Information Sciences 293 (2015) 163–181

Contents lists available at ScienceDirect

Information Sciences

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

CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations

Jia-Dong Zhang, Chi-Yin Chow *

Department of Computer Science, City University of Hong Kong, Hong Kong

ARTICLE INFO

Article history:
Received 15 October 2013
Received in revised form 8 September 2014
Accepted 13 September 2014
Available online 20 September 2014

Keywords:
Location-based social network
Location recommendation
Personalization
Geographical influence
Social influence

ABSTRACT

With the rapid growth of location-based social networks (LBSNs), location recommendations play an important role in shaping the life of individuals. Fortunately, a variety of community-contributed data, such as geographical information, social friendships and residence information, enable us to mine users' reality and infer their preferences on locations. In this paper, we propose an effective and efficient location recommendation framework called CoRe. CoRe achieves three key goals in this work. (1) We model a personalized check-in probability density over the two-dimensional geographic coordinates for each user. (2) We propose an efficient approximation approach to predict the probability of a user visiting a new location using her personalized check-in probability density. (3) We develop a new method to measure the similarity between users based on their social friendship and residence information, and then devise a fusion rule to integrate *the geographical influence* with *the social influence* so as to improve the user preference model on location recommendations. Finally, we conduct extensive experiments to evaluate the *recommendation accuracy*, *recommendation efficiency* and *approximation error* of CoRe using two large-scale real data sets collected from two popular LBSNs: Foursquare and Gowalla. Experimental results show that CoRe achieves significantly superior performance compared to other state-of-the-art geo-social recommendation techniques.

1. Introduction

In a location-based social network (LBSN) depicted in Fig. 1, such as Foursquare and Gowalla, users can establish social links with others, indicate their residence, check in some points-of-interest (POIs), e.g., restaurants, stores, and museums, and share their experiences of visiting specific locations by leaving some tips or comments [11]. These community-contributed data enable us to mine users' reality [39,41,45]. For example, in LBSNs it is prevalent to recommend locations for users based on their preferences inferred from the community-contributed data [43], which not only helps users explore new places and enrich their life but also enables companies to launch advertisements to potential customers and improve their profits.

The geographical feature of POIs distinguishes them from other non-spatial items, such as books, music and movies in conventional recommender systems [4], because physical interactions are required for users to visit POIs. Thus, **the geographical information (geographic coordinates) of locations plays a significant influence on users' check-in behaviors**,

* Corresponding author. Tel.: +852 3442 8679; fax: +852 3442 0503.
E-mail addresses: jzhang26-c@my.cityu.edu.hk (J.-D. Zhang), chiychow@cityu.edu.hk (C.-Y. Chow).

http://dx.doi.org/10.1016/j.ins.2014.09.014
0020-0255/

known as **geographical influence** for short, which gives new opportunities and challenges to infer users' preferences for making location recommendations for them.

Limitations in existing methods. There are a few studies that learn users' preferences through exploiting the geographical influence for location recommendations [8,28,35,55]. Unfortunately, these studies use the one-dimensional geographical distance between locations for location recommendations. For example, the literatures [28,35] simply assume that the propensity of a user for a POI is inversely proportional to the distance between the POI and her visited locations; more sophistically, two techniques [8,55] model the distance between locations visited by the same user as a power-law distribution or multi-center Gaussian distribution. In general, these studies suffer from two major limitations. (1) **One-dimensional geographic distance influence.** Existing work simplifies the geographical influence as a one-dimensional and monotonous distance distribution, but it is more reasonable and intuitive to model the geographical influence as a two-dimensional and multimodal check-in probability density. The reason is twofold. (a) The probability of a user visiting a location is not simply monotonous respecting their distance, because the visiting probability is not only affected by the distance but also the location's intrinsic characteristics. For example, in reality the check-in locations of a user are usually distributed in several areas. (b) It is hard to compute a visiting probability for a location based on a distance distribution, since it needs to find a reference location to derive a reasonable distance for the location in the first place. Conversely, it is considerably intuitive to employ a two-dimensional check-in probability density to compute a visiting probability for any location with latitude and longitude. In addition, it is rational to use the two-dimensional density on the spherical dimensions because the altitude is a dependent variable (i.e., the altitude is usually determined based on a pair of latitude and longitude) and the earth is almost a sphere (i.e., the altitude of a location is negligible). (2) **Non-personalized geographical influence.** They apply the same inversely proportional relation between the propensity and the distance for all users [28,35] or universally estimate a common distance distribution for all users [8,55], but in practice the geographical influence of locations should be personalized since the user's check-in behavior is unique [61].

Real-world motivating examples. To observe users' unique check-in behaviors and multimodal check-in probability densities over the two-dimensional geographic coordinates, a spatial analysis is conducted on two publicly available real data sets collected from Foursquare [14] and Gowalla [10], which are two popular LBSNs. Specifically, we focus on three users who are randomly chosen from each data set. Fig. 2 depicts their individual check-in probability density over the two-dimensional geographic coordinates. The probability density is estimated based on the kernel density estimation (KDE) technique [48]. We have the following two findings. (1) The geographical influence of locations on these three users' check-in behaviors is unique since their check-in probability densities are distinct from each other. (2) These check-in probability densities are usually multimodal rather than unimodal or monotonous.

Our approach. In this paper, we propose a new location recommendation framework, called CoRe, that achieves three key goals. (1) We infer users' preferences by exploiting the personalized two-dimensional geographical influence. Specifically, we estimate a personalized two-dimensional check-in probability density over the latitude and longitude coordinates for each user rather than using a common one-dimensional distance distribution for all users. In addition, since the check-in probability densities of users are diverse, we cannot assume their forms. We thus use a nonparametric density estimation method, i.e., the popular kernel density estimation [48]. (2) We can use a personalized check-in probability density to predict the probabilities of each user visiting each of all new locations (that have not been visited by the user). In particular, we contrive an efficient approximation approach to derive these visiting probabilities based on the fast Gauss transform [19], in order to improve system efficiency and scalability. (3) We integrate *the geographical influence* with *the social influence* in the conventional social networks to enhance the user preference model and obtain better quality of location recommendations. First of all, to consider the social influence, we develop a new method to measure the similarity between users based on their social friendships and residence information, because nearby friends share more commonly visited locations than others [10]; the similarity measure between users is employed to estimate the rating of a user to a new location. We then design a new fusion rule, called the union rule, to integrate the geographical probability and social rating of a user to a new location into a unified score, instead of applying the simple product or sum rule. Eventually, we can make a location recommendation for each user by returning top-k locations with the highest scores to her.

This study is significant different from our previous work [61]. The main differences and contributions of this paper can be summarized as follows:

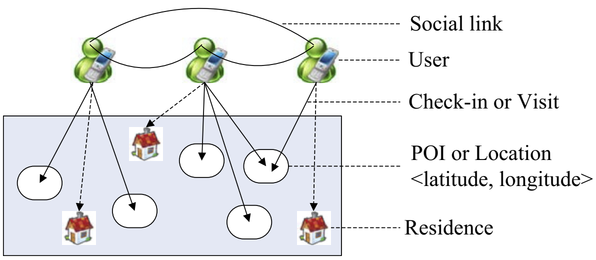

Fig. 1. A location-based social network [61].