DP-POIRS: A Diversified and Personalized Point-of-Interest Recommendation System

Xiangfu Meng
School of Electronic and Information Engineering
Liaoning Technical University
Huludao, China
marxi@126.com

Yanhuan Tang
School of Electronic and Information Engineering
Liaoning Technical University
Huludao, China
1072201811@qq.com

Xiaoyan Zhang
School of Electronic and Information Engineering
Liaoning Technical University
Huludao, China
187389797@qq.com

***Abstract*—Diversity point-of-interest recommendation system benefits users to broaden their interests, access and discover new interest points. This paper describes a Diversified and Personalized Point-Of-Interest Recommendation System (DP-POIRS) by leveraging the geo-social relationships between POIs. The system consists of three components. The first component – geo-social distance measuring component is used to construct a correlation matrix to describe the geo-social distance between points of interests. The second component –point-of-interest partition component, divides the interest points into diverse clusters by using the spectral clustering algorithm over the correlation matrix. The third component –personalized sorting component, finds out the user's favorite interest points from each cluster, and then sorts them into a list of recommendation by the use of matrix factorization algorithms.**

Keywords—point-of-interest (POI) recommendation; geo-social distance; diversification; personalization

I. Introduction

With the popularity of mobile Internet, point-of-interest (POI) recommendation system has become a new research hotspot [1-3]. Most of the POI recommendation algorithms [4-6] mainly focus on matching the user's personalized needs, and striving to provide the list of POIs that are closely relevant to the user preferences. However, these methods ignore the diversity of the list of contents, which usually lead to the POIs in the recommendation list are very similar to each other. To solve this problem, we construct a diversified and personalized POI recommendation system based on the geo-social distance between POIs, by fusing spectral clustering and matrix factorization algorithms. In this demo paper, we will briefly introduce our point-of-interest recommendation method and then presents a demo system accordingly.

II. Approach

A. Geo-social Distance Model

Based on the geo-social model proposed in [7], we design the following model of geo-social distance.

- Definition 1. (Place set). The set of places can be denoted as $P=\{p_1, p_2, ..., p_n\}$ including all places $P_i=(lat_i, lon_i)$, where lat_i and lon_i represents the latitude and longitude of P_i respectively.
- Definition 2. (User social relation network graph). We use a undirected graph $G=(U, E)$ to denote a social network, where U is the set of all users, every edge $(u_i, u_j)\in E$ indicates that user u_i has a friend relationship with user u_j. Note that, a friend relationship here refers to the fact that each user is registered with each other in the list of friends on the social software, and $(u_i, u_j)\in E$.
- Definition 3. (User check-in records). The check-in records of a user in the user set U can be denoted as $CK=\{<u_i ,p_k ,t_r>|u_i\in U, p_k\in P\}$. The collection of users who have visited location p_k can be represented as $U_{Pk}=\{u_i|<u_i , p_k , *>\in CK\}$, where * denotes any time.

The geo-social distance between places (P_i, P_j) is defined as

$$D_{gs}\left(p_i, p_j\right)=\omega\cdot D_P\left(p_i, p_j\right)+(1-\omega)\cdot D_S\left(p_i, p_j\right) \tag{1}$$

where, $\omega\in[0, 1]$ is a parameter used to regulate the proportion of geographic distance and social relation between POIs for calculating the geo-social distance. $D_p(p_i, p_j)$ represents the geographic distance between place p_i and p_j, which can be defined as

$$D_P\left(p_i, p_j\right)=\frac{E\left(p_i, p_j\right)}{\max D} \tag{2}$$

where $E(p_i, p_j)$ is the Euclidean (or road network) distance between p_i and p_j, and $maxD$ is the maximum distance between any pair of two points in P.

$D_s(p_i, p_j)$ represents the social relation between place p_i and p_j, which can be defined as

$$D_S\left(p_i, p_j\right)=1-\frac{\left|CU_{ij}\right|}{\left|U_{p_i}\cup U_{p_j}\right|} \tag{3}$$

where

$$\begin{aligned}CU_{ij}=&\left\{u_a\in U_{p_i}\mid u_a\in U_{p_j} or \exists u_b\in U_{p_j},(u_a,u_b)\in E\right\}\\&\cup\left\{u_a\in U_{p_j}\mid u_a\in U_{p_i} or \exists u_b\in U_{p_i},(u_a,u_b)\in E\right\}\end{aligned} \tag{4}$$

Therefore, we can define the correlation in both location and social relations between a pair of places p_i and p_j,

978-1-5090-5004-8/17 $31.00 © 2017 IEEE
DOI 10.1109/DSAA.2017.24

$$S(p_i, p_j) = 1 - D_{gs}\left(p_i, p_j\right) \tag{5}$$

After this, a $n*n$ correlation matrix can be obtained by calculating the correlation between any pair of places in set P.

B. Spectral Clustering

Spectral clustering is a well-known classical clustering algorithm, which makes the points in the same cluster more closely linked, while the association between different clusters is sparse. The more details of spectral clustering can be found from paper [8]. In this paper, we user the spectral clustering algorithm to partition the POI set into several diversified clusters, where each cluster represents one kind of POIs and the POIs in each cluster are closely related to each other both in social and location aspects.

C. Matrix Factorization

As the classical algorithms, the spectral clustering, matrix factorization algorithms are widely circulated. In this system, we use three types of matrix factorization algorithms to estimate the number of visits to a POI by a user to evaluate the POI's popularity, which was well described in paper [9].

III. Framework and Demo System

A. Framework

The framework of our DP-POIRS system is showed in Figure 1. In this framework, we first use the geo-social distance model to calculate the geo-social distance between POIs and then the correlation matrix of POIs can be obtained. Second, we use matrix factorization algorithms to calculate the satisfaction degree of each user to each POI, and thus a satisfaction matrix can be formed. Based on the correlation matrix obtained in the first step, we divide all POIs into different clusters by using the spectral clustering algorithm, which makes the POIs in the same cluster are as similar as possible while the differences between clusters are as large as possible. Next, for a specific user, we can pick out one POI from each cluster by considering the satisfaction degree matrix to form a candidate list of POIs. Lastly, according to the user preferences, the POIs in candidate list would be sorted to form a final recommendation list.

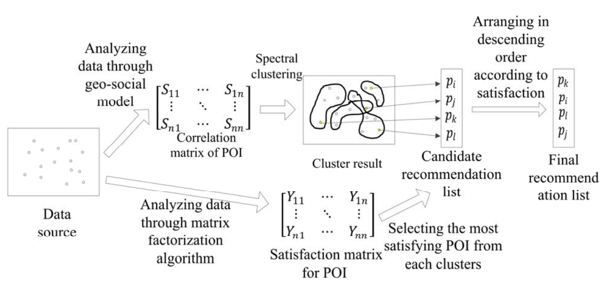

Fig. 1. General framework diagram.

B. Demo System

The demo system was developed by Python and ran on the Gowalla dataset. In this system, three parameters should be provided, the weighting value ω, the number of minimum eigenvalues to be used in spectral clustering, and the number of recommended POIs, respectively.

For example, given a user id 56878, we set the value of ω to 0.9, the number of eigenvalues to 2, and the number of POIs to 10. There are 1,078 POIs in check-in history. Figure 2 shows the recommendation list for the given user. As shown, the system returns 10 of the best POIs and their corresponding scores. The score here refers to the number of user attendance that is estimated by matrix decomposition.

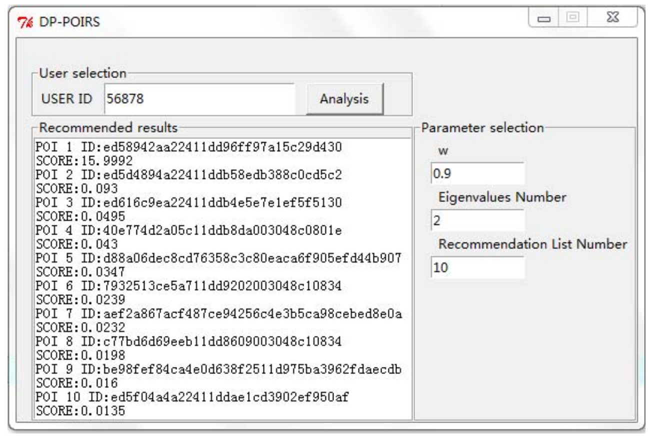

Fig. 2. The demo system of POI recommendation

Acknowledgment

This work is supported by the National Nature Science Foundation of China (61772249).

References

[1] Wang W, Yin H, Chen L, et al. Geo-SAGE: A Geographical Sparse Additive Generative Model for Spatial Item Recommendation[C] //Proc of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney: ACM, 2015: 1255-1264.

[2] Zhang J D, Chow C Y. GeoSoCa: Exploiting Geographical, Social and Categorical Correlations for Point-of-Interest Recommendations[C] // Proc of International ACM SIGIR Conference on Research and Development in Information Retrieval. Santiago: ACM, 2015: 443-452.

[3] Guo G, Zhang J, Yorke-Smith N. TrustSVD: Collaborative Filtering with Both the Explicit and Implicit Influence of User Trust and of Item Ratings [C] //Proc of the 29th AAAI Conf on Artificial Intelligence (AAAI'15). Austin: AAAI, 2015: 123-129.

[4] Li H, Hong R, Wu Z, et al. A Spatial-Temporal Probabilistic Matrix Factorization Model for Point-of-Interest Recommendation[C] //Proc of Siam International Conference on Data Mining. Miami: SIAM, 2016: 117-125.

[5] Zhao S, Zhao T, Yang H, et al. STELLAR: Spatial-Temporal Latent Ranking for Successive Point-of-Interest Recommendation[C] //Proc of Phoenix: AAAI, 2016: 315-322.

[6] Yuan Q, Cong G, Ma Z, et al. Time-aware point-of-interest recommendation[C] //Proc of International ACM SIGIR Conference on Research and Development in Information Retrieval. Dublin: ACM, 2013: 363-372.

[7] Shi J M, Mamoulis N, Wu D M, et al. Density-based place clustering in geo-social networks [C] //Proc of ACM SIGMOD Conference on Data Management (SIGMOD 2014). Snowbird: 2014: 99-110.

[8] Shi J, Malik J. Normalized cuts and image segmentation[C] //Proc of 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97). San Juan: IEEE, 1997: 731-737.

[9] Lee D D, Seung H S. Algorithms for Non-negative Matrix Factorization[C] //Proc of Neural Information Processing Systems. Denver: MIT, 2000: 556--562.