Multiple Similar Groups Based Information Technology for POI Recommendation in LBSNs

Yibo Ding[1*], Jinyu Bai[2], Yiqing Dai[3], Jinshuo Zhang[4], Yixuan Zheng[5], Qitai Xu[6] and Shengwen Yu[7]

Arts and Business College of London, London, British

[1]dingyibo2023@163.com; [2]baijinyu202209@163.com; [3]Yiqingdai2022@163.com; [4]zhangjinshuo202209@163.com; [5]zhengyixuan202210@163.com; [6]Qitaixu2022@163.com; [7]shengwx1008@163.com

Abstract. With the development of modern information technology, it becomes increasingly easier for users to obtain information through various type of services or applications. However, it also brings difficulty for users to search for really desired information among massive corresponding data. Thus, recommendation systems have become necessary for these services and applications to solve the information overload problem. Point-of-interest (POI) recommendation is famous in Location-based social networks (LBSN) for the ability of exploring users' preference and recommending interesting places. Most traditional POI recommendation algorithms are collaborative filtering (CF) based, and the key idea is to generate recommendation list for a target user by mining historical data of his similar users. We found that, different similarity measurement method may lead to different conclusions. Thus, we proposed a multiple similar group-based CF algorithm for POI recommendation in this paper. Given a target user, we first determined his similar users by using different similarity calculation methods and constructed corresponding groups. The recommended locations for the target user are determined by comprehensively considering the suggestions of our groups. We implemented our algorithm and compared with previous approaches by using the dataset. The experimental results show that, our algorithm performs best.

Keywords: Modern information technology; Point-of-interest; recommendation system; information overload; similarity measurement; collaborative filtering

1 Introduction

Location Based Social Network (LBSN) is a new type of social network which can link different users by using their visited locations. Through LBSN, users could share their experiences of the visited places and provide corresponding suggestions. However, as LBSN grows more and more popular, the increasing amount of historical visiting data bring difficulties for users to find the really desired places. Thus, point-of-interest (POI)

H. Wu et al. (Eds.): CSIEDE 2022, ACSR 103, pp. 397–404, 2022.
https://doi.org/10.2991/978-94-6463-108-1_45

recommendation becomes necessary for exploring interesting places among massive historical data. Therefore, how to achieve a precise POI recommendation has become a research hotspot. One major challenge in POI recommendation is data sparsity problem. Although data sparsity is a common problem in all recommendation systems, but this problem is extremely bigger in POI recommendation than recommendations of other areas, such as music recommendation or news recommendation [1]. Thus, in order to achieve precise POI recommendations, data sparsity problem should be considered first.

Collaborative Filtering (CF) is an effective method to solve the data sparsity problem and many previous POI recommendation algorithms are CF-based [2, 3]. The key ides of CF-based algorithm to solve data sparsity problem is to find similar users for a target user, and comprehensively consider historical data of similar users to generate recommendations for the target user. Thus, the similarity measurement between different users is a core issue for CF based algorithms. However, different similarity measurements methods may lead to even contrary results. Consider the example shown in Figure 1, given the check-in records of user 1 and user 2, the similarity between them is different by using Cosine similarity method and Euclidean distance similarity method. Thus, a single similarity measurement method may cause a bias determination of similarity users.

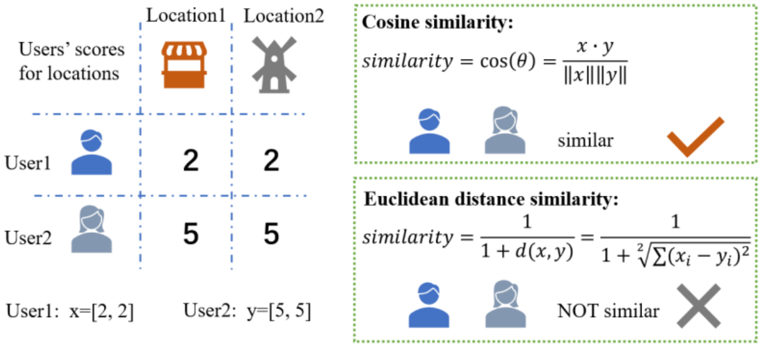

Fig. 1. An example of different similarity measurement [Owner-draw]

To address the problem causing by single similarity measurement, we propose a multiple similar groups-based CF algorithm for POI recommendation in this paper. For a target user, we construct multiple groups which contains similar users of the target user by using different similarity measurement methods. Each similar group generates recommendation candidates, and final recommended POIs are determined by integrating the candidates of all groups. In this paper, we propose a unified model to integrate the recommendations of multiple similar groups which constructed by using different similarity measurement methods.

The rest of paper is organized as follows: section 2 summarizes related work; section 3 describes the details of our multiple similar groups-based recommendation algorithm; section 4 shows the experimental results and section 5 concludes this paper.

2 Related work

CF is an effective method to solve data sparsity problem in POI recommendation [4]. Many previous CF-based methods focus on finding similar users. Song et al. [5] constructed a user-POI matrix and used Cosine similarity to calculate the similarities between different users. Ye et al. [6] explored multiple influences on POI recommendation and proposed a fusion framework to measure similarities between different users by comprehensively considering users' preferences, social influence and geographical influences. Although they considered multiple influences, the similarity calculation method is fixed and may lead to other results by using other similarity calculation method. Jiao et al. [7] pointed out that the order of check-ins of users should be considered to explore users' preference. Instead of user-POI matrix, they used trajectories of users to calculate similarities. Such trajectory contains rich information of the visited locations including longitude, latitude, category, check-in time and frequency. They defined similar users as the users who have similar trajectories. Then, they proposed a method the calculate the similarities between trajectories. Li et al [8] constructed a tensor to represent the historical data of users. The three dimensions of this tensor are user, location category and check-in time. Then, tensor decomposition method was used to fill the missing value of the tensor. Finally, Cosine similarity is used to measure similarities between different users.

3 Multiple similar group-based CF algorithm

In this section, we will illustrate our similar group-based CF algorithm for POI recommendation. We first construct similar user groups for a target user by using different calculation methods. Then, FunkSVD [9] is used to fill the missing values of user-POI matrix of each group. Finally, generate a recommendation list for the target user by considering all candidate provided by these groups. Our algorithm consists of two major components: similar user group construction and recommendation list generation. Figure 2 shows the framework of our algorithm. In the next subsections we describe the details of these two components.

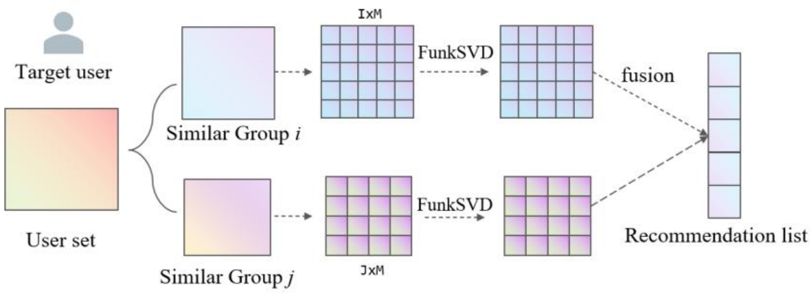

Fig. 2. Framework of our algorithm [Owner-draw]

3.1 Similar user group construction

In this subsection, we describe how to construct multiple similar groups and generate corresponding user-POI matrixes by using the historical data of these groups. We first user the entire dataset to construct a user-POI matrix. The value of each element of the matrix is normalized check-in frequency of a user. Then, FunkSVD method is employed to fill the empty values. Thus, we achieve a fully user-POI matrix for all users. The feature of each user u_i can be represents by a vector $[f_1, f_2, ..., f_j..., f_N]$, where N is the total number of POIs, and f_j denotes the normalized check-in frequency of u_i at POI p_j, calculated by Equation (1).

$$f_j = \frac{total\ number\ of\ check\text{-}ins\ at\ P_j}{N} \tag{1}$$

After achieve the features of users, we apply the following similarity calculation methods to construct different similar user groups. The details of similarity calculation methods are shown in Table 1.

Table 1. Similarity calculation methods

Method	Equation	Description
Euclidean distance	$sim_E(u_i, u_q) = \sqrt{\sum (f_j(u_i) - f_j(u_q))^2}$	$f_j(u_i)$ denotes the normalized check-in frequency of u_i at POI p_j.
Cosine	$sim_C(u_i, u_q) = \frac{\sum_{j=1}^{N}(f_j(u_i) \times f_j(u_q))}{\sqrt{\sum_{j=1}^{N}(f_j(u_i))^2} \times \sqrt{\sum_{j=1}^{N}(f_j(u_q))^2}}$	N is the total number of POIs.
Jaccard	$sim_J(u_i, u_q) = \frac{\lvert P(u_i) \cap P(u_q) \rvert}{\lvert P(u_i) \cup P(u_q) \rvert}$	$P(u_i)$ is a set which contains the POIs visited by u_i.

Give a target user u_i, we calculate similarities between u_i and each other by using three methos described in Table 1. For each similarity calculation method, we select top 20 users with the highest similarities between u_i to generate a set, respectively. Thus, we achieve three similarity group, denoted by $S_E(u_i)$, $S_C(u_i)$ and $S_J(u_i)$. Then, three sub user-POI matrixes can be constructed by only using the historical data of the target user and corresponding similar users. Finally, FunkSVD [9] is applied to predict the missing values of these three sub user-POI matrixes.

3.2 Recommendation list generation

Give similar groups of a target user, we next generate a recommendation list by respectively considering the historical check-in records of the users in different similar groups. Given a similarity group of a target user u_i, denoted by $S(u_i)$, we construct a set $P(u_i)$ to store the all the POIs visited by the users in $S(u_i)$. Such POIs can be regarded as the recommended candidates provided by corresponding similar group. The recommended scores of these POI can be calculated by using Equation (2).

$$socre(p_k) = \frac{\sum_{u_j \in S(u_i)} \frac{fre_k(u_j)}{NUM(u_j)}}{M(p_k)} \tag{2}$$

In Equation (2), *score*(p_k) is the recommended score of p_k by only considering the historical data of similar group. Obviously, higher score means higher recommended probability. $fre_k(u_j)$ denotes the check-in times of u_j at POI p_k, and $NUM\ (u_j)$ is the total number of check-in records of u_j. $M\ (p_k)$ represents the total number of users in $S\ (u_i)$ who visited POI p_k.

After calculating the scores of POIs provided by each single similar group, we integrate the scores all three similar groups by calculating average values of scores provided by different similar groups. The we select top K POIs to generate a recommendation list for the target user.

4 Experiments

We have implemented our algorithm and test the performance by using the public open-source dataset collected from Foursquare [10]. This dataset contains 226,428 check-in records in New York city and 573,370 check-ins in Tokyo. We randomly select 70% check-in records of each user to generate training set and the rest check-ins are used as testing set. The evaluation metrics are precision and recall as shown in Equation (3) and Equation (4).

$$precision = \sum_{all\ users} \frac{number\ of\ POIs\ correctly\ recommended}{number\ of\ recommended\ POIs} \tag{3}$$

$$recall = \sum_{all\ users} \frac{number\ of\ POIs\ correctly\ recommended}{number\ of\ POI\ actually\ visited\ in\ testing\ set} \tag{4}$$

We test the performance of top 5, 10 and 20 recommendations, respectively. The baseline methods are described as follows:

- TCF [11]: This is a temporal-aware POI recommendation approach by using collaborative filtering
- GPF [12]: This method represents a framework to capture geographical influences for POI recommendation.
- TRA [13]: This method compared the trajectories of different users to measure the similarities.
- RTP [14]: this method is a region transfer based collaborative filtering POI recommendation algorithm.

The experimental results are shown in Figure 3 – Figure 8. As shown in the figures, the algorithm proposed in this paper outperforms all baseline methods. The major reason is that we comprehensively consider different similarity calculation method to generate multiple similar user group and integrate the recommended candidates suggested by different groups to construct the final recommendation list. In this way, an appropriate similarity measurement between different users can be achieved and the recommendation results can be more accurate.

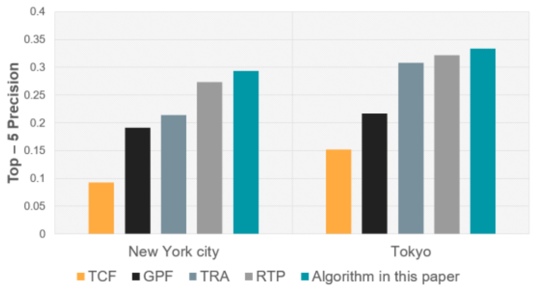

Fig. 3. Precisions of Top 5 recommendations [Owner-draw].

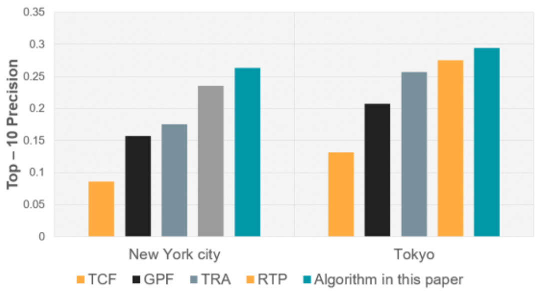

Fig. 4. Precisions of Top 10 recommendations [Owner-draw].

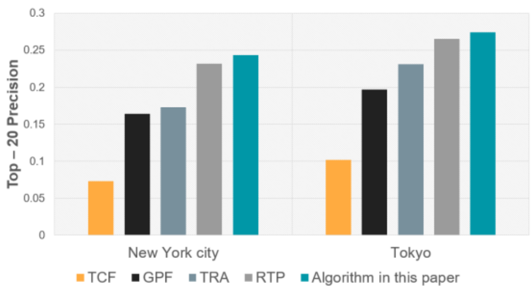

Fig. 5. Precisions of Top 20 recommendations [Owner-draw].

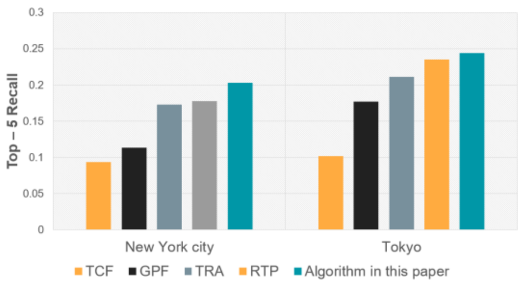

Fig. 6. Recalls of Top 5 recommendations [Owner-draw].

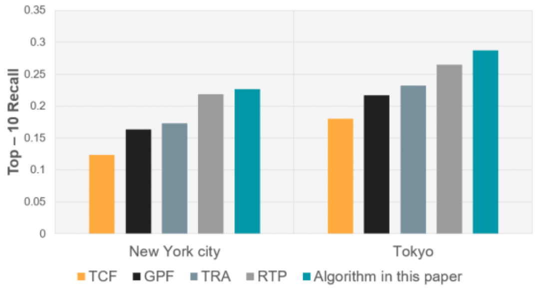

Fig. 7. Precisions of Top 10 recommendations [Owner-draw].

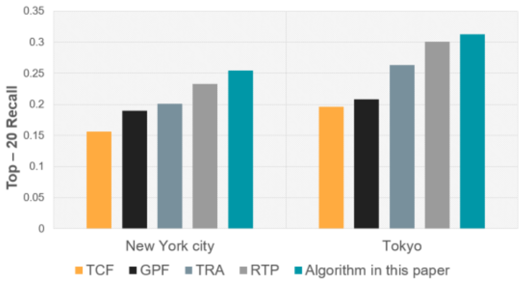

Fig. 8. Recalls of Top 20 recommendations [Owner-draw].

5 Conclusion

In this paper, we investigate POI recommendation problem, and propose a multiple similar group-based CF algorithm. The objective of this algorithm is to explore the historical check-ins of users and predict which locations will be visited by users in the future. Such algorithm can help user find their interested location among massive corresponding information. Besides, the algorithm can also help LBSN service to explore users' preference and provide effective suggestions. The major contribution of this paper is to construct multiple similar groups of a target user and comprehensively consider historical data of these groups to give suggestions. However, we only use the check-in frequencies of POIs to reflect users' preferences. Actuarially, comments of corresponding POIs are provided on LBSNs and contains rick information. In the future, we tend to investigate how to embed comments of users to POI recommendaion model.

References

1. Liu, B., Meng, Q., Zhang, H., Xu, K., Cao, J., VGMF, Visual contents and geographical influence enhanced point-of-interest recommendation in location based social network. Transaction on Emerging Telecommunications Technologies, 33(6), (2020)
2. Alshammari, G., Kapetanakis, S., Polatidis, N., Petridis, M.: A triangle multi-level item-based collaborative filtering method that improves recommendations. International conference on engineering applications of neural networks, pp. 145–157. Springer (2018)

3. Jiao X., Xiao Y., Zheng W., Wang H., Hsu C. H., A novel next new point-of-interest recommendation system based on simulated user travel decision-making process. Future Generation Computer Systems, 100: 982-993, (2019).
4. Vedavathi, N., and K. M. A. Kumar. SentiWordNet Ontology and Deep Neural Network Based Collaborative Filtering Technique for Course Recommendation in an E-Learning Platform. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (2022).
5. Song, C., Wen, J., Li, S.: Personalized poi recommendation based on check-in data and geographical-regional influence. In: Proceedings of the 3rd International Conference on Machine Learning and Soft Computing, pp. 128–133 (2019)
6. Ye M., Yin P., Lee W.C., Lee D.L., Exploiting geographical influence for collaborative point-of-interest recommendation, in: International Conference on Research and Development in Information Retrieval, pp. 325–334, (2011).
7. Jiao X., Xiao Y., Zheng W., Wang Y., C., Hsu C, H., A novel next new point-of-interest recommendation system based on simulated user travel decision-making process. Future Generation Computer Systems, 100: 982-993, (2019).
8. Li, M., Zheng, W., Xiao, Y., Jiao, X.: An adaptive poi recommendation algorithm by integrating user's temporal and spatial features in LBSNs. In: Proceedings of the 3rd International Conference on Data Science and Information Technology, pp. 135–139 (2020)
9. Jenatton, R., Roux, N. L., Bordes, A., Obozinski, G., A latent factor model for highly multi relational data. Advances in Neural Information Processing Systems. (2012).
10. Yang, D., Zhang, D., Zheng, V. W., Yu, Z., Modeling User Activity Preference by Leveraging User Spatial Temporal Characteristics in LBSNs. IEEE Transactions on Systems Man & Cybernetics Systems 45.1(2014):129-142.
11. Yuan Q., Cong G., Ma Z., Sun A., Thalmann N.M., Time-aware point-of-interest recommendation, in: International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 363–372, (2013).
12. B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, A general geographicalprobabilisticfactormodelforpointofinterestrecommendation, IEEE Trans. Knowl. Data Eng., vol. 27, no. 5, pp. 1167–1179, (2015).
13. Xu J., Xiao Y., Zheng W.,Xu L., Exploring Spatial and Mobility Pattern's Effects for Collaborative Point-of-Interest Recommendation. IEEE Access, 7: 158917-158930, (2019).
14. Liu, K., Zheng, W., Xiao, Y., Zhai, X., POI recommendation algorithm based on region transfer collaborative filtering. In IEEE 25th International Conference on Computer Supported Cooperative Work in Design (2022).