Electronic Commerce Research and Applications 29 (2018) 113–123

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

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

Do adjective features from user reviews address sparsity and transparency in recommender systems?

Xiaoying Xu[a], Kaushik Dutta[b], Chunmian Ge[c,*]

[a] Department of Decision Science, School of Business Administration, South China University of Technology, China
[b] Department of Information Systems and Decision Sciences, Muma College of Business, University of South Florida, United States
[c] Department of Financial Management, School of Business Administration, South China University of Technology, China

ARTICLE INFO

Keywords:
Adjective features
Recommender systems
Side information
Sparsity
Transparency
User reviews

ABSTRACT

Recommender systems have become increasingly essential in many domains for alleviating the problem of *information overload*, but existing recommendation techniques suffer from data sparsity and transparency issues. In this paper, we show that the adjective features embedded in user reviews can be used by the recommendation techniques to address the sparsity and transparency problems. We extend the standard *frequency-inverse document frequency* (TF-IDF) term weighting scheme by introducing *nearest neighbors frequency* (NNF) to automatically extract high-quality adjective features from user reviews, and incorporate the extracted adjective features into a specific recommendation technique to show effectiveness. The results of experiments conducted on real-world datasets show that the integrated method reduced the prediction errors of the state-of-the-art rating-based method by 19.5% in extremely sparse settings. When compared with the state-of-the-art tag-based method, the proposed method reduced the prediction errors by 11.3%, and increased the interest similarity in similar user identification by 7.1%.

1. Introduction

Recommender systems (RSs) are well-known artifacts in consumer marketing, having been utilized to great commercial success in iconic technological companies such as Amazon, TiVo and Netflix. Commensurate with their market impact, RS technology has enjoyed (and continues to enjoy) much attention from scientists and researchers. Over the past decade, countless papers have been published, systems have been released, and entire top-rated conferences have been established, backed by leading scientific and technological associations, on RS research. Suffice to say, in the domain of data mining, knowledge discovery and information retrieval, RSs stand out as one prominent example of the real-life impact of academic research.

Despite their popularity, RSs are associated with some problems. Arguably the two most major, and challenging weaknesses that permeate virtually every flavor of RSs, are the data sparsity and the transparency problems.

The data sparsity problem refers to the insufficiency of input data into recommendation algorithms (Adomavicius and Tuzhilin, 2005). We note that most common RSs, like *collaborative filtering* (CF) (Resnick et al., 1994; Sarwar et al., 2001) systems (powering virtually all commercial systems today), usually use rated items to represent a user. User interests are captured at the item level, making them very sensitive to rating sparseness. Unfortunately, in most domains studied (movies, books, restaurants etc.), a majority of items turn out to be unrated, resulting in *sparse* rating matrices (matrices with insufficient data), which adversely impact the quality of the recommendation (Adomavicius and Tuzhilin, 2005).

Having discussed sparsity, we turn our attention to the other major issue running through the bulk of recommendation systems work — *transparency*. A well-known drawback of virtually all CF techniques is the lack of transparency. Such methods generate recommendations by aggregating users' rating patterns at the item level, but the lack of awareness of item features and user taste aspects makes it difficult to interpret their underlying rationale intuitively to the users, which typically lowers the users' trust on the recommendations produced (Nilashi et al., 2016). As has been demonstrated conclusively, trust is a critical attribute impacting users' willingness to act upon recommendations (Guo et al., 2015).

In this context, we aim to create an approach to ameliorate the sparsity and transparency issues in RSs. A readily available source of opinion information that may cover most of the user's taste aspects for many consumer products (e.g., movies, books) are user reviews. Most of the existing research in this area has focused on how to employ opinion

* Corresponding author.
E-mail addresses: bmxyxu@scut.edu.cn (X. Xu), duttak@usf.edu (K. Dutta), bmgecm@scut.edu.cn (C. Ge).

https://doi.org/10.1016/j.elerap.2018.04.002
Received 25 October 2017; Received in revised form 21 April 2018; Accepted 21 April 2018
Available online 24 April 2018
1567-4223/

mining and sentiment analysis techniques to factorize and quantify textual user reviews (Lei et al., 2016; Lou et al., 2016; Luo et al., 2015; Zhang et al., 2012; Zhao et al., 2016), and how aspect ratings derived from user reviews are utilized for recommendation (Diao et al., 2014; Wang et al., 2010; Zhang et al., 2013, 2014). We note that besides the item aspects (generally in the form of nouns) and their corresponding sentiment features that have been studied though, there are other valuable elements embedded in user reviews that may be incorporated in the recommendation process.

We are interested in one specific kind of information from user reviews, namely *adjective features*, with this simple intuition: when asked to reveal why they like or dislike something, people often use adjectives to explain their preference. For instance, when asked why he/she likes the movie *Titanic*, a user's answer often includes words such as "romantic", "moving", "astounding", "beautiful" or "sad" — all being adjectives. These adjectives are different from those that simply reflect the general polarity of the sentiment, such as "good" and "bad". They are more informative and can be regarded as the descriptive features that truly reflect users' detailed perceptions towards items. Such adjective features can be found in abundance in user reviews, but they remain unexplored in recommendation research.

To bridge this gap, we focus on incorporating adjective features extracted from external user reviews into the recommendation process to generate more accurate and more explainable item recommendation. To automatically extract adjective features from user reviews, we employ well understood part-of-speech (POS) tagging methods. However, we quickly discover that a lot of adjectives are not helpful in discriminating tastes, that is, they are too general to be adequately representative of user tastes (lack of *representativeness*), for example, "good", while some adjectives are too specific to capture the users' general taste aspects (lack of *generalizability*), for example, "unsinkable" in the reviews of *Titanic*. To address this issue, we propose our own approach by extending the traditional *frequency-inverse document frequency* (TF-IDF) term weight (Cohen, 1995) to TF-IDF-NNF by introducing another unsupervised term weight measure, *nearest neighbors frequency* (NNF).

We incorporate the extracted adjective features into one specific recommendation technique: *singular value decomposition* (SVD) (Paterek, 2007). We call this integrated method *adjective feature vector* (AFV) method. The results of our work make substantial advances over existing recommendation techniques. In particular, our method reduces the prediction errors of state-of-the-art rating-based method by 19.5% in extremely sparse settings. It also outperforms the state-of-the-art tag-based method by reducing its prediction errors by 11.3% in item recommendation, increasing its interest similarity by 7.1% in similar user identification, and retaining full item and user-space coverage. The results also indicate that our method is effective for providing recommendation explanations.

The objective of this work is to alleviate the sparsity and transparency problems in recommender systems by utilizing the adjective features embedded in user reviews. By achieving this goal, we make these contributions:

(1) We show that adjective features, which have drawn scant attention from existing research, are useful for characterizing users' taste aspects, and can be utilized to improve the quality of recommendation. Theoretically, such findings enrich the literature on both user-generated content and recommender systems.
(2) Our proposed approach for adjective feature extraction, and the proposed extension of the SVD method, have theoretical implications for feature extraction and dimension reduction research.
(3) To implement the proposed solutions, we follow *design science research* (DSR) principles (Gregor and Hevner, 2013) to design a recommendation architecture, and conduct comprehensive experiments to show the effectiveness. Therefore, our work also has practical implications for recommender system design.

2. Literature review

Substantial research has been conducted on recommendation, most of which belongs to two main streams: *collaborative filtering* (CF) and *content-based* (CB) *recommendation*. We first introduce the two main streams of traditional recommendation algorithms respectively. Then we discuss the recent extensions with user reviews. Finally, we review the related works on adjective extraction, which is a primary task in our proposed method.

2.1. Collaborative filtering (CF) recommendation

CF has been explored in depth in the past ten years and represents the most popular recommendation algorithm. CF approaches can be classified into two types: user-based (Resnick et al., 1994) and item-based (Sarwar et al., 2001). Both the user-based and item-based CF work on the same fundamental principle. First the neighborhood relationship is built by measuring the similarity (typically cosine similarity or Pearson correlation) between users or items based on their rating vectors, and then the average ratings from the neighborhoods weighted by their similarities are used to predict the missing ratings.

CF methods are very sensitive to rating sparsity. Latent factor models, such as *probabilistic matrix factorization* (PMF) (Salakhutdinov and Mnih, 2007) and *singular value decomposition* (SVD) (Paterek, 2007), comprise an alternative approach to CF, with the aim of alleviating the sparsity problem by transforming both items and users to the same latent factor space. There have been many recent works based on latent factor models (Gedikli et al., 2011; Gopalan et al., 2013; Xie et al., 2016), and their underlying assumptions remain the same.

Another common approach adopted by recent studies to address the sparsity problem is to incorporate other side information in the recommendation process. For example, Sedhain et al. (2017) perform cold-start recommendation by incorporating social metadata. Similarly, Kalloori and Ricci (2017) improve cold-start recommendation by utilizing feature-based item comparisons. In addition to addressing the sparsity problem, incorporating side information is also a common way in addressing the transparency issue (Abdelkhalek, 2017; Zhang et al., 2014), which is another major limitation of CF recommendation. Our work is among those that aim at addressing the sparsity and transparency issues in recommender systems by using side information (i.e., the adjective features in our context).

2.2. Content-Based (CB) recommendation

In contrast to CF, the intuition of the CB approach is to recommend items that are similar to the ones that the user has liked in the past. The similarity between items is calculated based on the features associated with the compared items (Lops et al., 2011). CB techniques have a natural limit in the number and types of features associated with items.

Recently, due to the popularity of social tagging systems, tags as a specific kind of user-generated content have drawn attention from researchers (Belém et al., 2014; Xie et al., 2016). Tags are generated by users who collaboratively annotate and categorize resources of interests with freely chosen keywords (de Gemmis et al., 2008). Compared to descriptive attributes commonly used in CB recommendation, tags cover more features of the items and are more comprehensible to the users. Many tag-based approaches have been proposed. For example, Wei et al. (2011) and Ifada (2014) propose modeling the quaternary relationship among users, items, tags and ratings as a 4-order tensor and performed a multi-way latent semantic analysis. However, since tags are voluntarily given by the users, tag-based methods may have the problems of tag sparsity and diverging vocabulary.

2.3. Recommendation with user reviews

Recently, researchers have sought side information to extend