Electronic Commerce Research and Applications 21 (2017) 50–64

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

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

A hybrid multi-criteria recommender system using ontology and neuro-fuzzy techniques

Naime Ranjbar Kermany, Sasan H. Alizadeh *

Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

ARTICLE INFO

Article history:
Received 16 March 2016
Received in revised form 31 December 2016
Accepted 31 December 2016
Available online 3 January 2017

Keywords:
Recommender system
Multi-criteria
Collaborative filtering
Adaptive neuro-fuzzy inference system
Ontology
Semantic filtering

ABSTRACT

The importance of recommendation systems for business applications has led to extensive research efforts to improve the recommendations accuracy as well as to reduce the sparsity problem. Despite the success of both collaborative filtering and multi-criteria approaches, they still need to be further optimized to address the stated problems. In this paper, we propose a new hybrid method based on enhanced fuzzy multi-criteria collaborative filtering which incorporates demographic information and an item-based ontological semantic filtering approach for movie recommendation purposes. We use an adaptive neuro-fuzzy inference system to discover the relationship between each criterion and the overall rating. A fusion of fuzzy cosine and Jaccard similarities is further adopted to calculate the total similarity between users/movies with respect to the effect of co-rated item set cardinality on the reliability of similarity measures. To increase the robustness and reliability of the final similarity measure, especially in the case of cold start users, a convex combination of both user and movie based similarities is used; in which the convex weightings are determined through the gradient decent algorithm to ensure a minimum prediction error. Experimental results demonstrate the efficiency of the proposed method in reducing the sparsity problem and improving the prediction accuracy.

1. Introduction

Due to the success of Amazon.com as a frontier e-commerce website exploiting recommendation methods, recommender systems have been widely applied in a variety of business domain applications such as movie recommendations (Barragáns-Martínez et al., 2015; Briguez et al., 2014; Carrer-Neto et al., 2012; Colombo-Mendoza et al., 2015; Nilashi et al., 2014a; Wei et al., 2016), music (Bogdanov et al., 2013), books (Kim et al., 2011), e-books (Núñez-Valdéz et al., 2012), tourism (Al-Hassan et al., 2015), hotels (Korfiatis and Poulos, 2013), restaurants (Gallego et al., 2013), documentation (Hsu, 2011), TV (Véras et al., 2015), e-payments (Rosaci and Sarnè, 2014), and news (Cleger-Tamayo et al., 2012)). They assist users in finding only relevant information regarding their needs, rather than an indistinguishable mass of data. Because many users browse at most two pages (Jannach et al., 2010), it is crucial for recommender systems to have high predictive accuracy and to allocate the desired items at the top of the recommendation list in those pages, according to the users' preferences (Ekstrand et al., 2011b).

Content-based filtering (CBF) and collaborative filtering (CF) approaches have been mainly developed to gain insight into users' preferences (Adomavicius and Tuzhilin, 2005; Jafarkarimi et al., 2012). In CBF approach, items are recommended based on their similarity of content to the items previously rated by the user (Mooney and Roy, 2000). However, the CF approach exploits the similarity of users' preferences for recommendations (Liao and Lee, 2016; Melville and Sindhwani, 2010; Ricci et al., 2011). CF approach is divided into two approaches, user-based and item-based. In user-based CF approach, which has been the most popular and widely adopted CF technique (Nilashi et al., 2014a), the similarity between users is calculated based on co-rated items while in item-based CF (Sarwar et al., 2001), this similarity is measured between items rather than users. The users will be more interested in items they have previously liked. Compared to the individual approaches, the combination of the two filtering approaches (user- and item-based), called the hybrid approach, results in higher predictive accuracy (Jannach et al., 2012).

Sparsity problem is one of the major limitations in CF approaches, in which the user-item matrix is extremely sparse in that it contains a small percentage of non-zero elements (Nikolakopoulos et al., 2015). This problem will also lead to low predictive accuracy and low precision (Anand and Bharadwaj, 2011; Chen et al., 2011; Huang et al., 2004; Lika et al., 2014;

* Corresponding author.
E-mail addresses: naeemeh.ranjbar@gmail.com (N. Ranjbar Kermany), sasan.h.alizadeh@qiau.ac.ir (S.H. Alizadeh).

http://dx.doi.org/10.1016/j.elerap.2016.12.005
1567-4223/

Papagelis et al., 2005). Guo et al. (2014) incorporate trust, as a new source of information, to alleviate this problem. On the other hand, multi-criteria recommender systems employ the ratings given by users to various criteria of each item to compensate for the lack of information problems.

Users supply one overall rating for the explored items in the conventional single rating approach. Consequently, the other limitation—related to CF approaches—is that the attributes of items are not considered. In contrast, users rate the items according to various criteria in the multi-criteria rating approaches. As a result, a multi-criteria recommender system can be defined as a recommender system that incorporates preferences into multiple dimensions (Adomavicius et al., 2011), providing more information about the users and the items, which could be effectively used in the recommendation process. Some research works revealed the superiority of the latter (Adomavicius and Kwon, 2007; Adomavicius et al., 2011; Fuchs and Zanker, 2012; Lakiotaki et al., 2008; Liu et al., 2011; Nilashi et al., 2014c, 2015b) compared to the former (Guo and Lu, 2007; Ruiz-Montiel and Aldana-Montes, 2009; Wang and Kong, 2007)in terms of prediction accuracy. However, they still require significant study and optimization (Nilashi et al., 2014a,b, c). Often, a user's preference ratings are expressed by a set of linguistic terms such as "*Strongly Interested*", "*Interested*", and "*Not Interested*". Therefore, the ratings could be inaccurate, vague, and fuzzy, depending on the precision of the user's rating (Zenebe and Norcio, 2009). Extraction of knowledge from user's preferences and finding the relationship between user's overall rating and the ratings for each criteria constitute another problem. Although recent single rating approaches have diminished the sparsity problem to some extent, latest papers demonstrate the efficiency of incorporating additional semantic (Al-Hassan et al., 2015; Lu et al., 2013; Martín-Vicente et al., 2012) and demographic information (Ghazanfar and Prugel-Bennett, 2010; Vozalis and Margaritis, 2004, 2007) in recommender systems. It is also hypothesized that using semantic information and content of users can further enhance the predictive accuracy and reduce the sparsity problem in multi-criteria recommender systems (Nilashi et al., 2014a, 2015a); This idea is investigated throughout this paper. However, to our knowledge, no work has been done to study the effect of using both techniques simultaneously in multi-criteria recommender systems. In this paper, we propose a recommendation method by utilizing both ontological semantic filtering and user's demographic information in multi-criteria recommendations for movie recommendation purposes. A combination of Jaccard metric and cosine similarity is applied to calculate user-based, as well as item-based, similarities to achieve more reliable and accurate predictions. The weighted combination of these similarities is used to reduce the sparsity problem, where the weights are determined through gradient descent algorithm. In addition, we use fuzzy techniques to handle vagueness, uncertainty, and fuzziness, and to model humans' behavioral complexity in expressing their opinions. Finally, we employ ANFIS to model the relationship between users' overall ratings and their ratings for each [individual] criterion.

This paper is divided into 6 sections: Section 2 reviews related works. Section 3 provides the preliminaries of our research. The proposed method is explained in Section 4. Section 5 is devoted to the experiments and results. Finally, Section 6 presents the conclusion and proposed future works.

1.1. Multi-criteria recommender system specification

In a mathematics explanation, the utility function evaluates the suitability of the recommending item ($i \in Items$) for the user ($u \in Users$). It is often defined as

$$R: \ Users \times Items \rightarrow R_o, \tag{1}$$

where "*Users*" represents a set of all the users and "*Items*" symbolizes a set of all possible items that can be suggested to the *Users*. R_0 is normally indicated by integers in a bounded interval (Lu et al., 2015). The utility function is mostly a subspace of the *Users* × *Items* space because, as mentioned before, a user might consider any arbitrary subset of items while rating. Recommender system has to predict the utility function $R(u, i)$ for user u on item i, and then recommend an item which maximizes this utility function. More formally,

$$\forall u \in Users, i = \arg\max_{i \in Items} R(u, i), \tag{2}$$

where *arg* max stands for maximum elements. Applying this procedure iteratively, recommender system is able to provide more than one recommendation in the recommendation list. In a multi-criteria approach, the rating function is modified to consider the additional ratings as follows (Adomavicius et al., 2011):

$$Users \times Items \rightarrow R_0 \times R_1 \times \ldots \times R_k, \tag{3}$$

where R_0 represents the overall rating given by user to item i and R_j ($j \in \{1, \ldots, k\}$) is devoted to the rating expressed for each criterion.

2. Related works and our contribution

In this section, we first provide a literature study on the semantic-based recommender systems, followed by the literature on demographic information. Next, we review some previous studies on the hybrid recommendation algorithms and some recent works on multi-criteria recommender systems. Last, we compare our proposed recommendation method with some related studies.

Semantic techniques are applied to the recommender systems with ontology domain knowledge in order to improve the recommendation accuracy and to reduce the sparsity problem (Shambour and Lu, 2011). Semantic approach is applied in different contexts including business (Martínez-López et al., 2010), e-tourism (Daramola et al., 2009), online marketing (Blanco-Fernández et al., 2011), online recommendation (Wang and Kong, 2007), product category (Albadvi and Shahbazi, 2009) and others (Carrer-Neto et al., 2012; Hawalah and Fasli, 2014; Li et al., 2014; Martín-Vicente et al., 2014; Moreno et al., 2013). Several techniques have been incorporated with the semantic approach, including clustering methods (Wang and Kong, 2007), genetic algorithm (Cimino et al., 2012), Ant Colony (Salehi et al., 2013), neural networks and support vector regressions (Li et al., 2014), and fuzzy techniques (Gonzalez-Carrasco et al., 2012; Cimino et al., 2012), with the aim of improving the predictive accuracy of the recommendation systems. Among these, fuzzy techniques are the most extensively applied approach (Abraham, 2005; Anand and Mampilli, 2014; Cheng and Wang, 2014; Ranjbar Kermany and Alizadeh, 2013; Thong, 2015; Yager, 2003). Nilashi et al. (2014a) has reported a detailed study on the application of fuzzy techniques in recommendation systems. However, his work does not include the demographic information of users; which can be helpful to improve the prediction results.

Demographic information of users helps recommender systems to identify users more precisely (Montaner et al., 2003). Demographic information describes user's characteristics such as age, gender, education, occupation, and so on. This kind of information is reported to be useful to enhance the predictive accuracy of recommendations (Chikhaoui et al., 2011; Ghazanfar and Prugel-Bennett, 2010; Schiaffino and Amandi, 2009; Vozalis and Margaritis, 2004, 2007). Unlike CF approaches, the recommender systems which utilize demographic information are able to provide recommendations even when no user's rating exists. However, a demographic information technique on its own does not significantly improve the precision (Chikhaoui et al., 2011).