2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

T-LENs: A Tile-Assisted Prompt Framework for Next Location Prediction via Large Language Models

Yifei Luo[1], Yuhang Wang[1], Ningyun Li[1], Lin Zhang[1] ✉, Haichen Xu[1], Yu Liu[2], Rui Luo[3], and Lin Zhang [3]

[1] Beijing Big Data Centre, Beijing
[2] Beijing University of Posts and Telecommunications, Beijing
[3] China Telecom Corporation Beijing Branch, Beijing

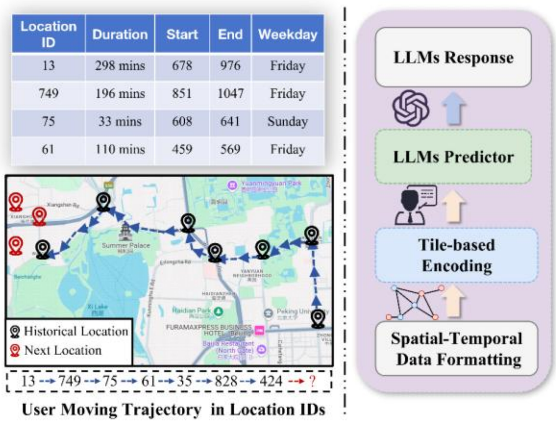
Abstract. Next location prediction is critical for personalized recommendations, transportation planning, and emergency responses. However, the sparsity of mobility data and the stochastic nature of individuals' daily activities make accurate forecasting still a significant challenge. Existing next location prediction methods often rely on discrete location IDs from limited-scale datasets, limiting interpretability and generalization across regions. To address these issues, we propose T-LENs, a prompt-based framework that combines continuous tile-assisted spatial encoding with the interpretive and reasoning capabilities of Large Language Models (LLMs). Our proposed tile-assisted encoding integrates seamlessly with existing methods and enhances privacy preservation by avoiding exposure of sensitive raw coordinates, while also mitigating noise from ultra-precise geolocation data. Furthermore, T-LENs models human mobility by jointly capturing long-term trends and short-term dependencies through a variable-length window, enabling LLMs to identify complex mobility patterns with high accuracy. Our experiments demonstrate that T-LENs significantly outperforms state-of-the-art baselines, achieving superior prediction accuracy with a 50% improvement in Acc@1 and 8% in nDCG@10, while requiring no dataset-specific training. To comprehensively assess the framework's adaptability, we further evaluate its performance across diverse LLMs, highlighting their potential and limitations in mobility modeling.

Keywords: Next location prediction, tile encoding, LLMs.

1 Introduction

Human mobility, referring to the movement of people within geographic areas such as cities, regions, or countries [1,2], has become increasingly available with the rapid development of smart city infrastructure and Location-Based Services (LBS) technologies [3]. Accurate next location prediction based on human mobility data serves as a significant foundation for various critical domains [4,5,6].

However, next location prediction remains a significant challenge due to the stochasticity and complex spatial-temporal dependencies [7]. To overcome these challenges, deep learning methods have emerged [8,9,10], using neural networks to automatically extract patterns from raw mobility data. While these methods offer improved accuracy and adaptability, their application is still constrained by several factors. One major limitation lies in their reliance on embedding tables based on discrete location IDs to represent locations, as illustrated in Fig 1.(a). This representation inherently overlooks the physical spatial relationships between locations, making it challenging to capture the true geographic and contextual connections. Additionally, these methods require extensive training on large-scale mobility datasets to achieve generalization, yet existing datasets are often limited in scale and diversity [2,11-13]. Furthermore, the results produced by these methods often lack interpretability, hindering their application in real-world scenarios.



(a) Discrete Location ID Representation (b) Our Tile-assisted Prediction

Fig. 1. Comparison of discrete location ID representation and our tile-assisted prediction framework. (a) Traditional methods represent user trajectories with discrete location IDs. (b) Our proposed T-LENs framework leverages tile-assisted spatial encoding for improved next-location prediction using LLMs.

In contrast, pre-trained on vast text, LLMs eliminate the need for dataset-specific training. Its prior knowledge enables them to reason about mobility patterns even with sparse or limited training data. Through carefully designed prompt strategies, LLMs can generate interpretable predictions with their reasoning processes. In this paper, we introduce T-LENs, a prompt-based LLMs prediction framework that integrates tile encoding instead of discrete location IDs for next location prediction. As shown in Fig. 1.(b), in spatial-temporal data formatting stage, the raw human mobility data is processed to extract relevant spatial-temporal feature. Then, T-LENs utilizes a tile-based representation to preserves spatial relationships between different locations. The encoded data are passed to the LLM-based predictor to generate accurate *Top-k* next location predictions. Finally, the LLMs provide a detailed response that includes both predictions and reasons, enhancing interpretability and reliability. Specifically, we organize mobility data into historical stays and context stays, enabling the model to account for long-term and short-term dependencies in human movements, respectively. In summary, our main contributions are as follows.

- We introduce tile-based encoding schemes that not only preserves geographic relationships between locations but also offers a simple, flexible design that can be seamlessly integrated into existing prediction models, overcoming the limitations of discrete location ID representations commonly used in existing methods.
- We propose T-LENs, a prompt-based framework that integrates both spatial and temporal features with LLMs for next location prediction. Temporally, we model user trajectories as a temporal sequence of stays. Spatially, we incorporate location IDs and tile IDs to better capture geographic relationships.
- We conduct comprehensive experiments to evaluate the performance of our proposed framework, demonstrating its effectiveness in accurate next location prediction. Additionally, we assess the suitability of different LLMs within our framework, providing valuable insights into their strengths, limitations, and adaptability for next location prediction.

2 Related Work

2.1 Human Mobility Prediction

Next location prediction relies on constructing mathematical models, which can be broadly categorized into pattern-based methods, Markov chain-based methods, and deep learning-based methods. Early researchers widely employed specific pattern-based methods for next location prediction, such as sequence patterns, periodic patterns, and probabilistic patterns [14,15]. Markov chain-based models, such as the Hidden Markov Model (HMM), treat state sequences as latent variables and observation sequences as visible variables to estimate trajectory patterns and predict next locations [16]. With the increasing prevalence of deep learning, current learning-based methods can be broadly subdivided into sequence-based [10,17] and graph-based approaches [18,19]. However, while these models perform exceptionally well on specific datasets, their generalization capability remains a significant challenge, which can be attributed to the lack of spatial information [20,21]. To address the lack of spatial information in traditional approaches, Yao *et al.* [10] combined geographic embedding, multi-layer attention, and Bi-LSTM to effectively integrate geographic information into their model. Besides, Liu *et al.* [3] encodes locations using continuous spatial coordinates instead of discrete IDs, allowing it to better capture spatial relationships between locations. However, these approaches often come with significant complexity, such as the need for additional embedding layers, attention mechanisms, or normalization steps, which may increase computational overhead.

2.2 Large Language Models

LLMs have shown great promise in high-fidelity human mobility simulation and forecasting [22-26]. Recently, researchers have also begun leveraging powerful language understanding and reasoning capabilities for next location prediction [2,3,27].

Wang *et al.* [2] are among the first to explore the use of LLMs for next location prediction. They employ a prompt-based strategy, structuring user mobility data into a format suitable for LLMs to predict. Beneduce *et al.* [27] evaluated the performance of LLMs under zero-shot settings, demonstrating their generalization and reasoning capabilities. These findings suggest that LLMs have the potential to serve as zero-shot next-location predictors. However, both approaches treat discrete location IDs as the prediction target, which fails to capture geometric relationships between locations. To address this limitation, Liu *et al.* [3] moved away from prompt-based strategies and instead modified the architecture of pre-trained LLMs to embed the coordinates of the prediction points as features. While this approach improves spatial representation, it introduces significant complexity and makes integration with prompt-based methods impractical. This highlights the need for simpler yet effective solutions to leverage LLMs for spatially aware next location prediction.

3 Problem Formulation

Next location prediction is commonly defined as forecasting the next destination an individual will visit based on their historical spatial-temporal trajectories [27]. In this paper, we model a trajectory as a combination of historical stays and contextual stays, where historical stays capture long-term mobility patterns, and contextual stays provide recent activity information.

Definition 1: Trajectory. A spatial-temporal point is represented as a tuple $p = (t, l)$, where t denotes the timestamp, and l represents the geographic location id. A trajectory $P = \{p_1, p_2, \cdots, p_n\}$ consists of n spatial-temporal points arranged in chronological order, reflecting the locations visited by an individual. Each user can have multiple trajectories $P_1, P_2, \cdots, P_k$, with all points in P_i preceding those in P_{i+1} in time.

Definition 2: Historical Stays. The trajectory P of a user is composed of historical stays $\mathcal{H}$ and context stays $\mathcal{C}$. Historical stays represent the long-term mobility patterns of the user and are defined as the spatial-temporal points that occur prior to the context stays. Formally, historical stays are expressed as:

$$\mathcal{H} = \{ h_i \mid h_i \in P_k, h_i \prec c_1 \} \tag{1}$$

where:

- $h_i \in P_k$: h_i is a spatial-temporal point in trajectory P_k.
- $c_1 = \min(\mathcal{C})$: c_1 represents the earliest point within the context stays.
- $h_i \prec c_1$: h_i occurs before c_1 in temporal order.

Definition 3: Context Stays. Context stays represent the user's short-term mobility patterns, consisting of the spatial-temporal points visited immediately prior to the target location p_{n+1}. These stays provide the recent activity context leading up to the prediction.

Formally, contextual stays are defined as:

$$\mathcal{C} = \{ c_j \mid c_j \in P_k \text{ and } c_j \prec p_{n+1} \}, \tag{2}$$

where:

- $c_j \in P_k$: c_j is a spatial-temporal point in the trajectory P_k.
- p_{n+1}: the target location being predicted.
- $c_j \prec p_{n+1}$: c_j occurs before p_{n+1} in temporal order.

Task Objective. Next location prediction involves determining the next spatial-temporal point p_{n+1} based on an individual's trajectory P_k. The provided data includes:

- The current trajectory of an individual $P_k = \{p_1, p_2, \ldots, p_n\}$, which contains at least two spatial-temporal points.
- The historical trajectories of the individual $\mathcal{L} = \{P_1, P_2, \ldots, P_{k-1}\}$, representing their past movement patterns.

Formally, the goal of next location prediction is to predict $p_{n+1} \in P_k$, by utilizing both long-term patterns from historical trajectories and recent contextual information. This task can be formalized as a mapping:

$$\mathcal{M}: (\mathcal{H}, \mathcal{C}) \rightarrow p_{n+1}, \tag{3}$$

where:

- $\mathcal{H}$: The historical stays, extracted from the earlier parts of P_k and previous trajectories $\{P_1, \ldots, P_{k-1}\}$, representing long-term mobility trends.
- $\mathcal{C}$: The context stays, composed of the most recent points in P_k before p_n, reflecting short-term movement patterns.

4 Methodology

In this section, we present the proposed methodology for next location prediction, which incorporates tile-based spatial encoding and spatial-temporal dependencies. As illustrated in Fig. 2, the T-LENs framework comprises three key components: the Tile Encoding Module, Spatial-Temporal Dependencies modeling, and Spatial-Temporal Instruction Prompt.

4.1 Tile Encoding Module

Inspired by video encoding and transmission techniques, where video frames are divided into smaller tiles for efficient processing [28,29,30], we adopt a similar approach for spatial data representation. By dividing the geographical area into tiles, we aim to segment the space into manageable units that facilitate effective encoding and analysis. As illustrated in Fig. 2, we propose two different tiling schemes, *i.e.*, Sequential Tiling and 2D Tiling, to represent spatial data in a structured manner, preserving both spatial relationships and continuity. The division is based on a grid structure, where R represents the total number of rows and C represents the total number of columns in the grid. Each tile is uniquely identified either by a sequential ID in a row-major order (as in Sequential Tiling) or by its 2D coordinates (x, y), where x is the column index and y is the row index (as in 2D Tiling). The values of R and C are determined by the geographic extent of the area being studied and the size of each tile. For example, if the height and width of the area are H and W, respectively, and each tile covers an area of $s_h \times s_w$, then the number of rows and columns can be calculated as:

$$R = \left\lceil \frac{H}{s_h} \right\rceil, C = \left\lceil \frac{W}{s_w} \right\rceil, \tag{4}$$

where H and W are the height and width of the study area, and s_h, s_w is the size of each tile.

This grid-based tiling ensures that both spatial continuity and adjacency relationships are preserved, which are critical for modeling mobility patterns effectively. By choosing an appropriate tile size, we ensure a balance between spatial granularity and computational efficiency. Furthermore, this method can be seamlessly integrated into existing frameworks, making it a versatile and practical approach for enhancing the representation of spatial data.

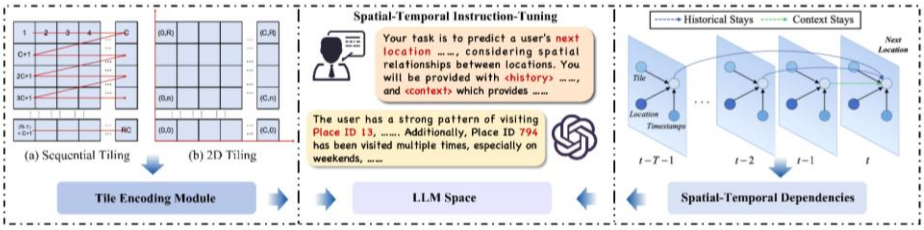


Fig. 2. Overview of the T-LENs framework: (a) Tile Encoding Module includes sequential tiling and 2D tiling to encode spatial information into continuous representations. (b) Spatial-Temporal Dependencies module models the interactions between tiles, locations, and timestamps to capture both short-term and long-term dependencies. (c) Spatial-Temporal Instruction-Tuning leverages task-specific prompts and contextual information to guide the LLM space, integrating historical and contextual mobility patterns.

4.2 Spatial-Temporal Dependencies Modeling

To effectively capture the spatiotemporal dynamics of human mobility, we model each stay as a tuple $S = (st, dow, dur, pid, tid)$, where st represents the start time of the stay, dow denotes the day of the week, dur indicates the duration of the stay, pid is the unique identifier of the place where the stay occurred and tid represents its tile id.

The inclusion of temporal features such as st, dow, and dur allows LLMs to capture users' time-dependent movement patterns, such as daily routines, weekly trends, and the duration of activities at specific locations. The spatial information, encoded through pid and tid, adds critical geographic context to the model. pid uniquely identifies specific places, enabling the model to recognize frequently visited or significant locations, while tid captures the broader spatial region or tile where the stay occurred, preserving geographic relationships between locations.

As illustrated in Fig. 3, the timeline represents the user's mobility history, divided into historical stays (M) and context stays (N). The task is to predict the next location

(pid_{n+1}) that a user will visit, based on a sequence of their previous stays, denoted as $\mathbb{S} = (S_{n-Q+1}, \cdots, S_n)$. Therefore, Equation 3 can be rewritten as a mapping:

$$\mathcal{M}: \mathbb{S} \rightarrow pid_{n+1}. \quad (5)$$

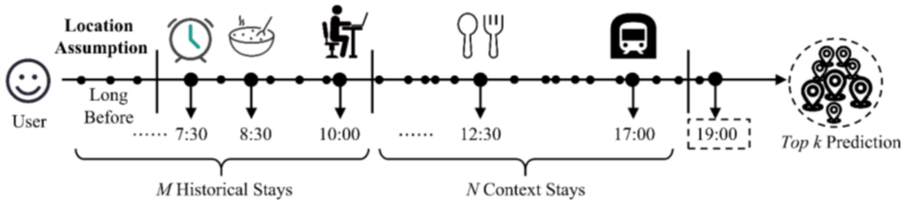


Fig. 3. Illustration of data organization into historical stays and context stays for next location prediction.

4.3 Spatial-Temporal Instruction Prompt

To effectively leverage LLMs for next location prediction, the proposed prompt design uniquely integrates structured spatial and temporal information. Drawing inspiration from established prompting strategies such as Chain-of-Thought (CoT) and Plan-and-Solve (PS), the proposed prompt incorporates rich contextual information, such as comprehensive descriptions of input data, to guide LLMs in understanding and reasoning about human mobility patterns [2]. As shown in Fig.4, the prompt is designed with clear instructions and detailed descriptions of the data format to ensure LLMs can comprehend and reason effectively.

The first distinguishing feature of this prompt design is the inclusion of tile-based spatial encoding in the data description. By introducing the concept of $tile\_id$, which represents the unique ID of a rectangular tile in the geographical area, the prompt preserves spatial relationships while maintaining simplicity. Second, the prompt explicitly emphasizes context-aware reasoning by guiding the model to analyze both historical and contextual stays. It instructs the LLM to focus on recurring patterns in historical data $< history >$, spatial transitions between nearby tiles, and recent activities in the contextual data $< context >$. This explicit mention of spatial tiles and their temporal associations ensures the model considers both fine-grained spatial patterns and broader temporal trends. Third, it explicitly instructs the model to generate the *Top-k* most probable next locations, ranked by probability, and justifies each prediction with a concise explanation. This explanation step not only enhances interpretability but also aligns with the reasoning generation strategy of Plan-and-Solve prompting. The explicit instructions to analyze patterns and justify predictions make the outputs more reliable and actionable. By integrating instructions, structured data, and contextual information, the proposed prompt makes it feasible for complex spatiotemporal tasks.

Instructxion | Specify the task Your task is to predict a user's next location based on his/her activity pattern, considering spatial relationships.

Data | Describe the data You will be provided with <history> which contains the user's historical stays and <context> which provides contextual information about recent stays. Stays in both <history> and <context> are in chronological order.

Each stay takes the form (start_time, day_of_week, duration, place_id, tile_id). The detailed explanation of each element is as follows:

start_time: the start time of the stay in 12h clock format.

day_of_week: indicating the day of the week.

duration: an integer indicating the duration (in minutes) of each stay.

place_id: an integer representing the unique place ID, indicating where the stay occurred.

tile_id: an integer representing the unique ID of a rectangular tile in the geographical area.

The area is divided into <R> rows and <C> columns, and tile IDs are assigned sequentially from top to bottom and left to right.

Instruction | Specify the number of output Please infer the 10 most probable next places (<next_place_id>), ranked in descending order of probability.

Instruction | Guide the model to "think" To predict the <next_place_id>, consider the followings:

1. Analyze patterns in <history>, including repeated visits to specific places and their associated time and spatial tiles (tile_id).

2. Consider <context>, which reflects recent activities and might indicate a spatial transition to nearby tiles or frequently visited tiles.

Instruction | Format the output and ask for explanations Provide your answer in the following JSON format:

"prediction": the IDs of the ten most probable places in descending order of probability.

"reason": a concise explanation, following the reasoning process step-by-step to identify recurring patterns, recent activities, and transitions between spatial tiles. Do not use line breaks in the reason.

Fig. 4. The prompt instruction used in our study. The red text emphasizes the corresponding prompt design about the tile.

5 Experiments

5.1 Experiment Implementation

All experiments are conducted on a server equipped with two NVIDIA RTX 4090 GPUs (24GB VRAM each) and an Intel Xeon W-2295 CPU. We implement T-LENs using Python 3.9 for LLMs inference.

The dataset used in this study is sourced from the Microsoft Research Asia Geolife project [31], which collected GPS trajectory data from 182 users over a span of more than three years (from April 2007 to August 2012). The dataset includes 17,621 trajectories, covering a total distance of approximately 1.2 million kilometers with a total duration of 50,176 hours.

In this work, we preprocess the raw GPS data to generate stay points and corresponding locations for next location prediction tasks. A stay point is defined as a location where a user remains stationary for at least 30 minutes within a 200-meter radius. These stay points are then clustered using DBSCAN to assign unique location IDs. Additionally, temporal features (*e.g.*, start time, duration, day of the week) and spatial features (*e.g.*, location ID, and tile-based representations) are extracted for modeling. This structured representation allows the model to capture both spatial and temporal dynamics, enabling accurate next location predictions.

5.2 Experiment Results

Prediction Performance.

Some baseline results are cited from previous studies [2] and their origin papers, providing a comprehensive understanding of the model performances under varying conditions.

The results presented in Table 1 demonstrate the performance comparison of various models on the Geolife dataset using multiple metrics. Our proposed T-LENs framework significantly outperforms existing baselines in terms of accuracy and ranking-based metrics.

Accuracy Improvements. The most notable improvement is observed in the Acc@1 metric, where sequential tiling (*wot*: 49.6%, *wt*: 55.6%) and 2D tiling (*wot*: 49.6%, *wt*: 53.7%) both exhibit significant gains compared to the best learning-based baseline (MHSA, 31.4%). This improvement can be attributed to the tile-based encoding, which captures spatial relationships more effectively than traditional location ID-based approaches. Specifically, by encoding locations based on their spatial relationships, the model can better understand and predict frequently visited or adjacent locations, especially for top-ranked predictions.

Weighted F1 and nDCG@10 Enhancements. The T-LENs framework also excels in ranking-based metrics like weighted F1 and nDCG@10, which consider not only the correctness of predictions but also their order of relevance. Sequential tiling achieves a weighted F1 score of 0.473 and nDCG@10 of 0.659, while 2D tiling achieves 0.458 and 0.656, respectively. These improvements reflect that the inclusion of tiling schemes helps in preserving the geographic proximity and contextual patterns, enabling better ranking of plausible next locations.

Impact of Target Time Information. Another factor contributing to the performance gains is the incorporation of target time (*wt*) information. For both tiling methods, adding time-aware features leads to consistent improvements across all metrics. This indicates that incorporating temporal patterns, such as time of day or day of the week, helps the model better align its predictions with user mobility behaviors.

Overall Insight. While both tiling methods show significant improvements over baselines, sequential tiling generally performs slightly better in terms of Weighted F1 and nDCG@10. This might be due to its simpler representation, which could better align with the spatial encoding capabilities of the LLMs. However, 2D tiling shows comparable performance, offering an alternative that preserves inherent two-dimensional spatial relationships, which can be advantageous for datasets with more irregular spatial distributions. Overall, the T-LENs framework demonstrates the improvements in accuracy and ranking-based metrics. The results also highlight the adaptability of the framework with different tiling schemes, making it a robust solution for next-location prediction tasks.

Table 1. Performance comparison of various models on Geolife datasets using multiple metrics (All metrics are better with higher value). Sequential tiling in the red column and 2D tiling in the blue column. *wt* and *wot* represent with and without target time information, respectively.

Dataset	**Metric**	**LSTM**[32]	**LSTM-SA**[33]	**Deep-Move**[34]	**Mob-Tcast**[35]	**MHSA** [36]	**LLM-Mob**[2]	**LLM-ZS**[27]	**T-LENS** (*k=10*)			
									wot	**wt**	**wot**	**wt**
Geolife	Acc@1(%)	28.4	29.8	26.1	29.5	31.4	35.8	44.8	49.6	**55.6**	49.6	53.7
	Acc@5(%)	55.8	54.6	54.2	51.3	56.4	68.7	65.5	71.5	**73.7**	73.0	72.6
	Acc@10(%)	59.1	58.2	58.7	56.2	60.8	72.1	70.7	58.7	75.2	**75.5**	**75.5**
	Weighted F1	0.193	0.213	0.189	0.173	0.218	0.344	0.403	0.428	**0.473**	0.388	0.458
	nDCG@10	0.447	0.450	0.426	0.434	0.465	0.607	0.585	0.636	**0.659**	0.613	0.656

Case Study

A concise case analysis is presented in Table 2, where the ground truth next location is *place_id* 13. Notably, LLM-Mob places location 13 in the third position of its prediction list, whereas both T-LENs (Sequential) and T-LENs (2D) identify location 13 as the top candidate. This difference in ranking is particularly significant for metrics such as Acc@1 and nDCG@10, both of which reward higher placement of correct predictions.

In LLM-Mob prediction, it emphasizes the user's weekend activities at place IDs 54 and 55, noting that place ID 13 is also frequently visited. However, the model ranks 13 only third, implying that while LLM-Mob does recognize 13's relevance, it deems 54 and 55 slightly more probable based on recent or temporal patterns alone. In our sequential tiling prediction, it attributes the higher rank of 13. By leveraging a sequential tiling scheme, T-LENs (Sequential) more precisely captures the user's preference for 13 and its surrounding locations (*e.g.*, tile_id 794), leading it to promote 13 to the top of the list. T-LENs (2D tiling) highlights spatial proximity, referencing tiles near

(16,13) and (13,18) that surround place ID 13. The 2D tiling preserves geographic coordinates, enabling the model to recognize 13's strong association with these closely situated points, further reinforcing 13 as the most likely next location. Consequently, T-LENs demonstrates both stronger predictive accuracy and better alignment with the user's actual movement patterns.

Impact Analysis of Parameter Settings and Model Variants.

To analyze the impact of different parameter settings for M (number of historical stays) and N (number of context stays) on the model's performance, we observe the trends in Table 3 across various evaluation metrics. The results highlight an interesting trade-off between these two parameters.

Table 2. Performance Evaluation Under Different M and N Setups

	Varying M with N=5				**Varying N with M=40**			
	40	**30**	**20**	**20**	**5**	**10**	**20**	**30**
Acc@1(%)	55.6	53.3	54.9	45.5	55.6	54.9	54.8	54.4
Acc@5(%)	73.7	72.5	73.4	70.4	73.7	73.3	71.8	72.6
Acc@10(%)	75.2	75.5	77.5	75.1	75.2	77.0	77.4	78.8
F1	0.473	0.467	0.488	0.410	0.473	0.481	0.473	0.469
nDCG@10	0.659	0.651	0.666	0.611	0.659	0.668	0.664	0.670

When the historical context M is large, such as at $M = 40$, the model achieves the highest Acc@1 of 55.6%, indicating that incorporating a larger amount of historical data is beneficial for predicting the most likely next location. However, as M decreases, the Acc@1 metric experiences a slight decline, which may be due to the loss of critical long-term mobility patterns that are essential for making accurate predictions based on past behavior. On the other hand, as the context size N increases, we observe improvements in metrics like Acc@10 and Weighted F1, with Acc@10 reaching its peak at 78.8% and Weighted F1 improving to 0.481. This suggests that a broader context of recent stays enhances the model's ability to rank the next most probable locations effectively. While this improvement comes at the expense of a slight reduction in Acc@1, it enhances the model's ability to consider a broader set of potential future locations, resulting in a higher cumulative prediction performance.

The nDCG@10 metric remains relatively stable across different values of M and N, showing that the quality of the rankings is consistently high, regardless of the length of historical or contextual stays. This stability suggests that the model is effective in prioritizing the most relevant locations for the next visit, even with varying historical and contextual input.

Besides, the results from different LLMs, including GPT-4o [37], Moonshot [38], Qianfan [39], and Doubao [40], are shown in Fig. 5. These models demonstrate varying levels of performance across key metrics, reflecting differences in their ability to capture spatial-temporal dependencies and make accurate next-location predictions. GPT-4o achieves the highest scores across most metrics. Its high nDCG@10 indicates that it ranks relevant predictions closer to the top more consistently than other models. Moonshot excels in Acc@5 (0.728) and maintains competitive scores in Acc@10 (0.741).

Table 3. Case study using user 1 test example. Sequential means using sequential tiling scheme, while 2D represents the 2D tiling scheme.

Case Study: Test Sample of User 1		
Ground Truth	13	
Target Time	(15:13, Friday)	
LLM-Mob	**Prediction**	[54, 55, 13, 60, 9, 48, 49, 50, 61, 35]
	Reason	Place IDs 54 and 55 were visited on Friday, indicating a preference for these locations. Additionally, the user has a strong pattern of visiting place ID 13, especially on Fridays around the same time (11:39 AM and 09:24 AM). The context shows recent visits to place ID 13, reinforcing its likelihood. The time of the target stay (1:19 PM) suggests a continuation of Saturday activities, making these places the most probable choices.
LLM-ZS	**Prediction**	[54, 55, 53, 13, 9, 48, 49, 50, 59, 60]
	Reason	The prediction is based on the user's historical stays, which show a strong preference for certain places on Fridays, particularly around midday. The context stays also indicate recent activity patterns that align with these preferences. The temporal details of the target stay (Friday afternoon) suggest that the user is likely to visit places they have frequented before on this day.
T-LENs (Seq)	**Prediction**	[13, 794, 492, 828, 61, 35, 9, 47, 48, 49]
	Reason	The user has a strong pattern of visiting place_id 13, particularly on weekends and Mondays, indicating it is a frequent destination. The recent context shows continued activity at place_id 13, reinforcing its likelihood as the next location. Additionally, place_id 794 has been visited multiple times, especially on weekends, suggesting it is a significant location for the user. Nearby places such as 492 and 828 have also been visited recently, indicating potential transitions. The user has shown a tendency to return to familiar places, especially those associated with tile_id 794, which is a central tile in their history. The prediction considers both the historical frequency and recent context, leading to a strong likelihood of returning to these places.
T-LENs (2D)	**Prediction**	[13, 23, 12, 60, 61, 35, 54, 55, 48, 49]
	Reason	The user has a strong pattern of visiting place_id 13, particularly on Tuesdays and Wednesdays, with multiple

long stays. The recent context also shows continued activity at place_id 13, indicating a high likelihood of returning there. Additionally, the user frequently visits nearby places such as 12 and 23, which are spatially close to 13. The user has also shown activity in nearby tiles, particularly around (16, 13) and (13, 18), which suggests that places in close proximity to these coordinates may also be relevant. Other places like 60 and 61 have been visited recently, but less frequently, making them less likely compared to the top candidates.

However, its Acc@1 (0.490) and Weighted F1 (0.390) are lower than GPT-4o. Moonshot and Doubao are better at broader predictions (*e.g.*, Acc@5, Acc@10), while GPT-4o excels in precise top-1 predictions. Qianfan exhibits the weakest performance across all metrics. The model seems to struggle with both capturing historical dependencies and effectively utilizing contextual information, leading to poor ranking and precision.

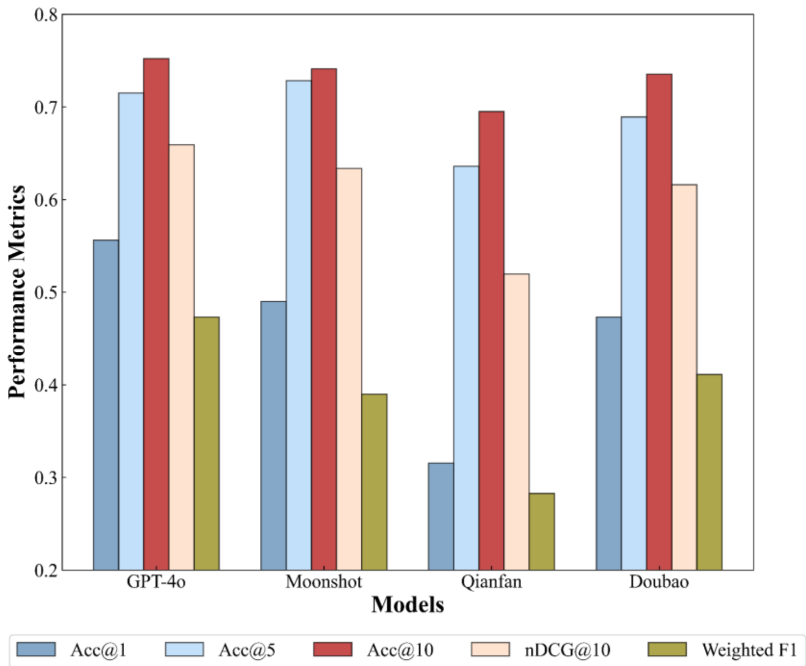


Fig. 5. Comparison of performance metrics across different LLMs for next location prediction.

6 Conclusion

In this paper, we introduced T-LENs, a framework that integrates tile-based spatial encoding with LLMs to advance next-location prediction. By leveraging either sequential or 2D tiling schemes, T-LENs efficiently captured the nuanced spatial relationships among locations and aligns them with temporal dynamics. Experimental results on Geolife dataset demonstrated that T-LENs outperforms established baselines, particularly in terms of ranking-based metrics such as nDCG@10 and Acc@1, thereby underlining the effectiveness of tile-based spatial encoding in enhancing predictive accuracy. Moreover, T-LENs provided interpretable insights into its predictions, as illustrated by the case study where it successfully identified the correct next location and offered clear justifications. These findings highlighted the potential of combining LLMs with tile-based spatial representations to address the complexities of human mobility data.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Pappalardo, L., Manley, E., Sekara, V., Alessandretti, L.: Future directions in human mobility science. Nature computational science 3(7), 588–600 (2023)
2. Wang, X., Fang, M., Zeng, Z., Cheng, T.: Where would i go next? large language models as human mobility predictors. arXiv preprint arXiv:2308.15197 (2023)
3. Liu, S., Cao, N., Chen, Y., Jiang, Y., Cong, G.: nextlocllm: next location prediction using llms. arXiv preprint arXiv:2410.09129 (2024)
4. Sánchez, P., Bellogín, A.: Point-of-interest recommender systems based on location-based social networks: a survey from an experimental perspective. ACM Computing Surveys (CSUR) 54(11s), 1–37 (2022)
5. Kraemer, M.U., Yang, C.H., Gutierrez, B., Wu, C.H., Klein, B., Pigott, D.M., Group†, O.C..D.W., Du Plessis, L., Faria, N.R., Li, R., et al.: The effect of human mobility and control measures on the covid-19 epidemic in china. Science 368(6490), 493–497 (2020)
6. Chen, C., Liu, Y., Chen, L., Zhang, C.: Bidirectional spatial-temporal adaptive transformer for urban traffic flow forecasting. IEEE Transactions on Neural Networks and Learning Systems 34(10), 6913–6925 (2022)
7. Katranji, M., Sanmarty, G., Moalic, L., Kraiem, S., Caminada, A., Selem, F.H.: Rnn encoder-decoder for the inference of regular human mobility patterns. In: 2018 International Joint Conference on Neural Networks (IJCNN). pp. 1–9. IEEE (2018)
8. Jiang, R., Cai, Z., Wang, Z., Yang, C., Fan, Z., Chen, Q., Song, X., Shibasaki, R.: Predicting citywide crowd dynamics at big events: A deep learning system. ACM Transactions on Intelligent Systems and Technology (TIST) 13(2), 1–24 (2022)
9. Jiang, R., Wang, Z., Tao, Y., Yang, C., Song, X., Shibasaki, R., Chen, S.C., Shyu, M.L.: Learning social meta-knowledge for nowcasting human mobility in disaster. In: Proceedings of the ACM Web Conference 2023. pp. 2655–2665 (2023)
10. Yao, Y., Guo, Z., Dou, C., Jia, M., Hong, Y., Guan, Q., Luo, P.: Predicting mobile users' next location using the semantically enriched geo-embedding model and the multilayer attention mechanism. Computers, Environment and Urban Systems 104, 102009 (2023)

11. Li, M., Li, J., Wang, Y., Liu, Y., Xu, H.: Dustnet: An unsupervised and noise-resistant network for martian dust storm change detection. IEEE Geoscience and Remote Sensing Letters (2025)
12. Li, J., Lv, W., Wang, J., Wang, Y., Liu, Y.: Unsupervised denoising with implicit noise mapping for single martian multispectral image. In: IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium. pp. 6120–6123. IEEE (2024)
13. Wang, J., Tian, H., Li, J., Hu, W., Li, X., Yue, W.: Towards clearer mars images: Self-supervised denoising with large vision model. In: IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium. pp. 7053–7056. IEEE (2024)
14. Monreale, A., Pinelli, F., Trasarti, R., Giannotti, F.: Wherenext: a location predictor on trajectory pattern mining. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 637–646 (2009)
15. Zheng, Y.: Trajectory data mining: an overview. ACM Transactions on Intelligent Systems and Technology (TIST) 6(3), 1–41 (2015)
16. Mathew, W., Raposo, R., Martins, B.: Predicting future locations with hidden markov models. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. pp. 911–918 (2012)
17. Yang, S., Liu, J., Zhao, K.: 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. pp. 1144–1153 (2022)
18. Rao, X., Chen, L., Liu, Y., Shang, S., Yao, B., Han, P.: Graph-flashback network for next location recommendation. In: Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining. pp. 1463–1471 (2022)
19. Xu, X., Jiang, R., Yang, C., Sezaki, K., et al.: Taming the long tail in human mobility prediction. Advances in Neural Information Processing Systems 37, 54748– 54771 (2024)
20. Comito, C.: Next: A framework for next-place prediction on location based social networks. Knowledge-Based Systems 204, 106205 (2020)
21. Jiang, Z.: A survey on spatial prediction methods. IEEE Transactions on Knowledge and Data Engineering 31(9), 1645–1664 (2018)
22. Tang, Y., Wang, Z., Qu, A., Yan, Y., Hou, K., Zhuang, D., Guo, X., Zhao, J., Zhao, Z., Ma, W.: Synergizing spatial optimization with large language models for open-domain urban itinerary planning. arXiv preprint arXiv:2402.07204 (2024)
23. Haydari, A., Chen, D., Lai, Z., Zhang, M., Chuah, C.N.: Mobilitygpt: Enhanced human mobility modeling with a gpt model. arXiv preprint arXiv:2402.03264 (2024)
24. Li, Z., Xia, L., Tang, J., Xu, Y., Shi, L., Xia, L., Yin, D., Huang, C.: Urbangpt: Spatio-temporal large language models. In: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. pp. 5351–5362 (2024)
25. Liang, Y., Liu, Y., Wang, X., Zhao, Z.: Exploring large language models for human mobility prediction under public events. Computers, Environment and Urban Systems 112, 102153 (2024)
26. Wang, J., Jiang, R., Yang, C., Wu, Z., Shibasaki, R., Koshizuka, N., Xiao, C., et al.: Large language models as urban residents: An llm agent framework for personal mobility generation. Advances in Neural Information Processing Systems 37, 124547–124574 (2024)
27. Beneduce, C., Lepri, B., Luca, M.: Large language models are zero-shot next location predictors. arXiv preprint arXiv:2405.20962 (2024)
28. Wang, Y., Li, J., Li, Z., Shang, S., Liu, Y.: Synergistic temporal-spatial user-aware viewport prediction for optimal adaptive 360-degree video streaming. IEEE Transactions on Broadcasting (2024)

29. Li, Z., Wang, Y., Liu, Y., Li, J., Zhu, P.: Just360: Optimizing 360-degree video streaming systems with joint utility. IEEE Transactions on Broadcasting (2024)
30. Li, J., Wang, Y., Liu, Y.: Meta360: Exploring user-specific and robust viewport prediction in360-degree videos through bi-directional lstm and meta-adaptation. In: 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 652–661. IEEE (2023)
31. Zheng, Y., Xie, X., Ma, W.Y., et al.: Geolife: A collaborative social networking service among user, location and trajectory. IEEE Data Eng. Bull. 33(2), 32–39 (2010)
32. Solomon, A., Livne, A., Katz, G., Shapira, B., Rokach, L.: Analyzing movement predictability using human attributes and behavioral patterns. Computers, Environment and Urban Systems 87, 101596 (2021)
33. Li, F., Gui, Z., Zhang, Z., Peng, D., Tian, S., Yuan, K., Sun, Y., Wu, H., Gong, J., Lei, Y.: A hierarchical temporal attention-based lstm encoder-decoder model for individual mobility prediction. Neurocomputing 403, 153–166 (2020)
34. Feng, J., Li, Y., Zhang, C., Sun, F., Meng, F., Guo, A., Jin, D.: Deepmove: Predicting human mobility with attentional recurrent networks. In: Proceedings of the 2018 World Wide Web Conference. pp. 1459–1468 (2018)
35. Xue, H., Salim, F., Ren, Y., Oliver, N.: Mobtcast: Leveraging auxiliary trajectory forecasting for human mobility prediction. Advances in Neural Information Processing Systems 34, 30380–30391 (2021)
36. Hong, Y., Zhang, Y., Schindler, K., Raubal, M.: Context-aware multi-head self-attentional neural network model for next location prediction. Transportation Research Part C: Emerging Technologies 156, 104315 (2023)
37. Openai: Chatgpt-4o (2025), https://chatgpt.com/, Accessed: 2025-01-15
38. MoonshotAI: Moonshot kimi (2025), https://www.moonshot.cn/, Accessed: 2025- 01-15
39. Baidu: Qianfan (2025), https://ai.baidu.com/ai-doc/WENXINWORKSHOP /hlvceiqyo, Accessed: 2025-01-15
40. ByteDance: Doubao (2025), https://team.doubao.com/en/, Accessed: 2025-01-15