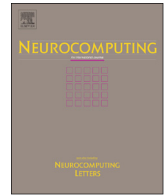




ELSEVIER

Contents lists available at ScienceDirect

## Neurocomputing

journal homepage: [www.elsevier.com/locate/neucom](http://www.elsevier.com/locate/neucom)

# Folksonomy-based personalized search by hybrid user profiles in multiple levels <sup>☆</sup>



Qing Du <sup>a</sup>, Haoran Xie <sup>b</sup>, Yi Cai <sup>a,\*</sup>, Ho-fung Leung <sup>c</sup>, Qing Li <sup>d</sup>, Huaqing Min <sup>a</sup>, Fu Lee Wang <sup>e</sup>

<sup>a</sup> School of Software Engineering, South China University of Technology, Guangzhou, China

<sup>b</sup> Department of Mathematics and Information Technology, The Hong Kong Institute of Education, N.T., Hong Kong Special Administrative Region

<sup>c</sup> Department of Computer Science and Engineering, Chinese University of Hong Kong, N.T., Hong Kong Special Administrative Region

<sup>d</sup> Department of Computer Science, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong Special Administrative Region

<sup>e</sup> Caritas Institute of Higher Education, N.T., Hong Kong Special Administrative Region

## ARTICLE INFO

### Article history:

Received 10 March 2015

Received in revised form

20 October 2015

Accepted 22 October 2015

Available online 8 April 2016

### Keywords:

Folksonomy

Social tagging

Web 2.0

User profiling

Personalized search

## ABSTRACT

Recently, some systems have allowed users to rate and annotate resources, e.g., MovieLens, and we consider that it provides a way to identify favorite and non-favorite tags of a user by integrating his or her rating and tags. In this paper, we review and elaborate on the limitations of the current research on user profiling for personalized search in collaborative tagging systems. We then propose a new multi-level user profiling model by integrating tags and ratings to achieve personalized search, which can reflect not only a user's likes but also a his or her dislikes. To the best of our knowledge, this is the first effort to integrate ratings and tags to model multi-level user profiles for personalized search.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Currently, collaborative tagging systems have become increasingly popular, and many social resource sites support a tagging mechanism. For example, bookmarks on Del.icio.us<sup>1</sup> may be tagged in terms of topics of interest by users, and in Flickr,<sup>2</sup> users can upload and annotate their own photos. The resources and tags posted by Web users on these systems are supposed to be highly dependent on their interests, and the tags given by users provide rich information for building more accurate and specific user profiles [2]. Further, the tags given by different users to a resource are useful for describing it. This provides a collaborative form of resource description, and such a description is considered to be more meaningful and acceptable from users' perspectives.

Given the characteristics of collaborative tagging systems, researchers consider that constructing user and resource profiles from collaborative tags is instrumental for personalized resource search. Some studies such as [3–5], have been conducted to construct user and resource profiles from tags in collaborative tagging systems for personalized search. However, there are several

limitations in the current user profiling methods, which include the following:

- All current work assumes that of a user's tags are the his or her favorite features. However, users may also use some tags to reflect their dislikes. For example, Alice may like science fiction movies but not dinosaurs. Thus, she may use the tag 'dinosaur' to annotate all the movies about dinosaurs to remind herself that these movies include dinosaurs. She also might give a very low rating to these movies to indicate her dislike. In other words, this assumption of the current research is not reasonable. This is because a user's tags include not only his or her like tags but also some dislike tags. Current work only models a user's positive preferences in user profiling, and ignores his or her negative preferences.
- Current work adopts a single vector to model a user profile. However, a user profile should include not only a user's most-favorite features but also least-favorite features as well as neutral features. Using only a single vector, it is difficult to reflect both most- and least-favorite features of a user simultaneously. In other words, current work lacks the building blocks to model a user's dislikes.

Recently, some systems have allowed users to both rate and annotate resources, e.g., MovieLens. Some researchers have integrated ratings and tags into recommender systems to calculate a feature's co-

<sup>☆</sup>A preliminary version of this manuscript has been published in [1].

\* Corresponding author. Tel.: +86 20 39380218.

E-mail address: [yicai.scut@gmail.com](mailto:yicai.scut@gmail.com) (Y. Cai).

<sup>1</sup> <http://delicious.com>

<sup>2</sup> <http://www.flickr.com>

occurrence effectiveness [6]. In these systems, some tags of resources that are rated by a user with high or low ratings directly indicate a user's likes and dislikes. Thus, we consider that integrating a user's ratings and tags provides a way to identify their like and dislike tags. In this paper, we review and elaborate on the limitations of the current work on user profiling for personalized search in collaborative tagging systems. To address these limitations, we propose a new multi-level user profiling (MUP) model for constructing user profiles to achieve personalized search. The contributions of our work are as follows.

- We clarify the limitations of current tag-based user profiling work.
- We hence propose a new three-level model for constructing a user profile by integrating tags and ratings that reflects not only the user's likes but also a his or her dislikes. Furthermore, we expand the three-level model to an n-level model and explore the relationship between the number of levels and personalized search performance.
- We employ the proposed multi-level user profile to enhance personalized search in collaborative tagging systems.
- We conduct experiments with a real data set from MovieLens.<sup>3</sup> The results show that our method outperforms other state-of-the-art methods in personalized resource search.

To the best of our knowledge, this is the first effort to integrate the ratings and tags to model multi-level user profiles for personalized search.

## 2. Related work and background

In this section, we first survey some existing work on semantic links, collaborative tagging, and personalized search. We then examine and discuss the limitations of these work in terms of user profiling.

### 2.1. Semantic links

Luo et al. [7] proposed a discovery algorithm of associated resources to build the original Association Link Network (ALN) for organizing loose Web resources. Liu et al. [8] utilized tags and the surrounding texts of multimedia resources, and integrated the Semantic Link Network and multimedia resources to organize multimedia resources by their semantics. In [9], Liu et al. identified three major problems that hinder efficient and reliable communications and proposed a novel in-middle recovery scheme that was achieved by designing and implementing a proliferation routing to address them. To construct more energy-efficient topologies, Liu et al. [10] proposed a novel opportunity-based topology control and designed a fully distributed algorithm called CONREAP based on reliability theory to realize it. Further, Liu et al. [11] considered that it is not necessary to handle every temporal violation in a scientific workflow system and proposed a novel adaptive temporal violation handling point selection strategy to avoid unnecessary temporal violation handling. In [12], Wang et al. presented the design and implementation of G-Hadoop, a MapReduce framework that aims to enable large-scale distributed computing across multiple clusters. Xu et al. [13] proposed a general approach to generate the temporal semantic annotation of semantic relations between entities by constructing connection entities, lexical syntactic patterns, context sentences, context graph, and context communities.

### 2.2. Personalized search in collaborative tagging systems

There are some existing studies that utilize resource and user profiles to facilitate personalized search in a folksonomy (also known as a collaborative tagging system). Noll and Meinel [4] proposed term frequency (TF) profiles to discover related tags for users and resources, to provide personalized ranking. Later studies follow the term frequency-inverse document frequency (TF-IDF), Best Matching 25 (BM25) [5], and hybrid [14] paradigms. In [15], TF-IDF was combined with user and resource profiles along with the positions of tags by considering two kinds of sources. In our earlier work [2], we proposed a normalized term frequency (NTF) to model user and resource profiles and compared it with previous methods. In [16], Gemmel et al. proposed a method to personalize a users experience within a folksonomy using clustering. Kim et al. [17] proposed a new model of tag-based personalized searches to enhance not only retrieval accuracy but also retrieval coverage. Han et al. [18] collected user tags from folksonomies and mapped them onto an existing domain ontology. By leveraging social tagging as preference indicator, they built two latent tag models: a preference model and an annotation model. Bouadjenek et al. [19] proposed a new approach to enhancing document representation using social annotations. Jin et al. [20] presented verbal context in folksonomy to capture a user's intention and addressed the irrelevant contextual factors for a verbal context model. In [21], Hsu proposed an approach to build a tag-based resource profile semantically. The primitive features extracted from images such as color histograms [22], centerlines [23] and texture patterns [24], can be very useful and powerful for various search applications.

Although there are several works that handle personalized search with tag-based user and item profiles, they have some limitations. In the following subsection, we examine and discuss these limitations.

### 2.3. Limitations of current user profiling methods

All current work uses a single vector of tags to represent user preferences and treats all resources tagged by a user as favorite features, even mistaking disliked features as favorite features. When constructing a user profile, a single vector is insufficient for reflecting the degree of like or dislike, as explained in Example 1 (shown in Fig. 1).

**Example 1.** Consider a user Alice who has eaten three dishes and given three ratings of 5, 3, and 1 to each of the three dishes. Alice uses three tags, "Pepper, Fish," and "Sichuan," to annotate the dish rated as 5, and these three tags indicate her favorite flavors. In addition, for the dish rated 1, she uses two tags, "Mushroom" and "Fish," to remind

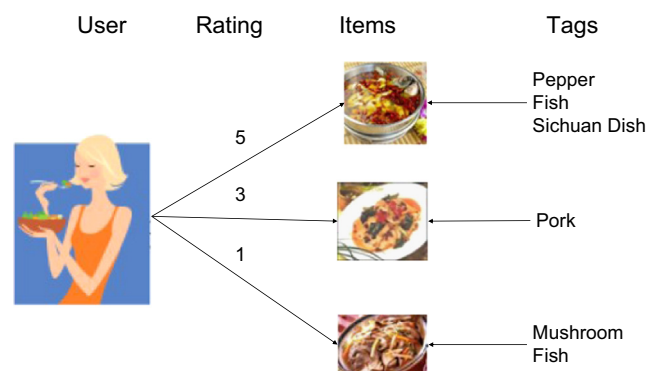


Fig. 1. Alice's ratings and tags for dishes.

<sup>3</sup> <http://www.grouplens.org/node/73>

Download English Version:

<https://daneshyari.com/en/article/405793>

Download Persian Version:

<https://daneshyari.com/article/405793>

[Daneshyari.com](https://daneshyari.com)