Indexed in
Scopus Compendex and Geobase Elsevier,
Geo-Ref Information Services-USA, List B of Scientific
Journals, Poland, Directory of Research Journals

International Journal of Earth Sciences and Engineering

www.cafetinnova.org

ISSN 0974-5904, Volume 09, No. 06

December2016,P.P.2765-2770

GPS-based Personalized Point-of-Interest Recommendation Algorithm

ZHIJUN ZHANG[1] AND HUALI PAN[2*]
[1]School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250101, China
[2]Business School, Shandong Normal University, Jinan 250014, China
Email: *zzjsdcn@163.com*

Abstract: The popularity of GPS makes it possible to record people's moving trajectories which contain the users' interests and preferences. Different moving trajectories reflect different interests of the users. In view of the low precision of the traditional tourism recommendation algorithm in tourist spot recommendation, the paper puts forward a personalized tourist spot recommendation algorithm based on GPS and LBSN. The algorithm can acquire the users' tourism intentions according to their position and time information and establish the user preference model based on the information, so as to generate multiple tourist spots for them to choose in real time and recommend them the appropriate ones. The experimental results on the real data set show that compared with the existing similar algorithms, the GPSRec algorithm in the paper has higher recommendation precision.

Keywords: *Location services, GPS, Personalization recommendation, POI, Time information*

1. Introduction

With the development of economy and transportation, tourism has gradually become a part of people's life. For the tourists, the most important part of preparations for tourism is to select the tourist destination and route. The traditional tourist spot recommendation algorithm usually recommends the users some popular tourist spots and routes with little consideration on their interests and preferences, leading to the low user satisfaction.

With the rapid development of mobile Internet and the development and popularization of GPS and other mobile positioning technologies, the location-based service has received more and more attention from the academic and business circles. In recent years, the fast progress of the Location Based Social Networks (LBSN) has prompted a lot of research work to focus on finding the interest points or mining the popular tourist tracks from GPS and check-in data. On this basis, some researchers have mined the tracks from LBSN data and automatically generate the tourist routes for the users according to their travel time and requirements on geographical area. Those studies mainly find out the popular visiting spots and routes in the data but do not take into account the users' personalized requirements. Considering a single tourist destination cannot satisfy the users' travel demands. At the same time, the tourist spots or routs mined from LBSN data often considers only the statistical properties of the data to get the most popular spots and routes, but the tourist tracks cannot adapt to the users' unique demands.

The recommendation based on the location information has become a research focus and important application scene of personalized recommendation. The solution of the problem needs to have a deep understanding of the users' moving patterns, including predicting the users' moving tracks and judging whether it is possible to carry out the tourist activities at the current locations of the users [1]. It also needs to define the similarity among the users and among the locations by the quantitative method [2].

In this paper, Chapter 2 reviews the relevant research work; Chapter 3 presents the definition and application of GPS; Chapter 4 puts forward the improved POI personalized GPSRec algorithm; Chapter 5 shows the experimental results and analysis; Chapter 6 makes a summary and prospects.

2. Related work

The users' GPS trajectory sequences record their moving routes in the real physical world, which implies their personal intentions, preferences and behavior patterns. It mines the users' behaviors, intentions, experience and life styles from personal data [3], integrates the group data to find out the hot spots and classic routes and even mine the correlation between people and the activity model of individuals between regions [4].

At present, deep research has been conducted based on the geographical and social information, including a lot on the location recommendation. Zhou et al. [5] put forward to convert the users' check-in times into an implicit score of the users' locations and then recommend the spots by the collaborative filtering thought, which is too simple. Levandoski et al. [6] proposed the personalized recommendation algorithm which recommends the tourist spots according to the users' locations. Since it does not consider the influence of other factors on the users' interests, its recommendation precision is very low. Cheng et al. [7] predicted the users' scoring of the locations based on the their check-in frequencies and recommended the tourist spots by the matrix decomposition thought, which can only partially alleviate the sparsity of the user's location scoring matrix. Zheng et al. [8] made use of the times of the users' access locations to find the similar friends and recommended the interested tourist spots of the friends to the target users, but it did not involve the users' activities and trust relationship. Ye et al. [9] used the collaborative scoring of the users' social friends and the similarity between the friends to recommend the tourist spots, which did not involve the users' activity preference information. Literature [10] analyzed the geographical factors of the locations in details and recommended the tourist spots after integration, but the experimental results were not ideal.

Some scholars added the time information in the recommendation research in the location-based social network which could reflect the change of the users' check-in over time, so as to make the more accurate location recommendation in the specific time. Koren [11] considered the change of the users' preferences over time and set the high weights for the preferences in the recent time. Xu et al. [12] studied the influence of the time factor on the personalized prediction. Gao et al. [13] put forward the new location preference framework by using the time property, divided the user's check-in matrix into the sub-matrices under different periods of time and integrated the users' check-in preferences at different periods of time to get the users' check-in preferences for the candidate locations. Literature [14] mapped the users' check-in records, geographic information and time information to the geographic time information perception diagram and put forward the preference propagation algorithm. Literature [15] put forward the location recommendation algorithm based on the check-in behaviors in the continuous time, mined the consecutive check-in pattern from the users' consecutive check-in location history, mapped the pattern into the dynamic change chart of the location time and predicted the probability that the users checked in at the same candidate locations in different periods of time by the use of Markov chain. Literature [16] studied the distribution function of time divisions of users' consecutive check-ins. Literature [17] added the time and space information in the user collaborative filtering-based model and put forward the interested point recommendation framework perceived by the time. Zhang et al.[18] put forward a user modeling method based on the wireless sensor network and ontology technology, designed a personalized recommendation model combining ontology and tourist information together and discussed the ontology-based personalized recommendation algorithm of tourist spots based on the scene sensitivity in the mobile Internet.

This paper uses the check-in data with the geographic information to study a personalized tourism recommendation algorithm GPSRec based on GPS. Considering the time factor, it establishes the preference model between users and tourist spots to recommend them multiple interested tourist spots. The experimental results on the real data set show that the GPSRec algorithm in the paper can improve the users' satisfaction with its high recommendation precision.

3. GPS introduction

3.1. GPS

GPS is a new generation of satellite navigation and positioning system which combines the satellite and communication development technologies and owns the air-sea-land real-time 3D navigation and positioning capabilities. It uses the navigation satellite to measure the time and distance. GPS consists of three parts: the space part - GPS constellation (GPS constellation is composed of 24 satellites, in which 21 are the working satellites and 3 are the backup satellites), the ground control part - ground monitoring and control system, the user equipment part - GPS signal receiver.

Here are several definitions related to GPS [19].

Definition 1: GPS track: GPS_t is a series of sequences of GPS track points related to the time. $GPS_t = (P_1, P_2, P_3, \ldots, P_n)$,

$P_i = (x, y, t), (1 \le i \le n)$, where (x, y) represents the longitude and latitude of the collected data, respectively; t represents the data-collection time and meets $P_i \cdot t < P_{i+1} \cdot t$, $(1 \le i \le n-1)$.

Definition 2: resident area: GPS resident area S_z refers to a set of adjacent or similar GPS track points in a certain period of time, $S_z = (P_i, P_{i+1}, \ldots, P_j)$, meeting $Dist(P_i, P_k) \le \theta_d$, $Int(P_i, P_j) \ge \theta_t$, in which θ_d is the diameter of the resident area; θ_t is the time in the resident area. $Dist(P_i, P_j)$ is the Euclidean distance between two points (P_i, P_j) and $Int(P_i, P_j)$ is the time interval between $P_i \cdot t$ and $P_j \cdot t$, $(i \le k < j)$.

Definition 3: User stagnation point $S_p = (x, y, t_{in}, t_{out})$ refers to the geometric center of the resident area, in which

$$S_p.x = \sum_{k=j}^{i} P_k.x / |P|,\ S_p.y = \sum_{k=j}^{i} P_k.y / |P|,$$

$S_P \cdot t_{in} = P_i \cdot t$, $S_P \cdot t_{out} = P_j \cdot t$, and $P_k \in S_z$.

3.2. GPS Application

Location service originated from North America. In May 2000, GPS was fully open to the business by the US government, which significantly improved the positioning precision and promoted the spread and application of the location service. The data collected by GPS are the main source of the users' moving trajectory data, in which the sampling precision and frequency have a great influence on the follow-up analysis. The direct application of the data with the interference factors to the user data mining often fails to get the desired effect.

With the characteristics of 24-hour service, high precision, automation and high benefit, GPS has been successfully applied in geodetic survey, engineering survey, aerial photography, vehicle navigation and control, crustal movement measurement, engineering deformation measurement, resource survey, earth dynamics and other disciplines and achieved good economic and social benefits.

3.3. Flow Chart of the Tourism Recommendation System

The general flow chart of the tourism recommendation system is shown in Fig.1

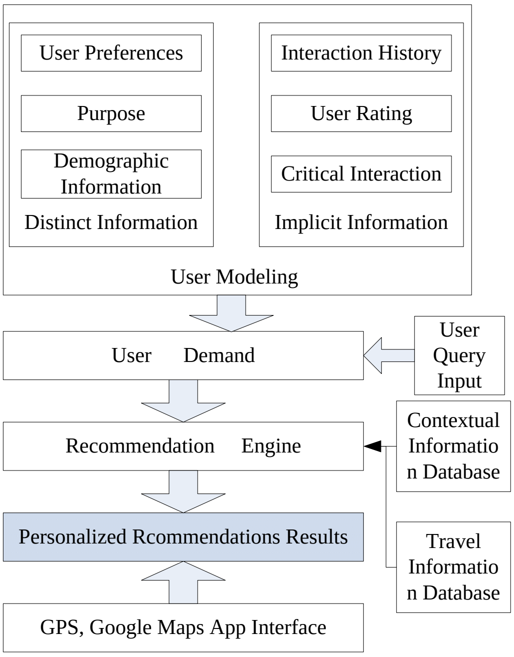

Fig.1 *A general framework of travel recommendation system*

4. Improved POI Personalized Recommendation Algorithm-GPSRec

4.1. The Map-matching Algorithm based on the Location Point [20]

The basic principle of the map-matching algorithm based on the location point is to project the GPS coordinates for matching to the nearby roads and analyze each GPS coordinate by the improved map-matching algorithm based on the location point. Four analytical conditions of the location Type can be matched as follows:

1. If the location Type is ROOFTOP or GEOMETRIC_CEN_TER, the location coordinates obtained through Google Geocoding API analysis are the requested coordinates.

2. If the location Type is APPROXIMATE, use the vertical projection method proposed in the paper to calculate the vertical coordinates and convert them into the corrected GPS coordinates. The northeast and southwest values of bounds analyzed by the vertical projection method are taken as the parameters of the secondary rectifying coordinates.

3. If the location Type is RANGE_INTERPOLATED, calculate the historical track coordinates to get the coordinates of the crossroads for linear interpolation and judge the moving direction of the user at the crossroads in real time, so as to match the tracks at the crossroads to the map.

4.2. Calculation of User Similarity and Location Similarity [21]

It is assumed that there are N_u users and N_l locations. M_{ul} is a user-location check-in matrix and $M_{ul}(u_i,l_j)$ represents the check-in numbers of the user u_i at the location l_j. $S_U(u_i,u_j)$ represents the similarity between u_i and u_j and $S_L(l_i,l_j)$ represents the similarity between l_i and l_j. It is assumed that each location resource implies a certain theme. If two locations are checked in by the users with the similar interest at the same time, it indicates that those two locations have the similar theme. Similarly, if two users often check in at some locations with similar themes, it indicates that those two users have the high interest similarity. The similarity among the users and among the locations can be solved by Formula (1) and (2).

For any user pair (u_i,u_j),

$$S_U^{k+1}(u_i,u_j)=\frac{C_U}{|P(u_i)||P(u_j)|}\sum_{m=1}^{|P(u_i)|}\sum_{n=1}^{|P(u_j)|}\frac{\min(M_{ul}(u_i,l_m),M_{ul}(u_j,l_n))}{\max(M_{ul}(u_i,l_m),M_{ul}(u_j,l_n))}S_L^k(L_m(u_i),L_n(u_j)) \quad (1)$$

Where, $P(u_i)$ is the number of locations where the user u_i has checked in; $L_m(u_i)$ is the mth location where the user u_i has checked in.

For any user pair (l_i,l_j),

$$S_L^{k+1}(l_i,l_j)=\frac{C_L}{|P(l_i)||P(l_j)|}\sum_{m=1}^{|P(l_i)|}\sum_{n=1}^{|P(l_j)|}\frac{\min(M_{ul}(u_m,l_i),M_{ul}(u_n,l_j))}{\max(M_{ul}(u_m,l_i),M_{ul}(u_n,l_j))}S_U^k(U_m(l_i),U_n(l_j)) \quad (2)$$

Where, $P(l_i)$ is the number of users who have checked in at l_i and $U_m(l_i)$ is the m user who has checked in at l_i.

According to the traditional collaborative filtering algorithm, the user u_i's interest R_{ik} in an unknown location l_k can be obtained by Formula (3).

$$R_{ik} = \frac{\sum_{u_m \in U} S_U(u_i, u_m) \times R_{mk}}{\sum_{u_m \in U} S_U(u_i, u_m)} \tag{3}$$

4.3. The Time-based Personalized Recommendation Algorithm

Literature [17] adds the time information in the model based on the user collaborative filtering and puts forward the time-based POI recommendation model. Given the user u and the time t, u's scoring of the location l at t can be calculated by Formula (4).

$$\hat{c}^{(T)}_{u,t,l} = \frac{\sum_v \omega^{(t)}_{u,v} \cdot c_{v,t,l}}{\sum_v \omega^{(t)}_{u,v}} \tag{4}$$

Where, $\omega^{(t)}_{u,v}$ is calculated by Formula (5).

$$\omega^{(t)}_{u,v} = \frac{\sum_{t=1}^{T} \sum_{l=1}^{L} c_{u,t,l} \cdot c_{v,t,l}}{\sqrt{\sum_{t=1}^{T} \sum_{l=1}^{L} c_{u,t,l}^2} \cdot \sqrt{\sum_{t=1}^{T} \sum_{l=1}^{L} c_{v,t,l}^2}} \tag{5}$$

4.4. Improved Personalized POI Recommendation Algorithm --GPSRec

The paper calculates the similarities among the users and among the locations according to the location information obtained by GPS and then puts forward an improved POI personalized recommendation algorithm GPSRec based on the time factor. The specific description of GPSRec algorithm is as follows:

Algorithm-GPSRec:

Input: the users' check-in data matrix M_{ul}, the target user u and the time t.

Output: the probability that the target user u checks in at the location l at the time t.

Step 1: analyze each GPS coordinate point by the location point-based map matching algorithm.

Step 2: calculate the similarities among the users by Formula (1) to get the final similar user set of the target user u.

Step 3: calculate the similarities among the locations to get the final similar user set of the target location l.

Step 4: Calculate the user's scores of the location l at different times by Formula (4).

Step 5: Calculate the probability that the user checks in at each candidate location and select Top-N spots to recommend.

5. Experimental Results and Analysis

5.1. Data Set and Evaluation Indexes

The paper selects the users' check-in information of Foursquare, Gowalla, and BrightKite based on the users' geographic information as the data set and the precision and recall rate as the evaluation indexes. The precision $@N$ and the recall rate $@N$ represent the number of the recommending result Top-N. $T_{u,t}$ represents the real location set where the target user u checks in at t. $R_{u,t}$ represents the location set recommended to the users at t. $N_{u,t,r}$ represents the numbers of locations contained in the intersection of $T_{u,t}$ and $R_{u,t}$. The number of locations contained in $T_{u,t}$ is written as $N_{u,t}$, and that in $R_{u,t}$ is written as $N_{u,r}$, so the recommendation precision and recall rate at t are expressed by Formula (6) and (7), respectively.

$$precision(t) = \frac{\sum_u N_{u,t,r}}{\sum_u N_{u,r}} \tag{6}$$

$$recall(t) = \frac{\sum_u N_{u,t,r}}{\sum_u N_{u,t}} \tag{7}$$

Since the paper adds the time characteristic, the precisions and recall rates of all times are averaged. The final recommendation precision and recall rate are calculated by Formula (8) and (9), respectively.

$$precision = \frac{1}{T} \sum_{t=0}^{T} precision(t) \tag{8}$$

$$recall = \frac{1}{T} \sum_{t=0}^{T} recall(t) \tag{9}$$

5.2. Characteristic Analysis of the Users' Check-in Locations

Through analyzing the location distribution of the users' check-in locations, the paper looks for the nearest users with high similarity for the target users, so as to improve the recommendation precision. It calculates the number of the same check-in locations between all users in Gowalla and BrightKite and the similar users to get the CDF (cumulative distribution function) of the number of the same locations between the similar users and the target users, as shown in Fig.2. The CDF represents the probability that the random variable is smaller than or equal to a value. It is expressed by Formula (10). $F(x)$ is the continuous strictly monotone increasing function.

$$F(x) = p(X \le x) \tag{10}$$

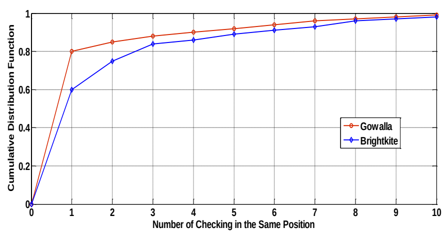

***Fig.2** Number of checking in the same position cumulative distribution function*

It is not hard to find from Fig.2 that with the increasing number of the same check-in locations between the similar users and the target users, the value of the cumulative distribution shows a rising trend. Therefore, according to the similarity characteristic of the users' check-ins, it is feasible to filter the users with low similarity to get the nearest similar ones.

5.3. Characteristic Analysis of the Users' Check-in Times

The paper analyzes the influence of the time factor on the users' check-in behavior through the experiment. It first discusses the time distribution law of the users' check-ins from 0:00 to 24:00 in the unit of h by extracting the check-in time column whose check-in location is not null from Four square data set. The analysis results are shown in Fig.3.

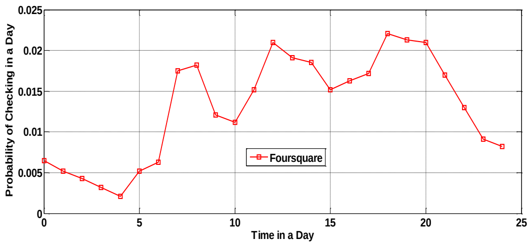

Fig.3 *Distribution of LBSN users' check-in time in a day*

According to Fig.3, the users' check-in frequency is very low from 0:00 to 4:00 and high from 5:00 to 8:00, from 12:00 to 14:00 and from 17:00 to 21:00. After 21:00, it drops significantly.

5.4. Comparison of the Tourist Spot Recommendation Precisions

The comparative experiment is set to compare the GPSRec algorithm in the paper with the traditional UserCF based on the user's collaborative filtering method, ItemCF based on the project's collaborative filtering method and Personalized Context - aware Rank (PCR) method [22]. PCR recommends the tourist spots through the weather information and simple user similarity calculation. Fig.4 shows the precision comparison of different recommendation algorithms when Top-N takes 5, 10 and 20 on BrightKite data set.

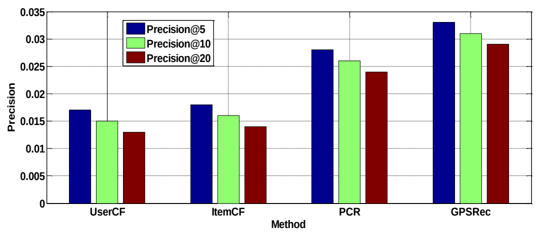

Fig.4 *Percision comparisons of recommendation algorithm*

It can be seen from Fig.4 that UserCF algorithm and ItemCF have the poorest recommendation performance. They are followed by PCR recommendation algorithm. Compared with UserCF, ItemCF and PCR, GPSRec in the paper has high recommendation precision. Compared with that at precision@5, the recommendation precisions at precision@10 and precision@20 drop significantly.

Fig.5 shows the recall rate comparison of different recommendation algorithms when Top-N takes 5, 10 and 20 on BrightKite data set.

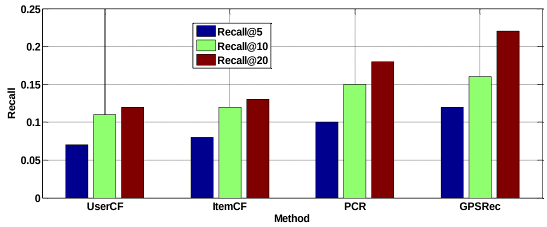

Fig.5 *Rcall comparisons of recommendation algorithm*

According to Fig.5, compared with UserCF, ItemCF and PCR, GPSRec has a good recall rate. Compared with that at precision@5, the recall rates at precision@10 and precision@20 increase significantly.

6. Conclusion and Prospects

The location-based social network is a new field. How to provide more and better services to the users in this field and explore more business opportunities is a problem that every mobile social Internet company and travel agency needs to think about. The good location social recommendation system is very important for LBSN to attract and retain the users and businesses. The paper conducts wide research on the LBSN recommendation algorithm and puts forward a GPS-based personalized tourist spot recommendation algorithm GPSRec. It integrates the users' preferences, time and geographical location information to make the personalized tourist spot recommendation. After the experimental and comparative analysis with other recommendation algorithms, it verifies the feasibility and validity of the improved algorithm. Although the improved algorithm proposed in the paper can better recommend the locations, it is not the optimal. It can also be further optimized by combining other aspects. For example, it can use the content containing the location label to make semantic modeling, so as to improve the recommendation effect of the algorithm, which is the direction of further research in the future.

7. Acknowledgements:

This study is supported by the Science and Technology Development Planning of Shandong Province (2014GGX101011, 2015GGX101018), A Project of Shandong Province Higher Educational Science and Technology Program (J12LN31, J13LN11, and J14LN14), and Jinan Higher Educational Innovation Plan (201401214, 201303001).

References

[1] C. Song and A.L. Barabási, “Limits of Predictability in Human Mobility”, *Science*, 327(5968), PP.1018-1021, 2010.

[2] V.W. Zheng, Y. Zheng and X. Xie, “Collaborative location and activity recommendations with GPS history data”, *International Conference on World Wide Web-2010*, PP.1029-1038, 2010.

[3] Y. Zheng, L. Liu, L. and Wang, “Learning transportation mode from raw gps data for geographic applications on the web”, *International Conference on World Wide Web-2008*, PP.247-256, 2008.

[4] Y. Zheng, L. Zhang,X. Xie and W.Y. Ma, “Mining interesting locations and travel sequences from GPS trajectories”, *International Conference on World Wide Web-2009*, PP.791-800, 2009.

[5] D. Zhou, B. Wang and S.M. Rahimi, “A Study of Recommending Locations on Location-Based Social Network by Collaborative Filtering”, *Canadian Conference on Advances in Artificial Intelligence-2012*, PP.255-266, 2012.

[6] J.J. Levandoski, M. Sarwat and A. Eldawy, “LARS: A Location-Aware Recommender System”, *ICDE-2012*, 41(4), PP.450-461, 2012.

[7] C. Cheng, H. Yang, I. King and M.R. Lyu, “Fused matrix factorization with geographical and social influence in location-based social networks”, *Proceedings of Twenty-Sixth Conference on Artificial Intelligence-2012*, PP.17-23, 2012

[8] Y. Zheng, L. Zhang and Z. Ma, “Recommending friends and locations based on individual location history”, *Acm Transactions on the Web*, 5(1), PP.99-111, 2011.

[9] M. Ye, P. Yin and W.C. Lee, “Location recommendation for location-based social networks”, *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, *ACM-2010*, PP.458-461, 2010.

[10] M. Ye, P. Yin and W.C. Lee, “Exploiting geographical influence for collaborative point-of-interest recommendation”, *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, ACM-2011*, PP. 325-334, 2011.

[11] Y. Koren, “Collaborative filtering with temporal dynamics”, *Communications of the ACM*, 53(4), PP.89-97, 2010

[12] J. Xu, P.N. Tan and L. Luo, “GSpartan: a Geospatio-Temporal Multi-task Learning Framework for Multi-location Prediction”, *Siam International Conference on Data Mining-*2016,PP.99-126, 2016.

[13] H. Gao, J. Tang, X. Hu and H. Liu, “Exploring temporal effects for location recommendation on location-based social networks”, *Proceedings of the 7th ACM Conference on Recommender Systems, ACM-2013* ,PP.93–100, 2013.

[14] Q. Yuan, G. Cong and A. Sun, “Graph-based point-of-interest recommendation with geographical and temporal influences”, *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*, *ACM-2014*, PP.659:668, 2014.

[15] J.D. Zhang, C.Y. Chow and Y. Li, “LORE: exploiting sequential influence for location recommendations”, *22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems -2014*, PP.103-112, 2014.

[16] W. Wei, X. Zhu and Q. Li, “LBSNSim: Analyzing and modeling location-based social networks”, *Proceedings - IEEE INFOCOM-2014*, PP.1680-1688, 2014.

[17] Q. Yuan, G. Cong and Z. Ma, “Time-aware point-of-interest recommendation”, *International ACM SIGIR Conference on Research and Development in Information Retrieval-2013*, PP.363-372, 2013.

[18] Z.J. Zhang G.W. Xu, P.F. Zhang and Y.K. Wang “Personalized Recommendation System Based on WSN”, *International Journal of Online Engineering*, 12(10), PP.91-96, 2016.

[19] W.Z. Chen, S. Lin, L. Wang and D.M. Li, “Personalized recommendation algorithm of health advice based on the user's mobile trajectory”, *CAAI Transactions on Intelligent Systems*, 11(02): PP.264-271, 2016.

[20] Z.M. Wang, O. Wang and R. Zhang, “Improved map-matching algorithm based on position points”, *Application Research of Computers*, 30(11), PP.3318-3319, 2013.

[21] S. Wang, “A location recommendation algorithm based on location-based social networks”, *Computer Engineering & Science*, 38(4), PP. 667-672, 2016.

[22] A. Majid, L. Chen and L. Chen, “A context-aware personalized travel recommendation system based on geotagged social media data mining”, *International Journal of Geographical Information Science*, 27(4), PP.662-684, 2013.