Spatial-Temporal Interplay in Human Mobility: A Hierarchical Reinforcement Learning Approach with Hypergraph Representation

Zhaofan Zhang[1*], **Yanan Xiao**[2,5*], **Lu Jiang**[3], **Dingqi Yang**[1,4], **Minghao Yin**[2,5], **Pengyang Wang**[1,4†]

[1]The State Key Laboratory of Internet of Things for Smart City, University of Macau, China
[2]School of Computer Science and Information Technology, Northeast Normal University, China
[3]Department of Information Science and Technology, Dalian Maritime University, China
[4]Department of Computer and Information Science, University of Macau, China
[5]Key Laboratory of Applied Statistics of MOE, Northeast Normal University, China
{mc25684, dingqiyang, pywang}@um.edu.mo, {xiaoyn117, ymh}@nenu.edu.cn, jiangl761@dlmu.edu.cn

Abstract

In the realm of human mobility, the decision-making process for selecting the next-visit location is intricately influenced by a trade-off between spatial and temporal constraints, which are reflective of individual needs and preferences. This trade-off, however, varies across individuals, making the modeling of these spatial-temporal dynamics a formidable challenge. To address the problem, in this work, we introduce the "Spatial-temporal Induced Hierarchical Reinforcement Learning" (STI-HRL) framework, for capturing the interplay between spatial and temporal factors in human mobility decision-making. Specifically, STI-HRL employs a two-tiered decision-making process: the low-level focuses on disentangling spatial and temporal preferences using dedicated agents, while the high-level integrates these considerations to finalize the decision. To complement the hierarchical decision setting, we construct a hypergraph to organize historical data, encapsulating the multi-aspect semantics of human mobility. We propose a cross-channel hypergraph embedding module to learn the representations as the states to facilitate the decision-making cycle. Our extensive experiments on two real-world datasets validate the superiority of STI-HRL over state-of-the-art methods in predicting users' next visits across various performance metrics.

Introduction

Human mobility refers to the patterns and behaviors associated with how and where individuals move or travel within physical spaces. Comprehending and accurately modeling human mobility offers valuable insights with applications spanning transportation (Wang et al. 2020c), urban planning (Wang et al. 2018, 2020a), public health (Qureshi and Xiong 2021), business (Antoniou et al. 2012), etc. Specifically, individual preferences regarding spatial and temporal factors are pivotal in modeling human mobility. From a decision-making standpoint, the selection of the next-visit site by humans is deeply influenced by the interplay between spatial and temporal constraints, which echo individual needs and preferences. This spatial-temporal dynamic is
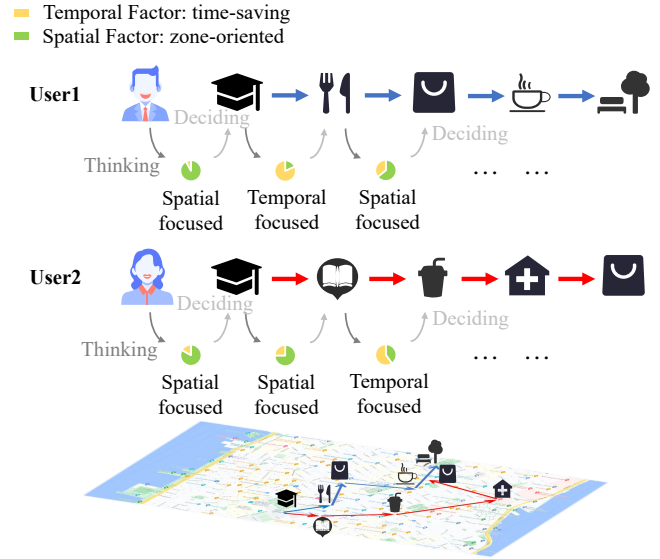

Figure 1: The two users with inconsistent spatial-temporal dynamics and the factors they prioritize when making decisions in various scenarios.

subject to change over time and can differ among individuals. For instance, as illustrated in Figure 1, User 1 prioritizes temporal factors, opting for a quick dinner post-school. However, post-dinner, his preference shifts towards a superior shopping experience, emphasizing spatial considerations over time. In contrast, when comparing the mobility decisions of User 1 and User 2, their spatial-temporal dynamics differ significantly.

Existing works on studying on spatial-temporal factors in human mobility fall into two categories: (1) Temporal-oriented approaches primarily emphasize the time dimension, modeling preferences that evolve over time, while incorporating spatial context as a secondary element. For instance, RNN-based methods (Wang et al. 2020d; Zhao et al. 2020) employ transition matrices or gates to merge spatial context with temporal factors within RNNs' recurrent hidden states. Some models (Lian et al. 2020; Luo,

Liu, and Liu 2021) further leverage attention mechanisms to enhance spatial-temporal integration. (2) Spatial-oriented strategies prioritize the spatial dimension, preserving spatial autocorrelation through graph representation learning and viewing the interplay as temporal graph dynamics. Recent studies harness high-order topological structures for graph representations and capture dynamics using sequential models (Wang et al. 2019; Yang et al. 2019; Wang et al. 2020b, 2022b, 2023). Despite promising results achieved, a common limitation of these methods is their predominant reliance on either the spatial or temporal dimension, potentially oversimplifying the intricate interactions between the two in decision-making. This presents a pressing research challenge: how can we effectively capture the spatial-temporal interplay inherent in mobility decision-making?

To address the challenge, we propose "Spatial-Temporal Induced Hierarchical Reinforcement Learning" (STI-HRL) framework, which formulates the human mobility as a two-layered decision-making process with a spatial-temporal Decoupling-Integration schema. Specifically, STI-HRL is built following the hierarchical setting (1) the low-level, that decouples spatial-temporal interplay with a spatial agent and a temporal agent to focus on each dimension, respectively; (2) the high-level, that integrates the insights from two agents in the low-level to make the final decision on mobility. To facilitate STI-HRL with a proper environment, we propose a Mobility Hypergraph to organize the multi-aspect semantics of mobility records. The hyperedge embeddings from this Mobility Hypergraph serve as the state to support the decision-making process in STI-HRL.

Our contributions can be summarized as follows:

- We introduce a novel formulation of human mobility modeling, conceptualizing it as a two-tiered decision-making process that emphasizes the spatial-temporal interplay in mobility decisions.

- To instantiate this formulation, we develop STI-HRL, offering a unique pipeline for modeling human mobility decisions.

- To preserve the multi-faceted semantics of mobility records, we devise the Mobility Hypergraph and introduce an effective hypergraph embedding method to represent the environment's state.

- We conduct extensive experiments to validate the effectiveness of our proposed STI-HRL. The results demonstrate the superior performance of STI-HRL, consistently outperforming baseline algorithms.

Definitions and Problem Formulation

Mobility Hypergraph

In order to capture the complicated spatial-temporal interplay of human mobilities, we propose to construct a hypergraph to represent historical records, called Mobility Hypergraph. Formally, the Mobility Hypergraph $\mathcal{G}$ is defined as $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ and $\mathbf{E}$ denote the vertex set and hyperedge set, respectively.

Vertices. Mobility Hypergraph aims to organize the check-in events with preserving multi-aspect semantics. Specifically, we decompose the check-in events records $\mathbb{H}$ into four
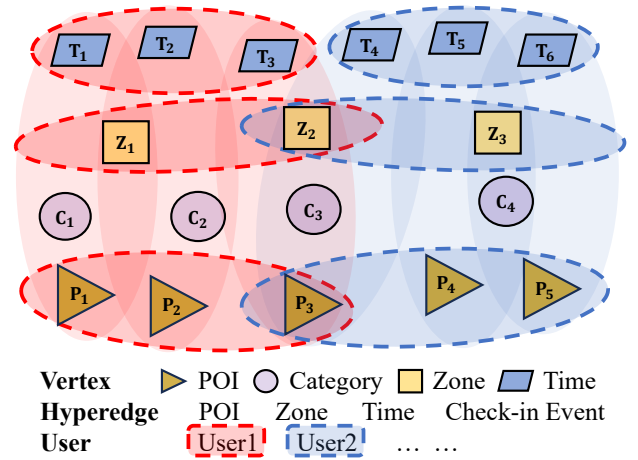

Figure 2: A toy example of Mobility Hypergraph.

semantic channels, including (1) POI channel, denoted as $\mathbf{P}$; (2) POI category channel, denoted as $\mathbf{C}$; (3) zone channel, denoted $\mathbf{Z}$; and (4) time channel, denoted as $\mathbf{T}$. In this work, we consider four types of vertices in corresponding to the four semantic channels. Then, the vertex set can be denoted as $\mathbf{V} = \mathbf{P} \cup \mathbf{C} \cup \mathbf{Z} \cup \mathbf{T}$.

Hyperedges. A hyperedge connects two or more vertices. We identify four types of hyperedges: (1) POI hyperedge, linking all POI vertices a user has visited; (2) zone hyperedge, connecting all the zone vertices visited by a user; (3) time hyperedge, associating all the time vertices visited by a user; (4) event hyperedge, which interlinks the POI, category, zone, and time vertices of a specific check-in event.

Figure 2 provides a toy example of the Mobility Hypergraph for two users. A user's historical records are structured into POI, zone, and time hyperedges, along with all the event hyperedges they have visited. Notably, while the first three hyperedge types are homogeneous (connecting vertices of the same semantic channel), the event hyperedges are heterogeneous, linking vertices across all semantic channels. This structure not only encapsulates the spatial-temporal nuances of each aspect but also establishes connections between users.

Problem Formulation

In this work, we formulate human mobility as a Markov Decision Process (MDP) that users make decisions on the next-visit location based on historical records reflecting individual preferences in certain spatial-temporal contexts. The key components of human mobility MDP are defined as (1) **States** S. Each state $\mathbf{s} \in S$ represents a specific spatial-temporal context derived from historical records, which are organized as a Mobility Hypergraph. (2) **Actions** A. Each action $a \in A$ corresponds to a potential next-visit location. (3) **Transition Probabilities** Γ. $\Gamma(\mathbf{s}'|\mathbf{s}, a)$ represents the probability of transitioning from state $\mathbf{s}$ to state $\mathbf{s}'$ when action a is taken. This probability can be estimated from historical records, reflecting how often a user transitions from one spatial-temporal context to another after choosing a par-
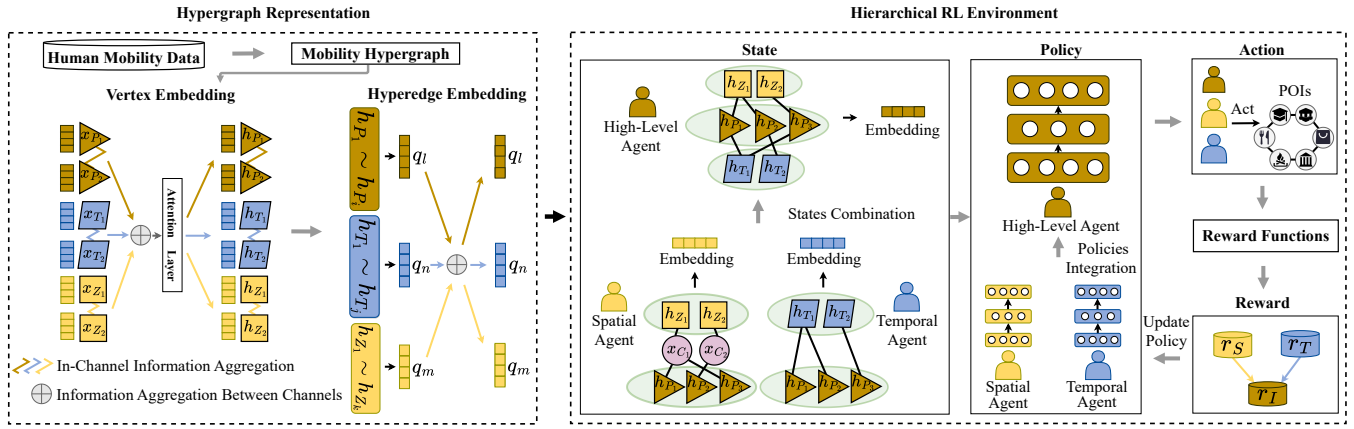
Figure 3: Spatial-Temporal Induced Hierarchical Reinforcement Learning.

ticular location. (4) **Rewards** R. $R(\mathbf{s}, a, \mathbf{s}')$ denotes the reward received after transitioning from state $\mathbf{s}$ to state $\mathbf{s}'$ due to action a. The reward can be designed to reflect user satisfaction or any other metric of interest. We will introduce the reward design later. (5) **Environment** E. The environment consists of all participants of check-in events. It responds to the user's action by providing a new state and a reward. The environment's dynamics are governed by the transition probabilities Γ and the reward function R. (6) **Policy** π. A policy π defines how users take action. Specifically, $\pi(\mathbf{s})$ gives the probability distribution over actions in state $\mathbf{s}$. The goal in the MDP is to find an optimal policy π^* that maximizes the expected cumulative reward over time.

Then, the human mobility MDP can be represented as a tuple: (S, A, Γ, R, E). Given the human mobility MDP formulation, we aim to develop a hierarchical reinforcement learning framework find the optimal policy π^* that guides users in making decisions on the next-visit location with considering spatial-temporal interplay.

Hypergraph Embedding for State Representation

In our study, human mobility records are systematically structured into a Mobility Hypergraph. Therefore, we employ hypergraph embeddings to represent the state to facilitate the hierarchical reinforcement learning framework. The process begins by learning vertex embeddings. Subsequently, these vertex embeddings are aggregated to form hyperedge embeddings. Along this line, the embeddings effectively capture the multi-aspect semantics of human mobilities, encompassing both spatial and temporal contexts. Next, we introduce the embedding procedure in detail.

Vertex Embedding. We denote the raw features of vertex $v_i \in \mathcal{V}$ as $\mathbf{x}_i \in \mathbb{R}^d$, and $\mathcal{N}_i$ represents vertex v_i's neighbors that are within the hyperedges. We employ the attention mechanism to capture the interrelationship between vertices and the respective neighbors in the same channel. Specifically, for the vertex v_i and its neighbor v_j ($j \in \mathcal{N}_i$), the attention coefficient α_{ij} can be represented as

$$\alpha_{ij} = \frac{\exp(\sigma(\mathbf{W}_\alpha^T[\mathbf{W}_v\mathbf{x}_i \| \mathbf{W}_v\mathbf{x}_j]))}{\sum_{j \in \{\mathcal{N}_i, i\}} \exp(\sigma(\mathbf{W}_\alpha^T[\mathbf{W}_v\mathbf{x}_i \| \mathbf{W}_v\mathbf{x}_j]))}, \quad (1)$$

where σ denotes the activation function, $\mathbf{W}_\alpha$ denotes the weight matrix for attention calculation, and $\mathbf{W}_v$ denotes the learnable weight matrix for feature transformation. Then, the embedding $\mathbf{h}_i$ of the vertex v_i can be represented by aggregating the neighbors' information as

$$\mathbf{h}_i = \sigma\left(\sum_{j \in \{\mathcal{N}_i, i\}} \alpha_{ij}\mathbf{W}_v\mathbf{x}_j\right). \quad (2)$$

Hyperedge Embedding. We define four types of hyperedges, among which, POI, zone, and time hyperedges are homogeneous, while event hyperedges are heterogeneous. Specifically, for the homogeneous hyperedges $e_i \in \mathbf{E}$, we first aggregate all the vertex embedding within the hyperedge to represent the hyperedge embedding. The hyperedge embedding $\mathbf{q}_i$ can be denoted as

$$\mathbf{q}_i = \sigma\left(\sum_{j \in |e_i|} \mathbf{h}_j\right), \quad (3)$$

where $|e_i|$ denotes all the associated vertices in e_i.

The event hyperedges serve as a bridge to link the semantics from different channels flowing through the hypergraph topology. We then update the hyperedge embedding $\mathbf{q}_i$ by aggregating information from hyperedges on other channels that are interlinked by the same event hyperedge:

$$\mathbf{q}_i = \sigma\left(\sum_{k \in \Phi(e_i)} \mathbf{W}_{\Psi(e_k)}\mathbf{q}_k\right), \quad (4)$$

where $\Phi(e_i)$ denotes the query function that retrieves hyperedges from alternate channels that are interconnected by the same event hyperedge as the given hyperedge, $\Psi(\cdot)$ is the function to return the type of the given hyperedge, and $\mathbf{W}_{\Psi(e_k)}$ denotes the aggregation weights for the given the type $\Psi(e_k)$.

The homogeneous hyperedge embedding $\mathbf{q}$ captures users' preferences from various channels across both spatial and temporal dimensions, which will be taken as the state to represent the status of the environment. In the next section, we detail the hierarchical reinforcement learning setting of our proposed STI-HRL. We explain how it integrates insights from the hypergraph representation to make optimized mobility decisions.

Spatial-Temporal Induced Hierarchical Reinforcement Learning

Low-level: Decoupling Spatial-Temporal Interplay

As the first phase of STI-HRL, the low level seeks to distinguish users' preferences within the spatial and temporal dimensions. STI-HRL employs a dual-agent setting, comprising a spatial agent and a temporal agent, to discern human mobility decision patterns within these dimensions, respectively. The subsequent sections detail the components and mechanics of the low-level decision-making in STI-HRL.

States. We take the hyperedge embeddings as the representation of states. Specifically, for the spatial agent, the state aims to capture the preferences and interactions on spatial factors. Therefore, we concatenate the associated POI hyperedge embeddings and zone hyperedge embeddings to represent the state. Formally, let $\mathbf{s}_{S,t}^{u}$ represent the spatial agent state for a user u at time t,

$$
\begin{aligned}
\mathbf{s}_{S,t}^{u} &= \text{CONCATENATE}(\mathbf{q}_l, \mathbf{q}_m) \\
&\text{s.t., } l = \Theta_{\mathbf{P}}(u) \ \& \ m = \Theta_{\mathbf{Z}}(u),
\end{aligned}
\tag{5}
$$

where $\Theta_{\mathbf{P}}(u)$ and $\Theta_{\mathbf{Z}}(u)$ denote the indexes of associated POI hyperedge and zone hyperedge for the user u, respectively. Similarly, for the temporal agent, we concatenate the associated POI hyperedge embeddings and time hyperedge embeddings to represent the temporal agent state as

$$
\begin{aligned}
\mathbf{s}_{T,t}^{u} &= \text{CONCATENATE}(\mathbf{q}_l, \mathbf{q}_n) \\
&\text{s.t., } l = \Theta_{\mathbf{P}}(u) \ \& \ n = \Theta_{\mathbf{T}}(u),
\end{aligned}
\tag{6}
$$

where $\Theta_{\mathbf{T}}(u)$ denotes the index of the associated time hyperedge for the user u.

Reward Design. For the spatial agent, we aim to evaluate the decision-making performance in terms of the following three aspects: (1) POI-POI geographic distance, denoted as r_d, which is the reciprocal of the geographic distance between the actual and predicted next-visited POI; (2) POI-POI category similarity, denoted as r_c, which is the cosine similarity between predicted and actual POI categories (computed using fastText embeddings (Bojanowski et al. 2017)); and (3) POI-POI relative ranking r_p, which is the ranking order of the actual visited POI in the predicted POI candidate list. Then, we define the reward for the spatial agent as the combination of the above three components as:

$$
r_S = w_d \cdot r_d + w_c \cdot r_c + w_{ps} \cdot r_p,
\tag{7}
$$

where w_d, w_c, and w_{ps} denote the weights for balancing the influence of r_d, r_c, r_p, respectively.

While for the temporal agent, we focus more on the time factors of mobility. In addition to the POI-POI relative ranking r_p mentioned above, we also include another measurement r_t, the Kullback–Leibler (KL) divergence between the visiting frequency distribution of the actual and predicted POIs across the time span. Then, we combine the two measurements to define the reward for the temporal agent as

$$
r_T = w_t \cdot r_t + w_{pt} \cdot r_p,
\tag{8}
$$

where w_t and w_{pt} denote the weights for balancing the influence of r_t and r_p, respectively.

High-level: Integrative Synthesis of Space-Time Dynamics

The high-level decision process employs a dedicated agent to integrate insights from the low-level agents, synthesizing spatial and temporal considerations in interaction with the Mobility Hypergraph to produce the final human mobility decision. In this section, we introduce the key components of the high-level agent, and outline the workflow.

State. To encapsulate the interplay of spatial and temporal dynamics, the state of the high-level process is defined by combining the states of the two low-level agents. Still taking the state for user u at time t for illustration, the state of the high-level integrative process $\mathbf{s}_{I,t}^{u}$ can be represented as

$$
\mathbf{s}_{I,t}^{u} = \lambda_T \cdot \mathbf{s}_{T,t}^{u} + \lambda_S \cdot \mathbf{s}_{S,t}^{u},
\tag{9}
$$

where λ_S and λ_T denote the weights for the spatial state $\mathbf{s}_{S,t}^{u}$ and temporal state $\mathbf{s}_{T,t}^{u}$, respectively.

Reward. Similarly, in terms of the high-level reward, we calibrate the interplay between spatial and temporal realms by amalgamating the rewards from the low-level agents. Specifically, let w_S and w_T denote the weights for spatial and temporal rewards, respectively. Then, the reward for the high-level agent r_I^{u} can be represented as:

$$
r_I^{u} = w_T \cdot r_T^{u} + w_S \cdot r_S^{u}.
\tag{10}
$$

Policy Learning

Workflow. In the low-level process of STI-HRL, the spatial agent, using hyperedge embeddings, derives its state $\mathbf{s}_{S,t}^{u}$ to predict the next location a user might visit, receiving feedback through the reward r_S. Similarly, the temporal agent uses embeddings to model its state $\mathbf{s}_{T,t}^{u}$, aiming to predict when the visit event occurs, with its efficacy gauged by r_T. Their states transit as:

$$
\mathbf{s}_{\triangleright,t+1}^{u} = \Gamma(\mathbf{s}_{\triangleright,t}^{u}, a_t, r_{\triangleright}),
\tag{11}
$$

where "$\triangleright$" denotes S or T, for the spatial or temporal agent, respectively. The high-level agent then integrates insights from the spatial and temporal agents. It synthesizes a holistic representation of user behavior and transitions as:

$$
\mathbf{s}_{I,t+1}^{u} = \Gamma(\mathbf{s}_{I,t}^{u}, a_t, r_I),
\tag{12}
$$

Optimization. We implement Multi-layer Perceptrons (MLP) for spatial, temporal and high-level policies, parameterized by θ_S, θ_T, and θ_I, respectively. To make a succinct

| Metrics | Recall | | | F1 | | | MRR | | | NDCG | | |
Method	@5	@10	@20	@5	@10	@20	@5	@10	@20	@5	@10	@20
New York												
ST-RNN	0.0037	0.0075	0.0121	0.0012	0.0013	0.0011	0.0017	0.0022	0.0025	0.0022	0.0037	0.0045
DeepMove	0.3857	0.4724	0.5462	0.1285	0.0859	0.0520	0.2501	0.2618	0.2670	0.2839	0.3120	0.3308
LSTPM	0.4039	0.5064	0.5958	0.1346	0.0920	0.0567	0.2698	0.2836	0.2989	0.3232	0.3364	0.3591
HST-LSTM	0.3158	0.3795	0.4345	0.1025	0.0690	0.0413	0.2139	0.2225	0.2264	0.2393	0.2600	0.2740
STAN	0.2754	0.3742	0.4688	0.0918	0.0680	0.0446	0.1619	0.1751	0.1817	0.1900	0.2220	0.2460
LBSN2vec	0.3559	0.4250	0.4544	0.0846	0.0622	0.0408	0.2041	0.2282	0.2551	0.2599	0.2815	0.3122
FPMC	0.2556	0.3437	0.4361	0.0852	0.0624	0.0415	0.1524	0.1642	0.1706	0.1780	0.2064	0.2298
KNN Bandit	0.2045	0.2552	0.2778	0.1414	0.0717	0.0466	0.1644	0.1944	0.2478	0.2522	0.2841	0.3015
STI-HRL	**0.4268**	**0.5227**	**0.6298**	**0.1467**	**0.1104**	**0.0638**	**0.2930**	**0.3066**	**0.3275**	**0.3444**	**0.3685**	**0.4018**
Improvement	**5.67%**	**3.22%**	**5.71%**	**8.90%**	**13.5%**	**12.52%**	**8.60%**	**8.11%**	**9.57%**	**6.56%**	**8.71%**	**10.36%**
Tokyo												
ST-RNN	0.0044	0.0068	0.0085	0.0014	0.0012	0.0008	0.0021	0.0024	0.0025	0.0027	0.0034	0.0039
DeepMove	0.4124	0.4879	0.5510	0.1374	0.0887	0.0524	0.2832	0.2935	0.2979	0.3155	0.3401	0.3561
LSTPM	0.4046	0.4864	0.5614	0.1348	0.0884	0.0534	0.2789	0.2899	0.2951	0.3102	0.3367	03557
HST-LSTM	0.3090	0.3783	0.4462	0.1030	0.0687	0.0424	0.2137	0.2229	0.2276	0.2374	0.2598	0.2769
STAN	0.1990	0.2574	0.3196	0.0663	0.0468	0.0304	0.1300	0.1378	0.1421	0.1471	0.1660	0.1817
LBSN2vec	0.3551	0.4122	0.4753	0.0944	0.0713	0.0412	0.2274	0.2589	0.2637	0.2476	0.2664	0.3067
FPMC	0.1086	0.1557	0.2209	0.0362	0.0283	0.0210	0.0623	0.0684	0.0729	0.0737	0.0888	0.1052
KNN Bandit	0.2388	0.2842	0.3104	0.0947	0.0764	0.0573	0.1441	0.1951	0.2174	0.2006	0.2504	0.2894
STI-HRL	**0.4425**	**0.5150**	**0.5904**	**0.1552**	**0.1005**	**0.0608**	**0.3031**	**0.3277**	**0.3349**	**0.3361**	**0.3615**	**0.3991**
Improvement	**7.30%**	**5.55%**	**5.17%**	**13.10%**	**13.30%**	**13.86%**	**7.03%**	**11.65%**	**12.42%**	**6.53%**	**6.29%**	**12.08%**

Table 1: Overall performance comparison. The best performance for each metric is highlighted in bold. The "Improvement" section denotes the percentage increase over the second-best models.

and unified description, we take "∗" to represent the spatial S, temporal T, or the high-level integration I for the following illustration. The STI-HRL method employs policy gradient (Agarwal et al. 2020) methods.

The primary objective is to maximize the expected rewards:

$$J(\theta_*) = \mathbb{E}[R_*|\theta_*], \quad (13)$$

where $J(\theta_*)$ is the expected reward under policy parameters θ_* and R_* represents the reward of the agents. To achieve this, agent policies are refined by:

$$\theta_*^{new} = \theta_*^{old} + \nabla_{\theta_*} J(\theta_*), \quad (14)$$

Policy Integration and Adaptation. The high-level agent seamlessly integrates policies from the low-level agents:

$$\pi_I(a|\mathbf{s}_I;\theta_I) = \beta\pi_S(a|\mathbf{s}_S;\theta_S) \\ + (1-\beta)\pi_T(a|\mathbf{s}_T;\theta_T), \quad (15)$$

where β is the weight to determine the balance between spatial and temporal inputs. Recognizing the ever-evolving nature of spatial-temporal dynamics, the high-level agent dynamically adjusts β based on the relative performance of the integrated policies:

$$\beta^{new} = \beta^{old} + \eta\nabla J(\theta_I), \quad (16)$$

where η represents a specific learning rate for weight adjustments. Moreover, the weights $\lambda_S, \lambda_T, w_T, w_S$ for the high-level agent in Equation 9 and Equation 10 also follow the same updating strategy with β in Equation 16.

Experiments

We evaluate the performance of STI-HRL in the next-location prediction task that is to predict the next location to visit given the historical mobility records. We aim to answer the following five primary research questions:

- **Q1:** How well does STI-HRL perform in the next-location prediction task?
- **Q2:** How significant are the roles of spatial and temporal factors in mobility decision-making, respectively?
- **Q3:** Is the design of STI-HRL reasonable in investigating spatial-temporal interplay?
- **Q4:** Is it necessary and reasonable to construct Mobility Hypergraph for organizing human mobility records?
- **Q5:** How does the reward design of each agent affect the performance of STI-HRL?

Experiment Settings

Datasets. We conduct the experiment on two widely-used check-in datasets of New York (NYC) and Tokyo (TKY) from April 2012 to February 2013 (Yang et al. 2014). Each entry in the dataset details user ID, POI ID, category, GPS coordinates, and the timestamp of each check-in event. In the experiment, we take the initial 70% serving as the training set, the subsequent 10% for validation, and the final 20% for testing.

Baseline Methods. We compare STI-HRL with eight baseline algorithms, including (1) **ST-RNN** (Liu et al. 2016), (2) **DeepMove** (Feng et al. 2018), (3) **LSTPM** (Sun et al. 2020), (4) **HST-LSTM** (Kong and Wu 2018), (5) **STAN** (Luo, Liu, and Liu 2021), (6) **LBSN2Vec** (Yang et al. 2019), (7)
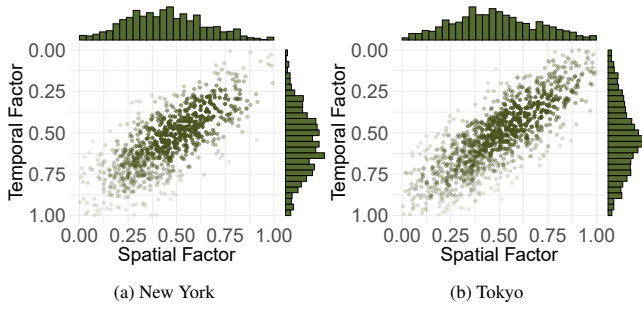
(a) New York (b) Tokyo

Figure 4: A visualization of the weight distribution for spatial (β) and temporal ($1-\beta$) factors in the final decision.
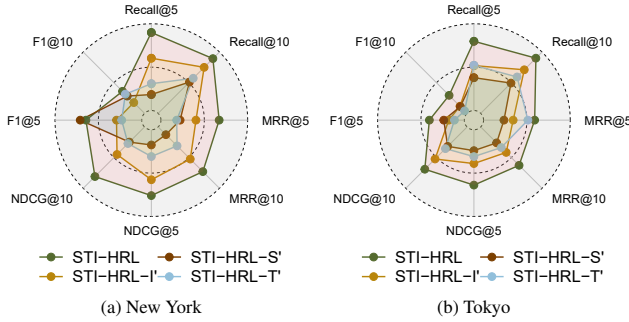

(a) New York (b) Tokyo

Figure 5: An ablation study of STI-HRL agent architecture. STI-HRL-S', STI-HRL-T', STI-HRL-I' denote the variants of STL-HRL that spatial, temporal, and high-level integration removed agents, respectively.

FPMC (Rendle, Freudenthaler, and Schmidt-Thieme 2010), and (8) **KNN Bandit** (Sanz-Cruzado, Castells, and López 2019). Detailed descriptions can be found in Appendix A.
Hyperparameter Settings. We set the hypergraph embedding dimension as 64, and employ 4 attention heads. We utilize Adam optimizer (Diederik, Jimmy et al. 2014) and set the learning rate as 0.001 for optimization. Due to the page limitation, we only present the key hyperparameters here. The full description of hyperparameter settings can be found in Appendix B.

Overall Performance (Q1)

We present the overall performance comparison in Table 1. The results indicate that our proposed STI-HRL consistently outperforms all the baselines on both datasets. Specifically, for the New York dataset, STI-HRL can overpass the second-best baseline by more than 5.6% on Recall, 8% on F1, 8.6% on MRR, and 6% on NDCG; for the Tokyo dataset, the improvement is more significant with more than 7% on Recall, 13% on F1, 7% on MRR, and 6.5% on NDCG. Such consistent improvement underscores the significance of spatial-temporal interplay in human mobility, and validates the effectiveness of our proposed STI-HRL.

The Analysis of Spatial and Temporal Factors (Q2)

We visualize the weights of policy integration for spatial (β) and temporal ($1-\beta$) factors in the prediction, and present the
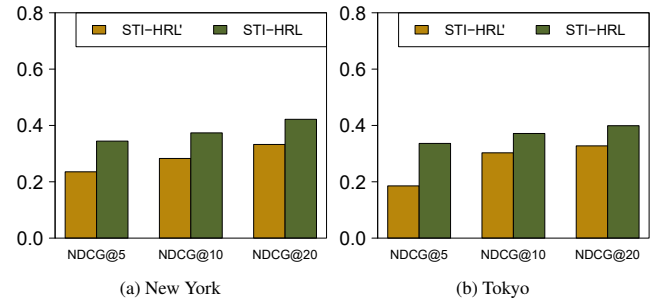

(a) New York (b) Tokyo

Figure 6: An ablation study on hypergraph.
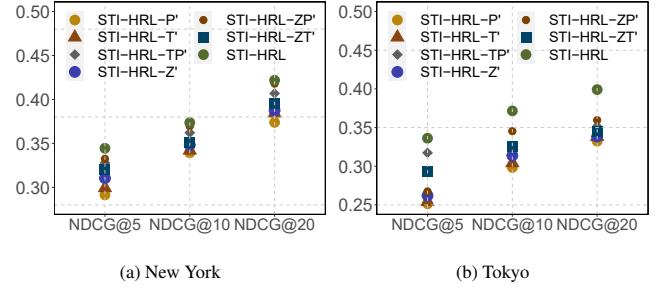

(a) New York (b) Tokyo

Figure 7: An ablation study on hyperedge types.

distribution in Figure 4. The results illustrate trade-off preferences towards spatial and temporal factors varies across users. It's evident that many users display a pronounced bias towards either time or space, suggesting these elements play pivotal roles in their decision-making. This indicates a discernible behavioral pattern among users when opting for mobility decisions.

The Study of STI-HRL Agent Architecture (Q3)

To investigate the design of the two-layered agents in STI-HRL, we introduce three variants of STI-HRL: STI-HRL-S', STI-HRL-T', and STI-HRL-I', denoting the variants that spatial, temporal, and high-level integration removed agents, respectively. We compare the performances of the three variants and STI-HRL, and present the results in Figure 5. The results indicate that STI-HRL performs the best, suggesting the necessity of two-layered decoupling-integration design of agents. Moreover, the performance of STI-HRL-S' and STI-HRL-T' shows that spatial and temporal factors are both essential for human mobilities, which further justifies the critical role of spatial-temporal interplay.

The Study of Mobility Hypergraph (Q4)

The design of Mobility Hypergraph. We conduct an ablation experiment to verify the necessity of the hypergraph architecture. In this experiment, we design a variant of STI-HRL called STI-HRL', which directly takes the user's sequence as input without using a hypergraph architecture. As shown in Figure 6, STI-HRL shows a clear performance advantage compared to STI-HRL'. This result highlights the effectiveness of the hypergraph architecture in dealing with such problems.
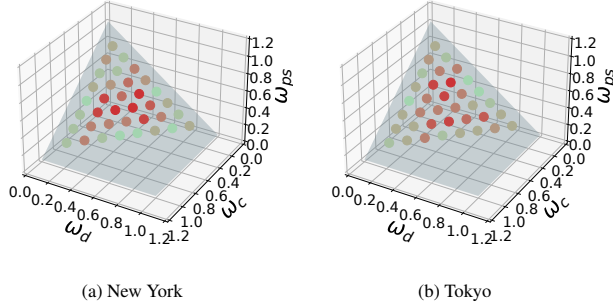
plot the results in Figure 9.

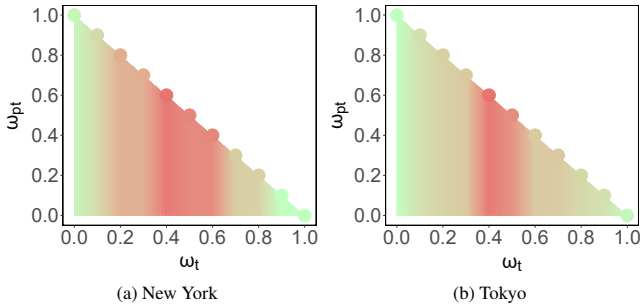Figure 8: The analysis of reward of the spatial agent.

Figure 9: The analysis of reward of the temporal agent.

The effects of hyperedge types. We take the embeddings of three types of hyperedges (POI, Time, and Zone) as the state. Therefore, we introduce six variants of STI-HRL to study the effects of hyperedge types, including (1) STI-HRL-P', POI hyperedges only; (2) STI-HRL-T', time hyperedges only; (3) STI-HRL-Z', zone hyperedges only; (4) STI-HRL-ZT', zone and time hyperedges; (5) STI-HRL-TP', time and POI hyperedges; (6) STI-HRL-ZP', zone and POI hyperedges. We conduct the ablation study on the six variants and present the results in Figure 7. The results demonstrate that varying permutations and combinations of hyperedge types yield different performances. This indicates the significance of the semantics conveyed by these hyperedge types in representing human mobility events. Only STI-HRL, which incorporates all hyperedge types, achieves optimal performance.

The Analysis of Reward Design (Q5)

We consider different combinations of weight setting of reward functions to test the performance of STI-HRL. Since the high-level reward settings (w_S and w_T) are automatically learned, which cannot be manually adjusted. Therefore, we only analyze the reward setting of spatial and temporal agents. Specifically, for the spatial agent reward r_S, three components (r_d, r_c, r_p) are considered. We project the performance of different combinations of w_d, w_c, and w_{ps} ($w_d + w_c + w_{ps} = 1$) to 3D space, and mark the performance in different scales of color, as shown in Figure 8. The better the performance, the darker the color. For the temporal agent reward r_S, two components (r_t and r_p) are considered. We

Related Work

Human Mobility Prediction

Human mobility modeling centers on predicting users' next-visit locations by harnessing insights from historical check-in data. The rise of deep learning introduced advanced techniques like Recurrent Neural Networks (RNNs) (Liu et al. 2016; Yang et al. 2020) and self-attention mechanisms (Lian et al. 2020; Luo, Liu, and Liu 2021). While they excel at modeling sequential regularities, especially with spatial-temporal data, they tend to oversimplify complex spatial-temporal interplay influencing user choices.

Hypergraph Learning

Compared to other graph models (Dong et al. 2023b; Jiang et al. 2023b; Dong et al. 2023a) that deal with heterogeneous information, hypergraph learning (Yang et al. 2019; Liu et al. 2022; Yan et al. 2023) offers a robust representation of the nuanced information within human mobility data, outshining traditional graph-based counterparts. Yet, even with advances in hypergraph embedding, many methods overlook daily behavioral similarities among unconnected individuals as seen in historical trajectories (Yang, Liu, and Zhao 2022), curbing the full potential of existing models.

Reinforcement Learning

At its core, reinforcement learning focuses on learning policy learning through environmental interactions and feedback (Mnih et al. 2015; Jiang et al. 2023a). The duality of exploration and exploitation in this learning paradigm aptly captures the evolving user preferences. Hierarchical Reinforcement Learning (HRL) extends it to a wide variety of comprehensive recommendation domains (Zhang et al. 2019; Yu et al. 2020; Xie et al. 2021; Du et al. 2022; Wang et al. 2022a) has a wide range of applications. Our study pioneers the application of HRL for the next visit in human mobility, approaching it from a unique disentangled viewpoint. Existing HRL models aren't directly transferrable to our challenge due to distinct goals and data structures.

Conclusion

In this study, we address the challenge of capturing the spatial-temporal interplay in human mobility. We frame this challenge as a two-tiered decision-making task: the low-level focuses on disentangling spatial and temporal preferences, while the high-level synthesizes these insights to produce a comprehensive mobility decision. Our approach leverages a hierarchical reinforcement learning framework. To effectively encapsulate the multifaceted semantics of spatial-temporal dynamics, we introduce a Mobility Hypergraph to structure the mobility records, utilizing hyperedge embeddings as states for policy learning. Comprehensive experiments validate the efficacy of our method. Its consistent superiority over baseline models underscores the importance of spatial-temporal interplay in human mobility modeling.

Acknowledgements

References

Agarwal, A.; Kakade, S. M.; Lee, J. D.; and Mahajan, G. 2020. Optimality and approximation with policy gradient methods in markov decision processes. In *Conference on Learning Theory*, 64–66.

Antoniou, A.; Theodoridis, E.; Chatzigiannakis, I.; and Mylonas, G. 2012. Human mobility trace acquisition and social interactions monitoring for business intelligence using smartphones. In *2012 16th Panhellenic Conference on Informatics*, 1–6.

Bojanowski, P.; Grave, E.; Joulin, A.; and Mikolov, T. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 135–146.

Diederik, K.; Jimmy, B.; et al. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 273–297.

Dong, H.; Ning, Z.; Wang, P.; Qiao, Z.; Wang, P.; Zhou, Y.; and Fu, Y. 2023a. Adaptive Path-Memory Network for Temporal Knowledge Graph Reasoning. *arXiv preprint arXiv:2304.12604*.

Dong, H.; Wang, P.; Xiao, M.; Ning, Z.; Wang, P.; and Zhou, Y. 2023b. Temporal Inductive Path Neural Network for Temporal Knowledge Graph Reasoning. *arXiv preprint arXiv:2309.03251*.

Du, Q.; Yu, L.; Li, H.; Leng, Y.; and Ou, N. 2022. Denoising-oriented deep hierarchical reinforcement learning for next-basket recommendation. In *ICASSP 2022 IEEE International Conference on Acoustics, Speech and Signal Processing*, 4093–4097.

Feng, J.; Li, Y.; Zhang, C.; Sun, F.; Meng, F.; Guo, A.; and Jin, D. 2018. Deepmove: Predicting human mobility with attentional recurrent networks. In *Proceedings of the 2018 World Wide Web Conference*, 1459–1468.

Jiang, L.; Liu, K.; Wang, Y.; Wang, D.; Wang, P.; Fu, Y.; and Yin, M. 2023a. Reinforced explainable knowledge concept recommendation in MOOCs. *ACM Transactions on Intelligent Systems and Technology*, 1–20.

Jiang, L.; Wang, Y.; Wang, J.; Wang, P.; and Yin, M. 2023b. Multi-view MOOC quality evaluation via information-aware graph representation learning. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence*, 8070–8077.

Kong, D.; and Wu, F. 2018. HST-LSTM: A hierarchical spatial-temporal long-short term memory network for location prediction. In Lang, J., ed., *Proceedings of the Twenty-*

Seventh International Joint Conference on Artificial Intelligence, 2341–2347.

Lian, D.; Wu, Y.; Ge, Y.; Xie, X.; and Chen, E. 2020. Geography-aware sequential location recommendation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2009–2019.

Liu, G.; Ge, H.; Su, S.; and Wang, S. 2022. Hypergraph regularized low-rank tensor multi-view subspace clustering via L1 norm constraint. *Applied Intelligence*, 1–18.

Liu, Q.; Wu, S.; Wang, L.; and Tan, T. 2016. Predicting the next location: a recurrent model with spatial and temporal contexts. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, 194–200.

Luo, Y.; Liu, Q.; and Liu, Z. 2021. STAN: Spatio-temporal attention network for next location recommendation. In Leskovec, J.; Grobelnik, M.; Najork, M.; Tang, J.; and Zia, L., eds., *Proceedings of the 2021 World Wide Web Conference*, 2177–2185.

Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A. A.; Veness, J.; Bellemare, M. G.; Graves, A.; Riedmiller, M.; Fidjeland, A. K.; Ostrovski, G.; et al. 2015. Human-level control through deep reinforcement learning. *Nature*, 529–533.

Qureshi, S.; and Xiong, J. 2021. Equitable healthcare provision: uncovering the impact of the mobility effect on human development. *Information Systems Management*, 2–20.

Rendle, S.; Freudenthaler, C.; and Schmidt-Thieme, L. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 2010 World Wide Web Conference*, 811–820.

Sanz-Cruzado, J.; Castells, P.; and López, E. 2019. A simple multi-armed nearest-neighbor bandit for interactive recommendation. In *Proceedings of the 13th ACM Conference on Recommender Systems*, 358–362.

Sun, K.; Qian, T.; Chen, T.; Liang, Y.; Nguyen, Q. V. H.; and Yin, H. 2020. Where to go next: Modeling long- and short-term user preferences for point-of-interest recommendation. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, 214–221.

Wang, D.; Fu, Y.; Wang, P.; Huang, B.; and Lu, C.-T. 2020a. Reimagining city configuration: Automated urban planning via adversarial learning. In *Proceedings of the 28th International Conference on Advances in Geographic Information Systems*, 497–506.

Wang, D.; Wang, P.; Fu, Y.; Liu, K.; Xiong, H.; and Hughes, C. E. 2023. Reinforced imitative graph learning for mobile user profiling. *IEEE Transactions on Knowledge and Data Engineering*, 12944–12957.

Wang, E.; Ding, R.; Yang, Z.; Jin, H.; Miao, C.; Su, L.; Zhang, F.; Qiao, C.; and Wang, X. 2022a. Joint charging and relocation recommendation for E-taxi drivers via multi-agent mean field hierarchical reinforcement learning. *IEEE Transactions on Mobile Computing*, 1274–1290.

Wang, P.; Fu, Y.; Xiong, H.; and Li, X. 2019. Adversarial substructured representation learning for mobile user profiling. In *Proceedings of the 25th ACM SIGKDD International*

Conference on Knowledge Discovery & Data Mining, 130–138.

Wang, P.; Fu, Y.; Zhang, J.; Li, X.; and Lin, D. 2018. Learning urban community structures: A collective embedding perspective with periodic spatial-temporal mobility graphs. *ACM Transactions on Intelligent Systems and Technology*, 1–28.

Wang, P.; Fu, Y.; Zhou, Y.; Liu, K.; Li, X.; and Hua, K. A. 2020b. Exploiting mutual information for substructure-aware graph representation learning. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 3415–3421.

Wang, P.; Liu, K.; Jiang, L.; Li, X.; and Fu, Y. 2020c. Incremental mobile user profiling: Reinforcement learning with spatial knowledge graph for modeling event streams. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 853–861.

Wang, Q.; Yin, H.; Chen, T.; Huang, Z.; Wang, H.; Zhao, Y.; and Hung, N. Q. V. 2020d. Next point-of-interest recommendation on resource-constrained mobile devices. In Huang, Y.; King, I.; Liu, T.; and van Steen, M., eds., *Proceedings of the 2020 World Wide Web Conference*, 906–916.

Wang, Z.; Zhu, Y.; Liu, H.; and Wang, C. 2022b. Learning graph-based disentangled representations for next POI recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1154–1163.

Xie, R.; Zhang, S.; Wang, R.; Xia, F.; and Lin, L. 2021. Hierarchical reinforcement learning for integrated recommendation. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*, 4521–4528.

Yan, X.; Song, T.; Jiao, Y.; He, J.; Wang, J.; Li, R.; and Chu, W. 2023. Spatio-temporal hypergraph learning for next POI recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 403–412.

Yang, D.; Fankhauser, B.; Rosso, P.; and Cudre-Mauroux, P. 2020. Location prediction over sparse user mobility traces using rnns. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 2184–2190.

Yang, D.; Qu, B.; Yang, J.; and Cudre-Mauroux, P. 2019. Revisiting user mobility and social relationships in LBSNs: a hypergraph embedding approach. In *Proceedings of the 2019 World Wide Web Conference*, 2147–2157.

Yang, D.; Zhang, D.; Zheng, V. W.; and Yu, Z. 2014. Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 129–142.

Yang, S.; Liu, J.; and Zhao, K. 2022. GETNext: Trajectory flow map enhanced transformer for next POI recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1144–1153.

Yu, J.; Wang, C.; Luo, G.; Hou, L.; Li, J.; Tang, J.; Huang, M.; and Liu, Z. 2020. Expanrl: Hierarchical reinforcement learning for course concept expansion in MOOCs. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 770–780.

Zhang, J.; Hao, B.; Chen, B.; Li, C.; Chen, H.; and Sun, J. 2019. Hierarchical reinforcement learning for course recommendation in MOOCs. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, 435–442.

Zhao, P.; Luo, A.; Liu, Y.; Xu, J.; Li, Z.; Zhuang, F.; Sheng, V. S.; and Zhou, X. 2020. Where to go next: A spatio-temporal gated network for next poi recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 2512–2524.