Information Sciences 385–386 (2017) 325–337

Contents lists available at ScienceDirect

Information Sciences

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

Combining tag correlation and user social relation for microblog recommendation

Huifang Ma*, Meihuizi Jia, Di Zhang, Xianghong Lin

College of Computer Science and Engineering, Northwest Normal University, Lanzhou, Gansu 730070, China

ARTICLE INFO

Article history:
Received 17 October 2015
Revised 28 November 2016
Accepted 31 December 2016
Available online 4 January 2017

Keywords:
Microblog recommendation
Tag retrieval
User-tag matrix
User-tag weight
Tag correlation
User–user social relation

ABSTRACT

With the development of social networking applications, microblog has turned to be an indispensable online communication network in our daily life. For microblog users, recommending high quality information is a demanding service. Some microblog services encourage users to annotate themselves with tags, which are used to describe their interests or attributes. However, few users are willing to create tags and available tags are not fully exploited for microblog recommendation. Besides, following/follower relationship in microblog is asymmetric, which can be used not only for communicating with friends or acquaintances but also for getting information on particular subjects. So far, there is no microblog recommendation algorithm which employs all the above mentioned information. This paper aims to investigate a joint framework to combine tag correlation and user social relation for microblog recommendation. Our approach identifies users' interests via their personal tags and social relations. More specifically, a user tag retrieval strategy is established to add tags for users without or with few tags, and the user-tag matrix is then built and user-tag weights are then obtained. In order to solve the problem of sparsity of the matrix, both inner and outer correlation between tags are investigated to update the user-tag matrix. Considering the significance of user social relation for microblog recommendation, a user–user social relation similarity matrix is constructed. Moreover, an iterative updating scheme is developed to get the final tag-user matrix for computing the similarities between microblogs and users. We illustrate the capability of our algorithm by making experiments on real microblog datasets. Experimental results show that the algorithm is effective for microblog recommendation.

1. Introduction

In recent years, online social networks witness a rapid development growth and become probably the main actor of the Web 2.0. It is universally acknowledged that social media is profoundly affecting not only the global economy but also every aspect of our daily life. Meanwhile, a growing number of users keep up with latest information by utilizing microblog tools, which provide a unique mechanism of information diffusion by allowing each user to receive messages from those he follows. Microblog not only amplifies interpersonal circles in social network, but also serves as a significant media for users to get the latest information. It is a momentary information publishing system based on Web 2.0 technology. According to Wikipedia, Sina Weibo, for example, owns over 500 million users, 46.2 million daily active users, and 100 million forwarding

* Corresponding author.
E-mail addresses: mahuifang@yeah.net, 35882614@qq.com (H. Ma).

http://dx.doi.org/10.1016/j.ins.2016.12.047
0020-0255/

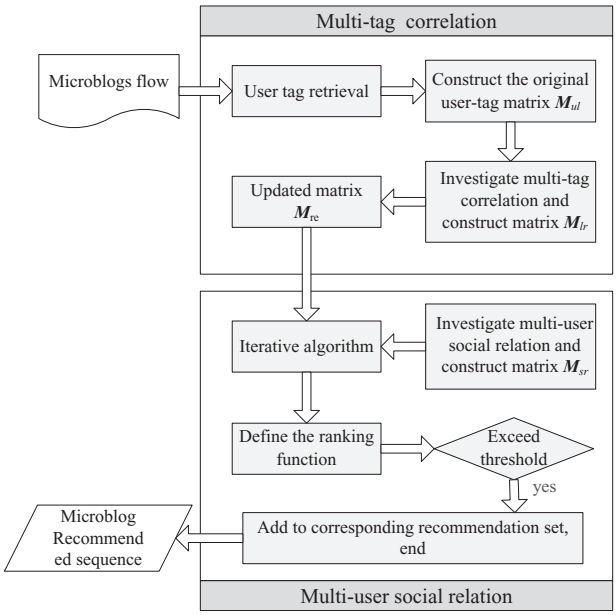

Fig. 1. Overview of the proposed microblog recommendation system.

message each day. However, the richness of online information also brings forth the 'information overload' problem. There is an urgent need to capture, understand and predict user interests for microblog application to provide better product designs, accurate targeted advertising, and personalized services [2,4,5,8]. How to identify users' interests and provide customized recommendation for each user becomes a challenging problem. Thereafter, a large number of recommendation algorithms based on user interest have been proposed [1,11,18,31].

Microblog contents can effectively reflect the public opinions, events in personal life, and personal interests [54]. Compared with the traditional blogs, the sparsity of content and the lack of contextual information in microblogs bring new challenges for recommendation. Most researchers try to extend text features and enrich semantic information for microblogs [33,46]. They usually take advantage of an external database to enrich the semantic information. Some researchers try to trim a short text representation to get a few most representative words for topical classification [44]. Besides, the distinctive features of microblog allow users to post their messages and to follow others, it is possible to utilize these features for better recommendations than other traditional applications.

In this paper, we propose a microblog recommendation algorithm based on tag correlation and user social relation. On the one hand, a user tag retrieval strategy is developed to assign tags for users and a user-tag matrix is created to provide the initial weights for users' tags; then the correlation matrix of multi-tags is constructed by investigating both inner and outer correlation between tags; finally the original user-tag matrix can be updated through multiplication of the above two matrices. On the other hand, a social relation matrix of multi-users is constructed by analysing following/follower relationship between users; then the final user-tag matrix can be obtained via iteratively multiplying social relation matrix with the updated user-tag matrix; finally the similarities between microblogs and users can be defined by the user-tag matrix for microblog recommendation. Our extensive experimental study shows the scalability and efficiency of our approach. Fig. 1 summarizes an overview of the proposed microblog recommendation system.

The contributions of this paper are summarized as follows:

1. We propose a novel tag retrieval method which is suitable for identifying the most representative words from a user's microblogs. Our model selects terms as tags for users by utilizing their significance reflected by importance and topical indication.
2. We investigate the semantic correlation of pairwise tags by involving both the inner correlation, reflecting the explicit correlation within tag pairs for each user and the outer correlation, capturing the implicit correlation between tags by considering the relation strength of their interactions with tags of other users. And we then update the original user-tag matrix, thus better representing users' interests.
3. We design a user–user similarity calculation method based on user–user social relation and provide an iterative approach for updating user-tag matrix to reveal the user's interests more accurately.

The rest of this paper is organized as follows: In Section 2, we discuss several related works. Section 3 presents user tag retrieval strategy, user-tag matrix construction method and multi-tags correlation calculation mechanism. Multi-users social relation matrix construction method, the iterative calculation mechanism and the recommendation algorithm are given in Section 4. The experimental results are demonstrated in Section 5. Lastly, we conclude our paper in Section 6.