Reasoning Over Space: Enabling Geographic Reasoning for LLM-Based Generative Next POI Recommendation

Dongyi Lv[1,2*], **Qiuyu Ding**[2], **Heng-Da Xu**[2], **Zhaoxu Sun**[2],
Zhi Wang[1†], **Feng Xiong**[2], **Mu Xu**[2],
[1]School Of Software Engineering, Xi'an Jiaotong University, [2]Amap, Alibaba Group
lvdongyi@stu.xjtu.edu.cn, zhiwang@xjtu.edu.cn,
{qiuyu.dqy,xuhengda.xhd,szx430629,xf250971,xumu.xm}@alibaba-inc.com

Abstract

Generative recommendation with large language models (LLMs) reframes prediction as sequence generation, yet existing LLM-based recommenders remain limited in leveraging geographic signals that are crucial in mobility and local-services scenarios. Here, we present REASONING OVER SPACE (ROS), a framework that utilizes geography as a vital decision variable within the reasoning process. ROS introduces a Hierarchical Spatial Semantic ID (SID) that discretizes coarse-to-fine locality and POI semantics into compositional tokens, and endows LLM with a three-stage Mobility Chain-of-Thought (CoT) paradigm that models user personality, constructs an intent-aligned candidate space, and performs locality informed pruning. We further align the model with real world geography via spatial-guided Reinforcement Learning (RL). Experiments on three widely used location-based social network (LBSN) datasets show that ROS achieves over 10% relative gains in hit rate over strongest LLM-based baselines and improves cross-city transfer, despite using a smaller backbone model.[‡]

1 Introduction

Generative recommendation with large language models (LLMs) (Rajput et al., 2023; Zhai et al., 2024; Senel et al., 2024; Qu et al., 2025; Gao et al., 2024; Lee et al., 2025) demonstrate a strong, promising alternative to conventional retrieval and ranking pipelines in practice. By framing recommendation as sequence generation, these models can be trained end-to-end, and often benefit from scaling in both data and model capacity(Deng et al., 2025; Chen et al., 2024; Han et al., 2025), which improves overall generalization and reduces reliance on task specific architectures.

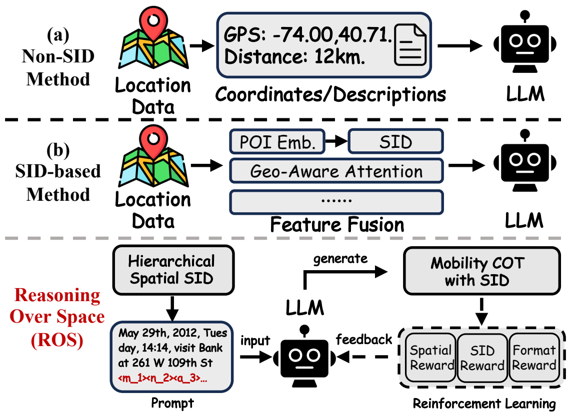

Figure 1: Comparison of different paradigms for incorporating geographic signals into LLM recommendation.

Without explicitly modeling geographic signals, a recommender struggles to capture core mobility patterns, such as users preference for short distance transitions and the frequent consecutiveness of spatially proximate Point-of-Interests[1](POIs). It also fails to exploit spatiotemporal feasibility priors, most notably that long distance moves are unlikely within short time intervals, which can lead to geographically implausible recommendations.

Despite their promise, current LLM-based generative recommenders remain limited in incorporating geographic signals that matter in mobility and local-services(Zhang et al., 2025; Feng et al., 2024a; Zhong et al., 2025). In mobility and local-services scenarios, next POI recommendation seeks to generate the next visited POI given a user's chronologically ordered check in history. Existing work typically incorporates location either by converting it into text inputs, such as appending coordinates and POI descriptions (Feng et al., 2024a), treating it as auxiliary numeric features (Wang et al., 2025; Liu et al., 2025; Jiang et al., 2025) that are fused into representations, or by fusing location via geo-aware self attention(Wei et al., 2025), see Figure 1. While these designs expose location con-

*Work done during the internship at Amap, Alibaba Group.
†Corresponding author.
‡Code available at https://github.com/amap-cvlab/Reasoning-Over-Space

[1]A POI is a place that a user can visit, typically represented by its geographic coordinates and semantic metadata.

Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 7322–7336
July 2-7, 2026 ©2026 Association for Computational Linguistics

text to the model, they do not leverage the model's reasoning capability to treat geography as a first class decision variable, making it difficult to compare candidates spatially, assess travel feasibility, and rule out geographically implausible options.

To bridge this critical gap, we propose a novel framework Reasoning Over Space (ROS) that incorporate geographical information in reasoning itself. Specifically, we (1) Propose a Hierarchical Spatial Semantic ID (SID) that encodes coarse-to-fine locality together with functional semantics as compositional tokens;(2) A recommender model is trained to apply a structured three stage reasoning paradigm that explicitly uses locality constraints for candidate elimination; and (3) Align the model with real world geography via spatial-guided Reinforcement Learning (RL).

2 Methodology

In this section, we first formulate the next POI recommendation problem, then introduce a Hierarchical Spatial SID for structured POI representation, describe a three-stage Mobility CoT reasoning paradigm, and finally detail our spatial-guided RL procedure. The overall framework of ROS is presented in Figure 2.

2.1 Problem Formulation.

Next POI recommendation aims to predict a user's next visited POI from their historical check-in behavior in location-based social networks (LBSN). For each user, we observe a chronologically ordered check-in trajectory $\mathcal{H} = \{(p_i, t_i, c_i, \mathbf{x}_i)\}_{i=1}^{n}$, where p_i denotes the visited POI, t_i is the timestamp, c_i is the POI category, and $\mathbf{x}_i \in \mathbb{R}^2$ is the GPS coordinate of p_i. Given history trajectories of users $\mathcal{H}$, the goal is to predict the next POI p_{n+1} that the user will visit. Each query is associated with a unique ground truth next POI, and the recommendation is considered correct if p_{n+1} exactly matches the ground truth.

2.2 Spatial Semantic Tokenization

2.2.1 Hierarchical Spatial SID

Human mobility[2] is governed not merely by temporal correlations, but by a latent structure that couples geography, semantics, and behavioral regularities. Some recent works introduce SID of POIs, but they still fail to exploit fine grained spatial structure while remaining high collision rate, weakening the model's ability to reason about mobility. To reveal and exploit the latent spatial semantic structure underlying human mobility, we construct a hierarchical spatial SID that transforms the raw POI into a symbolic, interpretable spatial semantic coordinate system, enabling LLMs to reason over location, function, and geographic granularity in a discrete and compositional manner. We decompose each POI p into three components:

$$\text{SID}(p) = [g(p)\ ;\ s(p)\ ;\ u(p)]$$

where $g(p)$ encodes geospatial locality, $s(p)$ captures functional semantics, and $u(p)$ is a lightweight discriminator that ensures uniqueness within each spatial semantic class.

Geospatial Prefix. Each POI coordinate is converted into an S2 Cell Id(Google Inc., 2023) and represented as a fixed hex string. Then we remove a dataset wide shared prefix and keep the next $2 \times B$ hex digits, which are grouped into B tokens:

$$g(p) = \left[\tilde{h}_{1:2}, \ldots, \tilde{h}_{2B-1:2B}\right]$$

Higher order tokens correspond to larger spatial regions, inducing a natural hierarchy in the discrete prefix that nearby POIs tend to share similar prefix. This hierarchy later allows the model to receive partial reward for predicting the correct coarse region even when the exact POI is wrong, and it naturally supports region level generalization.

Semantic Anchor. To encode functional roles, we embed each POI's category description using `Qwen-0.6B` and quantize the embedding via RQ-VAE(Lee et al., 2022). The resulting discrete tokens $s(p)$ clusters POIs with similar semantics, making semantic similarity explicitly accessible to the language model.

Differentiating Suffix. The suffix $u(p)$ carries no additional semantics, it only disambiguates POIs that share the same prefix and semantic anchor.

Together, the three components turn opaque POI IDs into an interpretable, hierarchical, and uniquely identifying code. This unified representation makes it easier for our model to acquire spatial semantic regularities during pretraining, and later provides a structured space on which we define SID level consistence rewards in reinforcement learning.

2.2.2 SID Grounding Pretraining

After constructing the SID representation, we pretrain the model to understand and utilize the SID

[2] Human mobility refers to the time resolved movement of individuals across geographic locations, commonly represented as a spatiotemporal trajectory of visited places and displacements(Gonzalez et al., 2008).

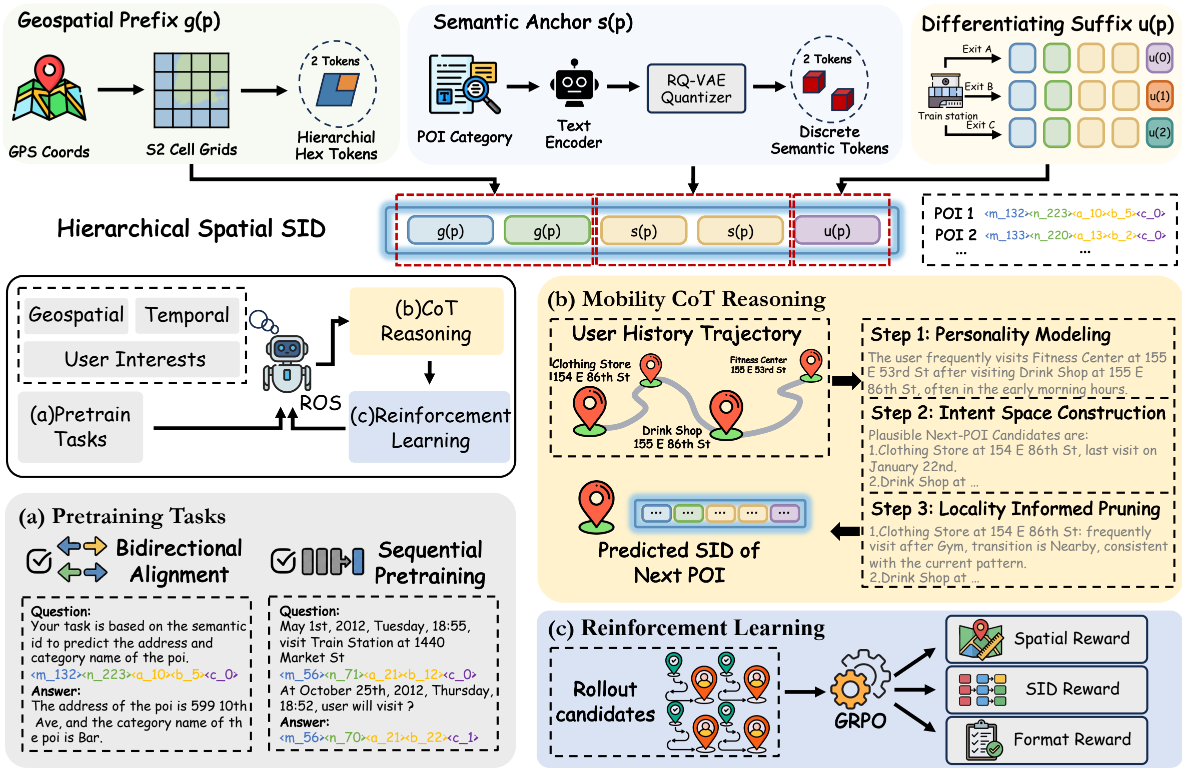

Figure 2: Overall Framework of our method.

hierarchy leveraging two objectives.

Bidirectional Text SID Alignment. We align each SID with its category and address via supervised fine-tuning (SFT) in both directions: the model is trained to predict the SID from text and to generate text from the SID. This bidirectional alignment grounds SID tokens in functional semantics and fine-grained neighborhood cues.

Sequence Pretraining for Mobility. In order to expose the model to temporal, semantic, and coarse spatial regularities such as recency, category repetition, and neighborhood continuity, we further train the model on an auto-regressive next POI recommendation task.

While this pretraining stage exposes the model to spatial semantic structure and generic mobility patterns, it still treats next POI recommendation as a single step mapping. In the next section, we explicitly structure the recommendation as a multi step reasoning process.

2.3 Mobility Chain-of-Thought Reasoning

Human mobility is not a flat classification problem over POI IDs, but a sequence of constrained decisions driven by time, intent, and geography. To reflect this, we model next POI recommendation as an explicit reasoning process rather than a one-shot label prediction. Instead of mapping a trajectory to the next POI in a single opaque step, we fine-tune the LLM to generate Mobility CoT traces that decompose the decision into three deterministic stages. Starting from the most recent check-ins, the model progressively explores, evaluates, and narrows down plausible candidates.

To make geographic reasoning accessible in natural language, we serialize each POI with its category name and street-level address. Compared with arbitrary POI IDs, street addresses provide fine-grained neighborhood cues, allowing the model to naturally perceive spatial locality and recurring areas directly from the given information.

Personality Modeling. In order to ground the recommendation in the user's mobility persona, we begin by distilling the most recent check-ins into a compact, evidence-based profile. Scanning the trajectory backward, the model extracts factual statements about last visit times, weekend or holiday indicators, repeated category or regional patterns, and recent movement characteristics. Since POIs are represented with both its category name and street-level address, the extracted evidence can also highlight locality patterns, which further supports geographic plausibility judgments in later steps.

Intent Space Construction. Next, the model aligns the candidate space with the user's intent by generating plausible next POI candidates conditioned on recent behavioral evidence. We encourage model to generate a recency-first candidate space, guiding the model to prioritize more recently visited POIs before older ones. This inductive bias introduces an implicit temporal structure that captures both short-term intent and habitual preferences, while also removes spurious positional cues introduced during data synthesis, reducing bias from arbitrary candidate permutations.

Locality Informed Pruning. In order to make geography a vital decision criterion, we finally evaluate and prune candidates using both locality cues and feasibility signals. Street-level addresses provide neighborhood-level anchors that allow the model to detect whether a candidate stays within the user's active area, while precomputed transition distances in the prompt offer a constraint on movement plausibility. By jointly checking temporal consistency, semantic fit, and spatial feasibility, the model deterministically discards incompatible candidates and retains only those that satisfy the extracted constraints.

Using a teacher model to generate Mobility CoT traces as supervision, the student model learns to emulate this three-step paradigm, enabling a structured and reproducible reasoning process for next POI recommendation rather than depending on shallow information from trajectories.

2.4 Spatial-Guided Reinforcement Learning

Supervised CoT training enable our model apply the three step Mobility CoT paradigm and verbalize its reasoning, but it does not by itself guarantee geographically faithful recommendations. To further align the model with real world geography and the hierarchical SID space, we apply spatial-guided RL with a composite reward that jointly encourages geographically plausible recommendations, SID level correctness, and format consistent reasoning.

2.4.1 Spatial Grounding Reward

Let d be the haversine distance(Sinnott, 1984) between predicted and ground-truth POIs. We transform it as $\tilde{d} = \log(1 + d)$ and map d to a clipped linear reward:

$$r_{\text{dist}}(d) = \text{clip}\Big(r_{\max} + \kappa(\tilde{d} - \tilde{d}_{\text{near}}),\, r_{\min},\, r_{\max}\Big)$$

where κ is the slope between two thresholds $\tilde{d}_{\text{near}}$ and $\tilde{d}_{\text{far}}$[3]. The *log transform* makes the reward sensitive to local deviations while saturating penalties for very distant recommendations.

2.4.2 Hierarchically Weighted SID Reward

Each POI is represented by SID, which composed by geospatial prefix $g(p)$, semantic anchor $s(p)$, and differentiating suffix $u(p)$.

$$\text{SID}(p) = [g(p); s(p); u(p)]$$

We decompose the geospatial prefix and semantic anchor into hierarchical sub-tokens and assign monotonically decreasing weights along the hierarchy, so that coarse location and category matches contribute more than fine-grained matches. The resulting score defines a base correctness reward r_{base}. Crucially, we further introduce a hard exactness bonus λ_u when the entire SID, including the suffix, is predicted exactly:

$$r_{\text{acc}} = \min\{1,\, r_{\text{base}} + \lambda_u \mathbb{I}[\text{SID} = \widehat{\text{SID}}]\}.$$

The format reward r_{fmt} is a binary signal that simply checks whether the generated trace conforms to the prescribed three step CoT paradigm.

2.4.3 Unified Reward Optimization

To jointly encode geographic preference, symbolic correctness, and format adherence into a single training signal, for each generated trajectory trace, we define the overall reward as:

$$r = r_{\text{fmt}} + \alpha r_{\text{acc}} + \beta r_{\text{dist}}.$$

Therefore, geographic preference, SID level symbolic correctness, and structured reasoning are optimized in a unified way. We then optimize the model with Group Relative Policy Optimization (GRPO)(Shao et al., 2024). As a result, the optimized model internalizes geography and SID hierarchy as actionable preferences, yielding recommendations that are not only symbolically correct but also spatially plausible, while maintaining trace-to-decision consistency under the three-stage Mobility CoT paradigm.

Table 1: **Next POI recommendation performance.** All results are HR@1.

Category	Method	NYC	TKY	CA
Traditional	PRME	0.1159	0.1052	0.0521
Neural-based	GETNext	0.2435	0.1829	0.1357
	TPG	0.2555	0.1420	0.1749
	MTNext	0.2620	0.2575	0.1453
	STHGCN	0.2734	0.2950	0.1730
	ROTAN	0.3106	0.2458	0.2199
LLM-based	ST-GR	0.2920	0.2610	0.1659
	LLM4POI	0.3372	0.3035	0.2065
	GNPR-SID	0.3618	0.3062	0.2403
	GA-LLM	0.3919	0.3482	0.2566
	CoAST	0.4027	0.3310	0.2721
	ROS	**0.4478**	**0.3864**	**0.3149**
	vs. SOTA (Δ%)	**+11.2%↑**	**+11.0%↑**	**+15.7%↑**

Table 2: **Ablation on SID design and pretraining tasks.** T1: sequence modeling pretraining; T2: text-SID alignment pretraining. All results are HR@1.

SID type	T1	T2	NYC	TKY	CA
Traditional	✓		0.3704	0.3310	0.2608
HS-SID	✓		0.3877	0.3338	0.2629
HS-SID	✓	✓	**0.3918**	**0.3372**	**0.2657**

3 Experiments

3.1 Experimental Setting

We thoroughly evaluate our method on three widely used LBSN benchmarks: Foursquare-NYC, Foursquare-TKY (Yang et al., 2015), and Gowalla-CA(Cho et al., 2011). We follow the LLM4POI(Li et al., 2024) preparation pipeline: (i) Filter users/POIs with fewer than 10 check-ins; (ii) Sort check-ins chronologically and segment trajectories using a time-gap threshold; (iii) Split data into 80%/10%/10% train/validation/test sets in temporal order while keeping validation/test users and POIs observed in training. We additionally apply reverse geocoding(Li, 2018) to obtain street-level addresses as auxiliary textual cues. For each experiment, we run inference for three times and report the mean score.

Baselines We compare our method against three categories of baselines: (i) Traditional baseline PRME (Feng et al., 2015); (ii) Neural-based baselines include GETNext (Yang et al., 2022), TPG (Luo et al., 2023), MTNext (Hu et al., 2018), STHGCN (Yan et al., 2023), and ROTAN (Feng et al., 2024b); and (iii) LLM-based baselines include SpaceTime-GR (Lin et al., 2025), LLM4POI (Li et al., 2024), GNPR-SID (Wang et al., 2025), GA-LLM (Liu et al., 2025), and CoAST (Zhai et al., 2025). Detailed descriptions of these baselines are provided in Appendix A.2.

Evaluation Metrics HitRate@1 (HR@1) is employed as the main evaluation metric. In our settings, each request is associated with a unique ground-truth POI, and the model recommends exactly one POI. Under this one-to-one recommendation regime, HR@1 directly measures whether the model makes a correct decision, providing an unambiguous and faithful reflection of recommendation accuracy.

Implementation Details All experiments are conducted on 8×NVIDIA H20 GPUs. We initialize the student model from `Qwen3-4B` (Yang et al., 2025). `Qwen3-235B-A22B` is employed as a teacher to generate three-stage CoT traces, and fine-tune the 4B student on these traces for 2 epochs. Additional training and hyperparameter details are provided in the appendix A.3.

3.2 Main Results

Table 1 reports next POI recommendation performance on three LBSN benchmarks. Despite using a compact 4B student backbone, our method consistently outperforms the strongest LLM-based baselines that are built on 7B models (CoAST and GA-LLM). On NYC and CA, we surpass CoAST with relative HR@1 gains of +11.2% and +15.7%, respectively; on TKY, we improve over GA-LLM by +11.0% HR@1. Notably, these improvements are achieved without resorting to backbone scaling. By turning mobility into a compositional spatial-semantic decision process, the model learns a stronger notion of "where" and "why" behind the next visit. This suggests that structured spatial semantics and explicit reasoning supervision can unlock more effective generalization than merely enlarging the backbone.

3.3 Ablation Study

Analysis of SID Expressivity. As shown in Table 2, we first replace our hierarchical spatial SID with traditional non-hierarchical SID while keeping all prompts and training settings identical. Compared with traditional SID, our hierarchical spatial SID yields a clear improvement on NYC and comparable performance on TKY and CA, indicating that the additional geospatial granularity primarily benefits regions with higher spatial density. Incorporating the text-SID alignment objective further im-

[3] We set d_{near}=0.1 km and d_{far}=3.0 km. These thresholds align with common practice in trajectory and POI mining for defining local tolerance, and trip and itinerary recommendation for neighborhood scale pruning (Zheng, 2015; Yeow et al., 2021; Gao et al., 2022; Halder et al., 2022).

Table 3: **Ablation on Mobility CoT stages.** Step 1: personality modeling; Step 2: intent space construction; Step 3: locality informed pruning; Each row enables a subset of stages and reports HR@1.

Step 1	Step 2	Step 3	NYC	Gain(Δ%)
			0.3918	–
✓			0.4105	+4.8%
	✓		0.4098	+4.6%
		✓	0.4029	+2.8%
	✓	✓	0.4119	+5.1%
✓		✓	0.4091	+4.4%
✓	✓		0.4140	+5.7%
✓	✓	✓	**0.4181**	**+6.7%**

Table 4: **Perturbing geographic cues after RL.** We report HR@1 on three datasets.

Variant	NYC	TKY	CA
Full model	**0.4478**	**0.3864**	**0.3149**
Random Address	0.4347	0.3721	0.2971
W/O Address	0.4195	0.3444	0.2842
W/O Distance	0.4416	0.3849	0.3139

proves model performance across all three datasets, showing that grounding discrete SID tokens in natural language descriptions enhances functional discrimination.

Analysis of Mobility CoT. Table 3 indicates the necessity of each Mobility CoT stage. While any single stage improves over direct recommendation, the best result is achieved only when all three stages are composed. Suggesting that Mobility CoT is not a collection of independent heuristics but a coupled reasoning scaffold, PERSONALITY MODELING distills trajectory evidence into stable constraints, INTENT SPACE CONSTRUCTION organizes a structured candidate space under these constraints, and LOCALITY INFORMED PRUNING provides the decisive verification signal that eliminates candidates inconsistent with the extracted evidence. Removing any stage weakens one link of this consecutive CoT, leading to a drop even when the remaining stages are present.

To better investigate how geographic cues support the above paradigm, we perturb the POI representation and distance annotations at inference time after reinforcement learning, as summarized in Table 4. Replacing each POI address with a random address leads to a consistent decline, indicating that street-level addresses act as a native locality prior that helps the model relate candidates through shared areas. Removing addresses entirely causes a larger degradation, consistent with the fact that our CoT relies on category name and street-level address as an anchor to disambiguate POIs during candidate pruning. Finally, removing explicit transition distances also hurts performance, suggesting that distances provide a complementary feasibility prior that sharpens the verifier when ruling out implausible long jumps, beyond what can be inferred from address tokens alone.

Table 5: **Effect of correctness bonus weight λ_u.** All results are HR@1.

λ_u	0.0	0.1	0.3	0.5	1.0
NYC	0.4423	**0.4478**	0.4444	0.4395	0.4381

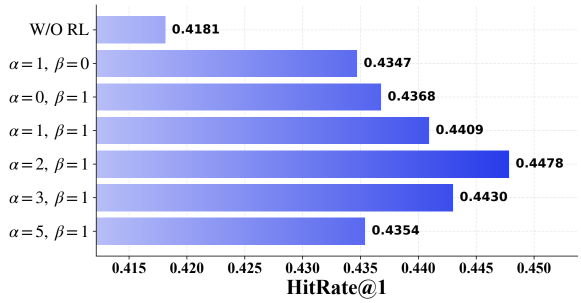

Figure 3: **Impact of reward weighting in spatial-guided RL.** α and β indicate the weights of the hierarchical correctness reward and the distance-based grounding reward, respectively. Results are reported in HR@1 on NYC.

Analysis of Spatial-Guided RL. Spatial-guided RL consistently improves next POI generation beyond supervised Mobility CoT fine-tuning, indicating that RL provides an additional alignment stage that better reflects mobility constraints. Figure 3 further shows that using either hierarchical correctness or distance grounding alone already brings noticeable gains, while combining them yields the strongest performance. This suggests the two signals are complementary, r_{acc} encourages SID-level semantic correctness, whereas r_{dist} injects a spatial-guided preference that discourages geographically implausible candidates. Importantly, the weighting study reveals a clear trade-off, moderate emphasis on correctness is beneficial, but overly large α degrades performance, likely because discrete SID matching starts to dominate optimization and weakens the influence of spatial constraints during candidate discrimination.

Table 5 further analyzes the correctness bonus weight λ_u, where a small bonus (λ_u=0.1) yields the best HR@1, while larger values gradually hurt performance. This indicates that a small exact match bonus can help training by sharpening the preference for fully correct predictions, whereas an excessive λ_u makes the optimization overly reliant on the sparse all correct signal. Based on these sensitivity analyses, we adopt α=2, β=1, and λ_u=0.1

Table 6: **Cross-city generalization performance.** Models are trained on one city and directly evaluated on other cities. All results are HR@1.

Model	Trained on	NYC	TKY	CA
LLM4POI	NYC	0.3372	0.2594	0.1885
	TKY	0.3463	0.3035	0.1960
	CA	0.3344	0.2600	0.2065
GA-LLM	NYC	0.3826	0.3018	0.2053
	TKY	0.4059	0.3429	0.2273
	CA	0.3670	0.3065	0.2499
Ours	NYC	**0.4478**	**0.3253**	**0.2556**
	TKY	**0.4202**	**0.3864**	**0.2650**
	CA	**0.4299**	**0.3294**	**0.3149**

Table 7: **Effect of history sequence length.** HR@1 with different maximum trajectory history lengths.

History Length	NYC	TKY	CA
30	0.4153	0.3403	0.2835
50	0.4202	0.3626	0.3024
100	0.4416	0.3742	0.3013
300	**0.4478**	**0.3864**	**0.3149**

as robust defaults that balance semantic correctness with geographic plausibility.

3.4 Generalization Analysis

Cross City Generalization. To evaluate cross city generalization, we train the model on one city and infer directly on the other two unseen cities. As shown in Table 6, our method consistently achieves the best transfer performance. We attribute this to the shared SID token space, including geospatial prefix and semantic anchor sub-tokens that are drawn from a global, city agnostic vocabulary and have been observed during pretraining. Consequently, cross city transfer does not require extrapolating to unseen symbols, but instead amounts to recombining familiar spatial and semantic primitives in new configurations. Mobility CoT and spatial-guided RL further encourage this compositional, geography-aware behavior, enabling the model to learn transferable mobility regularities instead of memorizing city-specific trajectories.

Effect of History Sequence Length We further analyze the effect of trajectory history length, see Table 7. Increasing the available history sequence length consistently improve the model's performance across all three datasets. Notably, the gains exhibit a clear marginal effect, when the trajectory history is short, extending the history leads to substantial performance improvements, while further increasing the history length beyond a moderate range yields only limited additional gains. Suggesting that most informative mobility signals are concentrated in recent and mid-range check-ins, whereas very long histories mainly provide redundant or weakly relevant information.

3.5 Analysis

Geographic Error Distribution Beyond accuracy, we evaluate the geographic plausibility of next POI recommendation by measuring the spatial deviation between the predicted POI and the ground-truth POI. For each test instance, we compute the Haversine distance error from the GPS coordinates of the predicted and true POIs, and report the cumulative distribution function (CDF) of distance errors in Figure 4. A method whose CDF rises faster yields smaller spatial errors for a larger fraction of cases. As shown, our approach consistently dominates the baseline across all three cities, indicating more localized and geographically feasible recommendations. In particular, the 50th, 75th, and 90th percentile distance errors are consistently reduced across datasets, suggesting that incorporating geographic reasoning not only improves ranking metrics, but also mitigates large-distance mismatches and better aligns recommendations with real world human mobility patterns.

Top-K Recommendation via Two-Stage Inference. Our main setting evaluates one-to-one next POI recommendation with HR@1. To test whether Mobility CoT also supports ranking, we introduce a lightweight Top-K variant that keeps PERSONALITY MODELING and INTENT SPACE CONSTRUCTION to form an evidence guided candidate set, and replaces the iterative LOCALITY INFORMED PRUNING with brief, ordered rationales that justify candidates in rank order. We then apply a two-pass inference procedure: we first generate the rationale-augmented CoT trace, and then condition on this trace to perform constrained beam search and output a ranked list. Notably, the ranked outputs often include POIs that do not appear in the user's historical sequence, indicating improved exploration beyond repeatedly visiting the same places. As shown in Table 8, this variant consistently improves HR and NDCG, outperforming the strongest baseline(Kang and McAuley, 2018; Sun et al., 2019; Hidasi et al., 2015; Tang and Wang, 2018; Zhou et al., 2020; Luo et al., 2023; Feng et al., 2024b; Rajput et al., 2023; Wang et al., 2025; Wei et al., 2025).

Analysis of CoT-Free Varient We further compare a CoT-free variant of our ROS method with

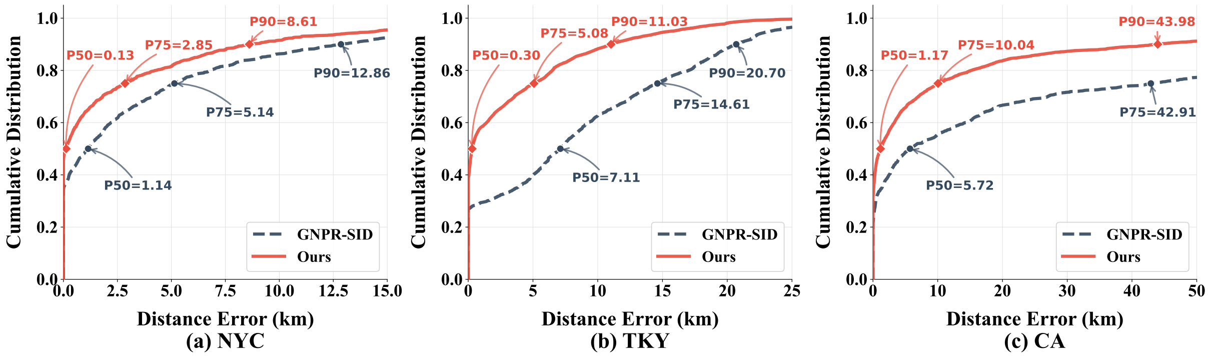

Figure 4: **Cumulative Distribution of three datasets.** Larger area under curve indicating better performance.

Table 8: **Top-K evaluation on NYC.** HR@K (Recall@K) and NDCG@K with $K \in \{5, 10, 20\}$.

Metric	Traditional					POI Related		Generative			
	SASRec	BERT4Rec	GRU4Rec	Caser	S^3-Rec	TPG	Rotan	TIGER	GNPR-SID	OneLoc	Ours
HR@5	0.3151	0.2857	0.1977	0.2883	0.3071	0.3551	0.4448	0.4965	0.5311	0.6107	**0.6372**
HR@10	0.3896	0.3564	0.2460	0.3570	0.3854	0.4441	0.5223	0.5514	0.5942	0.6563	**0.6897**
HR@20	0.4506	0.4130	0.2889	0.4135	0.4503	0.5121	0.5834	0.6001	0.6455	0.6977	**0.7284**
NDCG@5	0.2224	0.2074	0.1442	0.2044	0.2235	0.2464	0.3471	0.4131	0.4430	0.5355	**0.5398**
NDCG@10	0.2467	0.2304	0.1599	0.2267	0.2489	0.2755	0.3723	0.4276	0.4634	0.5504	**0.5570**
NDCG@20	0.2622	0.2448	0.1708	0.2410	0.2654	0.2927	0.3878	0.4443	0.4766	0.5608	**0.5668**

Table 9: **HR@1 of CoT-free variant.**

Training Setting	NYC	TKY	CA
Pretrain	0.3918	0.3372	0.2657
ROS†	**0.4181**	**0.3527**	**0.2874**

other baselines, see Table 9. Here we employ guided decoding to suppress the explicit think stage and let the model output only the final POI recommendation. Even without generating CoT traces at inference time, ROS† still outperforms all baselines on all three datasets. This gap shows that the proposed spatial semantic representation, structured Mobility CoT training, and spatial-guided RL not only improve performance when reasoning is explicitly generated, but also act as an effective inductive bias that strengthens the underlying model even in a pure recommendation setting.

4 Related Work

Next POI Recommendation. Early approaches model sequential transitions or personalized ranking, including LSTM (Hochreiter and Schmidhuber, 1997) and PRME (Feng et al., 2015). Later methods enhance trajectory modeling with spatiotemporal attention or graph structure, such as STAN (Luo et al., 2021), GetNext (Yang et al., 2022), and STHGCN (Yan et al., 2023). More recently, LLM-based POI recommenders describe places in natural language and generate next POI outputs, represented by LLM4POI (Li et al., 2024), while SID further improve generalization beyond numeric IDs, as explored in GNPR-SID (Wang et al., 2025) and CoAST (Zhai et al., 2025).

Generative Recommendation. Generative recommendation replaces retrieval-and-rank with autoregressive generation conditioned on user context (Rajput et al., 2023; Senel et al., 2024; Gao et al., 2024; Qu et al., 2025). Recent work highlights the importance of semantic item identifiers for open-vocabulary generation (Deng et al., 2025; Chen et al., 2024; Han et al., 2025; Wang et al., 2025). For generative next POI recommendation, geographic cues are commonly injected via coordinates and addresses in prompts or via auxiliary continuous location features (Feng et al., 2024a; Liu et al., 2025; Wei et al., 2025).

Chain-of-Thought. CoT reasoning improves LLM decision making (Wei et al., 2022), and structured CoT stabilizes reasoning with staged templates (Wang et al., 2023). Prior CoT for recommendation often uses free-form preference summaries or rationales (Yue et al., 2025; Tsai et al., 2024), which can be informative but may not enforce feasibility. In contrast, our Mobility CoT explicitly uses locality informed pruning as a hard reasoning criterion for next POI selection.

5 Conclusion

We present ROS, a generative next POI recommendation framework that treats geography as an explicit decision variable. ROS introduces a Hierarchical Spatial Semantic ID and a three-stage Mobility CoT paradigm to construct and verify candidates under spatial and semantic constraints, then refines the model via spatial-guided RL. Extensive experiments show that ROS consistently improves hit rate and produces more geographically plausible predictions, while maintaining strong cross-city transfer. In future work, we will explore integrating abundant transportation priors for more realistic mobility reasoning.

Limitations

While ROS demonstrates promising advancements in next POI recommendation, it also exhibits limitations in its geo-temporal feasibility modeling. By relying on a simplifying monotonic prior that long distance transitions are unlikely within short time intervals, ROS may underweight plausible rapid long range moves enabled by transportation networks such as metro rail or highways, particularly when both origin and destination are near transit hubs. Future work could incorporate network aware impedance signals such as estimated travel time or route based distance and adopt soft feasibility constraints to better preserve probability mass for such rare but realistic transitions.

References

Junyi Chen, Lu Chi, Bingyue Peng, and Zehuan Yuan. 2024. Hllm: Enhancing sequential recommendations via hierarchical large language models for item and user modeling. *arXiv preprint arXiv:2409.12740*.

Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, page 1082–1090, New York, NY, USA. Association for Computing Machinery.

Jiaxin Deng, Shiyao Wang, Kuo Cai, Lejian Ren, Qigen Hu, Weifeng Ding, Qiang Luo, and Guorui Zhou. 2025. Onerec: Unifying retrieve and rank with generative recommender and iterative preference alignment. *arXiv preprint arXiv:2502.18965*.

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*, volume 15, pages 2069–2075.

Shanshan Feng, Haoming Lyu, Fan Li, Zhu Sun, and Caishun Chen. 2024a. Where to move next: Zero-shot generalization of llms for next poi recommendation. In *2024 ieee conference on artificial intelligence (cai)*, pages 1530–1535. IEEE.

Shanshan Feng, Feiyu Meng, Lisi Chen, Shuo Shang, and Yew Soon Ong. 2024b. Rotan: A rotation-based temporal attention network for time-specific next poi recommendation. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*, pages 759–770.

Qiang Gao, Wei Wang, Kunpeng Zhang, Xin Yang, Congcong Miao, and Tianrui Li. 2022. Self-supervised representation learning for trip recommendation. *Knowledge-Based Systems*, 247:108791.

Shen Gao, Jiabao Fang, Quan Tu, Zhitao Yao, Zhumin Chen, Pengjie Ren, and Zhaochun Ren. 2024. Generative news recommendation. In *Proceedings of the ACM Web Conference 2024*, pages 3444–3453.

Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. 2008. Understanding individual human mobility patterns. *nature*, 453(7196):779–782.

Google Inc. 2023. S2 geometry library. `https://s2geometry.io/`. Accessed: 2025-03-01.

Sajal Halder, Kwan Hui Lim, Jeffrey Chan, and Xiuzhen Zhang. 2022. Efficient itinerary recommendation via personalized poi selection and pruning. *Knowledge and Information Systems*, 64(4):963–993.

Ruidong Han, Bin Yin, Shangyu Chen, He Jiang, Fei Jiang, Xiang Li, Chi Ma, Mincong Huang, Xiaoguang Li, Chunzhen Jing, and 1 others. 2025. Mtgr: Industrial-scale generative recommendation framework in meituan. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, pages 5731–5738.

Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

Guangneng Hu, Yu Zhang, and Qiang Yang. 2018. Mtnet: a neural approach for cross-domain recommendation with unstructured text. *KDD deep learning day*, pages 1–10.

Hao Jiang, Guoquan Wang, Donglin Zhou, Sheng Yu, Yang Zeng, Wencong Zeng, Kun Gai, and Guorui Zhou. 2025. Llm-aligned geographic item tokenization for local-life recommendation. *arXiv preprint arXiv:2511.14221*.

Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*, pages 197–206. IEEE.

Doyup Lee, Chiheon Kim, Saehoon Kim, Minsu Cho, and Wook-Shin Han. 2022. Autoregressive image generation using residual quantization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11523–11532.

Sunkyung Lee, Minjin Choi, Eunseong Choi, Hye young Kim, and Jongwuk Lee. 2025. GRAM: Generative recommendation via semantic-aware multi-granular late fusion. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 33294–33312.

Dapeng Li. 2018. 1.08 - geocoding and reverse geocoding. In Bo Huang, editor, *Comprehensive Geographic Information Systems*, pages 95–109. Elsevier, Oxford.

Peibo Li, Maarten de Rijke, Hao Xue, Shuang Ao, Yang Song, and Flora D Salim. 2024. Large language models for next point-of-interest recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1463–1472.

Haitao Lin, Zhen Yang, Jiawei Xue, Ziji Zhang, Luzhu Wang, Yikun Gu, Yao Xu, and Xin Li. 2025. Spacetime-gr: A spacetime-aware generative model for large scale online poi recommendation. *Preprint*, arXiv:2508.16126.

Zhao Liu, Wei Liu, Huajie Zhu, Jianxing Yu, Jian Yin, Wang-Chien Lee, and Shun Wang. 2025. Geography-aware large language models for next poi recommendation. *arXiv preprint arXiv:2505.13526*.

Yan Luo, Haoyi Duan, Ye Liu, and Fu-Lai Chung. 2023. Timestamps as prompts for geography-aware location recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 1697–1706.

Yingtao Luo, Qiang Liu, and Zhaocheng Liu. 2021. Stan: Spatio-temporal attention network for next location recommendation. In *Proceedings of the web conference 2021*, pages 2177–2185.

Haohao Qu, Wenqi Fan, Zihuai Zhao, and Qing Li. 2025. Tokenrec: Learning to tokenize id for llm-based generative recommendations. *IEEE Transactions on Knowledge and Data Engineering*.

Shashank Rajput, Nikhil Mehta, Anima Singh, Raghunandan Hulikal Keshavan, Trung Vu, Lukasz Heldt, Lichan Hong, Yi Tay, Vinh Tran, Jonah Samost, and 1 others. 2023. Recommender systems with generative retrieval. *Advances in Neural Information Processing Systems*, 36:10299–10315.

Lütfi Kerem Senel, Besnik Fetahu, Davis Yoshida, Zhiyu Chen, Giuseppe Castellucci, Nikhita Vedula, Jason Ingyu Choi, and Shervin Malmasi. 2024. Generative explore-exploit: Training-free optimization of generative recommender systems using llm optimizers. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5396–5420.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, and 1 others. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.

Roger W Sinnott. 1984. Virtues of the haversine. *Sky and telescope*, 68(2):158.

Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 1441–1450.

Jiaxi Tang and Ke Wang. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of the eleventh ACM international conference on web search and data mining*, pages 565–573.

Alicia Tsai, Adam Kraft, Long Jin, Chenwei Cai, Anahita Hosseini, Taibai Xu, Zemin Zhang, Lichan Hong, Ed H Chi, and Xinyang Yi. 2024. Leveraging llm reasoning enhances personalized recommender systems. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 13176–13188.

Dongsheng Wang, Yuxi Huang, Shen Gao, Yifan Wang, Chengrui Huang, and Shuo Shang. 2025. Generative next poi recommendation with semantic id. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pages 2904–2914.

Lei Wang, Wanyu Xu, Yihuai Lan, Zhiqiang Hu, Yunshi Lan, Roy Ka-Wei Lee, and Ee-Peng Lim. 2023. Plan-and-solve prompting: Improving zero-shot chain-of-thought reasoning by large language models. *arXiv preprint arXiv:2305.04091*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

Zhipeng Wei, Kuo Cai, Junda She, Jie Chen, Minghao Chen, Yang Zeng, Qiang Luo, Wencong Zeng, Ruiming Tang, Kun Gai, and 1 others. 2025. Oneloc: Geo-aware generative recommender systems for local life service. *arXiv preprint arXiv:2508.14646*.

Xiaodong Yan, Tengwei Song, Yifeng Jiao, Jianshan He, Jiaotuan Wang, Ruopeng Li, and Wei Chu. 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*, pages 403–412.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, and 1 others.

2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Dingqi Yang, Daqing Zhang, Vincent W. Zheng, and Zhiyong Yu. 2015. Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(1):129–142.

Song Yang, Jiamou Liu, and Kaiqi Zhao. 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*, pages 1144–1153.

Lih Wei Yeow, Raymond Low, Yu Xiang Tan, and Lynette Cheah. 2021. Point-of-interest (poi) data validation methods: An urban case study. *ISPRS International Journal of Geo-Information*, 10(11):735.

Weiqi Yue, Yuyu Yin, Xin Zhang, Binbin Shi, Tingting Liang, and Jian Wan. 2025. Cot4rec: Revealing user preferences through chain of thought for recommender systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 13142–13151.

Jiaqi Zhai, Lucy Liao, Xing Liu, Yueming Wang, Rui Li, Xuan Cao, Leon Gao, Zhaojie Gong, Fangda Gu, Jiayuan He, Yinghai Lu, and Yu Shi. 2024. Actions speak louder than words: trillion-parameter sequential transducers for generative recommendations. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org.

Penglong Zhai, Jie Li, Fanyi Di, Yue Liu, Yifang Yuan, Jie Huang, Peng Wu, Sicong Wang, Mingyang Yin, Tingting Hu, and 1 others. 2025. Cognitive-aligned spatio-temporal large language models for next point-of-interest prediction. *arXiv preprint arXiv:2510.14702*.

Qianru Zhang, Peng Yang, Junliang Yu, Haixin Wang, Xingwei He, Siu-Ming Yiu, and Hongzhi Yin. 2025. A survey on point-of-interest recommendation: Models, architectures, and security. *IEEE Transactions on Knowledge and Data Engineering*.

Yu Zheng. 2015. Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(3):1–41.

Lin Zhong, Lingzhi Wang, Xu Yang, and Qing Liao. 2025. Comapoi: A collaborative multi-agent framework for next poi prediction bridging the gap between trajectory and language. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1768–1778.

Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of the 29th ACM international conference on information & knowledge management*, pages 1893–1902.

A Appendix

A.1 Dataset Statistics

Table 10 summarizes dataset statistics after preprocessing. We also report the numbers of trajectories in the train/test splits used throughout our experiments. Overall, the three datasets cover diverse scales in both user activity and POI catalog size, providing a comprehensive testbed for evaluating next POI recommendation methods.

A.2 Baselines

We compare against representative baselines spanning (i) *traditional* metric-embedding methods, (ii) *neural* sequence/graph models with explicit spatio-temporal inductive biases, and (iii) *LLM-based* generative recommenders that leverage textualized mobility context.

PRME (Feng et al., 2015): Learns a metric-embedding space where next-step transition probability increases with proximity between user and POI embeddings, encouraging geographically nearby POIs as plausible next visits.

GETNext (Yang et al., 2022): Constructs a global flow map from collective mobility to provide collaborative priors, and integrates it into a graph-enhanced Transformer for modeling personalized next POI transitions.

TPG (Luo et al., 2023): Treats time as a prompt-like conditioning signal and uses shifted geographic windows to capture time-specific locality, improving temporal sensitivity in location transitions.

MTNet (Hu et al., 2018): Combines memory networks with transfer learning to incorporate auxiliary textual/semantic signals, mitigating data sparsity and improving generalization to infrequent POIs.

STHGCN (Yan et al., 2023): Employs hypergraph convolution to model high-order spatio-temporal relations among users, POIs, and contexts, capturing group-level mobility patterns beyond pairwise transitions.

ROTAN (Feng et al., 2024b): Introduces rotation-based temporal encoding to represent periodic and asymmetric temporal effects, enhancing the modeling of time-evolving mobility dynamics.

SpaceTime-GR (Lin et al., 2025): A generative POI recommender that couples spatio-temporal en-

Table 10: **Dataset statistics after preprocessing.** Two consecutive check-ins are assigned to the same trajectory if their time gap is within 24 hour.

Data	Users	POIs	All Trajs	Valid Trajs	Test Trajs	Category	Check-ins
NYC	1048	4981	14130	1486	1447	318	103941
TKY	2282	7833	65499	7174	7079	291	405000
CA	3957	9690	45123	3744	2864	296	238369

coders with hierarchical geographic indexing, aiming to produce context-consistent next POIs under structured location representations.

LLM4POI (Li et al., 2024): Reformulates trajectories as natural language sequences, enabling LLMs to exploit rich contextual cues (e.g., time, category, and address text) for next POI generation.

GNPR-SID (Wang et al., 2025): Uses semantic discrete POI identifiers (IDs) so that functionally similar places share token-level structure, allowing generative models to transfer knowledge across related POIs.

GA-LLM (Liu et al., 2025): Injects GPS coordinates and transition structure into the LLM input/representation, strengthening spatial generalization and reducing over-reliance on pure textual priors.

CoAST (Zhai et al., 2025): Performs enriched pretraining followed by instruction tuning and reinforcement learning, aligning LLM outputs with real world constraints and improving robustness for location recommendation.

A.3 Implementation Details

SID construction and pretraining. We initialize the student model from `Qwen3-4B` and apply full-parameter fine-tuning throughout. Each POI is mapped to a hierarchical spatial-semantic identifier $\mathrm{SID}(p) = [g(p); s(p); u(p)]$. For the geospatial prefix $g(p)$, we convert each GPS coordinate to an S2 Cell ID at the maximum S2 level, compute the dataset-wide longest common prefix over all POI cell IDs, remove this shared prefix, and keep the next $2B$ hex digits with B=2, grouped into two byte-level locality tokens (g_1, g_2) and implemented as special tokens. For the semantic anchor $s(p)$, we encode the POI category name with `Qwen-0.6B` and take the last hidden state as the POI embedding, then quantize it via a 2-level residual quantizer with codebook size 28 per level, yielding two discrete semantic tokens (s_1, s_2). The differentiating suffix $u(p)$ is a single special token assigned sequentially within each (g, s) pair; empirically at most 8 suffix types are observed across all datasets. We pretrain the model with (i) Bidirectional text SID alignment (predicting SID from category/address text and generating text from SID), and (ii) Autoregressive next POI sequence pretraining under the SID representation.

Mobility CoT supervised fine-tuning. We use *Qwen3-235B* as teacher LLM to generate three-stage Mobility CoT traces and fine-tune the student to reproduce the full trace and output the final POI. During training, we truncate the trajectory history to at most 50 check-ins, while at evaluation we provide up to 300 check-ins. The number of candidates enumerated in INTENT SPACE CONSTRUCTION is not fixed and is decided by the model itself. We compute transition distances between adjacent check-ins (km) and bucket them into five bins: Adjacent (≤ 0.2), Nearby (≤ 1.2), Short hop (≤ 3.0), Far (≤ 10.0), and Long (> 10.0). For SFT, we train for 3 epochs with global batch size 8, learning rate 1×10^{-5}, and 100 warmup steps.

Spatial-Guided RL. We further optimize the model with GRPO using a unified reward $r = r_{\text{fmt}} + \alpha r_{\text{acc}} + \beta r_{\text{dist}}$, where α=2 and β=1. For the distance reward, let d be the Haversine distance (km) between predicted and ground-truth POIs, and $\tilde{d} = \log(1 + d)$. We use two distance thresholds d_{near}=0.1 km and d_{far}=3.0 km, and define a clipped linear reward $r_{\text{dist}}(d) = \text{clip}(1 + \kappa(\tilde{d} - \tilde{d}_{\text{near}}), 0, 1)$ with $\kappa = (0 - 1)/(\tilde{d}_{\text{far}} - \tilde{d}_{\text{near}})$, so predictions within $\sim$100m receive full credit and errors beyond 3km saturate to zero. For the hierarchical SID reward, we assign +0.3 if g_1 matches and +0.5 if both (g_1, g_2) match, and +0.25 if s_1 matches and +0.4 if both (s_1, s_2) match; we additionally add an exactness bonus λ_u=0.1 when the full SID matches exactly, implemented as $0.3\,\mathbb{I}[g_1] + 0.2\,\mathbb{I}[g_1 \wedge g_2] + 0.25\,\mathbb{I}[s_1] + 0.15\,\mathbb{I}[s_1 \wedge s_2] + 0.1\,\mathbb{I}[\mathrm{SID} = \widehat{\mathrm{SID}}]$. The format reward r_{fmt} is set to 2 if a regex validator confirms the required three-stage template. We run GRPO with 4 rollouts per prompt for 3 epochs, learning rate 1×10^{-6}, and PPO-style clipping ϵ=0.2; rollout decoding uses temperature = 1.0 and top-p = 1.0.

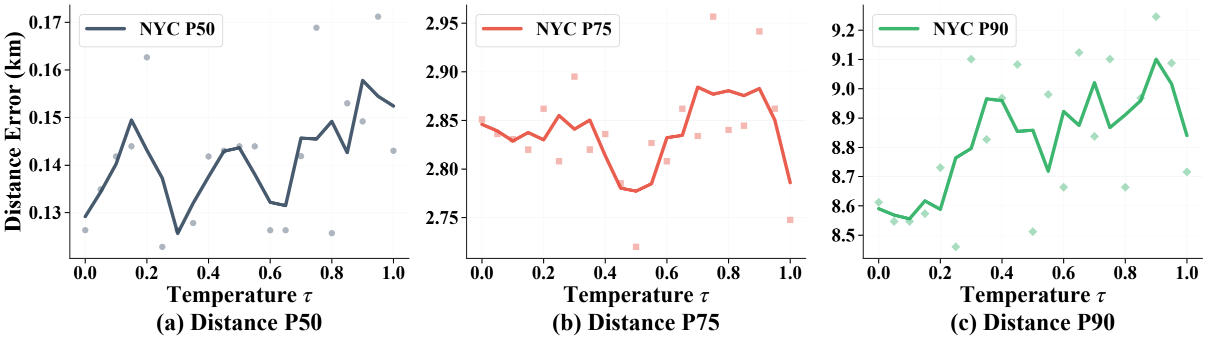

Figure 5: **NYC distance error percentiles versus sampling temperature τ.**

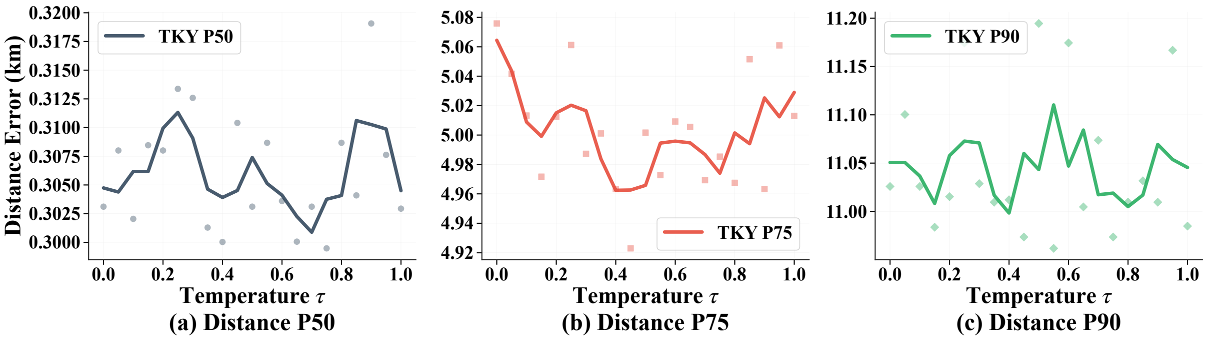

Figure 6: **TKY distance error percentiles versus sampling temperature τ.**

A.4 Hyperparameter Analysis

A.4.1 Impact of Temperature

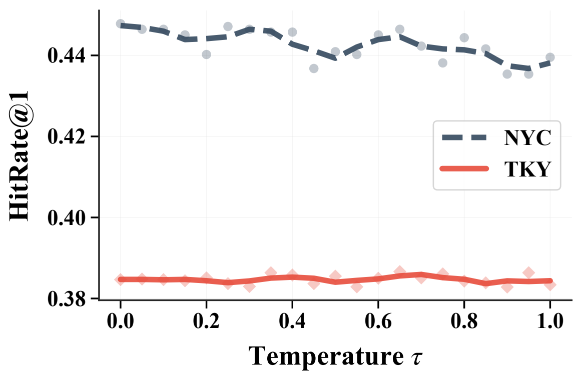

Figure 7: **HitRate@1 versus sampling temperature τ on NYC and TKY.**

As shown in Fig. 7, increasing the sampling temperature τ generally degrades next POI accuracy. This effect is most evident on NYC, the MA3-smoothed trend reveals a clear downward drift as τ increases, indicating that higher decoding randomness gradually pushes predictions away from the most likely destination. TKY appears less sensitive, but still shows no consistent gains from higher temperatures. Distance-error percentiles further support this observation (Figs. 5 and 6). As τ grows, the tail errors tend to increase or become more volatile, suggesting that higher temperature mainly introduces occasional long-range deviations rather than improving typical cases. Based on these results, we adopt deterministic decoding with $\tau = 0$ for all experiments to obtain the most stable and strongest overall performance.

A.4.2 Impact of Mobility CoT Data Scale

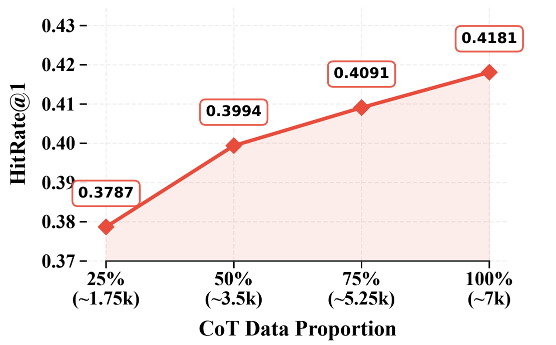

Figure 8: **HR@1 of scaling Mobility CoT supervision on NYC.**

To understand how much structured reasoning supervision is needed, we vary the amount of Mobility CoT traces used for NYC in the SFT stage while keeping the model, training recipe, and evaluation protocol unchanged. Specifically, we subsample the CoT training set with different ratios, where 1.0 denotes the full set with about 7k CoT traces.

Figure 8 shows a consistent monotonic improvement in HR@1 as more CoT data is used. The gain is most pronounced when increasing the ratio from 0.25 to 0.50, and gradually saturates afterwards, suggesting diminishing returns at higher supervision scales. Overall, these results indicate that Mobility CoT supervision is beneficial even at moderate scale, while using the full set provides the best accuracy in our setting.

A.5 Case Study

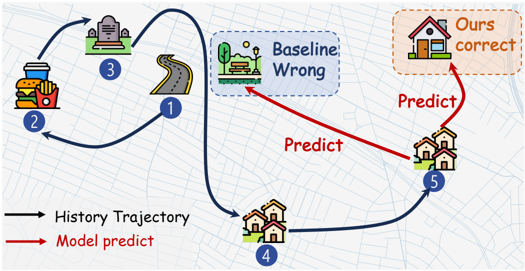

Figure 9: A representative case

Figure 9 shows a representative case where geographic feasibility matters. The user exhibits a strong local routine, with frequent visits to a small set of nearby POIs and mostly short transitions. The task is to predict the next POI one hour after the last check-in, making long-distance moves unlikely. Mobility CoT enforces this spatiotemporal prior by restricting candidates to nearby POIs and then choosing the one most consistent with recent habits, it predicts *Home* due to both proximity and the user's typical afternoon return pattern. By contrast, the pretrain-only model recommends a POI several kilometers away, which is travel-time implausible. Mobility CoT thus acts as a feasibility filter that mitigates long-jump artifacts.

A.6 Prompt Format and Examples

Input Prompt

```
<|im_start|>user
Here is a record of a user's POI accesses, your task is based on the history to predict the POI that the user is likely
to access at the specified time.
Given user historical data:
User Traj#1:
April 11th, 2012, Wednesday, 04:59, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>.
April 11th, 2012, Wednesday, 14:30, visit Office at 101 Broadway <m_161><n_115><a_12><b_7><c_0>, distance is Nearby.
April 12th, 2012, Thursday, 04:50, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>, distance is Nearby.
April 12th, 2012, Thursday, 05:44, visit Office at 101 Broadway <m_161><n_115><a_12><b_7><c_0>, distance is Nearby.
April 13th, 2012, Friday, 04:56, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>, distance is Nearby.
User Traj#2:
April 19th, 2012, Thursday, 04:58, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>.
April 19th, 2012, Thursday, 05:50, visit Office at 101 Broadway <m_161><n_115><a_12><b_7><c_0>, distance is Nearby.
April 20th, 2012, Friday, 04:55, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>, distance is Nearby.
April 20th, 2012, Friday, 06:07, visit Office at 101 Broadway <m_161><n_115><a_12><b_7><c_0>, distance is Nearby.
Given user behavior sequence:
April 23rd, 2012, Monday, 04:55, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>.
April 23rd, 2012, Monday, 05:27, visit Office at 101 Broadway <m_161><n_115><a_12><b_7><c_0>, distance is Nearby.
April 24th, 2012, Tuesday, 04:45, visit Parking at 85 Washington St <m_161><n_17><a_21><b_8><c_0>, distance is Nearby.
At April 24th, 2012, Tuesday, 04:58, user will visit <|im_end|>
<|im_start|>assistant
```

Reasoning Process and Recommendation

```
<think>
Step 1: On April 24th, 2012, Tuesday, the user visited Parking at 85 Washington St at 04:45. Historically, after visiting
 this parking location, the user consistently transitions to the Office at 101 Broadway, with visits occurring shortly
thereafter.
Step 2: Possible next-POI candidates are:
1. Office at 101 Broadway, last visited on April 23rd, 2012, at 05:27.
2. Parking at 85 Washington St, last visited on April 24th, 2012, at 04:45.
Step 3:
1. Parking at 85 Washington St: The user already visited this location on April 24th, 2012, at 04:45. Historical data
shows no pattern of revisiting Parking at 85 Washington St immediately after a recent visit.
2. Office at 101 Broadway: The user has repeatedly transitioned from Parking at 85 Washington St to the Office at 101
Broadway, especially after early morning visits, and the distance is Nearby, supporting this transition.
</think>
<m_161><n_115><a_12><b_7><c_0>
```