Textual-geographical-social aware point-of-interest recommendation

Ren Xingyi (✉), Song Meina, E Haihong, Song Junde

1. School of Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
2. Engineering Research Center of Information Networks, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract

The rapid development of location-based social networks (LBSNs) has provided an unprecedented opportunity for better location-based services through point-of-interest (POI) recommendation. POI recommendation is personalized, location-aware, and context depended. However, extreme sparsity of user-POI matrix creates a severe challenge. In this paper we propose a textual-geographical-social aware probabilistic matrix factorization method for POI recommendation. Our model is textual-geographical-social aware probabilistic matrix factorization called TGS-PMF, it exploits textual information, geographical information, social information, and incorporates these factors effectively. First, we exploit an aggregated latent Dirichlet allocation (LDA) model to learn the interest topics of users and infer the interest POIs by mining textual information associated with POIs and generate interest relevance score. Second, we propose a kernel estimation method with an adaptive bandwidth to model the geographical correlations and generate geographical relevance score. Third, we build social relevance through the power-law distribution of user social relations to generate social relevance score. Then, our exploit probabilistic matrix factorization model (PMF) to integrate the interest, geographical, social relevance scores for POI recommendation. Finally, we implement experiments on a real LBSN check-in dataset. Experimental results show that TGS-PMF achieves significantly superior recommendation quality compared to other state-of-the-art POI recommendation techniques.

Keywords   location-based social networks, POI recommendation, topic model, geographical correlations, social correlations

1 Introduction

With the rapid development of cities, along with the growth of the number of POIs, for example, stores, restaurants, museums, entertainments, hotels and so on, provide more opportunities for us to experience life than ever before. Generally, people are willing to explore the cities and the neighboring places in their daily life, and then according to their personal interests with regard to the kinds of choices of POIs to decide where to go. The huge volume of data contains valuable information about POIs and human preference, which can be exploited for POI recommendation. In the meanwhile, how to effectively make a satisfying decision for a user among the large number of POIs is a difficult problem, it is generally regarded as 'choice paralysis'. POI recommendation is a task that purposes to deal with this problem by helping users to filter out uninteresting POIs and reduce their decision making time.

As an increasingly popular application of location-based services, LBSNs provide an unprecedented opportunity to study human mobile behavior for POI recommendation in spatial, temporal, social, and content aspects. Typical location-based social networking sites, such as, Foursquare, Yelp, and Facebook Places, allow users to 'check in' on POIs with mobile devices like smartphones, and leave tips and share that experience with their online friends, resulting in a 'W[4]' (i.e., who, when, where, and what) information layout, corresponding to four distinct information layers. Indeed, the task of POI recommendation is to provide personalized recommendations of places of interest. It plays an important role in providing better location-based services in LBSNs.

Compared to the development of traditional recommender systems, the development of POI recommender systems is much more complicated. POI recommendation is challenging for some reasons. First, the user POI check-in matrix is highly sparse because users have only visited a very small proportion of POIs in a LBSN, and thus recommendation methods suffer from the data scarcity problem. Second, for POI recommendation, the interest of the user can vary dramatically at different time and locations. Third, in POI recommendation, different types of context information can be obtained, such as geographical coordinates of POIs, time stamps of check-ins, friendship of users, categories of POIs, etc. However, unlike traditional recommendation, the textual information associated with POIs is usually incomplete and ambiguous.

In light of the above challenges, in this paper, we propose a textual-geographical-social probabilistic matrix factorization method for POI recommendation. We are motivated to propose a new probabilistic matrix factorization approach for POI recommendation through exploiting and integrating topic model, geographical correlations and social correlations, called TGS-PMF.

Our contributions in this paper can be summarized as follows:

1) We propose a textual-geographical-social aware probabilistic matrix factorization method for POI recommendation called TGS-PMF, which incorporates the textual information, geographical information and social information.

2) We design an aggregated LDA model to learn the interest topics of users and to infer the interest POIs by mining textual information associated with POIs.

3) We devise a method to model the geographical correlations, which extend the kernel density estimation by applying an adaptive bandwidth that is learned from the historical check-in data of all users.

4) We develop a method to model the social correlations, which estimate the social check-in frequency or rating by a user's friends to a POI as a power-law distribution that is learned from the historical check-in data of all users.

5) To the best of our knowledge, our proposed TGS-PMF is the first study to incorporate interest, geographical and social relevance scores into probabilistic matrix factorization model for POI recommendation.

6) We conduct extensive experiment to evaluate the recommendation accuracy of TGS-PMF using the real world datasets. Experimental results show that TGS-PMF significantly outperforms other state-of-the-art POI recommendation techniques.

The rest of this paper is organized as follows. We review the related work on POI recommendations in Sect. 2. Our proposed the TGS-PMF model are presented in Sect. 3. The experimental evaluation of our proposed methods is shown in Sect. 4. Finally, we conclude this paper and put forward further work in Sect. 5.

2   Related work

2.1   POI recommendation using textual information

In order to better understand the patterns of LBSN and to improve the service of LBSN, more recent work began to explore textual information. Farrahi et al. [1] applied topic models to identified daily location-driven routines by mining text from mobile phone data. Ye et al. [2] proposed a work on semantic annotation for LBSNs to annotate places with category tags by exploring explicit patterns of individual places and implicit relatedness among similar places. Yin et al. [3] presented a latent geographical topic analysis method to explore both location and associated text of locations and found this can help to discover meaningful geographical topics. Ferrari et al. [4] analyzed Twitter posts and performed LDA on the data to extract urban patterns, such as hotspots and crowd behaviors. Agarwal et al. [5] associated the rating by regularizing both user and item factors simultaneously through user features and words associated with each item. Pennacchiotti et al. [6] explored topic models for item recommendation which used topic models to study social media user interests to recommend new friends with similar interests. Furthermore, the textual terms associated with POIs are usually incomplete and ambiguous. In this paper, we explore both textual information and context information through the topic model to address these challenges.

2.2   POI recommendation using geographical information

The geographical proximity significantly affects the check-in behaviors of users on the POIs, the geographical information has been intensively used in POI recommendations. One way is to simply consider the current locations of users to filter out the POIs that are far from the users [7–9]. Another way is to geographical latent