R2SIGTP: a Novel Real-Time Recommendation System with Integration of Geography and Temporal Preference for Next Point-of-Interest

Xu Jiao
Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin University of Technology
Tianjin, China
jiaoxu1999@sina.com

Yingyuan Xiao*
Key Laboratory of Computer Vision and System, Ministry of Education, Tianjin University of Technology
Tianjin, China
yyxiao@tjut.edu.cn

Wenguang Zheng
Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin University of Technology
Tianjin, China
wenguangz@tjut.edu.cn

Hongya Wang
Donghua University
Shanghai, China
hywang@dhu.edu.cn

Youzhi Jin
Tianjin University of Technology
Tianjin, China
jinyouzhi.lyg@gmail.com

ABSTRACT

With the rapid development of location-based social networks (LBSNs), point of interest (POI) recommendation has become an important way to meet users' personalized demands. The aim of POI recommendation is to provide personalized recommendation of POIs for mobile users. However, traditional POI recommendation systems cannot satisfy users' personalized demands. The reason is that the traditional POI recommendation system cannot recommend the next POI to a user based on the user's context information. Also, the traditional POI recommendation system provides no real-time guarantee on performance. In this demo, we propose a novel real-time next POI recommendation system named R2SIGTP which provides more personalized real-time recommendation compared with existing ones. Our system has the following advantages: 1) it has real-time performance; 2) it uses a unified approach to integrate geographic and preference information; 3) it considers the feedback of each single user to provide more personalized recommendation. We have implemented our system. R2SIGTP is easy to use and can be used by the mobile terminal's browser to recommend the next POI to the user in real-time based on the automatically identified user location and current time. The experimental results on real-world LBSNs show that R2SIGTP's performance is satisfactory.

CCS CONCEPTS

• **Human-centered computing → Social recommendation**; *User models*; Scenario-based design.

*Corresponding author

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.
WWW '19, May 13–17, 2019, San Francisco, CA, USA
© 2019 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.
ACM ISBN 978-1-4503-6674-8/19/05.
https://doi.org/10.1145/3308558.3314120

KEYWORDS

POI recommendation, tensor, sentiment analysis

ACM Reference Format:
Xu Jiao, Yingyuan Xiao, Wenguang Zheng, Hongya Wang, and Youzhi Jin. 2019. R2SIGTP: a Novel Real-Time Recommendation System with Integration of Geography and Temporal Preference for Next Point-of-Interest. In *Proceedings of the 2019 World Wide Web Conference (WWW'19), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3308558.3314120

1 INTRODUCTION

POI recommendation is one of the most important tasks in LBSNs, which helps users discover new interesting locations in the LBSNs. POI recommendations using users' check-in history are often influenced by multiple factors such as geography, time, sequence, etc. Geographical influences are the most essential feature that distinguishes POI recommendations from traditional recommendations. Since a user's check-in behavior presents a spatial clustering phenomenon, geographical influence can be modeled by power law distribution[5], Gauss distribution, Poisson distribution, and kernel density estimation[6]. Temporal influence in a POI recommendation system performs in two aspects: periodicity and non-uniformness. And the sequential influence[7] is the result of the interaction of temporal periodicity, adjacency of POIs in geographical space, property of POIs and human habits. At present, most existing POI recommendation systems use the methods based on content, link analysis[1], collaborative filtering[8] or matrix factorization[2] in the recommendation process.

Although POI recommendation system has achieved great success, there are still many challenges for it. 1) Most existing POI recommendation systems use offline modeling and online recommendation system architecture. As the system runs, a large amount of incremental data is generated in real-time. To take advantage of these incremental data, the system must recompute the offline model of the system, which is very time consuming and can seriously affect the operation of the online recommendation. Obviously, this strategy is not suitable for frequent updates and affects the accuracy of the recommendation system. 2) The user-POI check-in matrix is extremely sparse compared to the user-item matrix of the traditional recommendation system. Because the total number of POIs is quite large, and the number of POIs visited by a single user is very small. 3) Users' reviews are direct feedbacks to POI recommendation results. For a specific POI, most existing POI recommendation systems provide a general solution for all the users by analyzing all the corresponding reviews. Thus, it cannot provide personalized feedback for each single user. 4) Most existing POI recommendation systems integrate multi-aspect informations of check-ins. However, they do not construct a unified recommendation process to integrate these informations, that is, the implicit correlations between the various informations are ignored. 5) Most existing POI recommendation systems use Euclidean distance. This leads to deviations in actual applications and makes POI recommendation systems inaccurate.

Aiming to solve the aforementioned problems, a real-time personalized next POI recommendation system R2SIGTP was designed. R2SIGTP utilizes multi-aspect informations to model user temporal preferences. According to the current location of a user and time , R2SIGTP recommends the next POI which the user is most likely to access. R2SIGTP has the following advantages:

- It leverages a scalable parallel incremental approach to model users' check-in history, making the system real-time scalable.
- It classifies the POIs into users' preferences and uses the preferences to model users' check-in history such that the sparsity of check-in data can be overcome.
- It provides personalized feedback for each single user by using convolution neural network(CNN) to analyze the users' reviews.
- According to users' choices, it integrates geographic and preference information into the unified recommendation process.
- It effectively reduces the deviations by using road network distance for POI recommendation.

2 SYSTEM ARCHITECTURE

In this section, we introduce the architecture and background technologies of R2SIGTP. The architecture of R2SIGTP is shown in Figure 1. The detailed description is as follows.

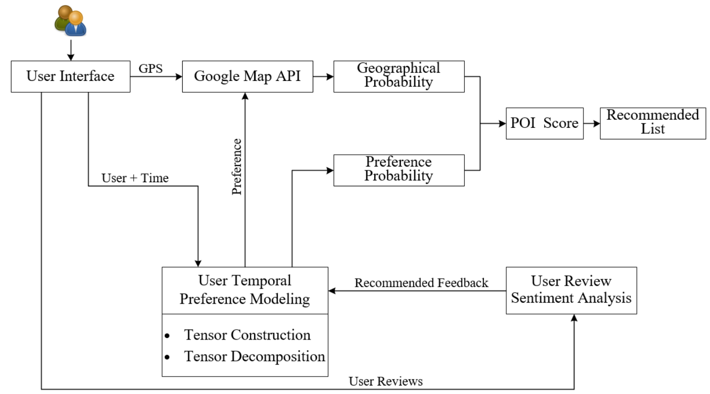

Figure 1: Architecture of R2SIGTP

2.1 User Temporal Preference Modeling

User temporal preference modeling is an important part of the system. The accuracy of users' preferences predicted in different time slots determines the accuracy of our system.Our system uses tensor to model user temporal preference for the following reasons:1) tensor is typically used to recover the missing/sparse data through tensor decomposition; 2) another usage of tensor is to isolate and analyze the patterns hidden in a dataset. Next, we will show the details of tensor construction and decomposition.

Tensor Construction: We define a 3 order-tensor $\chi \in \mathbb{R}^{I\times J\times K}$ to model user temporal preference, where I, J and K denote the size of user, time and preference dimension, respectively. Each entry of the tensor $\chi(i, j, k)$ is a sum of check-in frequencies of all the POIs which belong to preference k and are visited by user i in the time slot j. Note that we divide everyday into 6 time slots (i.e. 0:00-7:00, 7:00-9:00, 9:00-12:00, 12:00-14:00, 14:00-18:00, 18:00-0:00). Accordingly, a week is divided into 42 time slots. Each of time slots is represented by a Time ID. In particular, we select preferences instead of the venues used in most tensor models as a dimension, which can greatly overcome the sparsity of check-in data. Because the amount of POIs in the real world is enormous, but only a few hundreds of preferences can describe our real life in detail.

Tensor Decomposition: After the tensor is modeled, our task is to infer missing entries in the tensor. We can extract the latent features of each user, time slot and preference by decomposing the tensor χ into three matrixes A, B, C and a core tensor G. As is shown in Equation (1).

$$\chi \approx G \times_1 A \times_2 B \times_3 C = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} g_{pqr} a_p \circ b_q \circ c_r \qquad (1)$$

Here, $A \in \mathbb{R}^{I\times P}$, $B \in \mathbb{R}^{J\times Q}$, and $C \in \mathbb{R}^{K\times R}$ are the factor matrices. $G \in \mathbb{R}^{P\times Q\times R}$ is the core tensor. There are many methods for tensor decomposition. For our system, as a large number of users continue to access POIs, incremental data will continue to emerge. So our tensor is not static. This requires a fast and scalable tensor decomposition method for parallel processing of incremental data. Our system uses the Sambaten method [3] for tensor decomposition, which allows our system to update the user temporal preference model frequently without interrupting recommendations. Therefore, our system has real-time performance.

2.2 User Review Sentiment Analysis

Sentiment analysis of users' reviews requires knowledge in the natural language processing (NLP) field. Deep learning technology in the field of NLP has shown strong effectiveness for the reason that it satisfies various requirements such as multilayer classification representation and feature representation learning.

Our system uses CNN to analyze the users' reviews in TensorFlow. First, we do data preprocessing. The dataset our system used contains 9566 POI review sentences, half positive and half negative. Next, we construct the model(CNN). The first layer embeds words into low-dimensional vectors. The next layer performs convolutions over the embedded word vectors using multiple filter sizes. We max-pool the result of the convolutional layer into a long feature vector, add dropout regularization, and classify the result using a softmax layer. For details, please refer to Yoon Kim's article[4].

2.3 POI Score Calculation

Generally speaking, there are two main factors considered by users when choosing a POI, preference factor and geographical factor. Specifically, the preference factor denotes the probability of the activities to be chosen by a user in a certain time slot, and the geographical factor represents the influence of distances between the prefer POIs and the user's current location. Therefore, we define POI score as follows:

$$S_{POI} = \alpha * P_p + \beta * P_G \qquad s.t. \quad \alpha + \beta = 1 \tag{2}$$

where, S_{POI} denotes the POI score, P_p is the preference probability, P_G represents geographical probability. α, β are the weights which reflect the priority of preference and geographic factors. Priority can be chosen by the user himself. After obtaining the score of each POI, we will arrange them in a decreasing order and generate a recommendation list for a user. Next, we will show the details of preference probability and geographical probability.

Preference Probability: Given a particular user i and a specific time slot t, we obtain a preference vector $v_{i,t}$ through the approximate tensor $\hat{\chi}$ constructed in subsection 2.1. Normalize $v_{i,t}$ to get $v'_{i,t}$, and each entry $v'_{i,t,j} \in v'_{i,t}$ represents the probability of each preference j which is chosen by the user i in the time slot t.

Geographical Probability: In order to improve the accuracy of the recommendation results, our system uses road network distance instead of Euclidean distance. In reality, Euclidean distance cannot represent the actual distance between two locations. Consider two locations at the two sides of a river, the Euclidean distance is small but the road network distance is relatively large. Arrange entries of $v'_{i,t}$ in a descending order, we select the first x preferences with positive value in $v'_{i,t}$. Then, for each preference, we employ Google Map API to search the corresponding POIs around the user's current location and calculate the road network distances between these POIs and the user's current location. After normalized distance of these POIs, we can generate the geographical probability.

2.4 Recommended Feedback

Users' reviews are the most direct feedback to the POI recommendation results. Most previous systems only provide a binary decision (positive or negative) for a particular POI by analyzing the reviews of all the users. However, such approach is unfair for some users because the analysis results only reflect the preference of the majority and ignore some users with opposite opinions. Therefore, our system focuses on each single user and provides a feedback to the temporal preference model by analyzing his personal reviews, leading to a more personalized recommendation results.

Without loss of generality, a user's review of a POI k would affect the probability of the preference which the POI k belongs to. Thus, our system iteratively adjusts $v'_{i,t,j}$ according to the review of user i to a POI k which belongs to preference j by using Equation (3).

$$v'_{i,t,j} = v'_{i,t,j} \times (1 + s_k * \gamma) \tag{3}$$

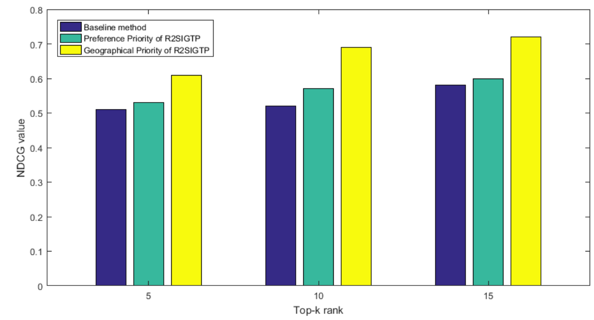

Figure 2: Performance of R2SIGTP and baseline method

where, s_k represents the result of a sentiment analysis of the user i to the POI k and γ is the tuning parameter. When the adjustment is finished, the vector $v'_{i,t}$ is normalized again.

3 EVALUATION

3.1 Dataset and Parameters

We use a collection of Foursquare check-ins of New York lasting for about 11 months (from 8 January 2012 to 27 December 2012). In order to make the recommendation results more accurate, we filter out noise and invalid check-ins. Our dataset contains 238972 check-ins, 37628 venues, and 300 preferences. Based on a lot of experimental tests and data analysis, the parameters of our system are determined as follows:

- $\alpha = 0.7, \beta = 0.3$, if preference is preferred by a user.
- $\alpha = 0.3, \beta = 0.7$, if geography is preferred by a user.
- γ is fixed to 0.1.

3.2 Baseline Method

Our system could provide more personalized POIs recommendation for each single user. The main reason is that our system integrates geographic and preference information into a unified recommendation process. To evaluate the effectiveness of R2SIGTP, we choose the baseline method as a comparison. The baseline approach also takes into account preference and geographical factors, but it does not integrate these two factors into a unified recommendation process. More specifically, the baseline method first ranks the recommendation POIs according to the preference ranking of the user's in current time, and then sorts POIs that belong to the same preference in ascending order of their distance from the current location.

3.3 Recommendation Results Evaluation

In order to accurately evaluate the final recommendation result, R2SIGTP and the baseline method use the same dataset. We selected 200 users to evaluate the performance of the two methods. We generate a Top-k (k=15) recommendation list for every recommendation of each user. We choose Normalized Discounted Cumulative Gain(NDCG) as an evaluation metric to evaluate the performance of the two methods. NDCG is widely used to measure ranking algorithms. Figure 2 shows the NDCG value of R2SIGTP and the baseline method. Since priority can be selected by users in

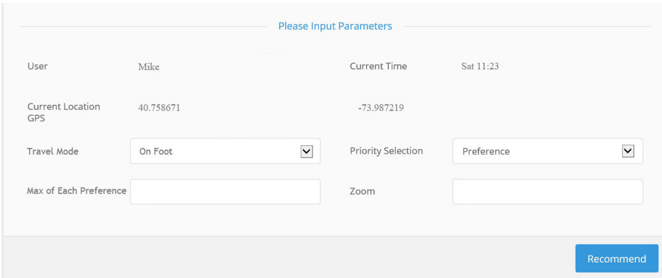

Figure 3: Start page of R2SIGTR

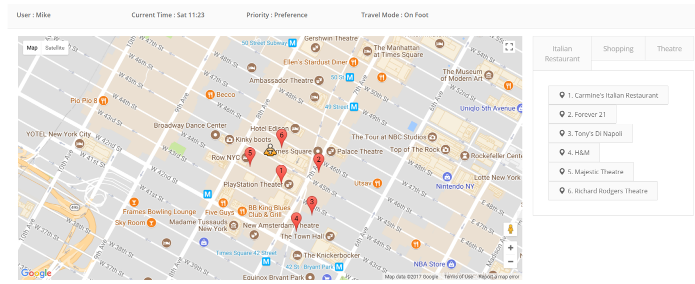

Figure 4: Interface of R2SIGTP

R2SIGTP, we compare the baseline method with R2SIGTP which selected preference priority and geographical priority, respectively. From the results, we can observe that R2SIGTP with preference priority is slightly better than the baseline method, and R2SIGTP with geographical priority is significantly better than the baseline method. This is because R2SIGTP with preference priority is more similar to the baseline method. When R2SIGTP meets the requirements of users to select geographical priority, the baseline method is powerless for users' requirements. Figure 2 demonstrates the final recommendation results of R2SIGTP.

4 DEMONSTRATION

R2SIGTP is composed of a browser frontend and a server backend. The web server is deployed with Apache Tomcat and PHP. Users can use R2SIGTP through a browser on any Internet-enabled terminal.

We show the start page of R2SIGTP in Figure 3. User column needs to fill in a user's account. R2SIGTP can automatically identify the current time and the user's current location and display it in the corresponding position. Priority(preference or geography) and travel mode(on foot or vehicle) can be selected by users. Max of each preference column is used to set the maximum recommended number of POIs for each preference. Zoom column is used to set the scale of the map.

After a user has completed filling and setting, the interface will change to another layout as shown in Figure 4. The user information, the current time, the priority and the travel mode are displayed at the top of the interface. Our system will set different search range for users according to different travel mode. For example, search range of on foot mode is small, and search range of vehicle mode is large. The map which can be zoomed in/out by a user is displayed in the middle of the interface. The map shows the current location of the user and the POIs in the recommendation list. Clicking on each POI on the map will present the corresponding introduction. The right side of the interface shows the top N1 (N1=3 here) preferences and the corresponding recommendation list in current time which shows the top N2 (N2=6 here) POIs.

5 CONCLUSION

In this demo paper, we present a novel real-time recommendation system called R2SIGTP to make next POI recommendation. Compared to the existing systems, our system could provide more personalized POIs recommendation for each single user. The main reason is that our system integrates geographic and preference information into a unified recommendation process. Furthermore, our system also considers the feedback of each single user to adjust the corresponding recommendation. The experimental results show the effectiveness of our system.

ACKNOWLEDGMENTS

This work is supported by the National Nature Science Foundation of China (61170174 and 61702368), major research project of National Nature Science Foundation of China (91646117), Natural Science Foundation of Tianjin (17JCYBJC15200) and Tianjin Science and Technology Correspondent Project (16JCTPJC53600).

REFERENCES

[1] Hakan Bagci and Pinar Karagoz. 2016. Context-Aware Friend Recommendation for Location Based Social Networks using Random Walk. In *Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16)*. April, 2016, Montreal, Canada, 531–536. https://doi.org/10.1145/2872518.2890466

[2] C. Cheng, H. Yang, I. King, and M. R. Lyu. 2012. Fused matrix factorization with geographical and social influence in location-based social networks. In *Proceeding of the 26th AAAI Conference on Artificial Intelligence (AAAI '12)*. July 22-26, 2012, Toronto, Ontario, Canada, 17–23.

[3] Ekta Gujral, Ravdeep Pasricha, and Evangelos E Papalexakis. 2018. SamBaTen: Sampling-based Batch Incremental Tensor Decomposition. In *Proceedings of the 2018 SIAM International Conference on Data Mining (SDM '18)*. November, 2018, Singapore, 387–395. https://doi.org/10.1137/1.9781611975321.44

[4] Yoon Kim. 2014. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP '14)*. October, 2014, Doha, Qatar, 1746–1751. https://doi.org/10.3115/v1/d14-1181

[5] Mao Ye, Peifeng Yin, Wang Chien Lee, and Dik Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information (SIGIR '11)*. July, 2011, Beijing, China, 325–334. https://doi.org/10.1145/2009916.2009962

[6] Jia Dong Zhang and Chi Yin Chow. 2013. iGSLR: personalized geo-social location recommendation:a kernel density estimation approach. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '13)*. November, 2013, Orlando, Florida, USA, 334–343. https://doi.org/10.1145/2525314.2525339

[7] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R Lyu, and Irwin King. 2016. STELLAR: Spatial-Temporal Latent Ranking for Successive Point-of-Interest Recommendation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI '16)*. February 12-âĂŞ17, 2016, Phoenix, Arizona, USA, 315–321. https://doi.org/10.1007/978-981-13-1349-3_5

[8] Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei Ying Ma. 2011. Recommending friends and locations based on individual location history. *ACM Trans. Web.* 5, 1, Article 5 (February 2011), 44 pages. https://doi.org/10.1145/1921591.1921596