

Geo-Pairwise Ranking Matrix Factorization Model for Point-of-Interest Recommendation

Shenglin Zhao^{1,2}(✉), Irwin King^{1,2}, and Michael R. Lyu^{1,2}

¹ Shenzhen Key Laboratory of Rich Media Big Data Analytics and Application, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China
{slzhao,king,lyu}@cse.cuhk.edu.hk

² Department of Computer Science and Engineering,
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Abstract. Point-of-interest (POI) recommendation that suggests new locations for people to visit is an important application in location-based social networks (LBSNs). Compared with traditional recommendation problems, e.g., movie recommendation, geographical influence is a special feature that plays an important role in recommending POIs. Various methods that incorporate geographical influence into collaborative filtering techniques have recently been proposed for POI recommendation. However, previous geographical models have struggled with a problem of *geographically noisy POIs*, defined as POIs that follow the geographical influence but do not satisfy users' preferences. We observe that users in the same geographical region share many POIs, and thus we propose the *co-geographical influence* to filter *geographically noisy POIs*. Furthermore, we propose the *Geo-Pairwise Ranking Matrix Factorization (Geo-PRMF)* model for POI recommendation, which incorporates *co-geographical influence* into a personalized pairwise preference ranking matrix factorization model. We conduct experiments on two real-life datasets, i.e., Foursquare and Gowalla, and the experimental results reveal that the proposed approach outperforms state-of-the-art models.

Keywords: POI Recommendation · Matrix factorization · Geographical influence · Pairwise ranking

1 Introduction

Point-of-interest (POI) recommendation has been being driven by soaring development of location-based social network (LBSN) services such as Foursquare and Facebook Places. A typical LBSN allows users to check-in at their locations, make friends, and share information. POI recommendation in LBSNs aims to help users explore new and interesting places in a city through an LBSN service. When you go shopping, for instance, you can easily find detailed downtown shopping mall information and nearby food shops using POI recommendation; and doing so not only improves users' experiences, but also provides merchants with new chances to target customers.

Due to the importance of POI recommendation, various methods have been proposed to tackle this task [1, 8, 11–13, 18, 20]. Inspired by the conventional recommendation systems, e.g., Netflix’s movie recommendation system, a user-POI matrix is constructed that treats POIs as items and users’ check-in frequencies as rating values. Then collaborative filtering techniques are used to recommend POIs. In addition, geographical influence has been incorporated as an important factor into the proposed POI recommendation systems to improve performance [1, 2, 11, 16, 19]. However, previous models designed to capture geographical influence have struggled with the problem of *geographically noisy POIs*.

Existing geographical influence models suffer from the problem of *geographically noisy POIs*, as they recommend new POIs that are close to those where the user has checked-in, depending solely on the user-POI geographical relationship. Here, we give an example of a *geographically noisy POI*. Suppose a user likes to visit shops and restaurants near his/her home, and as such generates many check-ins at these places. Meanwhile, a hotel is also located near the user’s home. According to previous geographical influence models, the hotel should be recommended as it is near the POIs where the user has checked-in. However, people live in their own houses, and do not typically want to visit a hotel nearby. Hence, the hotel is defined as a *geographically noisy POI*, which follows the geographical influence but does not satisfy the user’s preference.

In this paper, we propose the *co-geographical influence* to address the problem of *geographically noisy POIs*. We observe that users acting in the same region share many POIs. Two students attending the same university, for example, may not know each other, but may check into many of the same POIs, such as popular restaurants and night clubs around the university. Each user’s check-in behavior enhances each shop’s popularity, attracting more people. Inspired by this observation, we propose the *co-geographical influence*, which assumes that users follow similar visiting patterns in close areas.

Furthermore, we propose the *Geo-Pairwise Ranking Matrix Factorization (Geo-PRMF)* model to tackle the POI recommendation problem. Inspired by [19, 20], we treat users’ check-ins as implicit feedback and learn the system via personalized pairwise preference ranking. The preference is implicitly embedded in pairs (checked-in, unchecked-in), with users assumed to have stronger interest in the checked-in POIs than in the unchecked-in POIs. We exploit the *co-geographical influence* to refine the preference pair set, which reduces the complexity cost. Specifically, our model filters the *geographically noisy POIs*, which are unresolved in existing geographical influence models [1, 11, 16].

The contributions of this paper are summarized as follows. First, we propose the *co-geographical influence* to overcome the problem of *geographically noisy POIs* hindering previous geographical influence models. Moreover, we propose the *Geo-PRMF* model, which incorporates *co-geographical influence* into a personalized pairwise preference ranking model to learn user preference and performs better than state-of-the-art models.

2 Related Work

In this section, we first demonstrate the recent progress of POI recommendation. Then, we show how previous studies have modeled geographical influence. Finally, we explore how our proposed model relates to the prior work.

POI Recommendation. POI recommendation has attracted intensive academic attention recently. Most of the proposed methods have used collaborative filtering (CF) techniques, including the memory and model-based methods, to recommend POIs. The researchers in [11, 14, 15] employ the user-based CF to recommend POIs, whereas, other studies leverage the model-based CF, including the Matrix Factorization (MF) technique [1, 8, 9, 18]. Specifically, the researchers in [8, 9] model the check-ins as implicit feedback and use the weighted regularized MF for POI recommendation. Unlike the researchers in [8, 9], those in [3, 7, 19, 20] model implicit feedback via a pairwise ranking method, which exhibits better performance.

Geographical Influence. Geographical influence plays an important role in POI recommendation, as users' activity in LBSNs is limited by geographical constraints. To capture geographical influence, the researchers assume that the co-occurrence of POIs follows a specific distribution. On the one hand, studies in [1, 4, 16] suppose the checked-in POIs follow a Gaussian distribution and propose Gaussian distribution based models; those in [11, 14] employ the power law distribution model; and studies in [15] leverage the kernel density estimation model to learn the distribution. On the other hand, the researchers in [8, 9] incorporate geographical influence into a weighted regularized MF model. The work in [19] incorporate the geographical influence into a ranking model and propose a hierarchical geographical pairwise ranking for POI recommendation. The core idea of the proposed geographical influence models has based on the intuition that a user prefers the visit new POIs nearby where the user has checked-in.

Connection to Prior Work. Prior studies have captured the geographical influence to recommend new POIs, prioritizing proximity to the user's activity center or previous checked-in POIs. This creates the problem of *geographically noisy POIs*. We propose the *co-geographical influence* to overcome this problem. Moreover, due to the success of using pairwise preference ranking to model the check-in activity, we propose the *Geo-PRMF* model, which incorporates *co-geographical influence* into a pairwise preference ranking model to learn users' POI preferences.

3 Model

In this section, we first propose *co-geographical influence* to address the problem of *geographically noisy POIs*. Then, we propose the *Geo-PRMF* model, which incorporates *co-geographical influence* into a pairwise preference ranking model for recommending POIs.

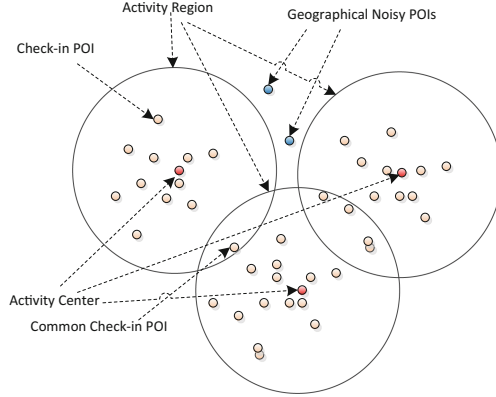


Fig. 1. Demonstration of user check-in pattern

3.1 Co-geographical Influence

For illustration purposes, we define several terms as follows.

Definition 1 (Geographical activity center) A geographical activity center is the POI with the highest check-in probability based on geographical influence.

Definition 2 (Geographical neighbors). Geographical neighbors are users who have close geographical activity centers.

Definition 3 (Geographically noisy POI). A geographically noisy POI is the POI near a user's geographical activity center but not preferred by the user.

Figure 1 demonstrates the user check-in pattern and the problem of *geographically noisy POIs*. Previous studies [1, 16] have shown that most people live and have fun in constrained activity regions. According to this kind of geographical characteristic, previous work constructs the user-POI geographical relation: a POI that is near a user's *geographical activity center* is geographically preferred [1, 16]. However, this assumption is easily affected by *geographically noisy POIs*, as shown in Fig. 1. Some POIs are geographically near a user's *geographical activity center* but they do not match the user's check-in pattern, such as the hotel example mentioned in Sect. 1.

Co-geographical influence depicts the user-user geographical relation instead of the user-POI relation. We observe that *geographical neighbors* share many POIs. Specifically, the Jaccard similarity between *geographical neighbors* is about 10 times higher than between random users. The model not only considers a user's geographical feature but also extracts geographical relation between two users. We follow the discovery that a user's checked-in POIs distribute around some activity center(s) [1, 16]. Hence, we expect the POIs in which a user is interested to be located in the range where the user's *geographical neighbors* have checked-in. This helps to filter out the *geographically noisy POIs*. As a result,

the candidate POI set for a user consists of POIs where the user’s *geographical neighbors* have checked-in but he/she has not yet. *Co-geographical influence* exploits the common check-in pattern among *geographical neighbors* to filter out *geographically noisy POIs*.

3.2 Geo-Pairwise Ranking Matrix Factorization (Geo-PRMF) Model

We propose the *Geo-PRMF* model, which incorporates *co-geographical influence* into a pairwise ranking model. Due to the success of pairwise preference ranking in modeling the check-in activity as implicit feedback in prior work [7, 18, 20], we utilize the Bayesian personalized ranking criteria [10] to learn user preference on POIs. Moreover, we exploit *co-geographical influence* to classify the unrated POIs as comparable POIs and unrelated POIs. We assume that the POIs where a user’s *geographical neighbors* have checked-in are comparable and others are unrelated. Therefore, we only make use of the comparable POIs to generate the pairwise preference set, and discard the unrelated ones, recommending POIs from the comparable POI candidate set. Based on this assumption, we extract the refined pairwise preference set and candidate POI set as follows:

1. We map a *geographical activity center* for a user and identify the top k *geographical neighbors* by nearby centers.
2. We consider only the POIs checked-in to by the user’s *geographical neighbors* but not yet by the user to be comparable and any others to be unrelated.
3. We generate triplets (user u , checked-in POI l_i , comparable POI l_j) as refined preference set \mathcal{P} , and comparable POIs as candidate set \mathcal{L}_u^c .

Then, we can learn the user’s preferences from the refined pairwise preference set and recommend POIs from the candidate POI set.

We formulate the POI recommendation problem as follows. Let \mathcal{U} be the set of users and \mathcal{L} be the set of POIs. The pairwise preference of user u prefers POI l_i over l_j , is defined as $l_i \succ_u l_j$. Then, we define the pairwise preference set $\mathcal{P} := \{l_i \succ_u l_j | l_i \in \mathcal{L}_u^+ \wedge l_j \in \mathcal{L}_u^c\}$, where \mathcal{L}_u^+ denotes the POIs where user u has checked-in, and \mathcal{L}_u^c denotes the POIs where *geographical neighbors* of user u have checked-in but u has not. Now training the POI recommendation system is to learn the pairwise preference relationships in \mathcal{P} ,

$$\arg \max_{\Theta} \prod_{(u, l_i, l_j) \in \mathcal{P}} p(l_i \succ_u l_j | \Theta), \quad (1)$$

where $p(l_i \succ_u l_j)$ is the probability of a user preferring POI l_i over l_j , and Θ denotes the model’s learning parameters.

We employ the biased MF to model the user preference on POI. Then, the preference score function of user u on POI l_i is formulated as,

$$f(u, l_i) = U_u^T L_{l_i} + b_{l_i}, \quad (2)$$

where $U_u, L_{l_i} \in R^d$ are latent feature vectors for user u and POI l_i respectively, and b_{l_i} is the estimation bias. Furthermore, we estimate the probability function of $p(l_i \succ_u l_j)$ via a sigmoid function, $p(l_i \succ_u l_j) = \sigma(f(u, l_i) - f(u, l_j))$, where σ is the sigmoid function $\sigma(x) = 1/(1 + \exp(-x))$. Thus, it is not hard to gain the objective function by minimizing the negative log likelihood

$$\begin{aligned} & \mathcal{O}(b_{l_i}, b_{l_j}, L_{l_i}, L_{l_j}, U_u) \\ &= - \sum_{u=1}^{|\mathcal{U}|} \sum_{l_i \in \mathcal{L}_u^+} \sum_{l_j \in \mathcal{L}_u^c} \ln \sigma(U_u^T(L_{l_i} - L_{l_j}) + b_{l_i} - b_{l_j}) + \frac{\lambda_1}{2} \|U_u\|^2 \\ & \quad + \frac{\lambda_1}{2} \|L_{l_i}\|^2 + \frac{\lambda_1}{2} \|L_{l_j}\|^2 + \frac{\lambda_2}{2} \|b_{l_i}\|^2 + \frac{\lambda_2}{2} \|b_{l_j}\|^2, \end{aligned} \quad (3)$$

where λ_1 and λ_2 are the regularization parameters.

We adopt the stochastic gradient decent (SGD) method to learn the parameters in Eq. (3). We define a common expression as $z = \frac{1}{1 + \exp(U_u^T(L_{l_i} - L_{l_j}) + b_{l_i} - b_{l_j})}$. Then, the parameters are updated as follows,

$$\begin{aligned} b_{l_i} &\leftarrow b_{l_i} + \gamma \cdot (z - \lambda_1 \cdot b_{l_i}), \\ b_{l_j} &\leftarrow b_{l_j} + \gamma \cdot (-z - \lambda_2 \cdot b_{l_j}), \\ L_{l_i} &\leftarrow L_{l_i} + \gamma \cdot (z \cdot U_u - \beta_1 \cdot L_{l_i}), \\ L_{l_j} &\leftarrow L_{l_j} + \gamma \cdot (-z \cdot U_u - \beta_2 \cdot L_{l_j}), \\ U_u &\leftarrow U_u + \gamma \cdot (z \cdot L_{l_i} - z \cdot L_{l_j} - \alpha \cdot U_u). \end{aligned} \quad (4)$$

After learning the parameters, the *Geo-PRMF* model predicts a user's check-in preference at a given POI according to the score computed by Eq. (2). We first rank the POIs in candidate set in terms of check-in preference, then recommend the top N POIs for a specific user. Algorithm 1 demonstrates how to recommend POIs through *Geo-PRMF* model.

Complexity Analysis. There are two steps to recommend POIs: model training and item recommendation. The complexity of training *Geo-PRMF* model is $O(d \cdot |S|)$, the same order as the BPR-MF model [10], where d denotes the latent factor vector dimensionality and $|S|$ denotes the number of samples. *Geo-PRMF* has an advantage over other models at the item recommendation step. For general MF-based recommendation models, the time complexity of the item recommendation step is $O(|\mathcal{U}| \cdot |\mathcal{L}| \cdot d)$. The item recommendation time complexity of the *Geo-PRMF* model is $O(|\mathcal{U}| \cdot |\mathcal{L}^c| \cdot d)$ with $|\mathcal{L}^c|$ denoting the average number of candidate POIs for a user. As $|\mathcal{L}^c|$ is much less than $|\mathcal{L}|$, the *Geo-PRMF* consumes less calculation than other models at the item recommendation step.

Algorithm 1. POI Recommendation via Geo-PRMF Model

Input: Preference set \mathcal{P} , learning parameters $b_{l_i}, b_{l_j}, L_{l_i}, L_{l_j}, U_u$, regularization parameters λ_1, λ_2

Output: recommended POI set S_N for each user

- 1: Randomly initialize parameters $b_{l_i}, b_{l_j}, L_{l_i}, L_{l_j}, U_u$
- 2: **repeat**
- 3: Draw $(u, l_i, l_j) \in \mathcal{P}$ uniformly
- 4: Update parameters with Eq. (4)
- 5: **until** convergence
- 6: **for** u in \mathcal{U} **do**
- 7: **for** l in \mathcal{L}_u^c **do**
- 8: Predict user preference score with Eq. (2)
- 9: Recommend N POIs S_N with higher preference score

4 Experiment

4.1 Data Description and Experimental Setting

Two real-world datasets are used in the experiment: Foursquare data in [5] and Gowalla data in [4]. We extract the data from March to October in 2010 from both datasets, filter the POIs checked-in to by less than 5 users, and then choose users who have checked-in more than 10 times as our samples. Table 1 shows the data statistics. We randomly choose 80% of each user’s check-ins as training data, and use the remaining 20% for test data. Following [8, 20], we use **precision** and **recall** to measure the model performance.

Table 1. Data statistics

Source	#users	#POIs	#check-ins	Density
Foursquare	18,106	17,145	1,330,247	0.001
Gowalla	107,091	144,034	2,581,093	0.0001

4.2 Baseline Methods

Given that the proposed method aims to construct an effective MF-based model for POI recommendation, we select BiasedMF [6] and BPR-MF [10] as the basic comparable models. Moreover, to show the advantage of our proposed model in capturing geographical influence, we compare it with fused MF with multi-center Gaussian model (MGMMF) [1] and joint model with geographical influence and MF (GeoMF) [8], which are state-of-the-art POI recommendation methods capturing geographical influence.

4.3 Experimental Results

In the following, we demonstrate the performance comparison on precision@ N and recall@ N between the baseline models and our proposed *Geo-PRMF* model. We set the latent factor vector dimension as 20 for all compared models.

We evaluate different models for both datasets on top-5 and top-10 POI recommendation tasks. Figure 2 shows the obtained results, from which we make the following observations. (1) The proposed *Geo-PRMF* model achieves the best performance, with advantages over the MGMMF and the GeoMF at capturing geographical influence by filtering out *geographically noisy POIs*. Compared with the best baseline competitor, the *Geo-PRMF* model achieves at least 5% improvements on precision@5 and recall@5, and at least 7% improvements on precision@10 and recall@10 for both datasets. (2) *Geo-PRMF*, MGMMF, and GeoMF perform better than BiasedMF and BPR-MF, which demonstrates the effectiveness of capturing geographical influence.

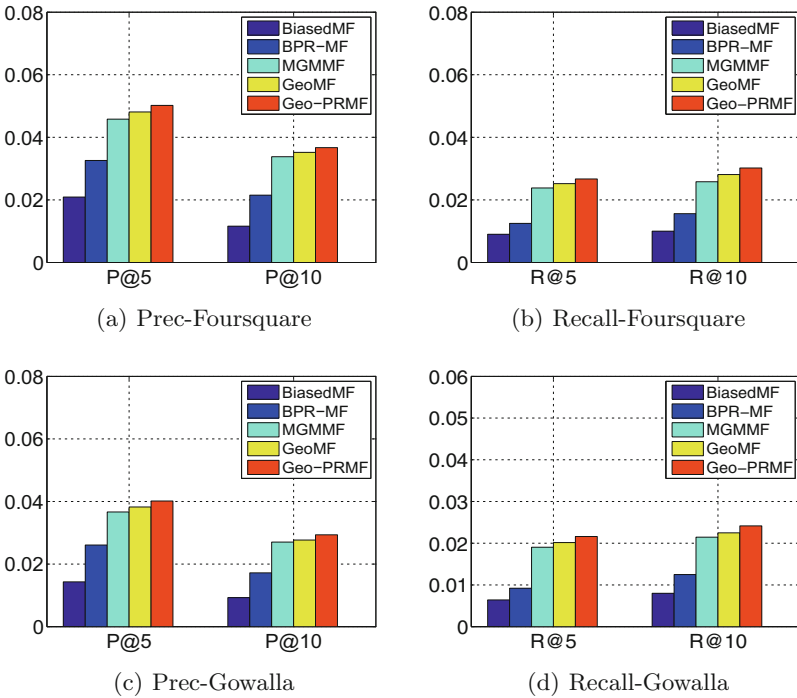


Fig. 2. Model comparison

5 Conclusion and Future Work

In this paper, we propose the *Geo-PRMF* model to tackle the POI recommendation problem. We first present *co-geographical influence*, which reduces

geographically noisy POIs and significantly shrinks the candidate set for a specific user. Moreover, we propose the *Geo-PRMF* model, which incorporates *co-geographical influence* into a pairwise ranking model. Finally, we conduct elaborated experiments on two real-life LBSN datasets to verify our proposed model. The experimental results show that our proposed *Geo-PRMF* model outperforms state-of-the-art models.

In the future, we will improve the *Geo-PRMF* model in the following aspects. We may design an adaptive way to select the number of activity centers to improve the performance. Furthermore, we may consider users' comments or location category features to further improve the overall recommendation performance. In addition, a new application in LBSNs [17] has appeared recently, which uses the check-in data to mine business opportunities. We will consider to exploit the check-in characteristics to enhance the business mining application.

Acknowledgments. The work described in this paper was partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Nos. CUHK 14203314 and CUHK 14234416 of the General Research Fund), and 2015 Microsoft Research Asia Collaborative Research Program (Project No. FY16-RES-THEME-005).

References

1. Cheng, C., Yang, H., King, I., Lyu, M.R.: Fused matrix factorization with geographical and social influence in location-based social networks. In: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, pp. 17–23. AAAI Press (2012)
2. Cheng, C., Yang, H., King, I., Lyu, M.R.: A unified point-of-interest recommendation framework in location-based social networks. *ACM Trans. Intell. Syst. Technol. (TIST)* **8**(1), 10 (2016)
3. Cheng, C., Yang, H., Lyu, M.R., King, I.: Where you like to go next: successive point-of-interest recommendation. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, pp. 2605–2611. AAAI Press (2013)
4. Cho, E., Myers, S.A., Leskovec, J.: Friendship and mobility: user movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1082–1090. ACM (2011)
5. Gao, H., Tang, J., Liu, H.: Exploring social-historical ties on location-based social networks. In: Sixth International AAAI Conference on Weblogs and Social Media. AAAI (2012)
6. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. *Computer* **42**(8), 30–37 (2009)
7. Li, X., Cong, G., Li, X.L., Pham, T.A.N., Krishnaswamy, S.: Rank-GeoFM: a ranking based geographical factorization method for point of interest recommendation. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 433–442. ACM (2015)

8. Lian, D., Zhao, C., Xie, X., Sun, G., Chen, E., Rui, Y.: GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 831–840. ACM (2014)
9. Liu, Y., Wei, W., Sun, A., Miao, C.: Exploiting geographical neighborhood characteristics for location recommendation. In: ACM International Conference on Conference on Information and Knowledge Management, pp. 739–748 (2014)
10. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: BPR: Bayesian personalized ranking from implicit feedback. In: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, pp. 452–461. AUAI Press (2009)
11. Ye, M., Yin, P., Lee, W.C., Lee, D.L.: Exploiting geographical influence for collaborative point-of-interest recommendation. In: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 325–334. ACM (2011)
12. Yin, H., Cui, B., Sun, Y., Hu, Z., Chen, L.: LCARS: a spatial item recommender system. *ACM Trans. Inf. Syst. (TOIS)* **32**(3), 11 (2014)
13. Yin, H., Cui, B., Zhou, X., Wang, W., Huang, Z., Sadiq, S.: Joint modeling of user check-in behaviors for real-time point-of-interest recommendation. *ACM Trans. Inf. Syst. (TOIS)* **35**(2), 11 (2016)
14. Yuan, Q., Cong, G., Ma, Z., Sun, A., Thalmann, N.M.: Time-aware point-of-interest recommendation. In: Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 363–372. ACM (2013)
15. Zhang, J.D., Chow, C.Y.: GeoSoCa: exploiting geographical, social and categorical correlations for point-of-interest recommendations. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 443–452. ACM (2015)
16. Zhao, S., King, I., Lyu, M.R.: Capturing geographical influence in POI recommendations. In: Lee, M., Hirose, A., Hou, Z.-G., Kil, R.M. (eds.) *ICONIP 2013 Part II*. LNCS, vol. 8227, pp. 530–537. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-42042-9_66](https://doi.org/10.1007/978-3-642-42042-9_66)
17. Zhao, S., King, I., Lyu, M.R., Zeng, J., Yuan, M.: Mining business opportunities from location-based social networks. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1037–1040. ACM (2017)
18. Zhao, S., Lyu, M.R., King, I.: Aggregated temporal tensor factorization model for point-of-interest recommendation. In: Hirose, A., Ozawa, S., Doya, K., Ikeda, K., Lee, M., Liu, D. (eds.) *ICONIP 2016 Part III*. LNCS, vol. 9949, pp. 450–458. Springer, Cham (2016). doi:[10.1007/978-3-319-46675-0_49](https://doi.org/10.1007/978-3-319-46675-0_49)
19. Zhao, S., Zhao, T., King, I., Lyu, M.R.: Geo-Teaser: geo-temporal sequential embedding rank for point-of-interest recommendation. In: Proceedings of the 26th International Conference on World Wide Web Companion, pp. 153–162. International World Wide Web Conferences Steering Committee (2017)
20. Zhao, S., Zhao, T., Yang, H., Lyu, M.R., King, I.: STELLAR: spatial-temporal latent ranking for successive point-of-interest recommendation. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pp. 315–321. AAAI Press (2016)