Exploiting Meta-Path with Attention Mechanism for Fine-Grained User Location Prediction in LBSNs

Zhixiao Wang\textsuperscript{1,2,3,4,a}, Wenyao Yan\textsuperscript{5,b}, Wendong Wang\textsuperscript{6,c}, Ang Gao\textsuperscript{7},
Lei Yu\textsuperscript{1}, Shaobo Yang\textsuperscript{1} and GaoYang Nie\textsuperscript{1}

\begin{itemize}
\item School of Computer Science and Engineering, Xi’an University of Technology, 710048 Xi’an, China
\item Telematics Group, the University of Goettingen, 37077 Goettingen, Germany
\item Shaanxi Key Laboratory of Network Computing and Security, 710048 Xi’an, China
\item School of Electronic and Information Engineering, Xi’an Jiaotong University, 710048 Xi’an, China
\item College of Xi’an Innovation, Yan’an University, 710100 Xi’an, China
\item Yan’an University, Yan’an 716000, China
\item National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081, China
\end{itemize}

E-mail: \textsuperscript{a}wangzhx@xaut.edu.cn, \textsuperscript{b}yadxywy@126.com, \textsuperscript{c}yadxwd@163.com

Abstract. Based on the huge volumes of user check-in data in LBSNs, users’ intrinsic mobility patterns can be well explored, which is fundamental for predicting where a user will visit next given his/her historical check-in records. As there are various types of nodes and interactions in LBSNs, they can be treated as Heterogeneous Information Network (HIN) where multiple semantic meta-paths can be extracted. Inspired by the recent success of meta-path context based embedding techniques in HIN, in this paper, we design a deep neural network framework leveraging various meta-path contexts for fine-grained user location prediction. Experimental results based on two real-world LBSN datasets demonstrate the best effectiveness of the proposed approach using various evaluation metrics than others.

1. Introduction

The rapid advancement of smart mobile terminals and communication technologies has witnessed the prosperity of Location-based Social Networks (LBSNs) in recent years. On their heels, the typical LBSN platforms such as Foursquare and Yelp are also attracting thousands of users joining in every day. Through the location-aware services provided by these platforms, a large amount of user data (e.g. comments on certain restaurants and check-in records at some scenic spots.) which is associated with geo-tagged identifiers has been generated. Such data is rich in temporal-spatial and semantic information, which can be exploited for mining user mobility patterns and thus predicting where a user will go for his/her next visit. The application of user location prediction has proved beneficial for intelligent travel route recommendation \cite{1, 2}, personalized product advertising \cite{3, 4}, to name a few.

Although predicting users’ next visit location has been one of hot issues in recent years and lots of embedding frameworks \cite{5–8} have been proposed to learn users’ distributed representation and locations in the latent space, these literatures mainly treat LBSNs as the homogeneous networks, which neglects the hidden semantics in interactions between different types of nodes.
Figure 1. The schematic illustration of LBSNs and meta-path contexts in a HIN manner. Specifically, in Figure 1(a), there are three types of nodes, i.e. user node, location node and attribute node, and there exist four types of relations, namely, User-User, User-Location, Location-Location and Location-Attribute. In Figure 1(b), we extract four types of meta-paths for the \( <u_i, l_j> \) pair, and an example for \( <u_1, l_2> \) pair is given.

Indeed, as noted in [9, 10], LBSNs are heterogeneous networks where different types of nodes and relationships can be observed. Therefore, the heterogeneous information network (HIN) embedding techniques, which are newly emerging learning frameworks with flexibility in characterizing diverse heterogeneous data [11], should be applicable for modeling interactions between user and location nodes in LBSNs.

To illustrate the motivation, we present a visual example in Figure 1. As we can see from Figure 1(a), there are three types of nodes in LBSNs, i.e. user nodes, location nodes and attribute nodes. Specifically, we use the venue category information associated with a venue in LBSNs as the corresponding attributes. LBSNs platforms such as Foursquare divide all Point-of-Interests (POI) into 10 categories. Following this convention, we give each category an abbreviation with one capital letter, where ‘A’ stands for ‘Arts & Entertainment’, ‘C’ stands for ‘College & University’, ‘F’ stands for ‘Food’, etc. Besides, there are four types of relations in LBSNs, i.e. user-user relation which represents social friendships, user-location relation which indicates check-in behavior, location-location relation which reflects successive transition pattern or geographical influence, and location-attribute relation which indicates what specific type a location belongs to. With regard to meta-path contexts, as we focus on the challenge of user location prediction, we only explore the interactions between users and locations, i.e. meta-paths for \( <u_i, l_j> \) pairs. Considering that long meta-paths are likely to introduce noise [12], we only select short meta-paths with sequence length up to four, and for a specific user-location pair \( <u_i, l_j> \), four types of meta-paths are extracted, i.e. ‘UUL’, ‘ULL’, ‘ULUL’ and ‘ULAL’. The detailed example of the meta-path contexts for \( <u_1, l_2> \) pair can be seen in Figure 1(b). As can be seen, we merge four meta-path contexts in an interaction between a user and a location, in other words, we use the triple-tuple \( <u_i, meta-paths, l_j> \) to characterize the contexts why user \( u_i \) visits venue \( l_j \). To address user location prediction problem, we conceive a deep neutral network framework incorporating meta-path contexts as well as an attention mechanism.

The remainder of this paper is scheduled as follows. Section 2 briefly reviews related works. Section 3 introduces the deep neutral network framework. In section 4, we evaluate the prediction performance, and finally section 5 concludes our work.

2. Related Works
The location prediction of users has been a tough and excited problem during these years. Quite a lot of previous works deal with this problem based on collaborative filtering (CF) strategies
such as matrix factorization [13–15]. The basic objective of matrix factorization methods is to factorize a user-location matrix into two low rank matrices, with each matrix representing the latent representation of users and locations, respectively. Then a user’s visit probability at a candidate location is fitted by the inner product of two vectors. However, CF based methods usually suffer from cold-start problem, which means they are generally infeasible for new users or new locations in the training data.

In recent years, embedding techniques, especially Heterogeneous Information Network (HIN) based embedding techniques [11, 12, 16] have drawn much attention from researchers. As it is able to depict semantic relations between different kinds of nodes using meta-paths [17], HIN embedding is becoming increasingly popular in venue recommendation, and thus can be effective for location prediction with minor modification. Hu et al. [12] propose to characterize a three-way interaction \(< u, \text{meta-paths}, l \rangle\) for semantic item recommendation. Particularly, they use a co-attention mechanism to capture the weights for each meta-path context. Inspired by [12], we conceive a deep neural network framework incorporating meta-path contexts as well as an attention mechanism for user location prediction. Unlike [12], we adopt a quite different random walk method to sample meta-path instances based on predefined rules, and simplify the attention mechanism for meta-path contexts learning. Besides, as we only have positive samples due to implicit feedback in user check-in data, we use a pairwise learning algorithm to maximize the margin between visited and unvisited locations, after which the involved parameters can be optimized.

3. Methodologies

3.1. Preliminaries

In this paper, we strive to predict the next visit location for each user relied on his/her historical check-in records in LBSNs. For a given LBSN dataset, suppose we have \(M\) users, i.e. \(\mathcal{U} = \{u_1, u_2, ..., u_i, ..., u_M\}\), and we have \(N\) locations, i.e. \(\mathcal{L} = \{l_1, l_2, ..., l_i, ..., l_M\}\). As a social network dataset, it also contains social relations (e.g. friendship relation) among some users. A venue in LBSN is defined as a uniquely identified site such as a cinema or a coffee shop, and it is associated with geographical coordinates as well as the category information indicating which type of location it belongs to. Depended on the data, we can construct a user-venue check-in matrix \(X \in \mathbb{R}^{M \times N}\), where each entry \(x_{u,i} \in \{0, 1\}\) represents whether user \(u\) has visited location \(i\). Given the above preliminaries, we can formulate the location prediction problem to be solved in this work.

**Definition 1. User Next Location Prediction.** Given the implicit feedback matrix \(X \in \mathbb{R}^{M \times N}\) with check-in records of each user \(u\) being \(C_u = \{< u, l_1, t_1 >, < u, l_2, t_2 >, < u, l_3, t_3 >, ..., < u, l_n, t_n >\}\), where each \(< u, l_i, t_i >\) denotes user \(u\) check-in at location \(l_i\) and time \(t_i\). We strive to predict the next venue where user \(u\) will visit after time \(t_n\). To do so, we sort all the \(N\) feasible locations, thus the precise location that \(u\) will be visited next is sorted at the supreme position of the ranking list.

3.2. Our Model

Following the main idea of [12], we also leverage meta-path contexts to model the semantics in a user-location pair \(< u, l_j >\). Specifically, we extract four meta-paths (see Figure 1(b)) and fuse the meta-path contexts with user and location embeddings through an attention mechanism. Due to the incorporation of such semantic information, the model is expected to deliver more interpretability than previous works. The overall framework of the deep neural network can be seen in Figure 2.
Figure 2. Framework of the deep neural network for user location prediction, where ‘⊕’ means the vector concatenation operation, and $\hat{x}_{i,j}$ represents the approximated real value of user $u_i$’s preference for location $l_j$. Besides, $m_{i\rightarrow j}$ indicates the aggregate meta-path contexts embedding for the input $< u_i, l_j >$ pair.

4. Experiments

4.1. Datasets

We utilize two real world LBSN datasets for experimental evaluation, both of which are crawled from Foursquare. The first Foursquare dataset is provided by Xu et al. [18] and it includes 10,901 users who live in New York (NY) with more than 764,328 check-ins in total. The second Foursquare dataset comes from Yin et al. [19] and it contains the check-in history of 4,163 users who live in California (CA), USA. Both of the datasets contain explicit undirected friendship relations. To filter the noisy data in both datasets, we eliminate users who have less than 10 check-ins, as well as venues with fewer than 10 visitors. The basic statistics of both datasets are shown in Table 1. Note that for each user in the datasets, we use his/her first 80% check-ins as the training data, the following 10% check-ins for validation, and the last 10% check-ins for testing. The overall framework is implemented using Keras with TensorFlow as the backend, where all hyper-parameters are set according to [12].

| Dataset | # Users | # Venues | # Check-ins | # 'U $\rightarrow$ U' | # 'U $\rightarrow$ L' | # 'L $\rightarrow$ L' | # 'L $\rightarrow$ U' | # 'A $\rightarrow$ L' |
|---------|---------|----------|-------------|----------------------|----------------------|----------------------|----------------------|---------------------|
| NY      | 7,678   | 13,045   | 983,984     | 48,917               | 595,743              | 170,172,025          | 595,743              | 13,045              |
| CA      | 2,259   | 3,244    | 111,817     | 5,266                | 68,439               | 10,523,536           | 68,439               | 3,244               |

4.2. Evaluation Metrics

In this work, we adopt two commonly used evaluation metrics in user location prediction [9], one of which is $\text{Acc}@N$ and the other one is $\text{APR}$. With regard to $\text{Acc}@N$, it’s often employed to evaluate the recommendation or predication accuracy. Once sorting that location in the top-$N$ list, we will be able to predict the next visit location for each user successfully. Herein, $\text{Acc}@N$ is calculated by

$$\text{Acc}@N = \frac{\#\text{hit}@N}{|L|},$$

where $\#\text{hit}@N$ indicates the hits number over the all test set. We take the mean value in the whole test cases as the final result.
With regard to APR, it’s usually used to estimate the extent of how one coming check-in location can be precisely sorted. Herein, \( PR \) is calculated by
\[
PR = \frac{|L| - \text{rank}(k)}{|L|},
\]
where \(|L|\) is the total number of candidate locations and \( \text{rank}(k) \) is the rank of the ground-truth location in the list. We again take the mean value in all the prediction instances as the final one, i.e. Average PR (APR).

4.3. Comparison Methods
We select two models as comparison, i.e. (1) PRME [20], which embeds user and location into the same latent space to capture the user transition patterns. The geographical influence is incorporated through a simple coefficient. (2) GE [6], which jointly learns the embeddings of locations, regions as well as venue categories.

4.4. Result Analysis
First of all, we compare the location prediction performance over all methods. To distinguish our model from other comparison models, we denote it as MP_Loc_Pre, which is short for Meta-Path based Location Prediction. Table 2 summarizes the numerical results in terms of Acc@10, Acc@50 as well as APR. As we can see, for one thing, location prediction models considering meta-path contexts achieve the highest prediction performance at all facets.

|          | NY Acc@10 | NY Acc@50 | NY APR | CA Acc@10 | CA Acc@50 | CA APR |
|----------|------------|------------|--------|------------|------------|--------|
| MP_Loc_Pre | 0.283      | 0.411      | 0.773  | 0.248      | 0.347      | 0.830  |
| PRME     | 0.226      | 0.336      | 0.762  | 0.174      | 0.242      | 0.755  |
| GE       | 0.257      | 0.373      | 0.768  | 0.220      | 0.305      | 0.824  |

5. Conclusion
In this paper, we investigate the issue of user next location prediction depended on user check-in data in LBSNs. Inspired by the recent success of HIN embedding techniques, we conceive a deep neural network framework incorporating meta-path contexts and a simple attention mechanism. To deal with the implicit feedback, we optimize the model with negative sampling and use a pair-wise training method to maximize the margin between visited and unseen locations. Experimental results verify the effectiveness of the proposed approach as it can achieve comparable and even better performance than PRME and GE.

Acknowledgments
The work is supported in part by the grants which include: Youth Fund of the Ministry of Education of China (No.16YJCZHI09), Scientific Research Program Funded by Shaanxi Provincial Education Department (Program No.16JS080), College Innovation Research and Development (No. 2016JXKY-20), YANAN Science and Technology Project (No. 2018KG-02), Xian Science technology (No.2017080CG/RC043/XALG027), the CSC201608615074, NSFC (No.61772407, No.61771387), Shaanxi Social Science (No.2014M18), Shaanxi Department of Education (No.18JK1216), CERNET Innovation Project (NGII20171202, NGII20170303), Ph.D Innovation (No.112-451115006), College Innovation Research and Development (No.310/252051835).
References

[1] Jiang Shuhui, Qian Xueming, Mei Tao and Fu Yun 2016 Personalized travel sequence recommendation on multi-source big social media IEEE Transactions on Big Data 2 43-56.

[2] Arain Qasim, Memon Hina, Memon Imran, et al. 2017 Intelligent travel information platform based on location base services to predict user travel behavior from user-generated GPS traces. International Journal of Computers and Applications, 39 155-68.

[3] Zhu Yu and Li Hao and Liao Yikang, et al. 2017 What to do next: modeling user behaviors by time-lstm. In Proc. of IJCAI 3602-08.

[4] Meng Xiangwu, Li Ruichang, Zhang Yujie and Ji Weiyu 2018 Survey on mobile recommender systems based on user trajectory data. Ruan Jian Xue Bao/Journal of Software(in Chinese), Chinese Academy of Sciences, 29 3111-33.

[5] Jarana Manotumruksa, Craig MacDonald and Iadh Ounis 2016 Modelling user preferences using word embeddings for context-aware venue recommendation. CoRR http://arxiv.org/abs/1606.07828.

[6] Xie Min, Yin Hongzhi, Xu Fanjiang, et al. 2016 Graph-based metric embedding for next poi recommendation. In Proc. of the International Conference on Web Information Systems Engineering 207-22.

[7] Xie Min, Yin Hongzhi, Wang Hao, et al. 2016 Learning graph-based poi embedding for location-based recommendation. In Proc. of the 25th ACM International on Conference on Information and Knowledge Management, 15-24.

[8] Long Yan, Zhao Pengpeng, Sheng Victor S, et al. 2017 Social personalized ranking embedding for next poi recommendation. In Proc. of the International Conference on Web Information Systems Engineering, 91-105.

[9] Cao Jiuxin, Xu Shuai, Zhu Xuelin, et al. 2018 Effective fine-grained location prediction based on user check-in pattern in lbsns. Journal of Network and Computer Applications, Elsevier, 108 64-75.

[10] Cao Jiuxin, Xu Shuai, Zhu Xuelin, et al. 2017 Efficient fine-grained location prediction based on user mobility pattern in lbsns. In Proc. of the 5th International Conference on Advanced Cloud and Big Data, 238-43.

[11] Shi Chuan, Binbin Hu, Wayne Xin Zhao and Philip S. Yu 2019 Heterogeneous information network embedding for recommendation. IEEE Transactions on Knowledge and Data Engineering, 31 64-75.

[12] Hu Binbin, Shi Chuan, Wayne Xin Zhao and Philip S. Yu 2018 Leveraging meta-path based context for top-n recommendation with a neural co-attention model. In Proc. of the 24th ACM SIGKDD international conference on knowledge discovery and data mining, 1531-40.

[13] Lian Defu, Zhao Cong, Xie Xing, et al. 2014 GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In Proc. of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, 831-40.

[14] Manotumruksa Jarana, Macdonald Craig and Ounis Iadh 2016 Regularising factorised models for venue recommendation using friends and their comments. In Proc. of the 25th ACM International on Conference on Information and Knowledge Management, 1981-84.

[15] Zhao Shenglin, Zhao Tong, Yang Haiqin, et al. 2016 STELLAR: spatial-temporal latent ranking for successive point-of-interest recommendation. In Proc. of the 13th AAAI conference on artificial intelligence, 315-21.

[16] Fu Taoyang, Lee Wang-Chien and Lei Zhen 2017 HIN2Vec: explore meta-paths in heterogeneous information networks for representation learning. In Proc. of the 26th ACM International on Conference on Information and Knowledge Management, 1797-806.
[17] Dong Yuxiao, Nitesh V. Chawla and Ananthram Swami 2017 Metapath2Vec: scalable representation learning for heterogeneous networks. In Proc. of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 135-44.

[18] Xu Shuai, Cao Jiuxin, Zhu Xuelin, et al. 2018 Community discovery based on social relations and temporal-spatial topics in lbsns. In Proc. of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, 206-17.

[19] Yin Hongzhi, Hu Zhiting, Zhou Xiaofang, et al. 2016 Discovering interpretable geo-social communities for user behavior prediction. In Proc. of the 32nd International Conference on Data Engineering (ICDE), 942-53.

[20] Feng Shanshan, Li Xutao, Zeng Yifeng, et al. 2015 Personalized ranking metric embedding for next new poi recommendation. In Proc. of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015), 2069-75.