three well-known datasets. Finally, some concluding comments are discussed in Section 5.

Please cite this article as: S. Ahmadian et al., A social recommendation method based on an adaptive neighbor selection mechanism, Information Processing and Management (2017), http://dx.doi.org/10.1016/j.ipm.2017.03.002