Knowledge-Based Systems 75 (2015) 124–140

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Item recommendation in collaborative tagging systems via heuristic data fusion

Hao Wu, Yijian Pei, Bo Li *, Zongzhan Kang, Xiaoxin Liu, Hao Li

School of Information Science and Engineering, Yunnan University, No. 2 North Green Lake Road, Kunming 650091, China

ARTICLE INFO

Article history:
Received 4 July 2014
Received in revised form 19 November 2014
Accepted 21 November 2014
Available online 28 November 2014

Keywords:
Collaborative tagging systems
Recommender systems
Data fusion
Item recommendation
Performance comparison

ABSTRACT

Collaborative tagging systems have been popular on the Web. However, information overload results in the increasing need for recommender services from users, and thus item recommendation has been one of the key issues in such systems. In this paper, we examine if data fusion can be helpful for improving effectiveness of item recommendation in these systems. For this, we first summarize the state-of-the-art recommendation methods which are classified into several categories according to their algorithmic principles. Then, we experiment with about 40 recommending components against the datasets from three social tagging systems-Delicious, Lastfm and CiteULike. Based on these, several heuristic data fusion models including rank-based and score-based are used to combine selected components. We also put forward a hybrid linear combination (HLC) model for fusing item recommendation. We use four kinds of evaluation metrics, which respectively consider accuracy, inner-diversity, inter-diversity and novelty, to systematically assess quality of recommendations obtained by various components or fusion models. Depending on experimental results, combining evidence from separate components can lead to performance improvement in the accuracy of recommendations, with a little or without loss of recommendation diversity and novelty, if separate components can suggest similar sets of relevant items but recommend different sets of non-relevant items. Particularly, fusing recommendation sets formed from different combinations of profile representations and similarity functions in user-based and item-based collaborative filtering can significantly improve recommendation accuracy. In addition, some other useful findings are also drawn: (i) Using the tag to represent users profiles or items profiles maybe not as good as profiling users with the item or profiling items with the user, however, exploiting tags in the topic models and random walks can notably improve the accuracy, diversity and novelty of recommendations; (ii) Generally, user-based collaborative filtering, item-based collaborative filtering and random walks methods are robust for the task of item recommendation in social tagging systems, thus can be chosen as the basic components of data fusion process; and (iii) The proposed method (HLC) is more flexible and robust than traditional data fusion models.

1. Introduction

Collaborative tagging systems [18,19], such as Delicious, Flickr, Lastfm, Connotea and CiteUlike, have become a kind of booming business on the Web. Beyond being a kind of resource discovery tools [32], such systems also aggregate heterogeneous informative elements and offer a wealth of information. The tags play an essential role in such systems. On one hand, they represent users' understandings of information resources, as well express users' needs; on the other hand, the set of tags derived from multiple users forms a description of information resources. Besides the tags, implicit communities emerge from users with similar interest, and explicit social network is built based on both shared interests and trusts of users. These elements can be used to enhance traditional recommendation methods, such as collaborative filtering, to solve the cold-start recommendation, data sparsity and other issues [55]. At the same time, the size and complexity of these systems can unfortunately lead to information overload and reduced utility. Too many resources can make users helpless in their process of finding useful contents. Consequentially, the increasing need for recommender services from users has arisen. For these reasons, applying the techniques of recommendation to deliver personalized views in the collaborative tagging systems has become a hot research spot [33,55].

* Corresponding author.
E-mail addresses: haowu@ynu.edu.cn (H. Wu), pei3p@ynu.edu.cn (Y. Pei), libo@ynu.edu.cn (B. Li), kangzongzhan@gmail.com (Z. Kang), lxxs369@163.com (X. Liu), lihao0730@sina.com (H. Li).

http://dx.doi.org/10.1016/j.knosys.2014.11.026
0950-7051/

There have been various approaches based on different principles to address recommendation problems in collaborative tagging systems. For example, collaborative filtering use tags for profiling users and items [33], social collaborative filtering considers networking influence in recommendation [6], random walks consider relevance propagation on a graph structure as well as the problem of sparsity [26], and semantic models integrate explicit or implicit community structure in recommendation [53]. Different recommending methods have distinctive advantages and disadvantages, and there is no all-time winner [20]. Therefore in practice, a hybrid recommending model is often used [10,2]. A hybrid recommender system is one that combines multiple techniques together to achieve some synergy between them. For instance, one simplest hybrid recommender is to use a combination of collaborative filtering and content-based filtering [33]. As far as the possible hybrid recommenders concerned, Burke [11] has summarized a taxonomy of them, including weighted hybrid, switching hybrid, mixed hybrid, feature combination hybrid, feature augmentation hybrid, cascade hybrid, and meta-level hybrid. According to Burke [11], a staged cascade, or feature argumentation of recommending components can significantly improve the quality of recommendations. A well-designed cascade model or feature argumentation model should be easily achieved as the number of features or components is quite small. Nevertheless, when available features and components become larger in amount, the selection of optimal hybrid schemes must be extremely complicated. For building a hybrid recommender system, a simple yet efficient model would always be a right choice. Motivated by this, we may be able to change our look to find a solution from the information retrieval area where many issues (such as ranking) can share the same solutions with recommendation systems.

Data fusion, also known as rank aggregation [16], accepts two or more ranked lists and merges these lists into a single ranked list with the aim of providing better effectiveness than all systems used for fusion [50]. Data fusion has been successfully used in many retrieval scenarios, such as meta-search [4] and expert finding [31], and leads to performance improvement with combined evidence from different systems or algorithms. Although data fusion shares some concepts with aforesaid hybrids of recommendation systems, e.g. weighted combination, most of the fusion models are easily-implemented, free of specific combinations (or cascades) of components, and scalable to be fed with abundant components. According to previous observations [40,27], CombSUM and CombMNZ are two basic ones and seem to be better than others (e.g. CombMAX, CombMIN, CombANZ and CombMED). Nevertheless, for data fusion techniques to improve retrieval effectiveness, the retrieved results set from each component being fused must retrieve similar sets of relevant documents, but retrieve different sets of non-relevant documents, such a prerequisite may result in some vulnerability to CombSUM/CombMNZ [27], where strong and weak components working together may be difficult to reach the condition. In contrast, a linear combination can mitigate the vulnerability by assigning lower weights to weak components, thus enables the hybridized one to be more robust [51]. However, ideal weights of components being combined have to be obtained by a learning-based process, such as logistic regression, in which building a training model is usually a hard job. To cope with existing defects in data fusion models, we propose a hybrid linear combination (HLC), in which CombSUM is first used to combine the relevance scores of components sharing the similar algorithmic principle, and then a weighted linear model is exploited to combine these scores. Such a compromising solution would make the hybridized recommender a simple yet robust one. On one hand, the components of the same type tend to be similar in performance, thus are applicable to the rationale of evidence combination; On the other hand, the fused results with different performance are more suitable for a post-process based weighted combination.

We conduct experiments and discuss the results on three well-known collaborative tagging systems to validate the proposed models. Also, to determine which ones of components work well for the task of item recommendation, we summarize the state-of-the-art recommendation methods and their variants for collaborative tagging systems, and use four kinds of evaluation metrics, which respectively consider accuracy, inner-diversity, inter-diversity and novelty, to systematically assess quality of recommendations obtained by various components and fusion models. Our experiments show some useful findings and conclusions to respond to the above-mentioned questions, such as "whether data fusion helps improve the quality of recommendations or not".

The rest of this paper is structured as follows. Section 2 presents some backgrounds concerning our methods and related works. In Section 3, we introduce several categories of recommendation methods, such as collaborative filtering, random walks, semantic models, retrieval models and popularity-based models. Also, we set the definition of data fusion models used in experiments. In Section 4, we present and analyze the experimental results according to three folksonomy-based data collections. Finally, we conclude the paper and point out future directions.

2. Related works

Data fusion at first occurred in multi-sensor processing, and has been practiced by researchers in information retrieval area to combine multiple document lists for the same information need for more than 20 years. One particular is that several different retrieval systems or several different settings or retrieval strategies in the same system to retrieve the same collection of documents, then merging these results into a single list for higher effectiveness [50]. In our recommendation scenario, this corresponds to combining different categories of recommending components (such as collaborative filtering, content-based filtering, random walks, semantic models and retrieval models) and their variations (e.g. collaborative filtering based on different combinations of similarity functions and representations of profiles) and fusing those predicted items into a single recommendation list.

Many data fusion models such as score-based, for instance, CombSUM and CombMNZ [40], and rank-based, for example, BordaCount, Condorcet fusion [34] have been proposed, and many experiments in different application scenarios have been conducted to evaluate them [3–5,31]. These experimental results have been showed that using data fusion models provides better effectiveness than all units used for data fusion. Also, fusion models are proved as an effective and scalable framework to be fed with numerous components. However, most of the proposed data fusion models compete in performance and there is no all-time winner. The linear combination can achieve the best runs compared with other fusion models [50]; nevertheless, the ideal weights of components have to be obtained by a learning-based process, such as logistic regression. On the contrary, heuristic model needs not building a training model which is usually a hard job, but to satisfy the prerequisite of evidence combination. In this paper, we thus put forward a hybrid linear combination model to exploit the advantages of both worlds, and compared with other heuristic fusion models, such as BordaCount, CombSUM, CombMED, CombMIN, CombMAX, CombMNZ, and CombANZ.

On applying data fusion models in recommender systems, there are few works presented. Wang et al. [47] reformulate the memory-based collaborative filtering problem in a generative probabilistic framework, treating individual user-item ratings as predictors of missing ratings. The final rating is estimated by fusing