Adversarial Point-of-Interest Recommendation

Fan Zhou*
University of Electronic Science and Technology of China
fan.zhou@uestc.edu.cn

Ruiyang Yin
University of Electronic Science and Technology of China
yinruiyang94@gmail.com

Kunpeng Zhang
University of Maryland, College park
kzhang@rhsmith.umd.edu

Goce Trajcevski
Iowa State University, Ames
gocet25@iastate.edu

Ting Zhong
University of Electronic Science and Technology of China
zhongting@uestc.edu.cn

Jin Wu
University of Electronic Science and Technology of China
wj@uestc.edu.cn

ABSTRACT

Point-of-interest (POI) recommendation is essential to a variety of services for both users and business. An extensive number of models have been developed to improve the recommendation performance by exploiting various characteristics and relations among POIs (e.g., spatio-temporal, social, etc.). However, very few studies closely look into the underlying mechanism accounting for why users prefer certain POIs to others. In this work, we initiate the first attempt to learn the distribution of user latent preference by proposing an *Adversarial POI Recommendation (APOIR)* model, consisting of two major components: (1) the *recommender* (R) which suggests POIs based on the learned distribution by maximizing the probabilities that these POIs are predicted as unvisited and potentially interested; and (2) the *discriminator* (D) which distinguishes the recommended POIs from the true check-ins and provides gradients as the guidance to improve R in a rewarding framework. Two components are co-trained by playing a minimax game towards improving itself while pushing the other to the boundary. By further integrating geographical and social relations among POIs into the reward function as well as optimizing R in a reinforcement learning manner, APOIR obtains significant performance improvement in four standard metrics compared to the state of the art methods.

KEYWORDS

POI recommendation, adversarial learning, policy gradient

ACM Reference Format:
Fan Zhou, Ruiyang Yin, Kunpeng Zhang, Goce Trajcevski, Ting Zhong, and Jin Wu. 2019. Adversarial Point-of-Interest Recommendation. In *Proceedings of the 2019 World Wide Web Conference (WWW '19), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3308558.3313609

1 INTRODUCTION AND MOTIVATION

Location-based recommendation systems primarily aim at suggesting spatial entities (i.e., point-of-interest (POI)) to users, and have recently spurred a significant research interests in both academia and industry. Different variations of POI recommendations have been explored, including next POI recommendation [4, 7], time-aware POI recommendation [13] and out-of-town POI recommendation [21].

A body of existing works has focused on improving the POI recommendation performance by exploiting various implicit context features embedded in check-ins, such as their spatial information [13, 14, 17]; semantics [9]; social relations [5, 11, 36]; temporal characteristics [18, 33]; and sequential dependence [6, 19, 27, 35]. Complementary to these, approaches that model user preference through mining various features for POI recommendation have been proposed – e.g., Matrix Factorization (MF) [11, 13, 14, 17], Context Embedding (CE) [15, 27, 33] and Pairwise Ranking (PR) [4, 12, 22, 33]. A recent work [16] summarized that geographical information and social influence are the two most effective factors for modeling user preference, while MF based methods such as GeoMF and RankGeoFM exhibit superior performance on POI recommendation. Some of the existing methods make simple but proven-to-be effective assumptions for POI recommendation.

At the heart of our motivation is the observation that many of the existing studies lack formal underlying mechanisms to understand essential user check-in behavior, which may often lead to elusive results. Inspired by recent advances in deep generative models [8, 25] that are successfully and widely used in the areas of computer vision and information retrieval, we attempt to learn user latent preference in a generative way, rather than extracting different features and quantitatively analyzing their impact on POI recommendation as done in most of the existing works.

Towards that, we propose a novel POI recommendation approach, called *Adversarial POI Recommendation (APOIR)*, which learns the underlying check-in distribution in an adversarial manner by simultaneously training two synergistic components. Specifically, we model a user u and his check-in locations l in a generative way, with two neural network components: *recommender* (R) and *discriminator* (D) – being co-trained alternatively, to optimize the generative process of $u \rightarrow l$. R recommends POIs based on the currently learned user preference distribution, while D, acting as a catalyzer, judges whether the recommended POIs are true locations visited by that user and provides guidance to improve R. Two effective features (geographical and social influence) are also incorporated into the APOIR to further improve the performance. Note that our model can be generalized to include other features.

*Corresponding author: Fan Zhou (fan.zhou@uestc.edu.cn)

WWW '19, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6674-8/19/05.
https://doi.org/10.1145/3308558.3313609

To train APOIR, we alternate the updates of the recommender and the discriminator as follows. Given the currently learned preference distribution, a set of POIs are sampled to update the R using the policy gradient method. Subsequently, D is updated using user u's true check-ins along with the ones sampled from the updated R. After an equilibrium is reached – which is, the true user preference distribution is close enough to the empirical learned distribution, a list of POIs with high probabilities based on the learned distribution are generated and eventually recommended. Our main contributions can be summarized as follows:

- We address the POI recommendation in a generative way, which is a novel approach in the area of spatial mining and POI recommendation.
- We propose a method – APOIR – to learn underlying user preference distribution, which significantly boosts the recommendation performance. In addition, APOIR successfully unifies reinforcement learning and matrix factorization methods into an adversarial learning framework for POI recommendation.
- We evaluate our method on three public location-based social network (LBSN) datasets and compare it to several state-of-the-art models. The results show that the APOIR approach performs well, e.g., achieving 10.17%, 11.7% and 10.0% improvement over the best baseline on Gowalla, Foursquare and Yelp datasets in terms of Precision@5.

2 RELATED LITERATURE

ML based POI Recommendation Matrix Factorization (MF) and its variants are prevalent techniques in traditional POI recommendation. Typically, a user-POI matrix is factorized to identify user latent references. For example, in [14] user and POI latent factors in the factorization model are augmented with activity area vectors of users and influence vectors of POIs, to deal with the challenge of matrix sparsity. User interest and check-in behavior [12], user preference ranking and metric embedding [4], as well as POI semantic categories [9] are also taken into account for POI recommendation in the MF-based framework.

An extensive experimental evaluation [16] has been recently conducted to compare 12 representative POI recommendation models and draw relevant findings, among which are: (a) geographical information [13, 14]; and (b) implicit feedback of user preference [17], are the most relevant factors to the recommendation performance. According to their results, RankGeoFM [13], IRenMF [17] and GeoMF [14] are the three state-of-the-art POI recommendation models. Our work differs from these methods in that we propose a generative model to learn underlying user preference distribution in an adversarial manner.

GAN and POI Recommendation Generative adversarial networks (GANs) have gained recent popularity due to their successes in natural image generation [8]. GANs have also been used for modeling sequential data, such as discrete token generation [29] and have been applied in information retrieval [25]. However, few efforts have been made towards recommending POIs in LBSNs with adversarial networks due to the following challenges: (1) data sparsity inherent in LBSN data, e.g., the density of the check-ins used in POI recommendation is usually around 0.1% [16]; (2) complexity. Various complicated, latent, and mixed-typed factors, including geographical, spatio-temporal, sequential, social influence, etc., need to be combined into a unified framework for addressing POI recommendation. Our proposed method APOIR provides a fundamentally different way for POI recommendation by learning the user preference distribution, and thus opens up a new perspective for addressing POI recommendation problems – e.g., adversarial time-aware POI recommendation and next POI recommendation.

3 PROBLEM FORMULATION AND PRELIMINARIES

Given a set of POIs $\mathcal{L}$ ($|\mathcal{L}| = M$) and a set of users $\mathcal{U}$ ($|\mathcal{U}| = N$), each with associations to multiple historical check-ins $\mathcal{L}^{u_i}$, POI recommendation aims at recommending each user $u_i \in \mathcal{U}$ with *top-K new* POIs in the set of $\widetilde{\mathcal{L}}^{u_i} = \mathcal{L} - \mathcal{L}^{u_i}$ that u_i is likely to be interested in but has never visited before.

Matrix Factorization (MF): [10] decomposes the user check-in matrix $\mathbf{C} \in \mathbb{R}^{N\times M}$ into a user matrix $\mathbf{U} \in \mathbb{R}^{N\times Q}$ and a POI matrix $\mathbf{L} \in \mathbb{R}^{M\times Q}$ with Q-dimensional latent factors by:

$$\arg\min_{\mathbf{U},\mathbf{L}} = \|\mathbf{C} - \mathbf{U}\mathbf{L}^{\intercal}\|_F^2 + \kappa^u\|\mathbf{U}\|_F^2 + \kappa^l\|\mathbf{L}\|_F^2 \tag{1}$$

where κ^u and κ^l are regularization coefficients, N is the number of users and M is the number of check-ins. The probability of recommending a POI l_j to user u_i is thus derived based on the inner product between the latent factor of user $\mathbf{u}_i$ and that of POI $\mathbf{l}_j$ (denoted as $\mathbf{u}_i\mathbf{l}_j^{\intercal}$). The preference score vector of u_i over all POIs is denoted $\mathbf{u}_i\mathbf{L}^{\intercal}$.

Gated Recurrent Units (GRU): [3] is a variant of recurrent neural network (RNN) models consisting of gating mechanisms that control the influence of the hidden state of previous unit h_{t-1} on the state h_t at time step t, i.e., learning to ignore the previous units if necessary. Specifically,

$$\begin{aligned}
g_t &= \sigma(W\mathbf{l}_j + Uh_{t-1})\\
s_t &= \sigma(W\mathbf{l}_j + Uh_{t-1})\\
\tilde{h}_t &= \tanh(W\mathbf{l}_j + U(s_t \odot h_{t-1}))\\
h_t &= (1 - g_t)h_{t-1} + g_t\tilde{h}_t
\end{aligned} \tag{2}$$

where g_t and s_t are update and reset gates, respectively; $\tilde{h}_t$ is a candidate hidden state; W and U are parameter matrices of respective units; σ and tanh are sigmoid and hyperbolic tangent functions, respectively; $\odot$ is element-wise product.

Generative Adversarial Nets (GAN): [8] aims at obtaining the Nash equilibrium between a discriminator D and a generator G by optimizing the following minimax objective :

$$\mathcal{J}_{GAN} = \mathop{\mathbb{E}}_{x\sim p_x}[\log(D(x))] + \mathop{\mathbb{E}}_{\tilde{x}\sim p_g}[\log(1 - D(\tilde{x}))] \tag{3}$$

where p_x is the data distribution and p_g is the model distribution implicitly defined by $\tilde{x} = G(z)$, and $\mathcal{J}_{GAN}$ is maximized w.r.t $D(x)$ and minimized w.r.t. $D(G(\tilde{x}))$. The generator G takes a noise prior distribution $z \sim p(z)$ (e.g., uniform or Gaussian) as input and upon which a sample is generated using a deep neural network. The discriminator D_ϕ – usually another neural network – plays the role of classifier and distinguishes that a certain sample coming from the true distribution p_x or the generator G. It has been demonstrated that this game achieves global equilibrium if and only if $p_g(x) = p_x(x)$, where p_g is the defined distribution and the optimal discriminator is $D^*(x) = p_x(x)/(p_x(x) + p_g(x))$ [8].

4 OUR PROPOSED APPROACH: APOIR

We now discuss in detail our proposed method, and feature modeling and training details.

Temporal & sequential preference modeling. Before presenting the APOIR method, we first leverage a variant of GRU, combined with MF, to capture both temporal and sequential preference of users. Given a sequence of POIs $l_1, l_2, \cdots, l_t$, each associated with a check-in time $\tau_1, \tau_2, \cdots, \tau_t$, we compute the time interval between adjacent POIs as $\Delta_{\tau_t} = \tau_i - \tau_{i-1}$, $i \in [1, t]$. Then, we can modify the candidate hidden state of Eq.(2) with a time gate T_t as:

$$
\begin{aligned}
T_t &= \sigma_t(W\mathbf{l}_j + \sigma([\Delta_{\tau_t}; \tau_t]W_t)) \\
\tilde{h}_t &= \tanh(W\mathbf{l}_j + U(s_t \odot T_t \odot h_{t-1}))
\end{aligned} \quad (4)
$$

where $[\Delta_{\tau_t}; \tau_t]$ is a concatenation of time interval between two successive check-ins and the current check-in time. Now time gate T_t captures the temporal preference of users, as well as POI representation $\mathbf{l}_j$, and is used to control the influence of previous hidden state h_{t-1} in Eq.(4). Finally, a user's temporal and sequential preferences are coded in the last hidden state h_t, which is then used to update user representation with an element-wise product as $\hat{\mathbf{u}}_i = \mathbf{u}_i \odot h_t$. The new user latent factor representation $\hat{\mathbf{u}}_i$ would be used in the following adversarial learning.

Note that above temporal GRU is similar to recent advances on recurrent unit modification [18, 19, 37] towards capture contextual information associated with input data in RNN, except that we only consider temporal features here because we would learn other important factors in an adversarial manner later.

4.1 Adversarial Learning

In this work, we use a generative process $u_i \rightarrow \mathcal{L}$ to model the relationship between a user u_i and the set of POIs $\mathcal{L}$. The underlying true distribution of the user preferences over POIs (expressed as a conditional probability $p_{true}(\mathcal{L}|u_i)$) will be learned by alternatively optimizing two competing components: *Recommender* and *Discriminator*.

4.1.1 Recommender. $R_\theta(l^R|u_i)$, parameterized by θ, is used to recommend a set of unvisited but potentially interested POIs l^R for u_i. The recommender R_θ here is a generator analogue in GAN where it mainly fits a true distribution of data $p_{true}(\mathcal{L}|u_i)$. Similar to this, the POIs recommended by R_θ are based on a process of sampling from the learned empirical distribution $R_\theta(\mathcal{L}|u_i)$. Such a user-POI preference distribution is approximated using the pairwise bayesian personalized ranking (BPR) [22], where user representation is obtained via above described temporal GRU. The objective is to learn a R_θ where the true distribution is close enough to the empirical one so that it is difficult for the discriminator to decide whether the POI is generated by R_θ or from the true distribution.

4.1.2 Discriminator. $D_\phi(u_i, l^R)$, parameterized by ϕ, is used to discriminate whether the recommended POI l^R matches the true preference $\mathcal{L}^{u_i}$ of user u_i. In other words, it constructs a binary vector where 1 denotes that the recommended POI is exactly the true check-ins (positive example) while 0 means a mismatch (non-visited example). The goal of $D_\phi(u_i, l^R)$ is to improve its ability to distinguish historical check-ins from the recommended ones upon R_θ, for a given user u_i.

4.1.3 Objective. According to the GAN paradigm [8], the above two models can be unified into a minimax game: $R_\theta(l^R|u_i)$ maximizes the probability of $D_\phi(u_i, l^R)$ not being able to discriminate recommended POIs l^R from the truth. Correspondingly, $D_\phi(u_i, l^R)$ judges and improves the recommendation performance of $R_\theta(\mathcal{L}|u_i)$. Eventually, $R_\theta(l^R|u_i)$ recommends POIs for user u_i with high quality once an equilibrium reaches, in which $D_\phi(u_i, l^R)$ cannot distinguish the recommended POIs from the truth. Formally, we have the following objective trained alternatively between the recommender R_θ and the discriminator D_ϕ:

$$
\begin{aligned}
\mathcal{J}^{R^*, D^*} = \min_\theta \max_\phi \sum_{u_i \in \mathcal{U}} \Big(&\mathbb{E}_{l^+ \sim \mathcal{L}^{u_i}}[\log D_\phi(u_i, l^+)] \\
&+ \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)}[\log(1 - D_\phi(u_i, l^R))] \Big)
\end{aligned} \quad (5)
$$

where l^+ is the set of visited (positive) check-ins, and discriminator $D_\phi(u_i, l^R)$ estimates the probability of POIs l^R being preferred/visited by user u_i. We use the sigmoid function $\sigma(D_\phi(u_i, l^R))$ as the discrimination score, similar to [25].

As shown in Eq.(5), both the discriminator and the recommender are iteratively optimized in a minimax game, and we now describe them in details.

4.1.4 Training Discriminator. The objective of the discriminator is to maximize the probability of correctly distinguishing the true check-in locations from the generated recommended POIs by the recommender, given positive samples from true preference distribution and non-visited samples from the recommender. That is, the training objective of D_ϕ is to find an optimal ϕ^* by maximizing:

$$
\begin{aligned}
\phi^* = \arg\max_\phi \sum_{u_i \in \mathcal{U}} \Big(&\mathbb{E}_{l^+ \sim \mathcal{L}^{u_i}}[\log D_\phi(u_i, l^+)] \\
&+ \mathbb{E}_{l^R \sim R_{\theta^*}(l^R|u_i)}[\log(1 - D_\phi(u_i, l^R))] \Big)
\end{aligned} \quad (6)
$$

where $l^R \sim R_{\theta^*}(l^R|u_i)$ is the generated POIs by the current optimal R_θ , and $l^+ \sim \mathcal{L}^{u_i}$ are the positive samples. $D_\phi(u_i, l^R)$ can also be considered as the probability of R_θ assigning correct labels to recommended POIs l^R. Since function $D_\phi(\cdot)$ is differentiable w.r.t parameters ϕ, the above objective can be solved by stochastic gradient descent [25].

4.1.5 Training Recommender. Similarly to SeqGAN [29], the recommender R_θ generates (selects) a list of ranked POIs for user u_i. Specifically, given the current D_ϕ which is fixed after Eq.(6), we can minimize the following objective to find an optimal θ^*:

$$
\begin{aligned}
\theta^* &= \arg\min_\theta \sum_{u_i \in \mathcal{U}} \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)} \log(1 - \sigma(D_\phi(u_i, l^R))) \\
&= \arg\max_\theta \sum_{u_i \in \mathcal{U}} \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)} \log(1 + \exp(D_\phi(u_i, l^R)))
\end{aligned} \quad (7)
$$

which is the objective considered in previous works [25, 34]. However, it does not consider the contexts associated with items, or POIs in our case. To explicitly explore the rewards from POI contexts, we modify above objective as:

$$\theta^* = \arg\max_\theta \sum_{u_i \in \mathcal{U}} \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)} \log(\lambda + \exp(D_\phi(u_i, l^R)))$$
$$= \arg\max_\theta \underbrace{\sum_{u_i \in \mathcal{U}} \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)} \log \mathcal{R}}_{\mathcal{J}^R(u_i)}$$

where we add a factor λ which is a constant for each u_i, and a defined reward $\mathcal{R} = \lambda + \exp(D_\phi(u_i, l^R))$ will be used as the reward (will explain it later).

Since the POIs l^R sampled from the recommender R_θ is discrete, $\mathcal{J}^R(u_i)$ cannot be directly optimized with gradient descent as in continuous GANs [8]. Following the discrete GANs [25, 29, 34] where the gradient descent is not available, we use the probability of being sampled for each POI to replace the discrete POI as:

$$\nabla_\theta \mathcal{J}^R(u_i) = \nabla_\theta \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)} \log \mathcal{R} = \sum_{j=1}^{M} \nabla_\theta R_\theta(l^j|u_i) \log \mathcal{R}$$
$$= \sum_{j=1}^{M} R_\theta(l^j|u_i) \nabla_\theta \log \mathcal{R} = \mathbb{E}_{l^R \sim R_\theta(l^R|u_i)}[\nabla_\theta \log \mathcal{R}]$$
$$\backsimeq \frac{1}{K} \sum_{k=1}^{K} \nabla_\theta \log R_\theta(l^k|u_i) \log \mathcal{R} \quad (8)$$

where K is the number of POIs sampled by the recommender and l^k is the k^{th} sampled POI. Eq.(8) shows that we use policy gradient based REINFORCE algorithm [26] to derive the gradient. In the context of reinforcement learning [23], $\log(\lambda + \exp(D_\phi(u_i, l^k)))$ acts as the *reward* for the *policy* $R_\theta(l^k|u_i)$ when taking *action* of recommending POI l^k in the *environment* u_i.

4.2 Modeling Reward

We note that above reinforcement learning based likelihood sampling is also used in previous work [25, 29, 34] for sequential data generation and information retrieval. However, these work did not explicitly model the reward of Eq.(8). Instead, they can be considered as optimizations for various methods, e.g., IRGAN [25] is essentially an adversarial optimization method for MF in item recommendation.

In this section, we proceed with modeling the context of POIs and the implicit feedback into the reward function. In this work, we exploit two most important factors [16], i.e., geographical and social influence, for explicitly measuring reward of candidate POIs. Specifically, we consider following reward λ in Eq.(8):

$$\lambda = \alpha \mathcal{R}_{geo} + (1 - \alpha)\mathcal{R}_{soc} \quad (9)$$

where $\mathcal{R}_{geo}$ and $\mathcal{R}_{soc}$ are the reward from the geographical and social factors, respectively. Hyperparameter α is used for scaling the two factors. $\mathcal{R}_{geo}$ and $\mathcal{R}_{soc}$ are in fact two 1-D vectors ($\mathcal{R}_{geo}, \mathcal{R}_{soc} \in \mathbb{R}^{1\times M}$) constructed as follows:

– **Geographical reward** $\mathcal{R}_{geo}$ is initialized with 0 in each column. We set the j^{th} item $l^j \in \mathcal{R}_{geo}$ as 1 if POI l^j is within a distance d to *any* visited POIs for user u_i. That is, we are interested in including the nearby POIs $\mathcal{N}(l^j)$ for all check-ins of user u_i into the candidate list and magnify its importance since people are normally visiting the neighboring POIs [14, 16, 17, 31].

– **Social reward** $\mathcal{R}_{soc}$ is generated in a similar way by setting the j^{th} column $u^j \in \mathcal{R}_{soc}$ to 1 if corresponding POI has been visited by u_i's friend $u_j \in \mathcal{F}(u_i)$, where $\mathcal{F}(u_i)$ denotes the friends of u_i – motivated by the observation that people may visit the POIs where their friends have visited before [2, 11, 16, 31].

4.2.1 Overall Rewards.

Note that the second term $\exp(D_\phi(u_i, l^k))$ in the reward of Eq.(8), the function of the discriminator, is also a reward item ($1 \times M$ vector) . Now, we have the following overall reward:

$$\mathcal{R} = \beta(\alpha \mathcal{R}_{geo} + (1 - \alpha)\mathcal{R}_{soc}) + (1 - \beta)\mathcal{R}_D \quad (10)$$

where hyperparameter β controls the effect from POI context and the discriminator $\mathcal{R}_D = \exp(D_\phi(u_i, l^k))$, both of which can be learned from the data. Then, Eq.(8) can be reformulated as:

$$\nabla_\theta \mathcal{J}^R(u_i) \backsimeq \frac{1}{K} \sum_{k=1}^{K} \nabla_\theta \log R_\theta(l^k|u_i) \log(\beta(\alpha \mathcal{R}_{geo}$$
$$+ (1 - \alpha)\mathcal{R}_{soc}) + (1 - \beta) \exp(D_\phi(u_i, l^k))) \quad (11)$$

Essentially, reward $\mathcal{R}$ acts as a regularizer to the recommended POIs from the recommender $R_\theta(l^k|u_i)$ – which outputs the user preference probability over POIs (also a $1\times M$ vector). As the training process goes, $\mathcal{R}$ may gradually push recommender R_θ to produce the POIs matching the preference of the user. It has been proved that if we know the true preference distribution of users, the above adversarial minimax training of APOIR can achieve Nash equilibrium [8] – the recommender exactly fits the true distribution of the user preference, i.e., $R_\theta(\mathcal{L}|u_i) = p_{true}(\mathcal{L}|u_i)$, the discriminator cannot distinguish the recommended POIs from the truth, i.e., the probability of l^R being preferred by u_i based on $D_\phi(u_i, l^R)$ is close to 0.5.

The overall logic of adversarial POI recommendation is summarized in Algorithm 1.

4.3 Discussion

The complexity of the APOIR training is linear in the number of GAN iterations, each of which has a time complexity $O(NK|\widetilde{\mathcal{L}}^{u_i}|)$ in terms of the number of candidate POIs $\widetilde{\mathcal{L}}^{u_i}$. We note that possible improvements can result from reducing the size of $|\widetilde{\mathcal{L}}^{u_i}|$ through filtering the POIs by considering spatial and/or categorical ranking influence [9].

Although both APOIR and IRGAN [25] leverage REINFORCE method for training the generator (recommender for APOIR), we highlight their fundamental differences: (1) APOIR explicitly models the POI reward which can help better understanding and interpreting the recommendation methods; In contrast, IRGAN is a GAN based optimization method for MF; and (2) APOIR additionally incorporates a temporal GRU for modeling the user dynamic preference.

We also note that the reward $\mathcal{R}$ in APOIR only explicitly considers two POI contexts (arguably, the most important two [16]), whereas temporal and sequential factors have been incorporated in the context GRU units. Other factors such as categorical information of POIs have been used in the literature [9]. Incorporating them into the reward function Eq.(9) for better understanding of the POI context, along with incorporating POI embedding [27, 33]

Algorithm 1: Adversarial POI Recommendation.

Input: Recommender R_θ, Discriminator D_ϕ, training data S.
Initialize R_θ and D_ϕ with random parameters θ and ϕ.
Pretrain R_θ on S with MF.
repeat
    Sample K POIs l^R for each user via R_θ;
    foreach $l^k \in l^R$ **do**
        Calculate reward $\mathcal{R}$ of l^k using Eq.(10);
    end
    Update R_θ parameters via policy gradient Eq.(11);
    Recommend POIs for each user using updated R'_θ;
    Sample a set of true locations: $l^+ \sim \mathcal{L}^{u_i}$;
    Update D_ϕ parameters via Eq.(6) with l^+ and l^R;
until *converge*
Output: Recommend *top-K* POIs based on the optimal R_θ for $u_i \in \mathcal{U}$.

for capturing the context of POIs, are also potential sources of improving the Interpretability. However, it may introduce more hyperparameters to tune. How to learn the weights of different contextual factors in the reward function remains an open problem. Such explainable considerations regarding the recommendation model are subject of our future work.

5 EXPERIMENTAL OBSERVATIONS

In this section, we compare the performance of APOIR to baselines using three real-world datasets.

The hyperparameters of APOIR are empirically tuned as: $\alpha = 0.4$, $\beta = 0.7$ for all experiments. The geographical distance parameter d is set to $50km$ following [2].

Table 1: Statistics of three datasets used in experiments.

Dataset	#Users	#POIs	#Check-ins	Sparsity
Gowalla	18,737	32,510	1,278,274	99.865%
Foursquare	24,941	28,593	1,196,248	99.900%
Yelp	30,887	18,995	860,888	99.860%

Datasets: We conducted our experiments on three publicly available LBSN datasets[1]: Gowalla, Foursquare and Yelp. For all datasets, we filter out those POIs with fewer than 10 visitors and those users (usually aka. the colder-start users) with fewer than 15 check-in POIs. Since Foursquare data does not have social information, we only report results from those not considering social information. Therefore, we remove the social factor $\mathcal{R}_{soc}$ in Eq.(9) for comparison on Foursquare. And Yelp data does not have check-ins time. Thus, we replace the temporal GRU in APOIR with a matrix factorization based user representation. The datasets after pre-processing are described in Table 1.

Following previous works [16, 27], we partition each dataset into training set and test set. For each user, we use the earlier 75% check-ins as the training data and the most recent 25% check-ins as the test data. All datasets are very sparse (the frequency of most POIs being visited is extremely low). Since a POI recommender system typically aims at recommending POIs that a user has not visited before, we further merge repetitive check-ins and use the earliest one. This can also avoid a testing interaction appearing in the training set.

Baselines: We compare APOIR with 10 approaches, covering from the most popular/representative POI recommendation techniques to models using different kinds of context information:
– **USG** [28]: is a collaborative filtering-based recommendation with user preference, social influence and geographical influence simultaneously incorporated.
– **MGMPFM** [1]: combines Poisson factor model and a multi-center Gaussian based geographical modeling method.
– **LFBCA** [24]: is a link-based method that constructs a graph to model users and their relations.
– **iGSLR** [30]: exploits personalized geographical preference and social influence with FCF (friend-based CF) and KDE (kernel density estimation).
– **LORE** [32]: considers sequential influence in addition to social and geographical influence by FCF, KDE and MF.
– **IRenMF** [17]: incorporates characteristics of neighboring POIs in both individual level and region level into weighted matrix factorization for POI recommendation.
– **GeoMF** [14]: integrates spatial influence in user geographical regions and its propagation.
– **RankGeoFM** [13]: is a ranking based geographical factorization method incorporating the spatial-temporal factors.
– **GeoTeaser** [33]: a temporal POI embedding model to capture the contextual check-in information and the temporal characteristics using word2vec framework.
– **PACE** [27]: builds a word2vec-based architecture to jointly learn the embeddings of users and POIs to predict both user preference over POIs and context associated with users and POIs.

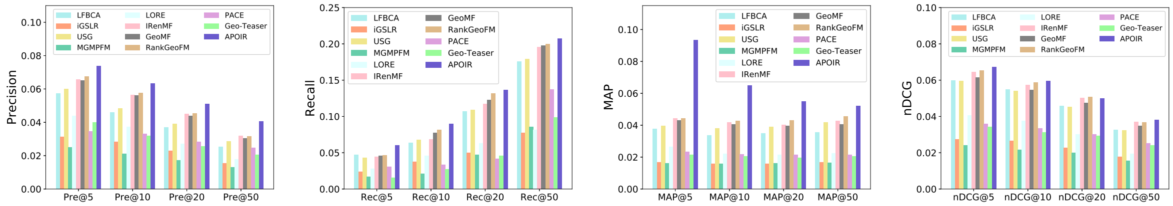
We exclude many MF-based approaches, since they have already been shown to be inferior to RankGeoFM [13]. Several recent approaches such as SG-CWARP [15] and ASMF [11] are also excluded because of their worse performance as compared to GeoTeaser [33] and PACE [27].

Metrics: We compare the model performance using four standard metrics in POI recommendation, i.e., Pre@K (precision), Rec@K (recall), nDCG@K (normalized discounted cumulative gain), and MAP@K (mean average precision) [16, 27]. They show different perspectives of the performance evaluation. Precision and recall measure the number of correct recommendations, while nDCG and MAP consider the rank of the recommendations by assigning higher score to hits at higher positions. We report the average score for all methods, and perform one-sample paired t-test to judge the statistical significance where necessary.

5.1 Results

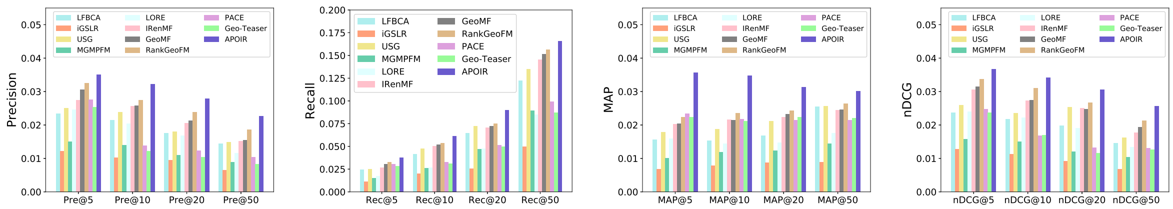
Overall Performance: Figures 1-3 illustrate the comparisons among different methods.From the results, we have following observations: **(1)** APOIR consistently performs the best and significantly improves the POI recommendation performance over the baselines on all metrics across datasets. Take the Yelp data for example (Figure 2), APOIR achives 10.0% on Pre@5, 16.0% on Rec@5, 23.3% on MAP@5 and 9.4% on nDCG@5 over RankGeoFM which performs

[1] http://spatialkeyword.sce.ntu.edu.sg/eval-vldb17/



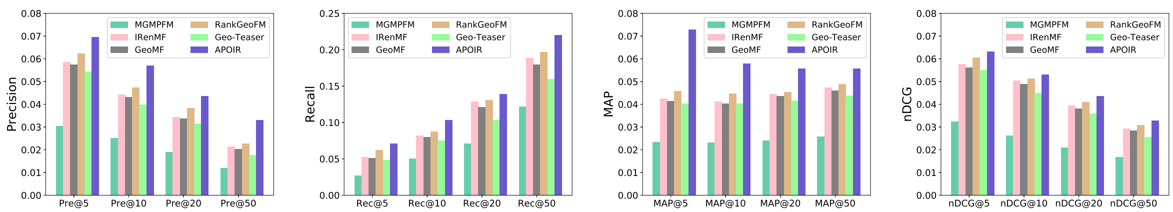
(a) Precision (b) Recall (c) MAP (d) nDCG

Figure 1: Comparisons Among Different Algorithms on Gowalla Data.



(a) Precision (b) Recall (c) MAP (d) nDCG

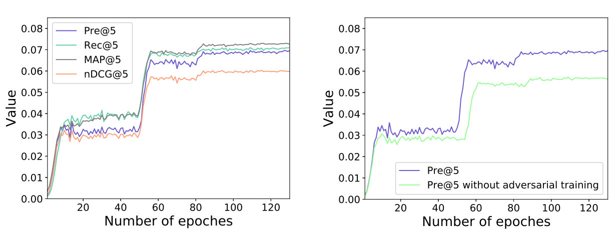
Figure 2: Comparisons Among Different Algorithms on Yelp Data.



(a) Precision (b) Recall (c) MAP (d) nDCG

Figure 3: Comparisons Among Different Algorithms on Foursquare Data.

the second best. **(2)** The strength of APOIR comes from its capability of capturing user preference with an adversarial learning process. Figure 4(a) plots the training process of APOIR. We observe in experiments that discriminator exhibits strong ability at the beginning of the training but deteriorates as the recommender grows to produce more competitive POIs – Figure 4(b) compares the performance with or without adversarial training. **(3)** MF-based methods, such as IRenMF, GeoMF and RankGeoFM, outperform other baselines which proves the effectiveness of MF in modeling latent features underpinning the complex interactions between users and POIs and the implicit feedback from the geographical information. **(4)** The most recent context embedding-based models such as GeoTeaser and PACE do not exhibit the performance as expected. We conjecture that the dataset used in their released implementation is relatively small and dense. Thus, it is difficult to obtain similar performance with larger and sparser datasets – having more POIs and more low-frequency visited POIs may result in worse embedding in word2vec[20].



(a) APOIR training (b) Effect of adversarial training

Figure 4: Training Process.

6 CONCLUSIONS

We proposed a novel approach – APOIR – for POI recommendation by learning the underlying user preference over POIs in an adversarial manner. APOIR leverages different contexts into the rewards in the reinforcement learning and adopts a generative framework for training two competing components: a recommender and a discriminator. By pushing each other to be close to its limit, the recommender may approach the true preference of users (upon which POIs are sampled and recommended), reaching an equilibrium where it becomes difficult for the discriminator to distinguish these generated POIs from the truly visited ones. Comprehensive experiments on three datasets have demonstrated the effectiveness of APOIR, with a significant performance improvement on POI recommendation when compared to existing methods.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (Grant No.61602097 and No.61472064), NSF grants III 1213038 and CNS 1646107, and ONR grant N00014-14-10215.

REFERENCES

[1] Chen Cheng, Haiqin Yang, Irwin King, and Michael R Lyu. 2012. Fused Matrix Factorization with Geographical and Social Influence in Location-Based Social Networks. *AAAI* (2012).

[2] Eunjoon Cho, Seth A Myers, and Jure Leskovec. 2011. Friendship and Mobility: User Movement in Location-Based Social Networks. In *KDD*.

[3] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv:1412.3555* (2014).

[4] Shanshan Feng, Xutao Li, Yifeng Zeng, Gao Cong, Yeow Meng Chee, and Quan Yuan. 2015. Personalized Ranking Metric Embedding for Next New POI Recommendation.. In *IJCAI*.

[5] Qiang Gao, Goce Trajcevski, Fan Zhou, Kunpeng Zhang, Ting Zhong, and Fengli Zhang. 2018. Trajectory-based Social Circle Inference. In *SIGSPATIAL*.

[6] Qiang Gao, Fan Zhou, Kunpeng Zhang, and Goce Trajcevski. 2017. Identifying Human Mobility via Trajectory Embeddings. In *IJCAI*.

[7] Qiang Gao, Fan Zhou, Kunpeng Zhang, Goce Trajcevski, Zhong Ting, and Fengli Zhang. 2019. Predicting Human Mobility via Variational Attention. In *WWW*.

[8] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *NIPS*.

[9] Jing He, Xin Li, and Lejian Liao. 2017. Category-aware Next Point-of-Interest Recommendation via Listwise Bayesian Personalized Ranking. In *IJCAI*.

[10] Yehuda Koren, Robert M Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *IEEE Computer* (2009).

[11] Huayu Li, Yong Ge, Richang Hong, and Hengshu Zhu. 2016. Point-of-Interest Recommendations: Learning Potential Check-ins from Friends. In *KDD*.

[12] Huayu Li, Yong Ge, Defu Lian, and Hao Liu. 2017. Learning User's Intrinsic and Extrinsic Interests for Point-of-Interest Recommendation: A Unified Approach. In *IJCAI*.

[13] Xutao Li, Gao Cong, Xiao-Li Li, and Shonali Krishnaswamy. 2015. Rank-GeoFM: A Ranking based Geographical Factorization Method for Point of Interest Recommendation. In *SIGIR*.

[14] Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. 2014. GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In *KDD*.

[15] Xin Liu, Yong Liu, and Xiaoli Li. 2016. Exploring the Context of Locations for Personalized Location Recommendations. *IJCAI* (2016).

[16] Yiding Liu, Tuan-Anh Pham, Gao Cong, and Quan Yuan. 2017. An Experimental Evaluation of Point-of-interest Recommendation in Location-based Social Networks. *PVLDB* (2017).

[17] Yong Liu, Wei Wei, Aixin Sun, and Chunyan Miao. 2014. Exploiting Geographical Neighborhood Characteristics for Location Recommendation. In *CIKM*.

[18] Jarana Manotumruksa, Craig Macdonald, and Iadh Ounis. 2017. A Deep Recurrent Collaborative Filtering Framework for Venue Recommendation. In *CIKM*.

[19] Jarana Manotumruksa, Craig Macdonald, and Iadh Ounis. 2018. A Contextual Attention Recurrent Architecture for Context-Aware Venue Recommendation. In *SIGIR*.

[20] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *NIPS*.

[21] Tuan-Anh Nguyen Pham, Xutao Li, and Gao Cong. 2017. A General Model for Out-of-town Region Recommendation. In *WWW*.

[22] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *UAI*.

[23] Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. 1999. Policy Gradient Methods for Reinforcement Learning with Function Approximation. *NIPS* (1999).

[24] Hao Wang, Manolis Terrovitis, and Nikos Mamoulis. 2013. Location recommendation in location-based social networks using user check-in data. In *SIGSPATIAL*.

[25] Jun Wang, Lantao Yu, Weinan Zhang, Yu Gong, Yinghui Xu, Benyou Wang, Peng Zhang, and Dell Zhang. 2017. IRGAN: A Minimax Game for Unifying Generative and Discriminative Information Retrieval Models. In *SIGIR*.

[26] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning* (1992).

[27] Carl Yang, Lanxiao Bai, Chao Zhang, Quan Yuan, and Jiawei Han. 2017. Bridging Collaborative Filtering and Semi-Supervised Learning: A Neural Approach for POI Recommendation. In *KDD*.

[28] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*.

[29] Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu. 2017. SeqGAN - Sequence Generative Adversarial Nets with Policy Gradient. In *AAAI*.

[30] Jia-Dong Zhang and Chi-Yin Chow. 2013. iGSLR: personalized geo-social location recommendation: a kernel density estimation approach. In *SIGSPATIAL*.

[31] Jia-Dong Zhang and Chi-Yin Chow. 2015. GeoSoCa: Exploiting Geographical, Social and Categorical Correlations for Point-of-Interest Recommendations. In *SIGIR*.

[32] Jia-Dong Zhang, Chi-Yin Chow, and Yanhua Li. 2014. LORE: exploiting sequential influence for location recommendations. In *SIGSPATIAL*.

[33] Shenglin Zhao, Tong Zhao, Irwin King, and Michael R Lyu. 2017. Geo-Teaser: Geo-Temporal Sequential Embedding Rank for Point-of-interest Recommendation. In *WWW*.

[34] Wei Zhao, Benyou Wang, Jianbo Ye, Yongqiang Gao, Min Yang, Zhou Zhao, and Xiaojun Chen. 2017. Leveraging Long and Short-term Information in Content-aware Movie Recommendation. *arXiv.org* (2017).

[35] Fan Zhou, Qiang Gao, Kunpeng Zhang, Goce Trajcevski, Zhong Ting, and Fengli Zhang. 2018. Trajectory-User Linking via Variational AutoEncoder. In *IJCAI*.

[36] Fan Zhou, Bangying Wu, Yi Yang, Goce Trajcevski, Kunpeng Zhang, and Ting Zhong. 2018. vec2Link: Unifying Heterogeneous Data for Social Link Prediction. In *CIKM*.

[37] Yu Zhu, Hao Li, Yikang Liao, Beidou Wang, Ziyu Guan, Haifeng Liu, and Deng Cai. 2017. What to Do Next - Modeling User Behaviors by Time-LSTM. In *IJCAI*.