The Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23)

Mobility Prediction via Sequential Trajectory Disentanglement (Student Abstract)

Jinyu Hong[1], **Fan Zhou**[1,3], **Qiang Gao**[2,3,*], **Ping Kuang**[1], **Kunpeng Zhang**[4]
[1]University of Electronic Science and Technology of China
[2]Southwestern University of Finance and Economics
[3]Kash Institute of Electronics and Information Industry
[4]University of Maryland, College park
qianggao@swufe.edu.cn,{202222090418@std.,kuangping@,fan.zhou@}uestc.edu.cn, kpzhang@umd.edu

Abstract

Accurately predicting human mobility is a critical task in location-based recommendation. Most prior approaches focus on fusing multiple semantics trajectories to forecast the future movement of people, and fail to consider the distinct relations in underlying context of human mobility, resulting in a narrow perspective to comprehend human motions. Inspired by recent advances in disentanglement learning, we propose a novel self-supervised method called **SelfMove** for next POI prediction. SelfMove seeks to disentangle the potential time-invariant and time-varying factors from massive trajectories, which provides an interpretable view to understand the complex semantics underlying human mobility representations. To address the data sparsity issue, we present two realistic trajectory augmentation approaches to help understand the intrinsic periodicity and constantly changing intents of humans. In addition, a POI-centric graph structure is proposed to explore both homogeneous and heterogeneous collaborative signals behind historical trajectories. Experiments on two real-world datasets demonstrate the superiority of SelfMove compared to the state-of-the-art baselines.

Introduction

Understanding human mobility enables exploring the informative semantics and mutual interactions behind human historical check-ins. Recent efforts based on deep neural networks such as recurrent networks have achieved significant success in learning human behavioral patterns and individual preferences (Guo et al. 2020; Wu et al. 2020). For instance, PLSPL (Wu et al. 2020) extends LSTM to model human short-term sequential preferences while VaNext (Gao et al. 2019) learns the contextual features of POIs behind human historical check-ins via a variational attention mechanism. However, there are still drawbacks in existing methods: (1) the implicit semantic entanglement poses the risk of uncovering individual human mobility habits that hardly change over time; (2) the sparse sequential check-ins raise the stakes for obtaining good representations of human mobility; and (3) the intricate heterogeneous semantics affiliated with the POIs are not exploited well. In this work, we present a novel approach called SelfMove that operates the sequential variational autoencoder with mutual information regularization to guide the training of evidence lower bound (ELBO). In addition, we provide two trajectory augmentation strategies to alleviate the sparsity issue of human check-in data. Besides, a POI-centric Graph (PGraph) structure is devised to capture both homogeneous and heterogeneous collaborative signals behind historical trajectories.

Methodology

Problem Definition. Given a user's historical trajectory $\mathcal{T}$ and her recent trajectory $T = \{l_1, l_2, \cdots, l_n\}$, our task is to predict the next POI l_{n+1}.

POI embedding with PGraph. We first leverage POI-centric Graph (PGraph) to obtain the informative POI representations. PGraph mainly preserves four contextual semantics, including consecutive, geographical, time-aspect, and activity-aspect interests underlying entire historical trajectories. Each aspect can be built into a correlation matrix $\boldsymbol{A}$. For heterogeneous semantics, the interest behind each POI (e.g., l_i) can be formulated as

$$\boldsymbol{e}_i = \tanh\left(\boldsymbol{A}_i \mathbf{W}\right), \tag{1}$$

where $\mathbf{W}$ is a trainable matrix. For the homogeneous semantics in PGraph, e.g., consecutive POIs, we exploit a graph neural network such as GAT (Veličković et al. 2018) to aggregate the features across all neighbors. Finally, we fuse both homogeneous and heterogeneous semantics into a unified representation for each POI.

As Fig. 1 shows, we learn the disentangled representations $z^r_{1:n}$ (i.e., $\{z^r_1, z^r_2, \cdots, z^r_n\}$) and z^s of a trajectory, where $z^r_{1:n}$ aims at exploring the dynamics of time-dependent intents while z^s is to learn time-independent periodicity.

Contrastively Disentangled Learning. According to variational Bayes, we choose Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{1})$ as $p(z^s)$, and use a standard GRU to model $\mathcal{N}(\mu(z_{<\tau}), \sigma^2(z_{<\tau}))$ as $p(z^r_{1:n})$. Then we employ another GRU to generate each posterior distribution. To generate clean disentanglement of $z^r_{1:n}$ and z^s, we exploit self-supervised trajectory learning by adding three Mutual Information(MI) terms to regularize the latent space, which can be defined as

$$\textbf{Objective: } \max_{p,q} \mathbb{E}_{l_{1:n}\sim p_D} \mathbb{E}_{q(z_{1:n}|l_{1:n})}[\log p\left(l_{1:n} \mid z_{1:n}\right) \tag{2}$$
$$-\alpha(KL\left[q\left(z^s \mid l_{1:n}\right) \| p(z^s)\right] + KL\left[q\left(z^r_{1:n} \mid l_{1:n}\right) \| p\left(z^r_{1:n}\right)\right])$$
$$+\beta(MI_q\left(z^s; l_{1:n}\right) + MI_q\left(z^r_{1:n}; l_{1:n}\right)) - \gamma MI_q\left(z^r_{1:n}; z^s\right),$$

*Corresponding Author (qianggao@swufe.edu.cn).

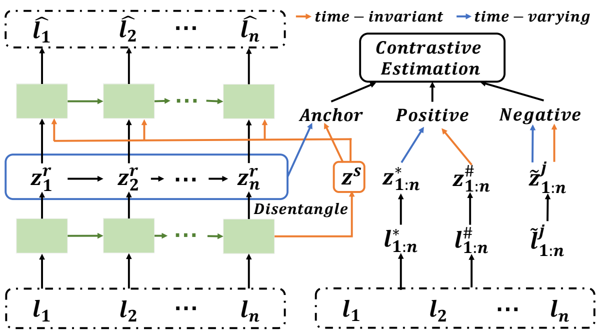

Figure 1: Overview of contrastively disentangled learning.

where α, β, and γ are weight coefficients and $z_{1:n} = (z^s, z^r_{1:n})$. In implementation, we employ contrastive estimation to optimize MI terms via a NCE loss. Notably, we treat $l_{1:n}$ as the positive trajectory, and the negative trajectories are sampled from other users. Since z^s is time-invariant, we randomly change the order of trajectory to generate positive augmented samples for time-invariant factors. For time-varying factors, we replace POIs with their neighbors of the same category within 300m to generate positive augmented trajectories where the temporal dynamics are not changed.

Task Learning. Finally, we fine-tune the model for the next POI prediction. Specifically, we use the pre-trained encoder with disentanglement capability as the kernel of short-term mobility encoder to compute $z^r_{1:n}$ and z^s of recent trajectories. In addition, we use a self-attention layer with position encoding to capture the long-distance dependencies of a historical trajectory $\mathcal{T}_{1:\mathcal{K}}$, and use the last state $\boldsymbol{H}_{\mathcal{K}}$ to represent $\mathcal{T}_{1:\mathcal{K}}$. Then we can take $z^r_{1:n}$, z^s, and $\boldsymbol{H}_{\mathcal{K}}$ as the input and employ a one-layer fully-connected network with *Softmax* to predict the next POI l_{n+1}.

Experimental Results

We use the Los Angeles data from Gowalla and New York data from Foursquare to evaluate the models. We compare our method with the following state-of-the-art baselines: ST-RNN (Liu et al. 2016), VANext (Gao et al. 2019), β-VAE (Higgins et al. 2017), SML (Zhou et al. 2021) and PLSPL (Wu et al. 2020). Following previous studies, ACC@K, area under the ROC curve (AUC) and mean average precision (MAP) are used as evaluation protocols.

Table 1 summarizes the results of all models on two datasets. SelfMove outperforms all baselines on two datasets, e.g., it achieved improvements of 6.8%, 9.7%, 0.69%, 9.4% in terms of ACC@1, ACC@5, AUC and MAP on Los Angeles than the best performance in baselines (underlined). These results verify the advantages of SelfMove in learning disentangled representations of semantic context of human trajectories. As Fig. 2 shows, we investigate whether our disentangled representations are well refined from original trajectories. Specifically, we randomly sample eight different trajectories and change their orders to generate several groups of trajectories. We find that β-VAE only disentangle the representations with a small margin. In contrast, the representations produced by SelfMove are grouped well, demonstrating that it can successfully separate the time-invariant factors to uncover the inherent preference of users.

Method	Gowalla (Los Angeles)			
	ACC@1	ACC@5	AUC	MAP
ST-RNN	10.11	19.05	78.07	4.97
VANext	14.36	27.91	86.22	7.73
β-VAE	14.39	27.43	85.95	7.72
SML	14.77	28.12	86.38	7.86
PLSPL	14.92	28.26	84.34	7.97
SelfMove	**15.94**	**31.02**	**86.98**	**8.72**
Method	**Foursquare (New York)**			
	ACC@1	ACC@5	AUC	MAP
ST-RNN	15.37	31.73	81.40	8.59
VANext	22.54	51.26	89.30	14.02
β-VAE	22.26	50.71	89.38	14.07
SML	22.62	52.16	90.17	14.74
PLSPL	23.02	53.33	89.21	14.83
SelfMove	**25.07**	**56.78**	**90.72**	**16.60**

Table 1: Performance Comparison.

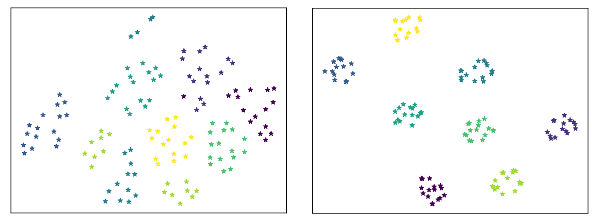
(a) β-VAE. (b) z^s in SelfMove.

Figure 2: The visualization of latent representations.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No.62102326.

References

Gao, Q.; Zhou, F.; Trajcevski, G.; Zhang, K.; Zhong, T.; and Zhang, F. 2019. Predicting human mobility via variational attention. In *WWW*, 2750–2756.

Guo, Q.; Sun, Z.; Zhang, J.; and Theng, Y.-L. 2020. An attentional recurrent neural network for personalized next location recommendation. In *AAAI*, 01, 83–90.

Higgins, I.; Matthey, L.; Pal, A.; Burgess, C.; Glorot, X.; Botvinick, M.; Mohamed, S.; and Lerchner, A. 2017. beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework. In *ICLR*.

Liu, Q.; Wu, S.; Wang, L.; and Tan, T. 2016. Predicting the next location: A recurrent model with spatial and temporal contexts. In *AAAI*.

Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. In *ICLR*.

Wu, Y.; Li, K.; Zhao, G.; and Xueming, Q. 2020. Personalized long-and short-term preference learning for next POI recommendation. *TKDE*, 34(4): 1944–1957.

Zhou, F.; Dai, Y.; Gao, Q.; Wang, P.; and Zhong, T. 2021. Self-supervised human mobility learning for next location prediction and trajectory classification. *Knowledge-Based Systems*, 228: 107214.