Decision Support Systems 54 (2012) 768–780

Contents lists available at SciVerse ScienceDirect

Decision Support Systems

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

A trust-semantic fusion-based recommendation approach for e-business applications

Qusai Shambour, Jie Lu *

Lab of Decision Systems and e-Service Intelligence Centre for Quantum Computation and Intelligent Systems School of Software, Faculty of Engineering and Information Technology University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia

ARTICLE INFO

Article history:
Received 16 March 2012
Received in revised form 19 June 2012
Accepted 5 September 2012
Available online 14 September 2012

Keywords:
Recommender systems
Collaborative filtering
Trust filtering
Semantic filtering
Information fusion
Cold-start
Data sparsity

ABSTRACT

Collaborative Filtering (CF) is the most popular recommendation technique but still suffers from data sparsity, user and item cold-start problems, resulting in poor recommendation accuracy and reduced coverage. This study incorporates additional information from the users' social trust network and the items' semantic domain knowledge to alleviate these problems. It proposes an innovative Trust–Semantic Fusion (TSF)-based recommendation approach within the CF framework. Experiments demonstrate that the TSF approach significantly outperforms existing recommendation algorithms in terms of recommendation accuracy and coverage when dealing with the above problems. A business-to-business recommender system case study validates the applicability of the TSF approach.

1. Introduction

Recommender systems are considered the most popular forms of web personalization and have become a promising and important research topic in information sciences and decision support systems [9,10,17,19,20,24,46,49]. Recommender systems are used to either predict whether a particular user will like a particular item or to identify a set of *k* items that will be of interest to a certain user, and have been used in different web-based applications including e-business, e-learning and e-tourism [8,22,31]. Currently, Collaborative Filtering (CF) is probably the most known and commonly used recommendation approach in recommender systems. CF works by collecting user ratings for items in a given domain and computing similarities between users or between items in order to produce recommendations [1,31]. CF can be further divided into user-based and item-based CF approaches. In user-based CF approach, a user will receive recommendations of items that similar users liked. In item-based CF approach, a user will receive recommendations of items that are similar to the ones that the user liked in the past [1]. Despite their popularity and success, the CF-based approaches still suffer from some major limitations; these include data sparsity, cold-start user and cold-start item problems [1,3,36,37]. The data sparsity problem occurs when the number of available items increases and the number of ratings in the rating matrix is insufficient for generating accurate predictions. When the ratings obtained are very small compared to the number of ratings that are needed to be predicted, a recommender system becomes unable to locate similar neighbors and produces poor recommendations. The cold-start (CS) user problem, which is also known as the new user problem, affects users who have none, or a small number of ratings. When the number of rated items is small for the CS user, the CF-based approaches cannot properly find the user neighbors using rating similarity, so it fails to generate accurate recommendations. The CS item problem, which is also known as the new item problem, affects items that have none, or only a small number of ratings. With few ratings for CS items, CF-based approaches cannot appropriately locate similar neighbors using rating similarity and would be unlikely to recommend them [1,33,36,37].

In view of these limitations, researchers have commonly decided to opt for trust-based [11,16,26,27,44] and semantic-based [2,15,22,35,45] recommender systems to tackle such limitations. These systems can deal with the trust relations between users and semantic features of items, which cannot be well handled in traditional CF-based recommendation approaches, to support the recommendation process. These systems have proved to be successful in solving some limitations of CF-based approaches by allowing the recommender systems to make inferences based on an additional source of knowledge. We believe that, by considering information extracted from the users' trust network and the items' semantic domain knowledge, a fusion-based recommendation approach that takes into account both trust and semantic information should provide more effective recommendations.

Based on this notion and following our previous work [22,23,38–41] where we addressed some limitations of CF-based recommendation approaches, this paper proposes a fusion-based recommendation approach that fuses the trust and semantic information

* Corresponding author. Tel.: +61 2 95141838; fax: +61 2 95144535.
E-mail addresses: Q.Shambour@gmail.com (Q. Shambour), Jie.Lu@uts.edu.au (J. Lu).

0167-9236/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2012.09.005

of users and items within the CF framework to achieve yet more effective results in terms of recommendation accuracy and coverage, especially when dealing with data sparsity, CS user and CS item problems. The proposed approach, called TSF (Trust Semantic Fusion), fuses two hybrid recommendation approaches; the user-based trust-enhanced CF, and the item-based semantic-enhanced CF. The user-based trust-enhanced CF approach utilizes the intuitive properties of trust and trust propagation to address the data sparsity and CS user problems. The item-based semantic-enhanced CF approach employs the underlying semantic relationships between items to help reduce the effect of data sparsity and CS item problems. We also define and introduce the notion of an item's reputation weight into the item-based semantic-enhanced CF to further improve the quality of predictions. This paper is organized as follows. In Section 2, research background and related work are described. Section 3 presents the components of the TSF approach. A case-based mathematical example for illustrating the procedure of the TSF is given in Section 4. Section 5 demonstrates the experimental evaluation and results using MovieLens and Yahoo! Webscope datasets. Section 6 describes a case study to validate the feasibility of applying the TSF approach into real e-business applications. Finally, the contributions of this study are summarized, and future research is presented in Section 7.

2. Background and related work

2.1. CF-based recommender systems

The CF approach is the most popular recommendation approach in current recommender systems. Typically, CF can be further divided into user-based and item-based CF approaches. The user-based CF approach produces recommendations for interesting items based on evaluations of users who have similar tastes. First, it analyzes the user–item matrix and creates a vector containing the user's ratings for each rated item. Then, it computes the similarity between the target user's vector and the vectors of the remaining users, using similarity measures such as the Pearson correlation and Cosine vector. These similarity measures compute the similarity between two users based only on the overlap items defined in their respective vectors. Next, the most similar users (*Top-n*) to the target user are selected as the user's nearest neighbors. Finally, predictions are generated using a weighted average of the neighbors' ratings of items that are contained in their profiles [1,37]. The item-based CF approach is the transpose of the user-based one. While the user-based CF approach produces predictions based on users' similarity, item-based CF approach produces predictions based on items' similarity [1,36,37].

2.2. Trust-based recommender systems

Trust-based recommender systems utilize a social network augmented with trust ratings, known as a trust network, to generate recommendations for users based on people they trust. A trust network is a directed graph where the nodes are users and the edges are weighted according to the degree of trust assigned by one user to another. By utilizing trust information, trust-based recommender systems allow users to be aware that the sources of recommendation were formed from people who are either directly trusted by the current user, or indirectly trusted by another trusted user through the trust propagation method. Trust propagation is often employed to infer the trust, and establish new relations between users who have no direct trust links between them [11,26]. In practice, trust-based recommender systems that exploit trust information can provide better recommendation effectiveness than conventional CF-based techniques, in particular, by alleviating issues concerning data sparsity or CS user problems [11,16,26,40,47]. Two main trust filtering methods have been adopted in the current literature: Explicit trust and Implicit trust filtering approaches.

Explicit trust filtering approaches obtain trust values from pre-existing social links between users [11,26]. Nevertheless, the use of explicit trust filtering approaches has exposed two major limitations: (1) they require additional manual labor and user effort from the end user (i.e. time consuming and expensive to get the explicit trust); (2) they suffer from the CS user problem because new users have to first build up their web of trust before the filtering is effective [16,47]. These limitations have limited the applicability of explicit trust filtering approaches in recommender systems, and makes the implicit trust filtering approaches more feasible to use [44,47]. Implicit trust filtering approaches derive trust values between users based on item ratings [16,27,47]. For example, O'Donovan and Smyth [27] acknowledged that user reliability in delivering accurate recommendations in the past is an important factor for influencing recommendation and prediction in the future. In particular, the more accurate predictions a given user has produced in the past, the more trustworthy he/she is. Hwang and Chen [16] developed an implicit trust filtering method where the trust values are directly derived from the user ratings data. Yuan et al. [47] proposed a novel implicit trust aware recommendation model (iTARS) based on the small-worldness of the implicit trust network, in which the implicit trust is generated from the user similarities. To sum up, most of the implicit trust filtering techniques we have explored share common features: (1) they use ratings or prediction errors between users' profiles as an indication of trust; (2) they operate on the intersection of users' profiles; as a result, they do not consider what has not been rated when computing trust.

2.3. Semantic-based recommender systems

Semantic-based recommender systems exploit the underlying semantic properties and attributes associated with users and items to generate recommendations. For instance, semantic information about items consists of the attributes of the items, the relationship between items, and the relationship between items and meta-information [30]. Taxonomies and ontologies as the major source of semantic information can be taken advantage of in recommender systems, since they provide a means of discovering and classifying new information about the items to recommend, about user profiles and even about their context [35]. For example, product taxonomies and ontologies have been presented in several recommender systems to utilize the relevant semantic information in order to help improve the recommendation quality [2,7,15,22,35]. In summary, most of the presented research provides two primary advantages. First, the semantic attributes for items provide additional explanations about why particular items have been recommended or not. Secondly, the additional source of semantic knowledge provides better recommendation effectiveness than current CF-based techniques, particularly in cases where little or no rating information is available.

3. Trust–semantic fusion-based recommendation approach

This section first describes the structure and each component of the TSF approach. Then, the TSF's recommendation computation process is demonstrated.

3.1. The structure of the TSF recommendation approach

The TSF approach (Fig. 1) obtains as inputs a raw user-item rating matrix R_{m*n} and item taxonomy, and produces as an output a user–item prediction matrix. R_{m*n} contains the rating values of m users and n items. The item taxonomy is represented in a tree hierarchy structure with two levels of nodes: the first level contains the item categories that items belong to; the second contains the items as leaf nodes.