Future Generation Computer Systems 78 (2018) 419–429

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

TCARS: Time- and Community-Aware Recommendation System

Fatemeh Rezaeimehr [a], Parham Moradi [a,*], Sajad Ahmadian [a], Nooruldeen Nasih Qader [b], Mahdi Jalili [c]

[a] Department of Computer Engineering, University of Kurdistan, Sanandaj, Iran
[b] Computer Science Department, University of Human Development, Iraq
[c] School of Engineering, RMIT University, Melbourne, Australia

HIGHLIGHTS

- We proposed a time-aware recommendation method called TCARS.
- A reliability measure is used to evaluate the quality of the initial predicted rates.
- A novel overlapping community detection method is proposed to group similar users.
- Experiments were performed on two real-world datasets.
- TCARS obtained accurate results compared to several state-of-the-art methods.

ARTICLE INFO

Article history:
Received 20 June 2016
Received in revised form
30 March 2017
Accepted 1 April 2017
Available online 9 April 2017

Keywords:
Recommender systems
Social networks
Network science
Overlapping community structure
Reliability

ABSTRACT

With the abundance of information produced by users on items (e.g., purchase or rating histories), recommender systems are a major ingredient of online systems such as e-stores and service providers. Recommendation algorithms use information available from users–items interactions and their contextual data to provide a list of potential items for each user. These algorithms are constructed based on similarity between users and/or items (e.g., a user is likely to purchase the same items as his/her most similar users). In this work, we introduce a novel time-aware recommendation algorithm that is based on identifying overlapping community structure among users. Users' interests might change over time, and accurate modeling of dynamic users' preferences is a challenging issue in designing efficient personalized recommendation systems. The users–items interaction network is often highly sparse in real systems, for which many recommenders fail to provide accurate predictions. The proposed overlapping community structure amongst the users helps in minimizing the sparsity effects. We apply the proposed algorithm on two real-world benchmark datasets and show that it overcomes these challenges. The proposed algorithm shows better precision than a number of state-of-the-art recommendation methods.

1. Introduction

With the explosion of online data, finding the right information at the right time has become a difficult and time-consuming challenge. This is known as information overload problem. Web users produce large-scale data from their daily activities, when they make online shopping, search keywords, watch movies, participate in social networks or play online games. Big data mining and analytics have become one of the core problems in various disciplines including computer science, information technology, mathematics and social sciences [1]. Recommender Systems (RSs) have been introduced to overcome this problem by providing the users with personalized recommendations. When users purchase (i.e. rate or click) items, such information can be efficiently used to provide a list of recommended items for them. RS uses historical ratings along with users and/or items contextual information (if available) to provide the recommendation lists. They have been successfully implemented in many systems such as e-commerce and movie recommendation websites. Motivated by many practical applications, researchers have developed several RS methods which can be generally classified into three categories including: Content-Based (CB), Collaborative Filtering (CF) and Hybrid (HB)

* Corresponding author.
E-mail addresses: fateme.rezaimehr@gmail.com (F. Rezaeimehr), p.moradi@uok.ac.ir (P. Moradi), s.ahmadian@znu.ac.ir (S. Ahmadian), nooruldeen.qader@uhd.edu.iq (N.N. Qader), mahdi.jalili@rmit.edu.au (M. Jalili).

http://dx.doi.org/10.1016/j.future.2017.04.003
0167-739X/

approaches. CF systems analyze only historical interactions, while CB methods are based on profile attributes, and HB techniques combine both of them.

CF algorithms have been widely applied in many applications, resulting in good performance. CF approach is based on the assumption that similar users share similar preferences and interests. In order to provide relevant recommendations for a target user, CF uses opinions of the users with similar preferences. To this end, the historical ratings given by users on items are used to find similar users (or items), and then make predictions for the target users. CF methods can be classified into memory-based and model-based approaches. In memory-based CF, the user–item interaction matrix (e.g., historical ratings) is used to find similarity between users (or items) and generate a prediction [2–5]. On the other hand, in model-based approaches, such as those based on matrix factorization [6–10], first a model is constructed in an offline learning phase, and then the model is applied online to generate recommendations.

Most traditional RS methods exploit the data as static and do not consider changes in the users' preferences or items popularity over time (i.e., the ratings made in different times are considered to have the same importance in the prediction/modeling process). However, in realistic scenarios, the preferences might change over time (often referred to as *drift* in the literature), and the recommender should take this into account [11–13]. A number of approaches, such as Association Rules (AR), Singular Value Decomposition (SVD) and Hidden Markov Models (HMM) have been proposed to deal with this kind of problems and to model users' drifts over time. However, most of the existing approaches model temporal drifts individually without taking into account the effects of related changes which occur in the users corresponding communities. Recently, an algorithm based on association rule and community identification approach has been proposed to handle the drift problem in RS systems [12]. Variations of users' interest can be effectively discovered using overlapping community detection methods to provide alternatives for representing the early efforts on community detection that assumed that communities are non-overlapping or disjointed. There are a number of algorithms in the literature to discover overlapping community structure in networked structures. Applying the community detection techniques on RS algorithms can help to increase the precision of the algorithms and better handle the cold-start and data sparsity problem. Cold-start user (or item) problem is a challenge in RS for which some users (or items) might not have enough rating history, and thus the similarity scores (which is required for CF) might be unreliable [14].

In [12], an overlapping community detection method is proposed for time-weighted recommender systems. This method ignores the users' similarities in constructing the temporal graph. In many practical RSs, each user rates only a few items, and thus the rating matrix is heavily sparse. In this paper, we propose a novel recommendation method that efficiently uses the time of ratings and an improved overlapping community detection method to complete the recommendation lists for users. We employ an innovative approach to show the reliability of predicted ratings and use it in the prediction period. Also, we develop a novel overlapping community detection method and integrate it with an association rule mining method to model temporal properties of users' interests. The general framework of the proposed recommender algorithm includes three steps, (i) a time-aware method based on a time-weighted similarity matrix between users to detect overlapping communities, (ii) a time-weighted mining algorithm based on association rule and temporal overlapping communities to model drift of users' interest over time, and (iii) a new recommendation model based on users' dynamic temporal interests and multi-memberships in their overlapping communities. The proposed method has made major improvements to the other state-of-the art methods. First, it employs an imputation technique that is integrated with a reliability measure to estimate the missing ratings. This method is used to enrich the user–item matrix, resulting in better performance against the data-sparsity problem. Second, the proposed method considers both user drifts and rating-based similarities in computing the time-weighted similarity values between the users. Third, the method uses a novel overlapping community detection to group users based on their interests. This property helps the system to consider more relevant users in its rating prediction process. The proposed community detection method employs subgraph finding algorithm in its process to identify a set of cluster seeds that are located in dense regions of the graph and with the maximum possible distance from one other. This seeding strategy leads to produce cohesive clusters and to identify the ground-truth communities. Fourth, the proposed method employs a rule mining approach to extract some suitable rules from the identified overlapping communities Our methods are extensions of our previous work [15]. We apply the proposed recommendation algorithm on two real-world benchmark datasets and show its superior performance over state-of-the-art algorithms.

The remainder of this paper is organized as follows. Time-aware RSs are reviewed in Sections 2 and 3 introduces the proposed method. In Section 4, the effectiveness of the proposed method is assessed through an experimental evaluation on two real-world datasets and Section 5 outlines our conclusions.

2. Related works

2.1. Time-aware recommender systems

Many of traditional recommender methods use the historical data without considering the changes in the users' preferences or items popularity over time. However, in realistic scenarios, the preferences might change over time (often referred to as drift in the literature), and the recommender should take this into account [11–13]. It is not crystal clear why the dynamic nature of the users' preferences takes place, thus the users' preferences are dynamic. Up to now, several recommendation methods have been proposed in the literature to consider user's drifts in their recommendation processes. For example the authors of [16] present a method which uses temporal features of items to scale down recommendation sets in the context of movie recommendations. Moreover, a series of different techniques such as discrete time windows [17] or continuous decay function [18] are used to adopt CF algorithms with temporal effects by boosting recent ratings and also penalizing those of older ratings. The authors of [19] transformed implicit ratings into another space in which higher weights are assigned to those of recent purchases and older preferences are considered as less valuable and penalized them. In [20], a time-dependent method is proposed by integrating four time effects on explicit ratings in order to demonstrate temporal dynamics and bias shift. As a similar work, the authors of [21] proposed three time effects and a model to track changing behavior over the time. There are also different studies such as [22–26] to exploit temporal dynamics in order to improve the temporal performance of RSs. In [19] two piecewise rating functions were proposed in order to compute the weights based on temporal information. The authors also analyzed a set of temporal information such as item launch time, user buying time, the time difference between the two and combinations of these three by performing empirical studies [27]. Moreover, the authors of [28] used a decay function to compute time weights for different items. Recently, an algorithm based on association rule and community identification approach has been proposed to handle the drift problem in recommender