DeepTrip: Adversarially Understanding Human Mobility for Trip Recommendation

Qiang Gao
University of Electronic Science and Technology of China, Chengdu, China
qianggao@std.uestc.edu.cn

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

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

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

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

Fengli Zhang
University of Electronic Science and Technology of China, Chengdu, China
fzhang@uestc.edu.cn

ABSTRACT

In this work we propose DeepTrip – an end-to-end method for better understanding of the underlying human mobility and improved modeling of the POIs' transitional distribution in human moving patterns. DeepTrip consists of: a *Trip Encoder* to embed a given route into a latent variable with a recurrent neural network (RNN); and a *Trip Decoder* to reconstruct this route conditioned on an optimized latent space. Simultaneously, we define an *Adversarial Net* composed of a *generator* and *critic*, which generates a representation for a given query and uses a critic to distinguish the trip representation generated from *Trip Encoder* and query representation obtained from *Adversarial Net*. DeepTrip enables regularizing the latent space and generalizing users' complex check-in preference. We demonstrate the effectiveness and efficiency of the proposed model, and the experimental evaluations show that DeepTrip outperforms the state-of-the-art baselines on various evaluation metrics.

CCS CONCEPTS

• **Information systems** → **Location based services**; **Geographic information systems**.

KEYWORDS

trip recommendation, auto-encoder, generative adversarial net

ACM Reference Format:

Qiang Gao, Goce Trajcevski, Fan Zhou, Kunpeng Zhang, Ting Zhong, and Fengli Zhang. 2019. DeepTrip: Adversarially Understanding Human Mobility for Trip Recommendation. In *27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '19), November 5–8, 2019, Chicago, IL, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3347146.3359088

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

SIGSPATIAL '19, November 5–8, 2019, Chicago, IL, USA

ACM ISBN 978-1-4503-6909-1/19/11…$15.00
https://doi.org/10.1145/3347146.3359088

1 INTRODUCTION

Recent popularity of Location based Social Network (LBSN) such as Twitter, Flickr and Instagram, enabled generation of massive check-in data with POIs, along with large amounts of short messages and other information relating space/location, time and other context. A typical trip recommendation system provides a list/sequence of ordered POIs for a given a query which includes, at a minimum, a starting location and a destination. Thus, the key task becomes to construct a proper model to capture the various transition distribution among POI pairs or a set of POIs.

Most of the existing works rely on developing statistical methods to identify top-k popular POIs from massive trajectories [17], applying Markov-based approaches to learn POI transition matrices [1], or adopting certain retrieval algorithms (e.g., Monte Carlo Tree) to make a personalized route recommendation [11].

The core motivation for our work is based on the observations that the existing studies on trip modeling/planning exhibit certain limitations: (1) Traditionally, trip recommendation problem is formulated as a search or statistical problem, which is usually not efficient in generating (2) It is difficult to exploit and integrate the semantics of sequential information among POIs by simply exploring tourist's interest and POIs' popularity. (3) Diversity of the POI sequences based on historical trajectories cannot be obtained through simple empirical analysis.

To tackle the aforementioned challenges, we propose *DeepTrip* – an end-to-end neural network in an adversarial spirit, to better understand various contexts affecting human mobility when making trip recommendations. Specifically, we first leverage RNN-based autoencoder as the basic framework, we then adopt an auxiliary neural network to better learn the sequential POIs distribution in an adversarial manner, so as to sidestep the efforts of explicitly modeling prior distribution.

2 METHODOLOGY

2.1 Problem Definition

Let L denote the set of POIs (check-ins left by the users) and for each POI l_τ ($l_\tau \in L$) we have its geographical coordinates $ð_\tau =< l_\tau^{lat}, l_\tau^{lon} >$ and a timestamp t_τ. The formulation of the trip recommendation problem follows the related works [1, 4, 11, 15]:
Trip Recommender: A tourist provides a query q that consists the desired start point l_s and start time t_s, the length of trip N

(i.e., the number of POIs) and the end point l_e at time t_e. A trip recommendation system returns a route ℓ =$(l_1, l_2, ..., l_N)$, where $l_1 = l_s$ and $l_N = l_e$ for denotation convenience.

2.2 Overview

The overall framework is shown in Fig. 1, and the model is designed with three main components: (1) **Trip Encoder (TE)**: $\mathcal{E}_\omega(\cdot)$; (2) **Trip Decoder (TD)**: $\mathcal{R}_\psi(\cdot)$; (3) **Adversarial Net** which further consists of a generator ($g_\theta(\cdot)$) and a critic ($f_\phi(\cdot)$).

In the recommendation process, we first obtain the *query code* $\tilde{\mathcal{X}}$ by feeding the given input query into the generator of the Adversarial Net, and then input $\tilde{\mathcal{X}}$ into the *TD* $\mathcal{R}_\psi(\cdot)$ to request a recommended route.

2.3 Trip Encoder *(TE)*

The *TE* applies a RNN to capture the POIs' long term dependencies. We first obtain the representation of each POI and then construct the route as a sequence of POIs, while generating a *trip code* $\mathcal{X}$ via this RNN.

2.3.1 POI embedding. We randomly initiate each POI l_τ with a $d-$dimensional vector $\mathbf{v}(l_\tau) \in \mathbb{R}^d$ by sampling from a truncated Gaussian distribution [13].

2.3.2 Trip Code Generation. After obtaining the distributed representations of POIs, we apply the Long Short Term Memory (LSTM) cells to encode each trip $(l_1, l_2, ..., l_N)$ and the current state h_τ^{enc} can be updated by:

$$\begin{aligned} i_\tau &= \sigma(W_i\mathbf{v}(l_\tau) + U_i h_{\tau-1}^{enc} + V_i c_{\tau-1} + b_i) \quad (1) \\ \mathfrak{f}_\tau &= \sigma(W_\mathfrak{f}\mathbf{v}(l_\tau) + U_\mathfrak{f} h_{\tau-1}^{enc} + V_\mathfrak{f} c_{\tau-1} + b_\mathfrak{f}) \\ o_\tau &= \sigma(W_o\mathbf{v}(l_\tau) + U_o h_{\tau-1}^{enc} + V_o c_{\tau-1} + b_o) \\ c_\tau &= f_\tau c_{\tau-1} + i_\tau \tanh(W_c\mathbf{v}(l_\tau) + U_c h_{\tau-1}^{enc} + b_c), \end{aligned}$$

where i_t, $\mathfrak{f}_t$, o_t and b_* are respectively the input gate, forget gate, output gate and bias vector; and matrices $W_* \in \mathcal{R}^{d \times d''}, U_*, V_* \in \mathcal{R}^{d'' \times d''}$ are parameters needed to be learned. Subsequently, each cell state c_τ will be transferred to the hidden state h_τ^{enc}:

$$h_\tau^{enc} = \tanh(c_\tau)o_\tau. \quad (2)$$

Following earlier work [6], the last output of LSTM will be leveraged as the *trip code* $\mathcal{X}$:

$$\mathcal{X} = h_N^{enc}, \quad (3)$$

which can be fed into the *TD* to recover the original route and to have $\mathcal{X}$ updated during the training process.

2.4 Trip Decoder *(TD)*

The *TD* actually has a dual role of a *decoder* and a *recommender* in the DeepTrip. As a decoder, its objective is to decode the *trip code* $\mathcal{X}$ to reconstruct the input trip. However, unlike the encoder, it generates ordered POIs one by one, predicting the next POI l_τ based on its hidden state h_τ^{dec} and the *trip code* $\mathcal{X}$:

$$l_\tau = W'[h_\tau^{\text{dec}}; \mathcal{X}] + b', \quad (4)$$

where W' and b' denote the weight matrix and bias, respectively. The overall probability of a route $(l_1, l_2, ..., l_N)$ can be computed as:

$$p((l_1, l_2, ..., l_N)) = \prod_{\tau=1}^{N} p(l_\tau | h_{\tau-1}^{\text{dec}}, \mathcal{X}). \quad (5)$$

2.5 Adversarial Net

Adversarial Net which we involve to jointly train the autoencoder neural network contains two components, as shown in Fig. 1. We employ the Wasserstein GAN [9] to construct our Adversarial Net. More specifically, the components can be described as:

- **Generator**: The generator g_θ produces a *query code* $\tilde{\mathcal{X}}$ given a user query $< l_s, N, l_e >$ which is modeled by another LSTM network. Unlike the *TE* and *TD*, it feeds the start-end pair $< l_s, l_e >$ into LSTM cells to generate $\tilde{\mathcal{X}}$ which has the same dimension as the *trip code*:

$$\tilde{\mathcal{X}} = g_\theta(< l_s, l_e >). \quad (6)$$

- **Critic**: The critic distinguishes the *query code* $\tilde{\mathcal{X}}$ from the *trip code* $\mathcal{X}$ via performing binary classification. We use a multi-layer fully-connected network as a critic function f_ϕ and, in particular, the critic parameters ϕ are restricted to an 1-Lipschitz function set. Thus, we set each parameter $\phi_i = [-\epsilon, \epsilon]$ ($\phi_i \in \phi$) following previous works [16, 19].

3 EXPERIMENTS

3.1 Settings

We now present the evaluation of the performance of DeepTrip in comparison with the state-of-the-art trip recommendation methods[1]. The datasets used in our experiments are shown in Table 1. The trips in Edinburgh, Glasgow and Osaka are extracted from Flickr photos and videos used in [12]. Following [1], we adopt leave-one-out cross validation to evaluate all methods.

Table 1: The Statistics of the Datasets

DataSet	Φ(Visits)	Φ(Trajectory)	Φ(User)
Edinburgh	33,944	5,028	1,454
Glasgow	11,434	2,227	601
Osaka	7,747	1,115	450

— *Baselines*: We compare DeepTrip to the following methods:
• **POIPopularity** [3]: Recommends the most popular or frequently visited POIs to user at each time instant.
• **Markov and Markov-Rank** [1]: Constructs a POI transition matrix, and recommends a trajectory based on Markov transitions and POI ranking.
• **PersQueue** [11]: A personalized tour recommendation method that considers the attraction popularity, user interests and queuing time to make tour recommendation based on Monte Carlo Tree Search (MCTS). Since we do not attempt to constrain explicitly the

[1] The source code of the implementations is publicly available at https://github.com/gcooq/DeepTrip

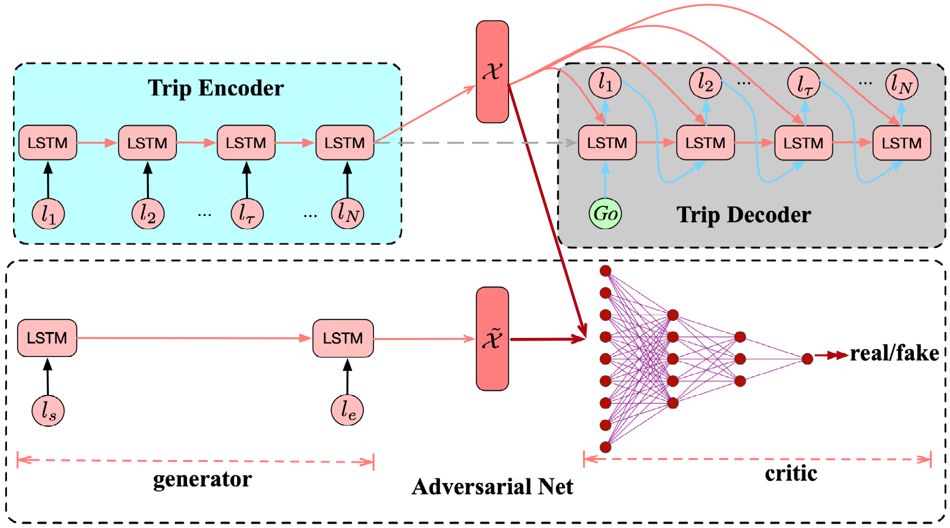

Figure 1: Basic architecture of DeepTrip.

queuing time of POIs, we employ the time interval between two adjacent POIs as the implicit queuing time for the next POI.
— *Metrics*: We choose two commonly used metrics for performance comparison, F_1 and pairs-F_1 scores.
F_1 score: We follow [1, 12] in using the F_1 score to evaluate a recommended trip – which is, the harmonic mean of Precision and Recall of POIs in a trip:

$$\mathrm{F}_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{7}$$

pairs-F_1 score: a specific metric proposed in [1]. It considers both POI correctness and sequential order by measuring the F_1 score of every pair of POIs, whether they are adjacent or not in a trajectory:

$$\text{pairs-F}_1 = \frac{2 \times \text{pairs-P} \times \text{pairs-R}}{\text{pairs-P} + \text{pairs-R}}, \tag{8}$$

where pairs-P and pairs-R denotes the Precision and Recall of ordered POI pairs respectively. The values of both pairs-F_1 and F_1 are between 0 and 1. The higher the value, the better the recommended results – e.g., a value of 1 means that both POIs and their visiting order in the planned trajectory are exactly the same as the ground truth.

3.2 Results

Table 2 shows the performance comparisons between the proposed DeepTrip and the baselines, in terms of F_1 and pairs-F_1 scores, respectively. The best results are indicated in boldface font.

We can observe that DeepTrip outperforms the baselines/state-of-the-art on all the three datasets, with an average improvements 11.99% and 49.16% over the best baseline method in terms of F_1 and pairs-F_1 score, respectively. Among the baselines, Markov-based methods, analogously, are also concerned with the transitional patterns among POI pairs, however, they do not perform as well in comparison with our proposed method. Therefore, the values in Table 2 demonstrates that it is effective to use the encoder-decoder framework jointly combining the adversarial net in trip recommendation.

4 RELATED WORK

The existing solutions for trip recommendation are based on variants of the orienteering based problem, where the main idea is to use a heuristic way to combine the POIs and trajectories [4]. Wei et al. construct top-k popular routes from uncertain trajectory [17]. Ge et al. consider the time cost in trip recommendation to learn tourist's interests and the travel cost [7]. A recent model called PersTour for recommending personalized tours aims at integrating the popularity of POIs and preference of user interest [12]. A learning model aiming at jointly exploiting the locations' and routes' preferences was proposed in [1], demonstrating that the learning-based methods outperform traditional heuristic trip recommendations. Complementary, the PersQ method incorporates attraction popularity, user interests and queuing times into consideration for personalized itinerary recommendation [11]. However, these works do not consider the long term dependency of POIs, arguably, it is difficult to make a diverse route recommendation by simply relying on statistical methods.

Deep learning techniques, especially RNNs have been widely adopted to capture the sequential influence and moving patterns. For example, [6] utilizes RNN-based methods to identify human mobility, and [18, 23] employs the RNN to model trajectories with topological constraints. Deep generative models, such as Variational Auto-encoder (VAE) [10] and Generative Adversarial Networks (GANs) [8], have been widely used in computer vision and natural language processing. VAE can capture the latent variability from complex high dimensional data and have been successfully used to tackle trajectory classification problem [21] and friendship inference from human mobility [5, 20]. GAN received broad attention due to the ability of generating high-quality image and fluent conversations. They have also been used for human mobility learning, e.g., WGAN [16] has been used to generate synthetic

Table 2: Comparisons Among Different Algorithms.

	F_1			Pairs-F_1		
	Edinburgh	***Glasgow***	***Osaka***	***Edinburgh***	***Glasgow***	***Osaka***
Popularity	0.701 ± 0.160	0.745 ± 0.166	0.663 ± 0.125	0.436 ± 0.259	0.507 ± 0.298	0.365 ± 0.190
Markov	0.645 ± 0.169	0.725 ± 0.167	0.697 ± 0.150	0.417 ± 0.248	0.495 ± 0.296	0.445 ± 0.266
Markov-Rank	0.659 ± 0.174	0.754 ± 0.173	0.715 ± 0.164	0.444 ± 0.263	0.545 ± 0.306	0.486 ± 0.288
PersQueue	0.470 ± 0.196	0.586 ± 0.231	0.507 ± 0.186	0.320 ± 0.243	0.384 ± 0.224	0.363 ± 0.283
DeepTrip	**0.765 ± 0.148**	**0.831 ± 0.172**	**0.834 ± 0.166**	**0.660 ± 0.246**	**0.782 ± 0.257**	**0.755 ± 0.265**

trajectories [14] for the purpose of privacy-preserving of human locations and to recommend POIs [22]. However, GAN cannot generate discrete data such as text directly due to lacking the ability of back-propagation through discrete latent variables. Similarly, it is difficult to apply GAN model directly to exploit human moving patterns. Recent popular approach is to firstly transfer or encode the discrete data into a continuous space, and utilize the discriminator to optimize such continuous space, to alleviate the limitations of traditional GANs [16, 19]. In this spirit, our DeepTrip model employs the encoder-decoder framework combining the regularization network [19], towards smoothly learning the latent representation of each trip, thereby improving the trip recommendation.

5 CONCLUSIONS AND FUTURE WORK

We presented DeepTrip, a method for learning human mobility based on adversarial encoder-decoder coupling. To the best of our knowledge, this is the first work that involves adversarial networks for mobility patterns of trips. As part of our future work, we will attempt to incorporate other generative model such as variational anto-encoder with Gaussian process [2], and involve more model interpretability.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (Grant No.61602097 and No.61472064), China Scholarship Council (No.201906070095), NSF grants III 1213038 and CNS 1646107 and ONR grant N00014-14-10215.

REFERENCES

[1] Dawei Chen, Cheng Soon Ong, and Lexing Xie. 2016. Learning Points and Routes to Recommend Trajectories. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management CIKM*. 2227–2232.

[2] Zhenwen Dai, Andreas Damianou, Javier González, and Neil Lawrence. 2015. Variational auto-encoded deep Gaussian processes. *arXiv preprint arXiv:1511.06455* (2015).

[3] 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 *Proceedings of the 24th International Conference on Artificial Intelligence (IJCAI)*. 2069–2075.

[4] Zachary Friggstad, Sreenivas Gollapudi, Kostas Kollias, Tamas Sarlos, Chaitanya Swamy, and Andrew Tomkins. 2018. Orienteering Algorithms for Generating Travel Itineraries. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining WSDM*. 180–188.

[5] Qiang Gao, Goce Trajcevski, Fan Zhou, Kunpeng Zhang, Ting Zhong, and Fengli Zhang. 2018. Trajectory-based Social Circle Inference. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems SIGSPATIAL*. 369–378.

[6] Qiang Gao, Fan Zhou, Kunpeng Zhang, Goce Trajcevski, Xucheng Luo, and Fengli Zhang. 2017. Identifying human mobility via trajectory embeddings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence IJCAI*. 1689–1695.

[7] Yong Ge, Qi Liu, Hui Xiong, Alexander Tuzhilin, and Jian Chen. 2011. Cost-aware Travel Tour Recommendation. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining KDD*. 983–991.

[8] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems NIPS*. 2672–2680.

[9] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. 2017. Improved training of wasserstein GANs. In *Advances in neural information processing systems NIPS*. 5767–5777.

[10] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

[11] Kwan Hui Lim, Jeffrey Chan, Shanika Karunasekera, and Christopher Leckie. 2017. Personalized Itinerary Recommendation with Queuing Time Awareness.. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR*. 325–334.

[12] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, and Shanika Karunasekera. 2015. Personalized tour recommendation based on user interests and points of interest visit durations. In *Proceedings of the 24th International Conference on Artificial Intelligence IJCAI*. 1778–1784.

[13] Jarana Manotumruksa, Craig Macdonald, and Iadh Ounis. 2018. A Contextual Attention Recurrent Architecture for Context-Aware Venue Recommendation. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval SIGIR*. 555–564.

[14] Kun Ouyang, Reza Shokri, David S Rosenblum, and Wenzhuo Yang. 2018. A Non-Parametric Generative Model for Human Trajectories.. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence IJCAI*. 3812–3817.

[15] Kendall Taylor, Kwan Hui Lim, and Jeffrey Chan. 2018. Travel Itinerary Recommendations with Must-see Points-of-Interest. In *Companion Proceedings of the WWW (WWW '18)*. 1198–1205.

[16] Ilya Tolstikhin, Olivier Bousquet, Sylvain Gelly, and Bernhard Schoelkopf. 2017. Wasserstein auto-encoders. In *International Conference on Learning Representations (ICLR 2018)*.

[17] Ling-Yin Wei, Yu Zheng, and Wen-Chih Peng. 2012. Constructing Popular Routes from Uncertain Trajectories. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining KDD*. 195–203.

[18] Hao Wu, Ziyang Chen, Weiwei Sun, Baihua Zheng, and Wei Wang. 2017. Modeling trajectories with recurrent neural networks. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence IJCAI*. 3083–3090.

[19] Junbo Jake Zhao, Yoon Kim, Kelly Zhang, Alexander M Rush, and Yann LeCun. 2018. Adversarially regularized autoencoders. In *Proceedings of the 35th International Conference on Machine Learning ICML*. 5897–5906.

[20] Ting Zhong, Fang Liu, Fan Zhou, Goce Trajcevski, and Kunpeng Zhang. 2019. Motion Based Inference of Social Circles via Self-Attention and Contextualized Embedding. *IEEE Access* 7 (2019), 61934–61948. https://doi.org/10.1109/ACCESS.2019.2915535

[21] Fan Zhou, Qiang Gao, Kunpeng Zhang, Goce Trajcevski, Zhong Ting, and Fengli Zhang. 2018. Trajectory-User Linking via Variational AutoEncoder. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence IJCAI*. 3212–3218.

[22] Fan Zhou, Ruiyang Yin, Goce Trajcevski, Kunpeng Zhang, Zhong Ting, and Jin Wu. 2019. Adversarial Point-of-Interest Recommendation. In *The World Wide Web Conference WWW*. ACM, 3462–34618.

[23] Fan Zhou, Xiaoli Yue, Goce Trajcevski, Ting Zhong, and Kunpeng Zhang. 2019. Context-aware Variational Trajectory Encoding and Human Mobility Inference. In *The World Wide Web Conference WWW*. 3469–3475.