ISSN 1533-9211

DOI 10.5281/zenodo.7106546

AN EFFICIENT POINT-OF-INTEREST RECOMMENDATION FOR LOCATION-BASED SOCIAL NETWORKS SYSTEM WITH SPATIO-TEMPORAL MODEL

[1]VENKATA PRAVEEN KUMAR VUPPALA, 2Dr. S.PAZHANIRAJA and [3]Dr. INDRANEEL SREERAM

[1]Research scholar, Department of CSE, Annamalai University, Annamalai, Tamil Nadu, India.
[2]Research Guide, Assistant Professor, Department of CSE, FEAT, Annamalai University, India.
[3]HOD & Professor, Department of Cyber security & IOT, St. Ann's College of Engineering and Technology, Chirala, India.

ABSTRACT:

The problem of personalized next point-of-interest (POI) recommendation is significant and of practical value in location-based social networks (LBSNs). Twitter together with other online social networks has begun to collect hundreds of millions of check-ins, which capture the spatial and temporal information of user movements and interests. Due to the sparsity of data about check-in, POI recommendations remain a challenging problem. An efficient point-of-interest recommendation system with Upstream Spatiotemporal Topic Model (USTTM) for LBSNs is proposed in this paper. This recommendation system contains purpose prediction phase that classifying the POIs in spatio-temporal database based on purpose and constructing a purpose ranking model to model, the selection of user's intended purpose for the trip. Second phase is scoring of each candidate POI that considers the properties of spatial and temporal information when calculating the score. USTTM is used in this system to model and analyze the spatio-temporal aspect of check-in data that can discover a user's choice of region. Extensive experiments were conducted and the results demonstrate that the recommendation accuracy of the model outperforms the state-of-the-art POI recommendation models with good runtime performance. In quantitative analysis, effectiveness of USTTM in terms of accuracy of POI recommendation and accuracy of user and time prediction are evaluated and results show that the USTTM achieves better performance than the state-of-the-art models.

KEYWORDS:point-of-interest (POI), Recommendation system, Spatiotemporal Topic Model, USTTM, purpose-ranking model.

I. INTRODUCTION

Recent years have witnessed the increased development of location-based social networking (LBSN) services, such as Foursquare, Facebook Places and Google Latitude [1]. LBSNs allow users to explore

Places-of-Interests (POIs) for better services through sharing check-in experiences and opinions on the POIs they have checked in [2]. In particular, LBSNs are a special kind of social Web systems where it is possible for users to register whenever they visit a specific Point-of-Interest (POI) through the so-called check-ins, or to establish social links with other users in the system. Indeed, the task of Point-of-Interest (POI) recommendation is to provide personalized recommendations of places of interest [3]. It plays an important role in providing better location based services in location based social networks.

The Seybold REPORT
ISSN 1533-9211
DOI 10.5281/zenodo.7106546

Point-of-interest (POI) recommendation has become an increasingly important sub-field of recommendation system research [4]. Users can categorize POI to help describe what type of places this POI is; tag a POI to let people know what they can expect from it; share their experiences of check-ins with others; know how many people have visited a specific POI and how much time they spent there [5]. Here a check-in comprises of which POI is visited and additional contextual information such as the time or geotag of the visit. Finding an efficient way to represent the POI and its contextual information is essential because this can improve the performance of the model and allow a better understanding of the seemingly complex inter-relationships of the heterogeneous properties of the POIs. It is worth mentioning that, even though the area of POI recommendation is of great interest to researchers because it allows studying the behavior and movement patterns of users, it is also appealing for companies and businesses in the tourism, leisure, and e-commerce domains, as they seek to attract and maintain customers by becoming popular and receiving good reviews [6]. It has been shown that human movements usually demonstrate strong patterns in both spatial and temporal domain. To take advantage of the spatiotemporal nature of check-ins, several recommendation systems have been proposed particularly for POIs [7].

In this paper, an efficient POI recommendation system with Upstream Spatio-Temporal Topic Model for LBSNs is presented. Time is a factor that influences and generates other factors in actual life, instead of being generated by other factors. Therefore, time should influence the topic distribution and region distribution of a user directly. In order to capture this property, an upstream model is used where the time variable is upstream of the region and topic variables, called Upstream Spatiotemporal Topic Model (USTTM). Each user has a unique topic distribution and region distribution per time, which enables USTTM to model user interests and moving preferences that change over time. The temporal property of user's check-in behaviour having memory property is considered in this especially, memory means that people will revisit the same location in near future. The memory property has an important characteristic in terms of the time interval w.r.t. events. So a new framework of purpose ranking model is also introduced in the POI recommendation system.

II. LITERATURE SURVEY

Due to the growing interest in the general recommendation area on this domain, there are a considerable number of surveys related to POI recommendation. On the one hand, several works were there which cannot be considered to be up to date anymore, since they were published 5 years ago, thus our survey should provide a novel overview of the works developed in this time. On the other hand, Christoforidis et al. [8], focus on deep learning techniques while neglecting the other types of recommendation algorithms. Finally, since our analysis is also tailored towards the evaluation aspects of the works, it is worth mentioning those reviews where this aspect has been considered. However, we must acknowledge that we could not find any survey that focused on this particular aspect; because of that, we consider our survey is very valuable in this domain now.

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7106546

In the literature, approaches like latent factor model and Markov chain have been widely applied for sequential data analysis and recommendation. For instance, in [9] Zheng et al. consider the problem of location prediction but only based on Twitter information. To somehow overcome this shortcoming, we believe it is important to mention the experimental comparison presented by Liu et al. in [10], where they compared 12 recommendation models under different evaluation protocols and using three datasets, which could help to analyze the behavior of those methods under the same and different conditions. Zhu et al. [11] proposed a variant of Long-Short Term Memory network (LSTM), called Time-LSTM, to equip LSTM with time gates to model time intervals for next item recommendation. He et al. [12] proposed a unified tensor-based latent model to capture the successive check-in behavior by exploring the latent pattern-level preference for each user. Recently, Recurrent Neural Networks (RNNs) have been successfully employed to model sequential data and become state-of-the-art methods. Hidasi et al.(2015) focused on RNN solutions for session-based recommendation task, where no user id exists, and recommendations are made on short session data.

Gavalas et al. present in [13] an overview of optimization approaches that aim to solve a problem with applications on related tasks: the Tourist Trip Design Problem (TTDP); this can be applicable to route recommendation, which, as we specify in the next section, is not completely in the scope of this survey. Additionally, we found some surveys that were too focused on specific subproblems. Cheng et al. [14] extended FPMC to embed personalized Markov chain and user movement constraint for next POI recommendation. Rendle et al. [15] proposed Factorizing Personalized Markov Chain (FPMC), which bridges matrix factorization and Markov chains together, for next-basket recommendation.

III. AN EFFICIENT POI RECOMMENDATION SYSTEM WITH USTTM MODEL

To implement an efficient recommendation system, a POI recommendation model by taking into consideration spatiotemporal effects based on USTTM and purpose ranking (STPR) which has two essential steps is considered. The first step is to construct a model to predict user's next trip purpose based on user's historical check-in data. The second step is to calculate the score of each candidate POI. The score function combines with spatial and temporal properties of user's check-in behavior. It is worth to note that the spatial and temporal properties are learned based on user's historical check-in data. Lastly, the top-k POIs are recommended to users.

3.1 Check-in Data Analysis

Each checkin record includes check-in time, coordinate of locations, marks of locations and user's ID. The users whose check-in records are more than 200 times are chosen.

3.2 USTTM

Different users have different interest that is topic preferences. On the other hand, different users have different activity spaces that are region preferences. The topic will affect the check-in POI of the user. The region will influence user's check-in position, and further

ISSN 1533-9211
DOI 10.5281/zenodo.7106546

influence the POI. Thus, the checkin POI will be different, when regions are the same but topics are different. Similarly, the check-in POI will be different, when topics are the same but regions are different. Instead of being generated by other factors, it is believes that time should be a factor that influences or generates other factors. Therefore, time should generate the topic distribution $p(z|\theta u)$and region distribution $p(r|\phi_u)$ per user directly. At different times, users will have different topic preferences. Moreover, user will also have different region preferences at different times. That means that the topic and region of a user will depend on time. $p(r|\phi_u)$or$p(z|\theta u)$ may be different at different times. The more often a user checks in at region r, the higher $p(r|\phi_u)$. Similarly, the more often a user checks in at topic z, the higher $p(z|\theta u)$. Thus, posteriors $\phi_u$and θu can obtain for each user.

In USTTM, each user has different topic distributions and region distributions at different times. However, continuous time would lead to an infinite number of topic distributions $p(z|\theta u, t)$ and region distributions $p(r|\phi_u,t)$, where t is the check-in time, which would make it very difficult to infer parameters. In order to solve this problem, a set of latent time intervals scan be learned. Given any check-in time, we sample a time interval using Gaussian distributions and Bayesian rules. There are two parameters for each latent time interval, i.e. the mean and variance of the distribution. The parameters of latent intervals are learnt from the data, instead of specified by user input. Specifically, the probability distribution of check-in time drawn by the i^{th} latent interval is a Gaussian distribution. Using Bayes rule, given a check-in time $t_{u,i}$, the probability that the check-in time belongs to the latent time interval sj. Therefore, θ and φ are three-dimensional matrixes with respect to time interval s and user u. In USTTM, s and u can be regarded as parameters of θ and φ.

ISSN 1533-9211

DOI 10.5281/zenodo.7106546

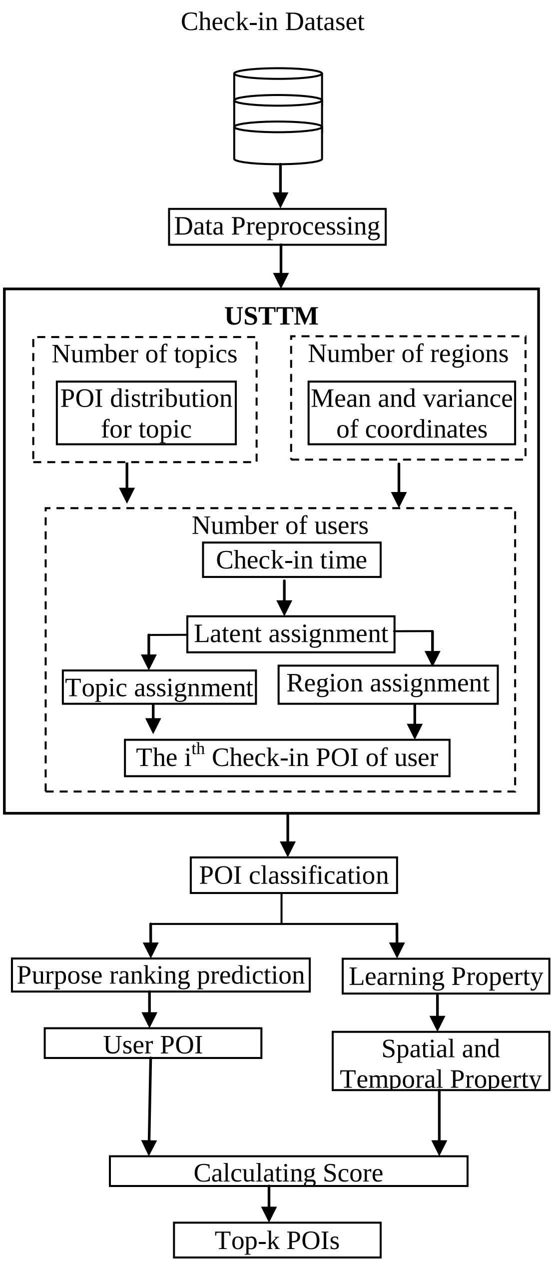

Fig. 1: Framework ofPOI Recommendation System Using USTTMmodel

The input of this model are the number of topics, the number of regions, the number of latent time intervals, and the check-in dataset, while the output of this model is the topic distribution of a user per time interval, region distribution per user per time interval, topics and regions with their POI distributions, together with the distribution of each latent time interval. To learn the various parameters of USTTM collapsed Gibbs sampling can be used to perform approximate inference by computing the distribution of regions and topics

ISSN 1533-9211

DOI 10.5281/zenodo.7106546

independently in the Gibbs sampling process. After the sampling has run for an appropriate number of iterations (until the chain has converted to a stationary distribution), the estimation of topic distribution θu, region distributionϕ_u, and POI distribution η_z, can be obtained. Note that USTTM can integrate out the time distribution for each topic or region. Specifically, each check-in will have its topic/region after Gibbs sampling. All check-in times and their sampled topic/region into the time distribution can be integrated.

3.3 Spatial and Temporal Analysis

In this section, the temporal and spatial property of user's check-in behavior is mainly analyzed. In terms of temporal property, the time interval of check-ins at the same POI is concentrated. This considers the information of one user's time interval at POI. Intuitively, the user visited this POI for several times. In addition, the user occasionally visits this place in a short period of time and sometimes visits it after a long time. This character is called "memory". The method of calculating the value of memory M∈[-1,1] of the time interval sequence is used. The memory of check-in sequences at POI is M<0 reflects the sequence has the property of weak memory where as M>0 corresponds to memory. In addition, the distribution time interval w.r.t. POI is used which is a lognormal distribution with μ=0.683 and σ=0.865 in logarithm.

The spatial property for user's check-in behavior can be analyzed based on visiting frequency and the distance. The visiting frequency of a POI is the times of visiting this place by a user. For each user in the dataset, the visiting frequencies are record for all POIs being visited, and find that some users' check-in locations have one center, some have two or three centers. It indicates that people's activities are often within an area. This phenomenon is called "boundness". In addition, visiting some new locations is also a common phenomenon for users. The central POIs are those POIs whose check-in frequency is in the top 5 percents among all visited POIs. The average distance which demonstrates that user would like to select nearby POIs.

3.4 Purpose Ranking

Based on the classification, the sequence of POIs can be viewed as a purpose sequence since each POI belongs to a purpose (category). Four categories correspond to four purposes. According to this idea, the problem of inferring the visiting probability of each POI can be transformed into calculating the visiting probability of each trip purpose. The time interval v_t^m of each trip purpose is calculated as follows. We assume the following application scenario related to user's travel behavior. A user already has his purpose before a trip. Then he chooses a POI to achieve the purpose. This is a purpose selection problem, it is model as a purpose ranking $m >_u n$, which implies user u prefers purpose m to n. The best purpose ranking for user u can be formalized as below:

$p(\Theta|m >_u n) \propto p(m >_u n|\Theta)p(\Theta)$-- (1)

Where, Θ is the set of parameters. The purpose ranking probability can be calculated by the following equation:

ISSN 1533-9211
DOI 10.5281/zenodo.7106546

$p(m >_u n|\Theta) = \sigma(p(v_t^m) - p(v_t^n))$-- (2)

Where σ is a logistic sigmoid function as follows:

$\sigma(x) = \frac{1}{1+e^{-x}}$ ------- (3)

Where$p(v_t^m)$and$p(v_t^n)$ are the probability of time interval v_t^m on purpose m and time interval v_t^n on purpose n, respectively. Specially, v_t^m is the time interval on m between $t-1$and t. The last check-in time slot is denoted as $t-1$, and t is the time slot at which we will make a POI recommendation. Owning to the weak memory property of check-ins, the next check-in time is associated with recent check-ins. Hence, for each purpose, the relationship between time interval v_t^m and the recent check-ins can be formalized by the following autoregressive model:

$v_t^m = \varepsilon^m + \sum_{i=1}^{w} a_i^m . v_{t-i}^m + c^m$ ----- (4)

Where a_i represent the i^th parameter, c^m is a constant, and ε m is Gaussian white noise with zero mean and variance $\sigma_m^2 . v_t^m$ is the sum of the random variableε^m and the constant $\sum_{i=1}^{w} a_i^m . v_{t-i}^m + c^m$. Thus, the probability of the case that a check-in record belongs to the category purpose) m occurs after v_t^m days is calculated by the following equation:

$p(v_t^m) = \mathcal{N}(\sum_{i=1}^{w} a_i^m . v_{t-i}^m + c^m, \sigma_m^2)$-- (5)

According to Eq. 5, the parameter set of purpose ranking model Θ={a^m, c^m, σ^m, where am is represented by a parameter vector $< a_1, a_2, .., a_w >$ w.r.t. purpose m.

IV. RESULT ANALYSIS

This section evaluates the effectiveness of USTTM and purpose ranking model in POI recommendation system on realworld data sets. The experimental results can be reported on Gowalla data set, where all check-ins contain a POI label, geographical coordinates, user id and check-in time. In our data sets, we select the first 80% of observed data for each user as the training data, and the remaining 20% as the test data, according to check-in time. We remove both POIs and users with only 1 check-in. Given the historical check-in POIs, the proposed model provides a list of top-k recommendation POIs to users. The following metric is used to evaluate the performance of different models.

$$\text{Precision@n} = \frac{N_r}{T_i}$$

Where, N_r denotes the number of correct predictions, and T_i is the number of testing cases. If the next check-in POI appears in the top-k recommendation POIs, the prediction is viewed as correct.

In this section, the comparison partners are evaluated for user prediction and for time prediction. Most competitive model USTTM is compared with STTM and DSTTM, as comparison partners. The results are reported in Figure table 1 which show the user

ISSN 1533-9211

DOI 10.5281/zenodo.7106546

prediction accuracy results of comparison partners for different numbers of topics on the Gowalla datasets, respectively.

Table1: User prediction accuracy for different number of Topics and regions

		STTM	DSTTM	USTTM
5	Topics	0.1	0.13	0.14
	Regions	0.08	0.11	0.12
10	Topics	0.15	0.135	0.154
	Regions	0.09	0.12	0.14
20	Topics	0.15	0.14	0.154
	Regions	0.12	0.145	0.15
50	Topics	0.122	0.145	0.16
	Regions	0.13	0.148	0.155

It was observe that our model USTTM consistently outperforms all other models. Compared to STTM, the most competitive method, USTTM improves the accuracy by 32.1%. It also observe the similar results of the impact of the number of regions in table 1. The accuracy of all models increases and then plateaus when the number of regions reaches 20 or 30. We note that all comparison partners are far less accurate for user prediction than for POI recommendation. Potential reasons are two-fold: on the hand, the metric of POI recommendation is top-10 accuracy, while the metric of user prediction is top1 accuracy; on the other hand, user prediction is more difficult. While the number of different POIs that a given user checks in at a given time on different days is typically small, the number of different users that check-in at a given POI at a given time can be large.

Figure 2 shows the impact of different numbers of latent time intervals in USTTM. As the number of latent time intervals increases, the accuracy reaches a peak at 4 or 6, and slightly decreases later.

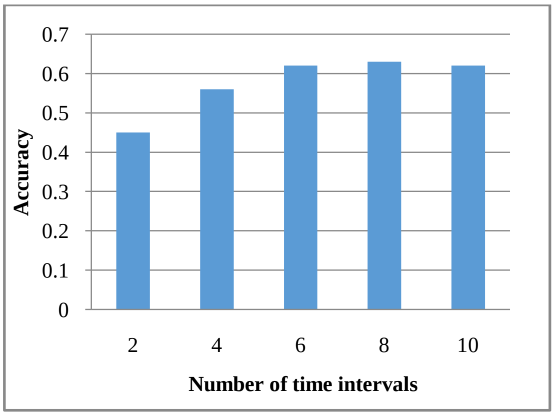

Fig. 2:Precision accuracy on Top-10 POI for different number of latent time intervals

ISSN 1533-9211

DOI 10.5281/zenodo.7106546

In order to demonstrate the impact of the number of check-ins in the training dataset, experiments are conducted for users with different numbers of check-ins. The number of recent check-ins is an important parameter for inferring the next POI. This comparison has not been done in related work. Figure 3 shows the precision analysis results on POI recommendation.

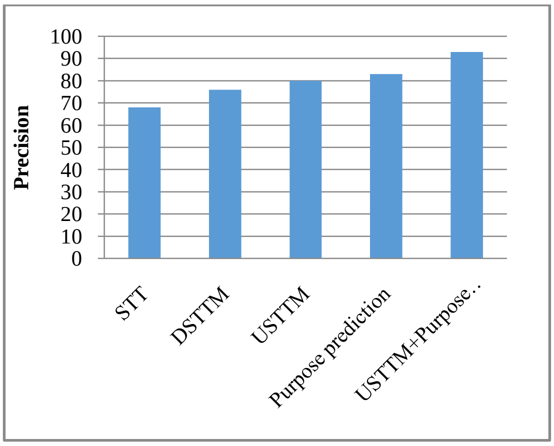

Fig. 3:precision analysis results on POI recommendation

Here analysis is performed on STTM (Spatio-Temporal Topic Modal), Downstream Spatio-Temporal Topic Model (DSTTM), USTTM and Purpose prediction model. In terms of the Gowalla dataset, USTTM performs much better than DSTTM and STTM where as the precision of purpose prediction reach 83% more than USTTM. The proposed USTTM+ purpose prediction acquired a 93% precision accuracy than all other models.

V. CONCLUSION

In this study, a POI recommendation model with spatiotemporal effects based on USTTM and purpose ranking is proposed to cope with the problem of personalized next POI recommendation. In this paper, firstly microscopic spatio-temporal topic models, i.e. USTTM is presented to capture the spatial and temporal patterns of user movements and interests. USTTM can discover temporal topics and regions. USTTM can capture the property that user’s interests and activity space will change over time, and users have different region and topic distributions at different times in this model. A purpose ranking model is also employed that take into full consideration the spatial property when calculating the score of each candidate POI. In our quantitative analysis, effectiveness of models is evaluated in terms of precision accuracy of POI recommendation, and accuracy of user and time prediction. Our results show that USTTM with purpose prediction achieves better performance than the state-of-the-art models, confirming that it is more natural to model time as an upstream variable affecting the other variables.

VI. REFERENCES

The Seybold REPORT

ISSN 1533-9211

DOI 10.5281/zenodo.7106546

[1] Qingbo Sun, “Key Research on Recommendation Algorithms Based on Spatio-temporal Relationships in Location Social Networks”, International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA), 2021

[2] Kang Yang;JinghuaZhu;XuGuo, “POI neural-rec model via graph embedding representation”, Tsinghua Science and Technology, Volume: 26, Issue: 2, 2021

[3] K. Papangelis et al., "Performing the digital self: Understanding location-based social networking territory space and identity in the city", ACM Trans. Comput. Human Interact. vol. 27, no. 1, pp. 1-26, 2020

[4] C. Chen, J. Zhou, B. Wu, W. Fang, L. Wang, Y. Qi, et al., "Practical privacy preserving POI recommendation", ACM Trans. Intell. Syst. Technol., vol. 11, no. 5, pp. 1-20, Sep. 2020.

[5] T.-N. Doan and E.-P. Lim, "Modeling location-based social network data with area attraction and neighborhood competition", Data Mining and Knowledge Discovery, vol. 33, no. 1, pp. 58-95, 2019

[6] Rashmi A. Wahurwagh, P. M. Chouragade, “Personalized POI Travel Recommendation with Multiple Tourist Information”, 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019

[7] T. Qian, B. Liu, Q. V. H. Nguyen and H. Yin, "Spatiotemporal representation learning for translation-based POI recommendation", ACM Trans. Inf. Syst., vol. 37, no. 2, 2019.

[8] Giannis Christoforidis, Pavlos Kefalas, Apostolos N. Papadopoulos, and YannisManolopoulos, “Recommending Points of Interest in LBSNs Using Deep Learning Techniques”, In INISTA, IEEE, 1–6, 2019

[9] Xin Zheng, Jialong Han, and Aixin Sun, “A Survey of Location Prediction on Twitter”, IEEE Trans. Knowl., Data Eng. 30, 9 (2018), 1652–1671

[10] Yiding Liu, Tuan-Anh Pham, Gao Cong, and Quan Yuan., “An Experimental Evaluation of Point-of-interest Recommendation in Location-based Social Networks”, PVLDB 10, 10 (2017), 1010–1021, 2017

[11] Zhu, Y.; Li, H.; Liao, Y.; Wang, B.; Guan, Z.; Liu, H.; and Cai, D., “What to do next: Modeling user behaviors by time-lstm”, In IJCAI-17, 3602–3608, 2017.

[12] He, J.; Li, X.; Liao, L.; Song, D.; and Cheung, W. K, “Inferring a personalized next point-of-interest recommendation model with latent behavior pattern”, In AAAI, 137–143, 2016.

[13] DamianosGavalas, CharalamposKonstantopoulos, Konstantinos Mastakas, and Grammati E. Pantziou, “A survey on algorithmic approaches for solving tourist trip design problems”, J. Heuristics 20, 3 (2014), 291–328, 2014.

[14] Cheng, C.; Yang, H.; Lyu, M. R.; and King, I, “Where you like to go next: Successive point-of-interest recommendation”, In IJCAI, volume 13, 2605–2611, 2013.

[15] Rendle, S.; Freudenthaler, C.; and Schmidt-Thieme, L., “Factorizing personalized markov chains for next-basket recommendation”, In WWW, 811–820. ACM, 2010.