IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 27, 2020, accepted March 8, 2020, date of publication March 10, 2020, date of current version March 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2979922

Points-of-Interest Recommendation Algorithm Based on LBSN in Edge Computing Environment

KEYAN CAO [1,2], JINGJING GUO [1], GONGJIE MENG [1], HAOLI LIU [1], YEFAN LIU [1], AND GUI LI [1]
[1]College of Information and Control Engineering, Shenyang Jianzhu University, Shenyang 110168, China
[2]Liaoning Province Big Data Management and Analysis Laboratory of Urban Construction, Shenyang 110168, China

Corresponding author: Keyan Cao (caokeyan@sjzu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602323, in part by the National Postdoctoral Foundation of China under Grant 2016M591455, in part by the Youth Seedling Foundation of Liaoning Province under Grant lnqn201913, and in part by the Natural Science Funds of Liaoning Province under Grant 2019MS264.

ABSTRACT With the advancement of the Internet of Everything era and the popularity of mobile devices, Location-based Social Networks (LBSN) have penetrated people's lives. People can take advantage of portable edge terminal devices and use the geographic information in LBSN to arrange or adjust their travel plans. However, due to the explosive growth of current Internet applications and users, it has brought greater pressure and operation and maintenance costs to cloud storage. It is a key research direction based on location recommendation to accurately obtain the places of interest of users and push them to clients in such a large amount of original data. In order to better process the data generated by edge devices, this paper firstly uses the Rank-FBPR matrix decomposition framework based on social network to analyze the user's personal preference function on the edge server. Then interact with the geographic information stored in the Cloud to cluster the POIs. And embeds the geographic information into the framework to get the candidate points of interest. Finally, the scores of candidate points of interest are predicted using the personal preference function and power law distribution, then a sorted list of points of interest is generated in descending order of scores, and the list is recommended to the target user. This algorithm effectively integrates the time information and geographic information of users' check-in in the LBSN, and proposes a POIs recommendation algorithm that comprehensively considers edge devices and Cloud. The experiments verify the effectiveness of framework from both cold start and non-cold start. The experimental results on the Foursquare and the Yelp datasets show that Rank-FBPR has higher recommendation accuracy and recall than other comparison models, and can adapt to cold start problems.

INDEX TERMS LBSN, edge computing, personal preference.

I. INTRODUCTION

With the rapid development of Internet technology and the continuous popularization of mobile communication devices, Location-based Social Networks (LBSN) have penetrated people's lives. People can use the portable terminal to access the Internet, and use the geographic information and social attributes in LBSN to define the geographic location preferences of users to access points of interest. Users can arrange or adjust work and travel plans in time to achieve the effect of intelligent perception and convenient use of all kinds of information. At the same time, the rapid arrival of the era of Cloud computing, big data and IoT has led to the explosive growth trend of network edge devices (such as smart phones, wearable smart devices, etc.) in the past decade, and the increasing demand of mobile users for matching service quality. In addition, the high computing power and accuracy required for points-of-interest recommendation are increasingly not guaranteed. In this case, the centralized processing mode is unable to process the data generated by the edge devices, so the edge calculation comes into being. The edge in edge calculation refers to the calculation and storage of network edge [1], which is opposite to the data center and closer to the user in terms of network distance or geographical distance [2]. In the edge computing model, some or all computing tasks from the original Cloud center are migrated

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaofei Wang.

to network edge devices. The edge computing model not only has lower communication cost, but also can reduce data transmission bandwidth to improve data transmission performance and ensure real-time processing. Meanwhile, it can better protect private data and reduce the risk of privacy leakage of terminal sensitive data.

The improvement of edge computing capability makes edge intelligence more and more common. The survey conducted by Wang *et al.* [3] shows the importance and practicability of the application of edge intelligence in life. Wang *et al.* [4] put forward the concept of edge artificial intelligence by combining deep learning technology and joint learning with mobile edge system, which is very forward-looking. Li *et al.* [5] uses a multiple Deep Reinforcement Learning (DRL) agents on the internet of things device to guide the calculation transfer decision, and uses Federated Learning (FL) to carry out distributed training on DRL agents to further reduce the transmission cost between the internet of things device and edge nodes.

In the research of Points-of-Interest(POI) recommendation, more and more researches are focused on POIs recommendation of LBSN service. Integrate edge computing in LBSN location service to build a points-of-interest access model for POIs recommendation is a measure to achieve efficient points-of-interest guidance service, and also one of the new directions of the development of POIs recommendation system. In LBSN, users can check-in, release their geographic location information, and share their experiences and comments about the check-in of the points-of-interest. Each website and platform providing location-based services (such as Foursquare, Yelp, etc.) collects user check-in information and analyzes it, and then provides location-based POIs recommendation services for users [6]. In recent years, most of the research results of POIs recommendation are based on multi-source heterogeneous information and user's check-in data in LBSN to recommend points of interest that meet the needs of users. For example, Ye *et al.* [7] used the user's personal preferences, social relationships, and geographic influences to make points of interest recommendations, and used a power-law distribution to model clustering in space. Zhang *et al.* [8] first used the Kernel Density Estimation (KDE) method to simulate the user's access behaviors, and then made points of interest recommendations. Koren *et al.* [9] decomposed the user's points-of-interest matrix into a user matrix and a points-of-interest matrix, and used Matrix Factorization (MF) to model the potential characteristics of each user and points-of-interest and predict the user's score on the POIs. Finally, a list of points of interest is recommended to users based on the rating. Lian *et al.* [10] introduced the Weighted Matrix Factorization (WMF) method as a basic framework to assign high weight to activities with high user participation in order to adapt to user behaviors. Finally, a method of integrating geographic factors, which is called factor augmentation model was proposed.

Although existing POIs recommendation algorithms can recommend a set of places with high similarity to users on the premise of considering the user's interest preferences, these recommendation results have good accuracy, but most methods only use a single contextual information to build model, ignoring the diversity of recommendation results, while facing the problem of sparse data, making it difficult to recommend a personalized set of points of interest to target users based on their real-time geographic location. In the past, the recommended system or model was to trigger the Cloud server through the client, and then the server would respond to the user's request. Such a mode will lead to the delay in the real-time perception of user behaviors, resulting in the butterfly effect, which makes the Cloud server unable to timely adjust the recommendation results to perceive the changes in user preferences, and problems such as a decrease in user dependence on the client. Based on this, combining the characteristics of edge computing, this paper proposes a novel POIs recommendation algorithm by embedding geographic information on the basis of Bayesian Personalized Ranking(BPR) method that integrates social information. Based on the framework of hybrid Cloud and edge computing, the algorithm aims to extend Cloud functions to edge servers with computing power to obtain lower processing delay and real-time feedback. Firstly, the computing task of terminal data is completed on the edge server, and the user's personal preference is inferred by using the Bayesian ranking method and integrating into the user's social relations. Then, the information of urban areas is processed in the Cloud, and the edge server clustering points-of-interest according to geographic information. Finally the algorithm is generated. This paper has three main contributions as follows:

(1) We propose a BPR learning model based on social information on edge servers. The POIs recommendation is regarded as a ranking problem, and the user's personal preference and social relationship are used to evaluate the points-of-interest.

(2) We integrate the geographical information of users into the BPR framework of social relations(FBPR), process the information of urban areas in the Cloud, and combine the framework with the information of points-of-interest clusters after clustering to form the recommendation list of POIs with both individuation and diversity.

(3) We conduct a large number of experiments on Foursquare and Yelp. The validity and superiority of the proposed method are verified by comparing the accuracy and recall of the experimental results of this algorithm with existing algorithms.

II. RELATED WORK

Different from the traditional recommendation method, the POIs recommendation needs to consider the influence of various factors such as social information, time information and spatial information, which makes the points-of-interest recommendation more complicated. The algorithm proposed in this paper mainly involves two aspects of points-of-interest clustering and POIs recommendation. The research status of these two aspects is introduced below.

A. CLUSTERING OF THE POINTS-OF-INTEREST

Clustering is a common method of data mining, data representation and data visualization. Classical clustering methods include partition-based clustering, density-based clustering and grid-based clustering. The typical clustering methods are K-Means clustering and K-MEDOIDS clustering. As a heuristic algorithm based on distance, these two algorithms avoid the disadvantage of exhausting all the partitions based on clustering, but they are not suitable for large-scale POIs clustering. DBSCAN is the most widely used algorithm based on density. The guiding principle of this method is to cluster according to the density of the points of the region. It overcomes the shortcoming that only regular shape clustering can be found based on distance algorithm, but only regions with large density in geographical space can be found, which is inconsistent with the clustering demand results of the points of interest. Grid-based clustering first divides the data space into a finite number of grid structures, and then uses a single grid as the object for processing. This kind of algorithm has high processing speed, but it needs a lot of resources to divide the grid, and its real-time performance is poor. As a new developing clustering method, spectral clustering is more efficient than the traditional K-Means algorithm, with more uniform clustering effect and less computation.

Zhong *et al.* [11] proposed a new method of geo-spatial data clustering called Multi-Reference Clustering (MRC). This method adds the concept of reference points to K-Means clustering to transform geo-spatial data. Data points are grouped into K clusters, and the local search approximation algorithm greatly reduces the time complexity of MRC, but because this method is more complicated, it is not suitable for clustering of POIs recommendations. Shi *et al.* [12] defined a new clustering method DCPGS (Density-based Clustering Places in Geo-Social Networks, DCPGS) on the Geo-Social Network (GeoSN) model. Based on the consideration of the spatial distance and social distance between regions, the clustering results are more comprehensive and more effective. However, there are two side effects at the same time: two users who have no social connection at all but are very close in geographical location may be clustered, or two users who are close in social distance but are geographically far away may be clustered.

B. RECOMMENDATION OF THE POINTS-OF-INTEREST

The POIs recommendation system is essentially to solve the problem of information overload and long tail effect. By mining the attributes and behaviors of users, the points of interest is accurately and efficiently recommended to target users who are interested in it. According to the research needs of this paper, the POIs recommendation systems are roughly divided into three categories: contextual information recommendation, implicit feedback recommendation, and embedded recommendation.

In the LBSN, the contextual information includes social information, geographic location information, and time information, etc. It is found that users are more inclined to visit places suggested by friends and the points of interest close to their own location. Ye *et al.* [7] integrated geographical influence, user preference and social influence into collaborative filtering recommendation, and designed a check-in probability prediction model for a given user's access to points of interest. Experimental results show that this prediction model is superior to traditional collaborative filtering recommendation methods. The random walk is not suitable for POIs recommendation due to some obviously different interest preferences among friends. Qian *et al.* [13] integrated temporal and geographic information, and proposed a Translation-based, Time and Location aware (TransTL) representation method that can successfully respond to real-time POIs recommendations, and can solve the problem of data sparsity.

Location selection is crucial in geo-social network, Zhong *et al.* [14] proposed the problem of sample location selection to maximize the influence of distance perception in the geo-social network. Transforming the query location from the whole two-dimensional space to a particular query area can simplify the query problem and improve the tightness of the target distance constraint. Li *et al.* [15] put forward the problem of maximum geographic spanning regions over location-aware social networks and found that social influence may improve the accuracy of location selection or recommendation. Haldar *et al.* [16] compared eight representative prediction models and tested them on four real-world large-scale datasets using five metrics. The location prediction framework they proposed can comprehensively evaluate the location prediction model. Analysis shows that the effectiveness of user location prediction depends on the richness of its neighbor information.

The core data required for points-of-interest recommendation is users' feedback information, which has two forms of explicit feedback and implicit feedback. If the recommendation is only based on the user's explicit feedback information such as ratings, the recommendation result may be too singular. Combining the recommendation with implicit feedback such as evaluation, browsing and collection can significantly improve the diversity of recommendation results. At present, most of the implicit feedback-oriented recommendation systems are modeled on matrix decomposition algorithms using weighting or sorting. Lian *et al.* [10] extended the matrix decomposition algorithm based on weighting, using active region vector and influence region vector to increase the potential feature factors of users and points of interest, the proposed Geography Matrix Factorization(GEOMF) model to combine geographic information and matrix decomposition, which improved the performance of recommendation. Rendle *et al.* [17] proposed a commonly used personalized ranking framework, which used BPR to

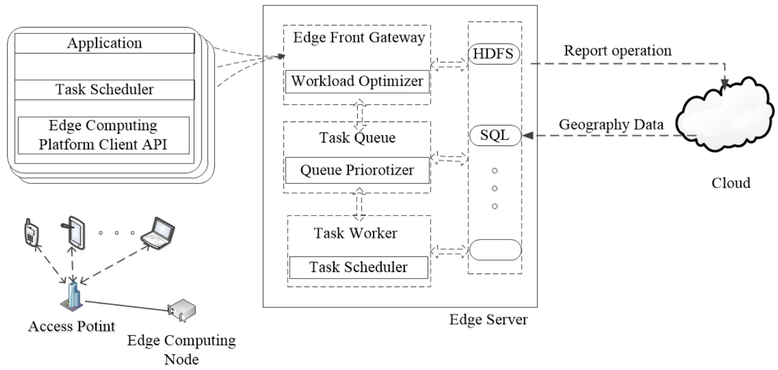


FIGURE 1. The architecture of Rank-FBPR algorithm.

analyze the maximum posterior estimator, so as to directly optimize the loss error of ranking. The prediction quality of the personalized ranking framework is not only determined by the model, but also affected by the optimization criteria.

With the development of distributed technology to the recommendation field, some researches show that using distributed embedded model can obtain more satisfactory recommendation results. Translation-based Recommendation (TranRec) constructed by He *et al.* [18] is a model based on graph embedding technology, which embeds items into a "transition space". Users are represented as relationship vectors on the sequence of points of interest to predict users' personalized sequential behavior. TranRec mainly focuses on the sequence effect between points of interest. If the position between POIs is too far or the last check-in time of users is too long, it is not meaningful to use this continuity effect. Guo *et al.* [19] first preprocessed the embedded model based on neural network to mine the deep information in social network and evaluation, and then embedded the pre-trained network into the weighted matrix decomposition method, linearly integrated the extracted factors and potential factors, and formed the collaborative filtering model which was considered more suitable for social recommendation.

C. OVERALL

The problem is divided into three steps in order to clearly describe the algorithm proposed, and the architecture of Rank-FBPR algorithm is shown in Figure 1.

Firstly, build a BPR framework based on social relationship on the edge servers. A Bayesian ranking learning model based on social relationships is used to extract user personal preferences based on the user's geographical location and social information, evaluate the matching degree of the target users' personal preferences and points of interest, and calculate and generate a BPR framework based on social relations.

Secondly, urban area information is processed in the Cloud, the Cloud interacts with the edge servers, which cluster POIs according to geographic information. The city blocks are divided into geographic units, and the Gaussian Mixture Model (GMM) is used to represent the distribution of points of interest. Chinese Restaurant Process(CRP) is used to cluster points of interest. Since the number of clusters generated by topic clustering captured by GMM may be different, it is important to find the appropriate number of clusters for each topic at this stage. To achieve this goal, CRP is used to generate clusters. Different from other clustering algorithms, CRP clustering does not need to specify the number of clusters in advance, but is automatically determined during the analysis process, and the inference algorithm is simple and easy to implement.

Finally, the personalized ranking list of points of interest is generated. Considering the personal preference and geographical location of the target user, the personal preference function and power-law distribution are used to predict the scores of candidate points, and the ranking list of points of interest is generated according to the descending order of scores.

III. BPR FRAMEWORK BASED ON SOCIAL RELATIONS

This section mainly introduces the BPR framework based on social relationships built on edge devices. The framework integrates user personal preferences and social relations by using the information of POIs location and user's social relations, and embeds Bayesian ranking into matrix decomposition to obtain a matrix decomposition framework that relaxes the BPR standard.

BPR is the target criterion based on pairwise ranking in the POIs recommendation method. Compared with traditional matrix decomposition, embedded BPR model is easier to mine the implicit feedback information of users, and matrix decomposition needs to calculate the global score and then optimize it. BPR model is optimized for the individual preferences of each user, reducing the calculation process of global score. In this paper, BPR is used to refine the user's personal preferences and distinguish between the negative samples (users have no interest in such points of interest) and the missing values (users do not find such points of interest). Bayesian model assumes that: (a) each user's preference behavior is independent of each other; (b) the partial order of the same user to different items is independent of each other.

In this paper, hypothesis (b) is relaxed, and it is considered that missing values are as important as positive samples, and user preferences are affected by users with social relations.

A. IMPACT OF SOCIAL RELATIONS

Cho *et al.* [20] found that about 38% of users didn't check-in by their friends; nearly 90% of users and their friends had less than 20% coincidence check-in rate. In this dataset, the overlap of two users with social relations is not obvious. Although the impact of social relations on users' attendance is limited, but not easy to be ignored.

B. CONSTRUCTION OF FBPR FRAMEWORK

The social relation is integrated into the matrix decomposition algorithm based on BPR to obtain the Friendship Bayesian Personalized Ranking(FBPR) framework.

BPR matrix decomposition model is based on the interest point set S to construct a training data set D_s. For user u, if there is a check-in behavior at the point of interest i, but no behavior at j, denoted as$(i, j \epsilon S)$. It is expressed by preference pair (u, i, j) or $i \succ_u j$. An $I \times I$ preference matrix is constructed for each user, and all user preference pairs constitute the training data set $D_s : U \times I \times I$, which is:

$$D_s = ((u, i, j)|u \in U, i, j \in S) \tag{1}$$

where U is the user set. BPR is based on the maximum posterior estimation method $P(W, H| \succ_u)$ to solve the model parameters W and H, and find the correct personalized ranking description for all points of interest $i\epsilon I$. Because the model assumes that the behaviors of all users are independent of each other, there are:

$$P(\Theta| \succ_u) \propto P(\succ_u |\Theta)P(\Theta) \tag{2}$$

where Θ is the parameter vector of the model.

The likelihood function of the user's partial order relations on the points of interest can be expressed as:

$$\begin{aligned} FBPR = &\prod_{u\epsilon U, i\epsilon I_u^+, k\epsilon I_u^*} P(r_{ui} \succ r_{uk}) \times 1 - P(r_{ui} \succ r_{uk}) \\ &\times \prod_{j\epsilon I_u^-, k\epsilon I_u^*} P(r_{uk} \succ r_{uj}) \times 1 - P(r_{uk} \succ r_{uj}) \end{aligned} \tag{3}$$

where (u, i, k) indicates that the user's preference for positive feedback is greater than the social relation feedback, (u, k, j) indicates that the user's preference for social relationship feedback is greater than the negative feedback, I_u^+ indicates that the user u has positive feedback to points of interest, I_u^* indicates points of interest that user u has not found, and I_u^- indicates that user u has negative feedback to points of interest. Since the partial order relation satisfies completeness and antisymmetry, (3) can be simplified as:

$$\begin{aligned} FBPR = &\sum_{u\epsilon U, i\epsilon I_u^+, k\epsilon I_u^*} P(r_{ui} \succ r_{uk}) \\ &+ \sum_{j\epsilon I_u^-, k\epsilon I_u^*} P(r_{uk} \succ r_{uj}) \end{aligned} \tag{4}$$

In order to facilitate the optimization calculation, according to BPR, a sigmod function $\sigma(x) = \frac{1}{(1+e^{(-x)})}$ is used to approximate the probability P $(\cdot)$ to obtain the maximum logarithmic posterior probability:

$$\begin{aligned} &\sum_{i\epsilon I_u^+} \sum_{k\epsilon I_u^*} \sum ln\,(\sigma)\, \frac{r_{ui} - r_{uk}}{1 + f_{uk}} \\ &\quad + \sum_{k\epsilon I_u^*} \sum_{j\epsilon I_u^-} ln(\sigma(r_{uk} - r_{uj})) - R(\Theta) \end{aligned} \tag{5}$$

Add the regularization term:

$$R\,(\Theta) = \sum_{u\epsilon U} \sum_{t\epsilon S} [\alpha_u \left\|U_u\right\|^2 + \alpha_v \left\|V_t\right\|^2 + \beta_v \left\|b_t\right\|^2] \tag{6}$$

In order to prevent overfitting during the learning process. Model parameter set $\Theta = \left\{U_u \epsilon R^{1\times d}, V_i \epsilon R^{1\times d}, b_i \epsilon R\right\}$, the sampling term $S = \{i, j, k\}$, U_u is a potential feature vector describing user u, V_i is a potential feature vector describing the point of interest i; b_i is the bias of the point of interest i; f_{uk} refers to the number of friends whose user u has not selected the point of interest k, but his friends choose k, and use the objective function $\frac{1}{1+f_{uk}}$ to weigh the preference distance between user u's positive feedback and social feedback, the larger f_{uk} is, the smaller the distance is between positive feedback and social feedback, indicating that friends of user u is interested in the point of interest k. Using matrix decomposition to model personal preference functions:

$$\begin{cases} r_{ui} = U_u V_i + b_i \\ r_{uk} = U_u V_k + b_k \\ r_{uj} = U_u V_j + b_j \end{cases} \tag{7}$$

IV. GEOGRAPHICAL DIVISION OF POINTS OF INTEREST

The advantage of coarse-grained division of a city according to administrative regions is that points of interest of each region can be trained in parallel to accelerate the clustering process. The geographic location of a city does not change too quickly, so storing the location of POIs in the Cloud can improve the servers availability.

This section first uses GMM to represent the distribution of points of interest, and then uses CRP to generate points of interest clusters.

Advanced hybrid model is an extension of the single Gaussian density function. GMM can approximate a density distribution of any shape. It can be applied to points-of-interest clustering to obtain good results. The point of interest i is defined as the multinomial distribution in cluster c, i.e., $\lambda_i = \left\{\lambda_{i,c} : c\epsilon C_i\right\}$, where $\lambda_{i,c}$, represents the probability of point of interest i in cluster c. In each cluster, the position distribution is captured by a mixed Gaussian distribution, i.e., $l \sim \mathbb{N}\left(\mu_c, \Lambda_c^{-1}\right)$. According to the generality of GMM, the POIs set is divided into K clusters, i.e.,

$$p(x) = \sum_{k=1}^{K} \pi_k \mathbb{N}(x; \mu_k \sigma_k) \tag{8}$$

where K is the number of Gaussian functions, and the CRP method is used to determine the number of K. The process of CRP is random. When a^{th} customer enters the restaurant, he or she can choose to sit at a table that is occupied, or a table without people. For customer y, the probability of choosing is

$$P = (z_y = m|z_{-y}, \alpha) \propto \begin{cases} \dfrac{n_m^{-y}}{\mathbb{N}+\alpha-1}, & \textit{Sit at the } m^{th} \textit{ table that is occupied} \\ \dfrac{\alpha}{\mathbb{N}+\alpha-1}, & \textit{Sit at a new table} \end{cases} \tag{9}$$

where n_m^{-y} is the number of people sitting at the m^{th} table except the a^{th} customer.

For the probability distribution P of customer a choosing a table, Sampling is conducted from the joint distribution according to Gibbs Sampling. The Sampling algorithm is shown in Algorithm 1. The sampling process first initializes the location of the POIs, and then iterates. Each iteration needs to calculate the conditional probability $P(z_i|\cdot, \alpha)$. After the algorithm converges, a sample of the probability distribution (i.e., $z^{(t)} \equiv (z_1^{(t)}, \ldots, z_{\mathbb{N}}^{(t)})$,..., $z_{\mathbb{N}}^{(t)}$) is obtained.

Algorithm 1 Gibbs Sampling

Require: $\mathbb{N}$ = number of customer,
T=number of iteration,
α = Dirichlet concentration,
b=burn-in.
Ensure: The last sample
Initialization: time t=0, $z_i : i = 1, \ldots, n$.
Start with one joint sample,e.g.$(z_1^{(0)} = 1 \ldots z_{\mathbb{N}}^{(0)} = 1)$
for t=1,…,T+b **do**
  for n=1,…, $\mathbb{N}$ **do**
    $z_1^{(t)} \sim P(z_1|z_2^{(t-1)}, z_3^{(t-1)}, \ldots, z_{\mathbb{N}}^{(t-1)}, \alpha)$
    $z_2^{(t)} \sim P(z_2|z_1^{(t)}, z_3^{(t-1)}, \ldots, z_{\mathbb{N}}^{(t-1)}, \alpha)$
    …
    $z_i^{(t)} \sim P(z_i|z_1^{(t)}, \ldots, z_{i-1}^{(t)}, z_{i+1}^{(t-1)}, \ldots, z_{\mathbb{N}}^{(t-1)}, \alpha)$
    …
    $z_{\mathbb{N}}^{(t)} \sim P(z_{\mathbb{N}}|z_1^{(t)}, z_2^{(t)}, \ldots, z_{\mathbb{N}-1}^{(t)}, \alpha)$
  end for;
  Generated a sample: $z^t \equiv (z_1^{(t)}, \ldots, z_{\mathbb{N}}^{(t)})$
end for;
Return the last sample after discard burn-in: $(z^{(1+b)}), \ldots, z^{(T+b)})$

V. SELECTION AND RANKING OF POIs

In order to ensure the quality of the points-of-interest ranking list, in this section, two factors that affect the user's selection of points of interest are considered in the selection of the points of interest and sorting. The first factor is the user's preference for POIs, which is the personal preference function proposed in Section 3. The second factor is the geographical location of the user, who is more willing to access the points of interest that are closer to him or the set of points-of-interest centered on the geographic location of a certain point of interest. Based on the clustering results in Section 4, this section uses the personal preference functions to select several interest points from each interest point cluster that meet the user's personalized needs and personal preferences. Then, uses power-law distribution to predict the score of each candidate interest point, and the score is sorted to form the final recommendation list of POIs, afterwards, the recommendation list is fed back to the target user.

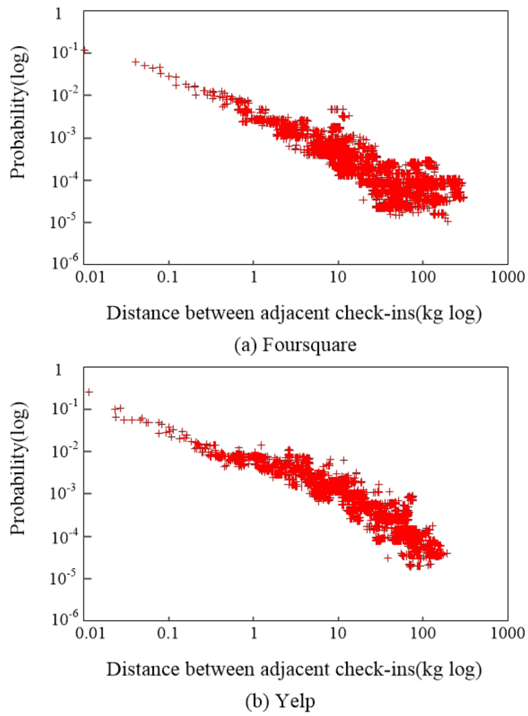


FIGURE 2. Distance distribution between two check-in points.

The distance is calculated between two adjacent points of interest that each user checked in in one day in the two datasets used, summarize the results of all users, and plot the number of check-ins as a function of distance as shown in Figure 2. A larger probability value indicates that the user is more willing to check in the points of interest at this distance. It can be seen from Figure 2 that the probability value distribution follows a power-law distribution [21], which shows that as the distance between two points of interest increases, the probability of users accessing decreases.

The pseudo-code of the algorithm for selecting points of interest and sorting is shown in Algorithm 2.

In order to incorporate geographic influence into POIs recommendations, a power-law distribution is used to simulate where users check-in to points of interest.

$$p(dis) = Cdis^k \tag{10}$$

Among them, $p(dis)$ represents the probability that a user visits points of interest that are kilometers away from his dis, C and k are parameters of a power law function respectively. Using Maximum Likelihood Estimation to estimate the

Algorithm 2 Selection and Ranking of POIs

Require: Personal preference function r_{ui}, r_{uk}, r_{uj}, Probability distribution sample $z^{(t)} \equiv (z_1^{(t)}, \ldots, z_{\mathbb{N}}^{(t)}), \ldots, z_{\mathbb{N}}^{(t)}$, Geographical location of users l_i.

Ensure: $list_{pois}$

1: Initialization: Ranking list of points of interest $list_{pois}$.
2: **for** u=1,...,n **do**
3: Using Equation(12) to calculate the probability of users check-in the POIs
4: Using Equation(13) to calculate the score of the candidate of POIs
5: Using Equation(14) to calculate the final ranking score
6: **end for**;
7: Return $list_{pois}$.

parameters of $p(dis)$, so we get:

$$ln(p(dis)) = ln(C) + kln(dis) \tag{11}$$

Assume that the geographic location of the user is l_i, the location of the POI is s_i, the distance between the two locations is $dis(l_i, s_i)$, and s_k is any other points of interest except s_i. According to the power-law distribution, the probability of accessing s_i is inversely proportional to $dis(l_i, s_i)$. The calculation method of conditional probability is as follows:

$$p(s_i|l_i)\frac{dis(l_i, s_j)}{\prod_{s_k \epsilon S, s_k \neq s_i} dis(l_i, s_k)} \tag{12}$$

Given user u and his historical check-in POIs set Sh, the Priori probability $P(l)$ of all users signing in the dataset. According to the Bayesian formula calculate the ranking score of each candidate POI, and then recommend the top-ranked POIs to the user. The score is calculated as follows:

$$\hat{r}_{ul}^{p} = P(l|S_h) \propto P(l)P(S_h|l) = P(l)\prod_{l \epsilon S_h} P(l'|l) \tag{13}$$

Next, the FBPR framework and geographic information are combined to generate a recommendation list. The final POIs ranking score is calculated as Equation (14).

$$\begin{aligned} Score &= Sr_u \times S\hat{r}_{ul}^{p} \\ Sr_u &= \frac{p_{ikj}^{u}}{max(p_{ikj}^{u})} \\ S\hat{r}_{ul}^{p} &= \frac{\hat{r}_{ul}^{p}}{max(\hat{r}_{ul}^{p})} \end{aligned} \tag{14}$$

where Sr_u and $S\hat{r}_{ul}^{p}$are personal preference and geographic location probability scores respectively, p_{ikj}^{u}represents the probability of user's personal preference, and $max(p_{ikj}^{u})$ represents the highest probability of user's personal preference.

TABLE 1. Statistical information of experimental datasets.

Dataset	Foursquare	Yelp
number of users	25379	30887
number of POIs	32623	18995
number of check-in records	1395856	860888
number of social-relations	118717	42163

VI. EXPERIMENTS

A. DATASETS

In our experiments, we utilize two popular social networks Foursquare and Yelp to evaluate our proposed method. The two datasets have different scales such as geographic ranges, the number of users, POIs, and check-ins. The Foursquare dataset is a service website based on geographic information and the Yelp dataset is one of the most influential review websites in the United States. Hence they are good for examining the performance of algorithms on various data types. Their statistics are listed in Table 1.

B. EVALUATION METHODS AND METRICS

To evaluate the performance of the algorithm, the experimental evaluation method uses two metrics Precision and Recall which widely used in Top-N recommendations, the calculation method is as follows:

$$Precision(@k) = \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|L_u^T \cap L_u^R|}{k} \tag{15}$$

$$Recall(@k) = \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|L_u^T \cap L_u^R|}{L_u^T} \tag{16}$$

where, Precision($@k$) represents the accuracy rate of recommending Top-K points of interest to the target user, represents the recall rate of recommending Top-K POIs to the target user. $|U|$ represents the number of all users, L_u^T represents the set of POIs recommended to users in the training set, and L_u^R represents the set of POIs that user U has checked in the test set. In the experiment, $k = \{5, 10, 15, 20\}$ was selected to calculate the accuracy and recall respectively.

C. COMPARISON MODELS AND PARAMETER SETTINGS

1) COMPARISON MODELS

In order to evaluate the recommendation performance of this algorithm, several classical models based on geographic location for points of interest recommendation are selected for comparative experiments.

(1) GeoMF [10]:Based on the weighted matrix decomposition model, the potential factors of users and points of interest are augmented to obtain an augmented model. At the same time, two-dimensional Kernel density estimation is used to cluster the spatial points of interest.

(2) MGM [13]:Geographic influences is captured by the multi-center Gaussian Model (MGM), and then geographic information and social influences are fused into matrix decomposition to recommend points of interest.

(3) BPRMF [17]: For implicit feedback, the user's interest preferences are learned from the user's paired item preferences, and a personalized ranking list of points of interest is recommended to the user based on matrix decomposition without using any contextual information.

(4) Rank-FBPR: The recommendation model proposed in this paper. First, integrates social relationships and BPR algorithms to build a FBPR framework of POIs, then uses CRP clustering method to obtain candidate POIs, and finally combines the personal preference and geographical location to recommend a Top-K POIs ranking list for users.

2) PARAMETER SETTINGS

The sparseness of user check-in of POIs has a great influence on the effectiveness and efficiency of the recommended model, which can be alleviated by adjusting the regularization coefficient in the model. This section will examine the effects of regularization coefficients and points-of-interest dimensions to ensure that each comparison model is compared with Rank-FBPR under optimal conditions. Figure 3 and Figure 4 show the effects of regularization coefficient and point of interest dimension on recommendation performance, respectively. It can be observed from Figure 3(a) and 3(b) that the accuracy of GeoMF, MGM and Rank-FBPR recommended results decreases with the increase of regularization coefficient. When the coefficient is 0.01, the accuracy of all recommended models is optimal. Since BPRMF has no regularization term, the coefficient has little effect on it; meanwhile, the recall increases as the coefficient increases. Figure 3(c) and 3(d) show that the accuracy and recall of the recommended model raise with the increase of the dimension of POIs. When the POIs dimension increases to 50, the increase of all models starts to slow down, which means that when the dimension is large enough, there will be enough information to characterize social and location relationships. Next observe the results on the Yelp dataset. As shown in Figure 4(a) and 4(b), similarly, the accuracy and recall decreased with the increases of regularization coefficient. Figure 4(c) and 4(d) show that the accuracy of all models starts to decrease when the points-of-interest dimension rises to 50. The reason may be that Yelp, as the largest review site in the United States, contains more points of interest and user information. The more sparse, the more difficult it is to model.

According to the different performance of the four models in the two datasets, it can be known that the recommended effect is best when GeoMF considers all negative feedback and sets the regularization coefficient γ to 0.01. At the same time, the POIs dimensions and the number of users in the MGM are also set to 10. For fair comparison, the coefficients in Equation(6) in this paper are set to users U = 10, POIs dimensions V = 10, and regularization coefficient b = 0.01.

D. EXPERIMENTAL RESULTS AND ANALYSIS

In order to verify the effectiveness of embedding geographic information in the FBPR framework, four recommended models are evaluated under cold start and non-cold start conditions. In the case of non-cold start, randomly select 80% of the data as the training set and 20% as the test set. In the case of cold start, select the same test set as in the case of non-cold start for testing, and randomly select 10%-80% data is used as the training set.

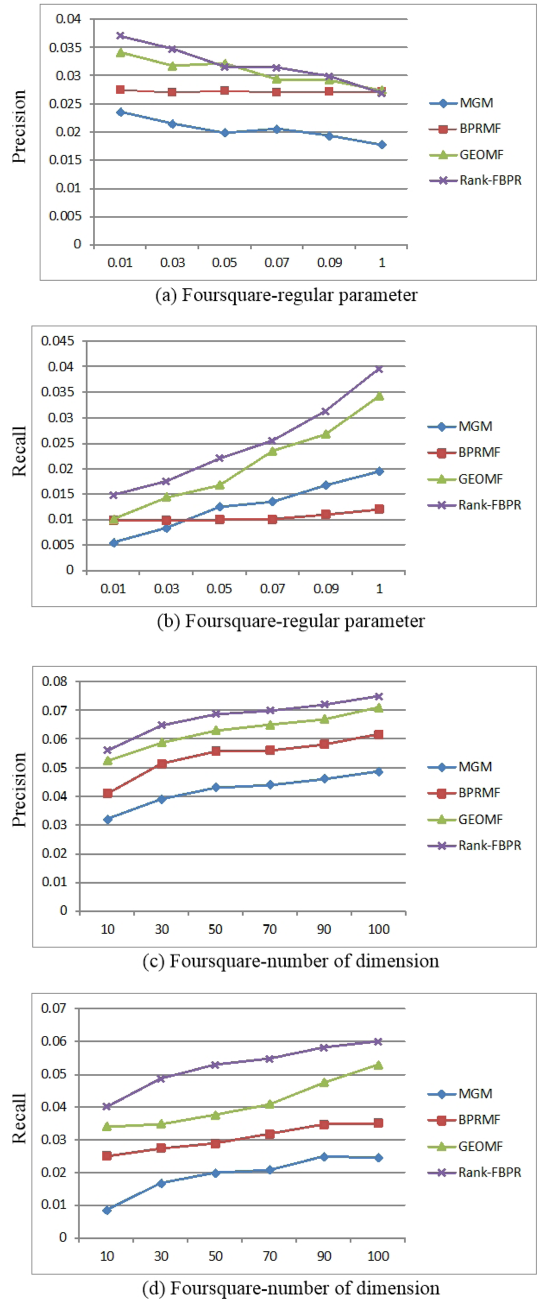
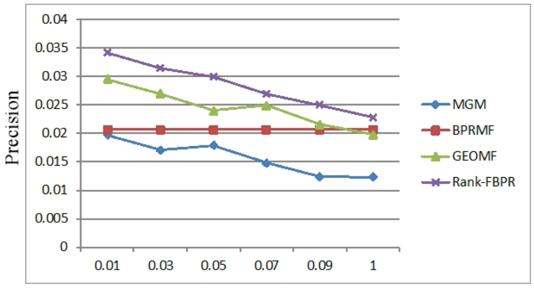


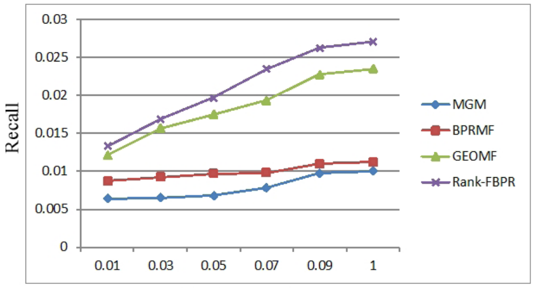
FIGURE 3. Influence of regularization coefficients and POIs dimensions on Foursquare.

1) EXPERIMENTAL RESULTS OF NON-COLD START CONDITIONS

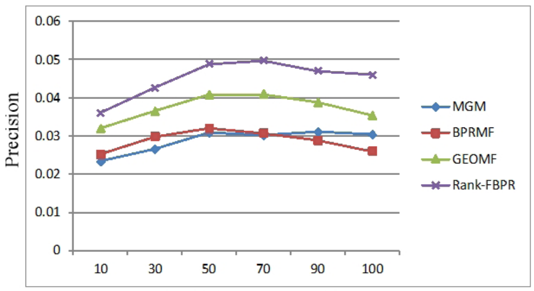
In Figure 5 and Figure 6, k indicates that users recommend k POIs. As can be seen from the figure, the Rank-FBPR model proposed in this paper has improved the precision



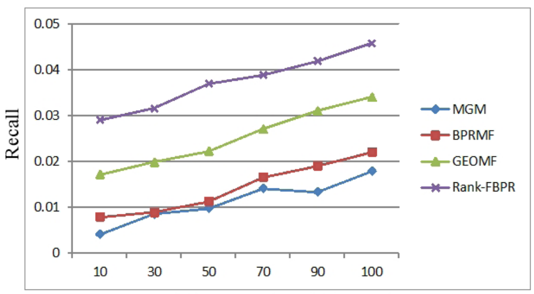
(a) Yelp-regular parameter



(b) Yelp-regular parameter



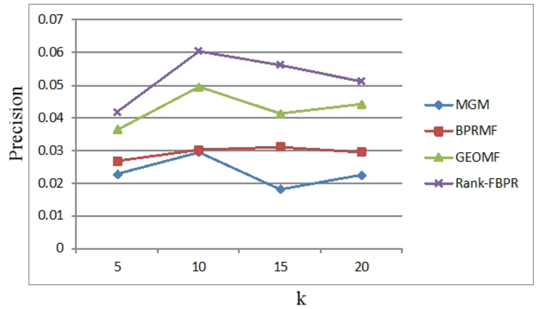
(c) Yelp-number of dimension



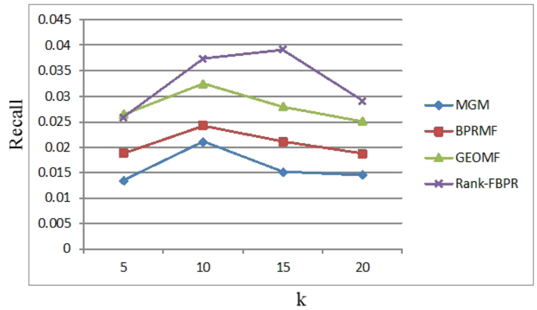
(d) Yelp-number of dimension

FIGURE 4. Influence of regularization coefficients and POIs dimensions on Yelp.

and recall performance indicators in general compared to GeoMF, MGM and BPRMF. Considering the influence of geographical location, the overall recommendation performance of MGM was less different from that of BPRMF, but the performance of MGM was unstable when the points of interest were different. BPRMF does not consider geographic information and social information, resulting in poor performance. GEOMF integrates geographic information into the weighted matrix factorization and performs cluster modeling on geographic space, which improves the overall recommendation performance compared to MGM and

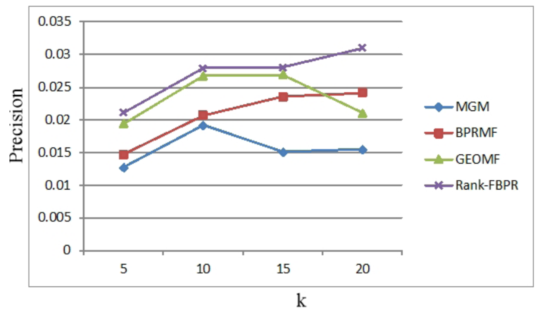


(a) Foursquare-comparison of Precision

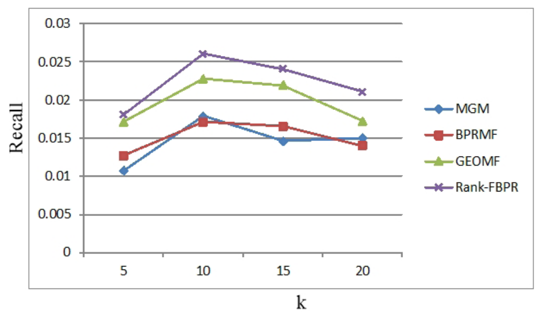


(b) Foursquare-comparison of Recall

FIGURE 5. The results of model comparison in non-cold start on Foursquare.



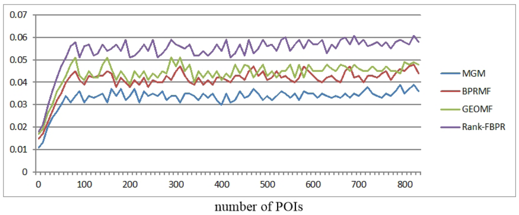
(a) Yelp-comparison of Precision



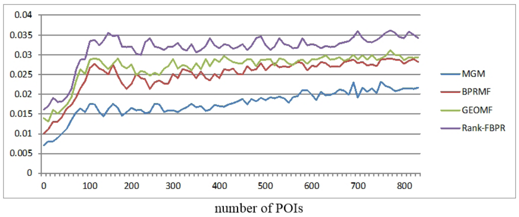
(b) Yelp-comparison of Recall

FIGURE 6. The results of model comparison in non-cold start on Yelp.

BPRMF. The Rank-FBPR proposed in this paper integrates both geographic information and social information. Compared with the other three algorithms, respectively, increased the accuracy rates by 30.9%, 30.2% and 10.9%, and increased the recall rates by 16.3%, 13.2% and 4.9% on Foursquare dataset (k=10), increased the accuracy rates by 8.7%, 7.2% and 1.2%, and increased the recall rates by 8.2%, 9% and

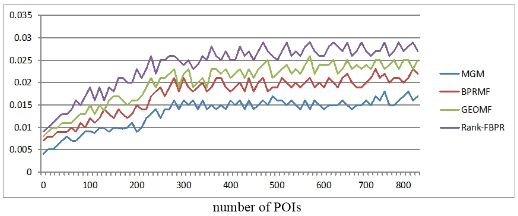


(a) Foursquare-comparison of Precision

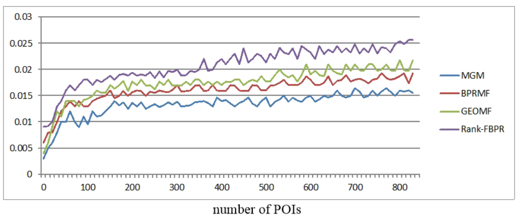


(b) Foursquare-comparison of Recall

FIGURE 7. The results of model comparison in cold start on Foursquare.



(a) Yelp-comparison of Precision



(b) Yelp-comparison of Recall

FIGURE 8. The results of model comparison in cold start on Yelp.

3.3% on the Yelp dataset (k=10). The results indicates that embedding geographic information and social information into the recommendation framework can significantly improve the accuracy and recall of recommendations at points of interest.

From Figure 5 and Figure 6, it can also be observed that the accuracy and recall of all POIs recommendation methods are low. The reason may be that the user check-in data is sparse in LBSN. In addition, compared with traditional movie recommendation or news recommendation, the users' check-in data cannot fully reflect their check-in POIs. For example, a user may make a check-in when passing a point of interest, which is not the POI that the user really prefers.

2) EXPERIMENTAL RESULTS OF COLD START CONDITIONS

According to the analysis of the experimental results in Figure 7 and Figure 8, it can be known that the recommended method Rank-FBPR can deal with cold start situations. If the target user has a cold start problem, Rank-FBPR can use social information(i.e., $r_{uk} = U_u V_k + b_k$) to make predictions for users, and also be able to use geographic information to mitigate cold start problems. The GEOMF model can use the user-interest point matrix to make predictions, and also takes into account geographical factors, which can also alleviate the cold start problem, but it does not incorporate social information, so it is not as good as the Rank-FBPR model in improving the quality of recommendations. BPR does not use any contextual information to cause poor recommendation results, but because it has advantages in mining user implicit feedback information, it has a small difference from GEOMF in general. MGM integrates the user check-in behavior into the Gaussian distribution model to predict the user check-in probability, which can enhance the accuracy of the recommendation. Considering the user's location information but not clustering the points of interest, the recommendation result is the worst compared to the other three models.

VII. CONCLUSION AND FUTURE WORK

In this work, we proposes an algorithm of Rank-FBPR framework for hybrid Cloud and edge computing innovatively, which provides real-time user awareness on the edge servers and interact with geographic information on the Cloud. Firstly, the algorithm integrates the user's social relationship into the BPR ranking criteria to obtain the user's personal preference function, divides the points-of-interest in the Cloud. Then, the points of interest are clustered based on the CRP process. After that, according to personal preferences and clustering results, combined with the user's geographical location, several candidate POIs that meet the user's personalized needs are selected, the candidate POIs scores are predicted and the scores is sorted in descending order to get the recommended list of interest points. Finally, the formed recommendation list is fed back to the target user. Experiments on real data sets show that the hybrid recommendation method adopted have great advantages in both the recommendation results and the performance of the framework, which reflects the high efficiency of the edge computing environment. And the method has more diversified POIs, higher accuracy and recall rate, and can alleviate the problem of data sparsity and better meet the personalized needs of users.

In future, we considers incorporating more contextual information into the framework for learning analysis, such as time effects, the diversity of users at the same points-of-interest and the consumption level of POIs. On the other hand, it is also necessary to provide continuous recommendation of POIs based on the correlation between points-of-interest, to improve accurate references for users, and to facilitate users' access to POIs. Furthermore, we hope to use a more hierarchical mechanism to further improve the recommendation performance and study a simpler, more scalable hybrid network structure.

REFERENCES

[1] M. Satyanarayanan, P. Simoens, Y. Xiao, P. Pillai, Z. Chen, K. Ha, W. Hu, and B. Amos, "Edge analytics in the Internet of Things," *IEEE Pervas. Comput.*, vol. 14, no. 2, pp. 24–31, Jun. 2015.

[2] H. Bangui, S. Rakrak, S. Raghay, and B. Buhnova, "Moving to the edge-cloud-of-things: Recent advances and future research directions," *Electronics*, vol. 7, no. 11, p. 309, 2018.
[3] X. Wang, Y. Han, V. C. M. Leung, D. Niyato, X. Yan, and X. Chen, "Convergence of edge computing and deep learning: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, to be published.
[4] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Netw.*, vol. 33, no. 5, pp. 156–165, Sep. 2019.
[5] X. Wang, Y. Han, and C. Wang, "STCS: Spatial-temporal collaborative sampling in flow-aware software defined networks," *IEEE J. Sel. Areas Commun.*, to be published.
[6] R. Gao, J. Li, X. Li, C. Song, and Y. Zhou, "A personalized point-of-interest recommendation model via fusion of geo-social information," *Neurocomputing*, vol. 273, pp. 159–170, Jan. 2018.
[7] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. (SIGIR)*, 2011, pp. 325–334.
[8] J.-D. Zhang and C.-Y. Chow, "IGSLR: Personalized geo-social location recommendation: A kernel density estimation approach," in *Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst. (SIGSPATIAL)*, 2013, pp. 334–343.
[9] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.
[10] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui, "GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2014, pp. 831–840.
[11] Y. Zhong, J. Li, and S. Zhu, "Clustering geospatial data for multiple reference points," *IEEE Access*, vol. 7, pp. 132423–132429, 2019.
[12] J. Shi, N. Mamoulis, D. Wu, and D. W. Cheung, "Density-based place clustering in geo-social networks," in *Proc. ACM SIGMOD Int. Conf. Manage. Data (SIGMOD)*, 2014, pp. 99–110.
[13] T.-Y. Qian, B. Liu, L. Hong, and Z.-N. You, "Time and location aware points of interest recommendation in location-based social networks," *J. Comput. Sci. Technol.*, vol. 33, no. 6, pp. 1219–1230, Nov. 2018.
[14] M. Zhong, Q. Zeng, and Y. Zhu, "Sample location selection for efficient distance-aware influence maximization in geo-social networks," in *Proc. Int. Conf. Database Syst. Adv. Appl.* Cham, Switzerland: Springer, 2018, pp. 355–371.
[15] J. Li, T. Sellis, J. S. Culpepper, Z. He, C. Liu, and J. Wang, "Geo-social influence spanning maximization," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 8, pp. 1653–1666, Aug. 2017.
[16] N. Al Hasan Haldar, J. Li, M. Reynolds, T. Sellis, and J. X. Yu, "Location prediction in large-scale social networks: An in-depth benchmarking study," *VLDB J.*, vol. 28, no. 5, pp. 623–648, Oct. 2019.
[17] S. Rendle, "BPR: Bayesian personalized ranking from implicit feedback," in *Proc. 25th Conf. Uncertainty Artif. Intell.* AUAI Press, 2009, pp. 452–461.
[18] R. He, W.-C. Kang, and J. McAuley, "Translation-based recommendation," in *Proc. 11th ACM Conf. Recommender Syst. (RecSys)*, 2017, pp. 161–169.
[19] L. Guo, Y.-F. Wen, and X.-H. Wang, "Exploiting pre-trained network embeddings for recommendations in social networks," *J. Comput. Sci. Technol.*, vol. 33, no. 4, pp. 682–696, Jul. 2018.
[20] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2011, pp. 1082–1090.
[21] Q. Yuan, "Time-aware point-of-interest recommendation," in *Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2013, pp. 363–372.

KEYAN CAO received the M.S. and Ph.D. degrees in computer science and technology from Northeastern University, China, in 2009 and 2014, respectively. She is currently an Associate Professor with the College of Information and Control Engineering, Shenyang Jianzhu University. Her current research interests include data management, cloud computing, and query process and optimization.

JINGJING GUO received the B.S. degree from the Dalian Neusoft University of Information, China, in 2018. She is currently pursuing the master's degree with the School of Computer Technology, Shenyang Jianzhu University. Her research interests include edge computing and recommender systems.

GONGJIE MENG received the bachelor's degree in software engineering from the Jincheng College, Nanjing University of Aeronautics and Astronautics, in 2019. He is currently pursuing the master's degree in software engineering with Shenyang Jianzhu University. He is also conducting research on pulmonary.

HAOLI LIU was born in Bozhou, China, in 1996. She received the B.S. degree from Shenyang Jianzhu University, China, in 2019, where she is currently pursuing the master's degree with the School of Computer Technology. Her research interest includes big data analysis.

YEFAN LIU received the B.S. degree from the Shenyang University of Technology, China. He is currently pursuing the master's degree with the School of Computer Science and Technology, Shenyang Jianzhu University. His current research interest includes big data processing.

GUI LI received the Ph.D. degree in computer science and theory from Northeastern University, China, in 1998. He is currently a Professor with the College of Information and Control Engineering, Shenyang Jianzhu University. His current research interests include data management, data mining, pattern recognition, and the IoT.