

A Unified Model for Human Mobility Generation in Natural Disasters

Qingyue Long, Huandong Wang, *Member, IEEE*, Qi Ryan Wang, Yong Li, *Senior Member, IEEE*

Abstract—Human mobility generation in disaster scenarios plays a vital role in resource allocation, emergency response, and rescue coordination. During disasters such as wildfires and hurricanes, human mobility patterns often deviate from their normal states, which makes the task more challenging. However, existing works usually rely on limited data from a single city or specific disaster, significantly restricting the model’s generalization capability in new scenarios. In fact, disasters are highly sudden and unpredictable, and any city may encounter new types of disasters without prior experience. Therefore, we aim to develop a one-for-all model for mobility generation that can generalize to new disaster scenarios. However, building a universal framework faces two key challenges: 1) the diversity of disaster types and 2) the heterogeneity among different cities. In this work, we propose a unified model for human mobility generation in natural disasters (named UniDisMob). To enable cross-disaster generalization, we design physics-informed prompt and physics-guided alignment that leverage the underlying common patterns in mobility changes after different disasters to guide the generation process. To achieve cross-city generalization, we introduce a meta-learning framework that extracts universal patterns across multiple cities through shared parameters and captures city-specific features via private parameters. Extensive experiments across multiple cities and disaster scenarios demonstrate that our method significantly outperforms state-of-the-art baselines, achieving an average performance improvement exceeding 13%.

Index Terms—Mobility trajectory, generative models, diffusion models.

I. INTRODUCTION

In recent decades, the frequency and intensity of natural disasters have increased notably, encompassing a wide range of extreme events such as hurricanes, wildfires, and floods. Facing these sudden or potential disasters, human mobility generation becomes critical for evacuation route planning [1], rescue resource allocation [2], and long-term social reconstruction [3]. However, disasters may induce mobility perturbations, breaking people’s routine travel patterns, as reflected in shifts in travel distances, changes in regular routes, and even evacuations to temporary shelters [4]. These complex and dynamic mobility patterns have also posed significant challenges to accurately generating human mobility under disaster scenarios.

By leveraging advanced generative AI techniques, numerous studies have focused on human mobility generation [5], [6],

[7]. However, the majority of these studies primarily focus on capturing the periodic patterns of human mobility under normal scenarios, which limits their applicability in disaster scenarios [8], [9], [10]. Therefore, a few studies have addressed mobility modeling in disaster scenarios. For example, DeepMob leverages deep neural networks to learn human mobility patterns during earthquakes from multi-source data of millions of users, enabling accurate trajectory prediction [11]. MemeSTN combines memory networks and meta-learning to integrate social media data with human mobility data, enabling human mobility nowcasting in disaster scenarios [12]. Although these studies have made significant progress in modeling human mobility under disaster scenarios, their effectiveness is often confined to specific disaster types within a single city. Moreover, their reliance on large-scale training data from diverse cities and disaster types further limits the flexibility and generalizability of such models.

In fact, disasters are highly sudden and unpredictable, and any city may encounter new types of disasters without prior experience. If models are trained exclusively on the limited trajectory data collected under specific disaster scenarios in individual cities, they often struggle to generalize to new cities and novel disaster types [13], [14]. As shown in Figure 1, we aim to develop a unified model for human mobility generation in natural disasters that captures shared spatiotemporal patterns across multiple cities and disasters. However, a unified model for human mobility generation in natural disasters is still an open problem with the following challenges:

- **Diversity of disaster types.** Different types of disasters exhibit significant differences in disturbance intensity, duration, and spatial extent. For example, hurricanes may trigger prolonged large-scale evacuations, whereas earthquakes often result in short-term but high-intensity changes in mobility patterns. Therefore, it’s challenging to model the impact of different disaster types on human mobility within a unified model.
- **Heterogeneity across different cities.** Differences across various cities in population density, spatial layout, and other factors lead to complex and diverse human mobility patterns. For example, when the target city and the source city differ significantly in geographical location and spatial structure, the resulting location embedding features are often substantially different. This heterogeneity makes it difficult to achieve cross-city transfer with a unified model.

To address the above challenges, we propose a unified model for human mobility generation in natural disasters named **Uni-**

Manuscript received July 20, 2025; revised xxxx, 2025.

The authors are with the Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing 100084, China and the Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115. (e-mail: longqy21@mails.tsinghua.edu.cn; wanghuandong@tsinghua.edu.cn; q.wang@northeastern.edu; liyong07@tsinghua.edu.cn)

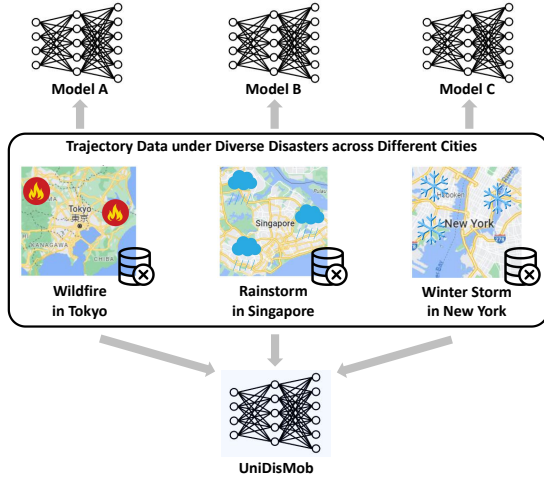


Fig. 1. The transition from single model to universal model.

DisMob. Firstly, we design a physics-informed neural network to achieve effective generalization across multiple disaster types. Specifically, we introduce physics-informed prompts that leverage spatiotemporal decay mechanisms derived from physical models [4] to capture the latent spatiotemporal patterns of mobility changes following different disasters, which serve as prior knowledge to guide trajectory generation. Moreover, we construct a physics-guided alignment mechanism. The alignment mechanism is implemented via a physics-guided loss, which enforces consistency between the generated mobility patterns and the expected spatiotemporal decay. This loss function incorporates constraints based on known spatiotemporal decay laws of mobility under disaster scenarios into the training objective, which enhances the interpretability and robustness of the unified model. Secondly, we adopt a meta-learning framework to address the heterogeneity across different cities. This framework extracts common knowledge from multiple cities while leveraging adaptation mechanisms to fine-tune the model for specific cities. In this way, it balances generalization and flexibility. Specifically, the framework consists of shared and private parameters. The shared parameters are responsible for learning spatiotemporal patterns that are consistent across cities and extracting knowledge applicable to various disaster scenarios. In contrast, the private parameters capture city-specific mobility characteristics, enabling the model to adapt to diverse urban structures. Therefore, UniDisMob achieves strong generalization across cities and disaster types by learning the common patterns of mobility changes under different disaster scenarios. Overall, our contributions can be summarized as follows:

- We explore the potential of a one-for-all model for mobility generation in natural disasters, achieving generalization across different cities and disaster types.
- We propose a unified model for human mobility generation in natural disasters. Specifically, we design a physics-informed neural network for cross-disaster generalization and a meta-learning framework to extract shared knowledge across cities for cross-city generalization.
- We conduct extensive experiments on seven datasets covering different cities and disaster types. The results demonstrate

that our method outperforms state-of-the-art models in mobility generation, achieving an average performance improvement exceeding 13%. Further analysis confirms UniDisMob's strong zero-shot capability, with an average performance improvement of over 8% on new cities.

II. RELATED WORK

A. Spatial-Temporal Foundation Models

In recent years, inspired by foundation models, particularly large language models, researchers have begun to explore spatio-temporal foundation models (STFMs) to enhance adaptability and generalization across a wide range of spatio-temporal tasks [15], [16], [17]. Existing approaches can be broadly categorized into two main types: The first category is LLM-based methods, which transform spatio-temporal data into text-like sequences and leverage the universal representation capabilities of LLMs to achieve cross-domain transfer and few-shot learning [18], [19], [20]. For example, Mai et al. [21] investigate how large pretrained language models can be utilized for representation learning on geo-spatial multimodal data, thereby improving remote sensing analysis and geographic information reasoning. Similarly, Time-LLM [6] reprograms raw time series into structured textual prototype representations, thereby directly leveraging the powerful representation and reasoning capabilities of LLMs for prediction and analysis. UrbanGPT [22] transforms complex urban spatio-temporal data into structured textual inputs and leverages instruction tuning to build a large language model for spatio-temporal tasks. The second category is a pretrained foundation model, which is pretrained directly on large-scale, multi-domain spatio-temporal data to build model architectures with efficient spatio-temporal representation capabilities [23], [24], [25]. For example, UrbanDiT [26] is designed for urban spatio-temporal tasks, providing a unified model that integrates diverse spatio-temporal data sources and types while learning general spatio-temporal patterns across different cities and scenarios. UniST [25] is a universal spatio-temporal foundation model that leverages large-scale pre-training and prompt adaptation to extract general spatio-temporal patterns across different domains. Although foundation models have made some progress in the spatiotemporal domain, there is still a significant gap in universal mobility modeling under disaster scenarios. In this work, we propose UniDisMob, the first attempt to build a universal model for mobility generation in disaster scenarios.

B. Human Mobility Generation

Human mobility generation can be divided into trajectory generation under normal scenarios and trajectory generation under disaster scenarios. Trajectory generation under normal scenarios focuses on people's travel patterns in daily life, capturing the periodicity and regularity of their mobility behaviors. Rule-based methods, such as EPR [27] and TimeGeo [28], model the spatio-temporal regularities of human mobility behaviors to generate trajectories. Deep generative models learn the distribution of trajectory data to generate human mobility trajectories. For example, PateGail [10] leverages an

adversarial learning framework to learn individual behavior policies from trajectories, generating high-fidelity mobility patterns. DiffTraj [7] employs a diffusion-based denoising process to model the complex spatiotemporal distribution of trajectories, enhancing the diversity of the generated outcomes. Different from trajectory generation under normal scenarios, trajectory generation under disaster scenarios focuses more on modeling human mobility behaviors that deviate from normal patterns during disasters. For example, Liang et al. [4] propose a spatial-temporal decay mechanism, which models how mobility perturbations after a disaster gradually diminish over time and with increasing spatial distance, offering a new perspective on mobility generation. Song et al. [29] develop an emergency mobility simulator based on large-scale disaster data to generate high-fidelity trajectories under disaster scenarios. Subsequently, Song et al. [11] design a framework integrating behavioral models and spatio-temporal dependencies to model human mobility patterns following natural disasters. However, these methods are typically trained for a specific disaster type or city and lack generalization across diverse scenarios. In this work, we develop a universal model with cross-city and cross-disaster generalization capability for trajectory generation under disaster scenarios.

III. PRELIMINARIES

A. Problem Definition

Definition 1: (Mobility Trajectory). The mobility trajectory of a user u is defined as an ordered sequence of locations recorded at uniform time intervals, denoted as $x_u = l_1, l_2, \dots, l_n$. Each location l_i corresponds to either a geographic coordinate (latitude and longitude) or a designated region identifier.

Definition 2: (Disaster Intensity). $N = \{N_j(t) \mid j \in \mathcal{L}, t \in \mathcal{T}\}$ denotes the disaster intensity, where $N_j(t)$ represents the disaster intensity at location j and time t . $\mathcal{L}$ denotes the set of all locations, and $\mathcal{T}$ denotes the set of all time slots.

Problem Definition: (Human Mobility Generation Problem in Disaster Scenario). Given disaster intensity N and a set of real-world mobility trajectories $X = \{x_u\}$, the goal is to learn a generative model that can generate mobility trajectories while preserving the fidelity of real-world data.

B. Conditional Denoising Diffusion Probabilistic Models

Diffusion models are a class of latent variable models characterized by the expression $p_\theta(x_0) := \int p_\theta(x_{0:T}), dx_{1:T}$, where the latent variables $x_1, \dots, x_T$ share the same dimensionality as the observed data $x_0 \sim q(x_0)$. These models employ two Markov chains: a forward process that gradually adds noise to the data, and a reverse process that aims to recover the original input. The following Markov chain defines the forward (diffusion) process:

$$q(\mathbf{x}_{1:T} \mid \mathbf{x}_0) := \prod_{t=1}^T q(\mathbf{x}_t \mid \mathbf{x}_{t-1}), \quad (1)$$

where $q(\mathbf{x}_t \mid \mathbf{x}_{t-1}) := \mathcal{N}(\sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I})$ and β_t are modest positive constants that reflect the noise intensity. x_t may

be represented in closed form as $x_t = \sqrt{\alpha_t} x_0 + (1 - \alpha_t) \epsilon$, for $\epsilon \sim \mathcal{N}(0, \mathbf{I})$, where $\alpha_t = \sum_{i=1}^t (1 - \beta_i)$.

We turn to the conditional extension [30], where the target variable x is generated given a set of observed conditions c . The objective of the conditional denoising diffusion probabilistic model is to approximate the conditional distribution $q(x_0 \mid c)$ using a parameterized model $p_\theta(x_0 \mid c)$.

In contrast, the conditional reverse process is defined by the following Markov chain, which iteratively denoises x_t in order to recover x_0 .

$$p_\theta(\mathbf{x}_{0:T} \mid c) := p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1} \mid \mathbf{x}_t, c), \quad (2)$$

where $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$. The conditional transition distribution $p_\theta(x_{t-1} \mid x_t, c)$ is modeled as a Gaussian with learnable parameters, given by:

$$p_\theta(\mathbf{x}_{t-1} \mid \mathbf{x}_t, c) := \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t \mid c), \sigma_\theta(\mathbf{x}_t, t \mid c) \mathbf{I}). \quad (3)$$

As shown by Ho et al. [31], the unconditional denoising diffusion model adopts a specific parameterization for the transition distribution $p_\theta(x_{t-1} \mid x_t)$, formulated as follows:

$$\begin{cases} \boldsymbol{\mu}_\theta(\mathbf{x}_t, t) = \frac{1}{\alpha_t} \left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(\mathbf{x}_t, t) \right), \\ \sigma_\theta(\mathbf{x}_t, t) = \tilde{\beta}_t^{1/2} \quad \text{where} \quad \tilde{\beta}_t = \begin{cases} \frac{1 - \alpha_{t-1}}{1 - \alpha_t} \beta_t & t > 1 \\ \beta_1 & t = 1 \end{cases} \end{cases} \quad (4)$$

where ϵ_θ is a trainable denoising function. In Eq. 4, we represent $\boldsymbol{\mu}_\theta(\mathbf{x}_t, t)$ and $\sigma_\theta(\mathbf{x}_t, t)$ as $\boldsymbol{\mu}^{\text{DDPM}}(\mathbf{x}_t, t, \epsilon_\theta(\mathbf{x}_t, t))$ and $\sigma^{\text{DDPM}}(\mathbf{x}_t, t)$.

Building on the unconditional diffusion model, a conditional denoising function ϵ_θ is introduced, which incorporates the condition c as an additional input. We now analyze the resulting parameterization involving ϵ_θ :

$$\begin{cases} \boldsymbol{\mu}_\theta(\mathbf{x}_t, t \mid c) = \boldsymbol{\mu}^{\text{DDPM}}(\mathbf{x}_t, t, \epsilon_\theta(\mathbf{x}_t, t \mid c)), \\ \sigma_\theta(\mathbf{x}_t, t \mid c) = \sigma^{\text{DDPM}}(\mathbf{x}_t, t). \end{cases} \quad (5)$$

To optimize the denoising function ϵ_θ , the model minimizes the following loss function:

$$\min_{\theta} \mathcal{L}(\theta) := \min_{\theta} \mathbb{E}_{\mathbf{x}_0 \sim q(\mathbf{x}_0), \epsilon \sim \mathcal{N}(0, \mathbf{I}), t} \|\epsilon - \epsilon_\theta(\mathbf{x}_t, t \mid c)\|_2^2, \quad (6)$$

where $\mathbf{x}_t = \sqrt{\alpha_t} \mathbf{x}_0 + (1 - \alpha_t) \epsilon$. The denoising function ϵ_θ aims to predict the noise vector ϵ added to the perturbed input $\mathbf{x}_t$. This formulation reduces the influence of tokens at very small t (i.e., low noise levels), and can be interpreted as a variational constraint weighted by a negative log-likelihood.

IV. METHOD

As shown in Figure 2, the overview architecture of UniDis-Mob consists of five modules: trajectory encoder, noise scheduler, physics-informed prompt, noise predictor, and physics-guided alignment. Among them, the trajectory encoder, noise scheduler, and noise predictor together serve as common components for diffusion-based trajectory generation. To achieve cross-disaster generalization, we design a physics-informed neural network that incorporates physical priors of human

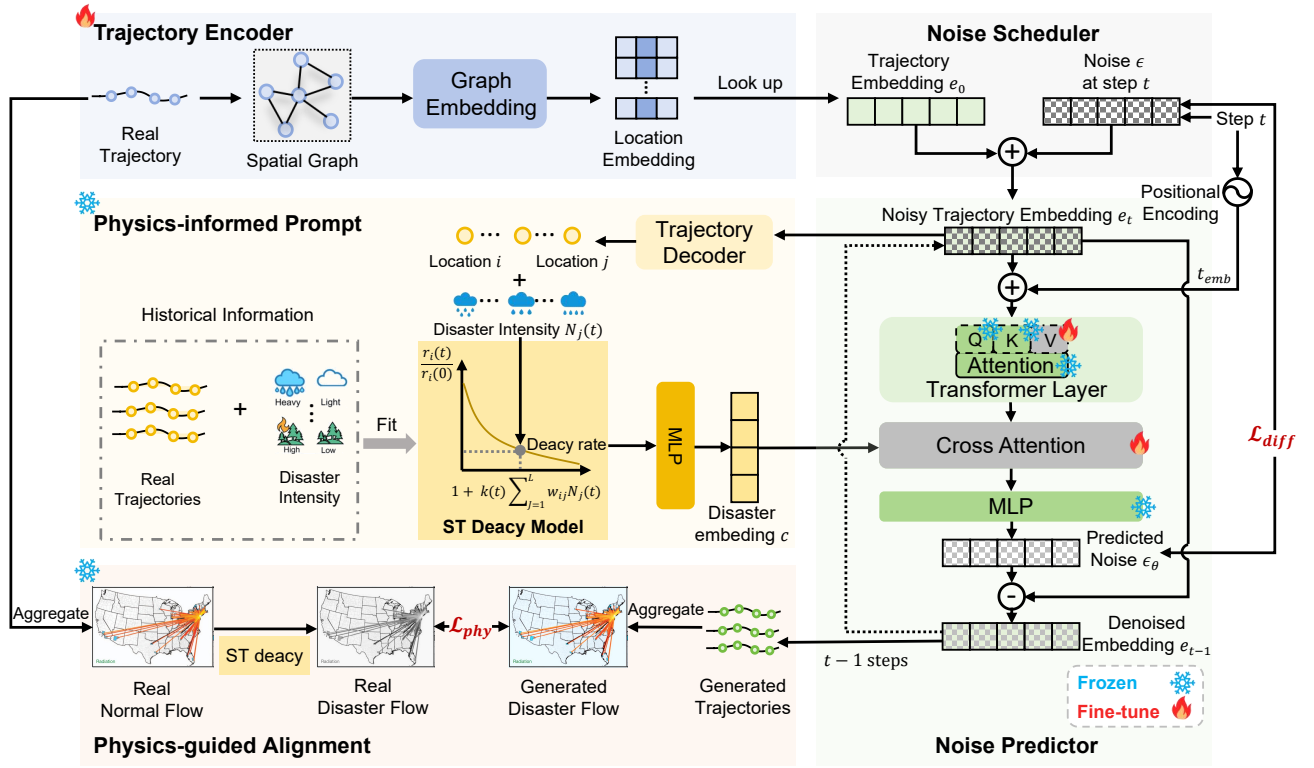


Fig. 2. The overview architecture of UniDisMob, which consists of five modules: (1) Trajectory Encoder, (2) Noise Scheduler, (3) Physics-informed Prompt, (4) Noise Predictor, (5) Physics-guided Alignment.

mobility dynamics after disasters to improve adaptability to new disaster scenarios. Specifically, we first introduce physics-informed prompts that leverage a spatiotemporal decay model to capture the shared patterns of mobility changes following different disasters, serving as conditions to guide the trajectory generation process. Subsequently, we employ a physics-guided alignment mechanism that integrates the spatiotemporal decay patterns under disasters into the training objectives, aligning the generated trajectories with the physical patterns to ensure more realistic outcomes. To address the heterogeneity across different cities, we adopt a meta-learning framework that extracts shared knowledge from multiple cities and quickly adapts to specific cities through an adaptation mechanism.

A. Trajectory Encoder

Following the common strategies adopted in previous works [32], [33], we perform diffusion and denoising processes on the trajectory embedding. Firstly, we construct two types of temporal embedding matrices: $P \in \mathbb{R}^{T_d \times w}$ and $Z \in \mathbb{R}^{T_h \times W}$, to characterize the temporal dependencies and periodic patterns of mobility behaviors. Specifically, the matrix P encodes the weekly periodic patterns, where $T_d = 7$ denotes the number of days in a week. The matrix Z represents the temporal features of different time intervals within a day (e.g., morning and evening peaks), where T_h denotes the number of time slots determined by the sampling granularity. Through temporal embeddings, the model can effectively capture the periodicity and regularity in the temporal dimension.

Secondly, to capture the spatial characteristics of mobility behaviors, we construct a spatial graph $G = (V, E)$ to model

the geographical continuity of trajectory points, ensuring that each transition spans a limited spatial range. In this graph, the node set V represents all visited locations, while the edge set E defines the topological relationships between locations, reflecting their spatial proximity and potential interactions. Each edge $e = (u, v)$ is an unordered pair associated with a positive weight w_{uv} , which quantifies the Euclidean distance between locations u and v , thereby characterizing the spatial closeness among locations. We adopt a graph embedding function $\mathcal{F}(\cdot)$ to derive the location embedding D , which is denoted as follows:

$$D = \mathcal{F}_\theta(G), \quad (7)$$

This process yields a spatial embedding matrix $D \in \mathbb{R}^{L \times W}$, where L denotes the total number of regions and W indicates the embedding dimension.

Finally, we concatenate the spatial and temporal embeddings to form the comprehensive trajectory representation, as follows:

$$e = [D; P; Z], \quad (8)$$

where e denotes the embedding vector of the trajectory, integrating both spatial and temporal features to provide high-dimensional, information-rich inputs for generation tasks.

B. Noise Scheduler

To enhance the robustness of trajectory representation learning and promote the generation of high-quality samples, we adopt a noise scheduling mechanism that progressively injects noise into the trajectory embeddings. Specifically, given an initial clean embedding e_0 , the noise injection at diffusion step

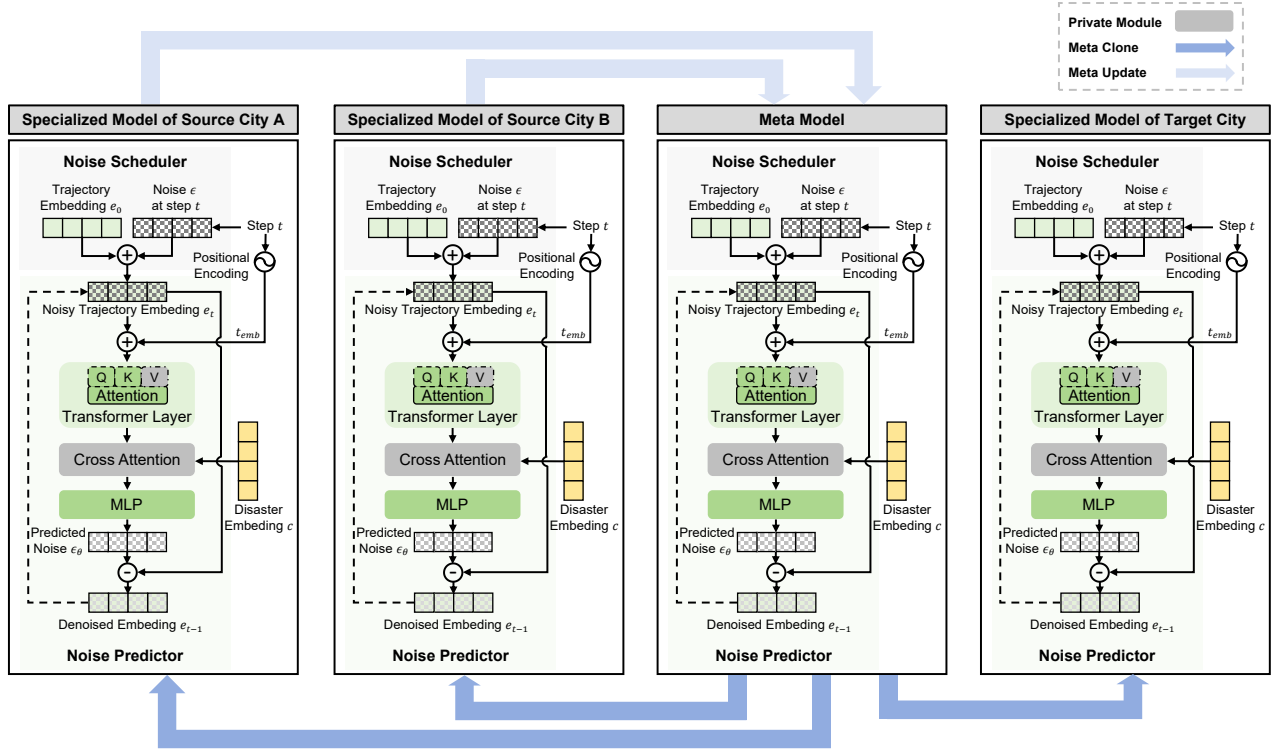


Fig. 3. The meta-learning framework, which extracts common patterns across multiple cities through shared modules, while capturing city-specific characteristics via private modules.

t is implemented by adding Gaussian noise to the embedding, formally defined as:

$$\mathbf{e}_t = \sqrt{\bar{\alpha}_t} \mathbf{e}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, \quad (9)$$

where $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$ denotes standard Gaussian noise, and $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$ represents the cumulative product of noise scaling factors. The sequence $\{\alpha_s\}$ controls the scheduling strategy for noise intensity, which can be smoothly adjusted using linear or cosine annealing.

C. Physics-informed Prompt

The extent and duration of mobility changes caused by different disasters vary significantly across disaster types. Therefore, it is necessary to rely on explicit physical interpretations to model the unified patterns of mobility changes. The spatiotemporal decay model (ST decay model) captures the latent spatiotemporal patterns of crowd mobility following various disasters and reveals the underlying consistencies within these changes [4].

The spatiotemporal decay model demonstrates that across various types of disasters, despite differences in their causes and durations, the resulting disturbances in human mobility consistently exhibit a unified hyperbolic decay pattern. On the one hand, there is a spatial decay effect: the closer a population is to the disaster's core area, the more pronounced the reduction in mobility, while those farther away experience relatively minor impacts. On the other hand, there is a temporal decay effect: as time progresses, the demand to resume normal activities gradually increases, leading to a recovery in mobility.

Based on these observations, this study proposes a hyperbolic model that jointly captures spatiotemporal decay dynamics as follows:

$$r_i(t) = \frac{r_i(0)}{1 + k(t) \sum_{j=1}^L w_{ij} N_j(t)}, \quad (10)$$

where $w_{ij} N_j(t)$ characterizes spatial decay, w_{ij} represents the spatial weight between location i and j , and $N_j(t)$ denotes the intensity of the disaster in location j at time t .

$k(t)$ characterizes temporal decay. As time progresses, the initial changes in mobility behavior gradually diminish until they eventually disappear. Therefore, $k(t)$ is a time-dependent dynamic function, expressed by the following equation:

$$k(t) = k_0 e^{-\alpha t}, \quad (11)$$

where α is the parameter that controls the rate of decay. $k(0)$ represents the initial rate of change in mobility behavior, which is assumed to reach its maximum at $t = 0$. As t becomes sufficiently large, $k(t)$ gradually approaches zero.

Based on historical trajectories collected under various types of disasters and their corresponding disaster intensities, we fit the aforementioned spatiotemporal decay model to characterize the underlying patterns across different disasters, as formulated below:

$$H_\theta(i, t) = \frac{r_i(t)}{r_i(0)} = \frac{1}{1 + k(t) \sum_{j=1}^L w_{ij} N_j(t)}, \quad (12)$$

At each step of the denoising process, the model uses the current noisy trajectory embedding e_t and decodes it into a sequence of locations through the trajectory decoder:

$$s_t = \{l_1, \dots, l_i\} = \arg \max \text{sim}(e_t, \mathbf{D}), \quad (13)$$

where $\text{sim}(\cdot)$ is a similarity function, and $\arg \max$ returns the index corresponding to the maximum value.

The decoded location i , the current time t , and the corresponding disaster intensity N are input into the fitted ST decay model $H_\theta(i, t)$ to obtain the mobility decay rate r_i for each location. Then, we use an MLP to obtain the disaster embedding c , which models the representation of the impact of the current disaster conditions on mobility. The formulation is as follows:

$$c_i = \text{MLP}(r_i), \quad (14)$$

In the denoising network, the disaster embedding c serves as a prompt that provides contextual information about the impact of disasters on mobility. Thus, it dynamically guides the generation process and ensures that the generated trajectories conform to disaster-specific patterns.

D. Noise Predictor

At each step of the denoising process, the model leverages the denoised embedding e_t , the current diffusion step t , and the disaster embedding d to achieve condition-guided denoising. The noise predictor takes these inputs to output the predicted noise:

$$\hat{\epsilon}_\theta = \epsilon_\theta(e_t, t, c). \quad (15)$$

To obtain the representation of the diffusion step t , we follow previous works [34], [35] and adopt a positional encoding approach, encoding t into a 128-dimensional vector t_{emb} as follows:

$$t_{emb} = \left[\sin \left(10^{0 \times 4/63} t \right), \dots, \sin \left(10^{63 \times 4/63} t \right), \right. \\ \left. \cos \left(10^{0 \times 4/63} t \right), \dots, \cos \left(10^{63 \times 4/63} t \right) \right]. \quad (16)$$

We add the positional encoding of the diffusion step t to the trajectory embedding e_t at step t :

$$z_t = e_t + t_{emb}. \quad (17)$$

Then, we utilize a transformer to capture the spatio-temporal dependencies of trajectories, formulated as:

$$h_t = \text{Transformer}(z_t). \quad (18)$$

To introduce dynamic guidance from disaster conditions during the trajectory denoising process, we fuse the disaster embedding d obtained from the physics-informed prompt module with the trajectory representation h_t using a cross-attention mechanism. Specifically, the disaster embedding d is used as the key and value input, while the current trajectory representation h_t serves as the query. The computation is as follows:

$$\tilde{h}_t = \text{CrossAttn}(h_t W_Q, d W_K, d W_V) \\ = \text{softmax} \left(\frac{h_t W_Q (d W_K)^\top}{\sqrt{d_k}} \right) (d W_V), \quad (19)$$

where h_t denotes the trajectory features obtained from the Transformer, d is the disaster embedding, d_k represents the dimensionality of the key vectors, and W_Q , W_K , and W_V are learnable projection matrices.

This conditional fusion mechanism allows each trajectory point to adaptively acquire relevant information from the disaster context according to its state, thus realizing the perception of the disaster context. The final output $\tilde{h}_t$ is fed into an MLP to predict the noise at the current diffusion step t :

$$\hat{\epsilon}_\theta = \text{MLP}(\tilde{h}_t), \quad (20)$$

The output $\hat{\epsilon}_\theta$ is of the same shape as e_t .

E. Physics-guided Alignment

To enhance the realism of trajectory generation under disaster scenarios, we introduce the physics-guided alignment module. This module aligns the generated disaster-induced human flow with the disaster flow modeled by the physical method, guiding the trajectory generation process to follow the macroscopic mobility trends influenced by disasters.

We aggregate the real trajectories under normal scenarios X_{normal} and the generated trajectories X_{gen} under disaster scenarios into crowd flow F_{normal} and F_{gen} , expressed as:

$$F_{normal} = \text{Aggregate}(X_{normal}), \quad (21)$$

$$F_{gen} = \text{Aggregate}(X_{gen}), \quad (22)$$

where $F \in \mathbb{R}^{L \times T}$, T denotes the number of time slots, and L represents the number of spatial locations.

We use the spatiotemporal decay model $H_\theta(i, t)$ fitted in Section IV-C to estimate the disaster-induced crowd flow F_{dis} based on the normal crowd flow F_{normal} :

$$F_{dis} = H_\theta(i, t) F_{normal}. \quad (23)$$

We align the crowd flow F_{gen} , obtained by aggregating the trajectories generated by UniDisMob, with the disaster-induced crowd flow F_{dis} derived from the spatiotemporal decay model, and define the physics-guided loss function $\mathcal{L}_{phy}$ as follows:

$$\mathcal{L}_{phy} = \|F_{gen} - F_{dis}\|_2^2. \quad (24)$$

This loss serves as an additional supervisory signal during training, constraining the generated trajectories to align with the decay patterns observed under disaster scenarios at the population level.

F. Training and Sampling

a) *Training*: As shown in Figure. 3, to address the heterogeneity across different cities, we adopt a meta-learning framework that extracts universal mobility patterns from multiple source cities and enables quick adaptation to each target city. The modules marked as frozen and fine-tuned in Figure. 2 reflect the division between shared and private components. The shared modules learn cross-city common mobility patterns for knowledge transfer, while the private modules capture city-specific features to enhance local adaptability.

Our loss function consists of the diffusion model loss and an additional physics-guided loss. Based on the derivation in

Section III-B, the diffusion model is trained using a simplified loss function:

$$\mathcal{L}_{diff} = \mathbb{E}_{e_0, \epsilon, t} \|\epsilon - \epsilon_\theta(e_t, t|d)\|_2^2. \quad (25)$$

Therefore, the overall loss function is defined as follows:

$$\mathcal{L}_{total} = \alpha \mathcal{L}_{diff} + \beta \mathcal{L}_{phy}, \quad (26)$$

where α and β are the weighting coefficients for the two loss terms. The total loss $\mathcal{L}_{total}$ encourages the model to preserve individual trajectory characteristics while generating mobility patterns that align with physical laws under disaster scenarios.

Meta-training on Source Cities: We extract universal mobility patterns from multiple source cities. For each source city $D^k = (D_{train}^k, D_{val}^k)$, we first clone the shared modules from the meta model and combine them with the city-specific private modules to form the initial model parameters:

$$\Theta_{src}^k = \Theta_{shared}^{meta} \cup \Theta_{private}^k. \quad (27)$$

Next, internal update is performed on the training set D_{train}^k of the city to minimize the total loss $\mathcal{L}_{total}$ and update both the shared and private modules:

$$\Theta_{src}^k \leftarrow \Theta_{src}^k - \alpha_{src} \nabla_{\Theta} (L_{total}). \quad (28)$$

Then, the updated model is evaluated on the validation set D_{val}^k of the corresponding city, and the gradients are computed to update only the shared modules (Meta Update):

$$\Theta_{meta} \leftarrow \Theta_{meta} - \alpha_{meta} \nabla_{\Theta_{shared}} (L_{total}). \quad (29)$$

Target Adaptation: After completing meta-training on the source cities, we select the target city dataset $\tilde{D} = (\tilde{D}_{train}, \tilde{D}_{test})$, clone the shared module parameters, and combine them with the private modules of the target city to obtain the initialized model parameters:

$$\Theta_{tgt} = \Theta_{shared}^{meta} \cup \Theta_{private}^{tgt}. \quad (30)$$

Then, fine-tuning is performed on the training set $\tilde{D}_{train}$ of the target city, aiming to minimize the total loss function:

$$\Theta_{tgt} \leftarrow \Theta_{tgt} - \alpha_{tgt} \nabla_{\Theta} (L_{total}). \quad (31)$$

Finally, the model is evaluated on the test set $\tilde{D}_{test}$ of the target city and used to generate trajectories under disaster scenarios.

b) Sampling: We sample using a linear combination of conditional and unconditional predictions [36]:

$$\tilde{e}_\theta = (1 + \omega) \epsilon_\theta(e_t, t|d) - \omega \epsilon_\theta(e_t, t|\emptyset), \quad (32)$$

where $\emptyset$ indicates a vector that effectively does not encode any information. When $\omega = 0$, the model is fully conditional; $\omega = 1$ yields an unconditional model. Increasing ω balances diversity and sample quality.

Denosing proceeds iteratively from step T to 1. At each step, noise is removed via:

$$e_{t-1} = \frac{1}{\sqrt{\alpha_t}} (e_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \tilde{e}_\theta) + \sigma_t z \quad (33)$$

where $\tilde{e}_\theta$ estimates the noise ϵ from e_t , and $z \sim N(0, I)$.

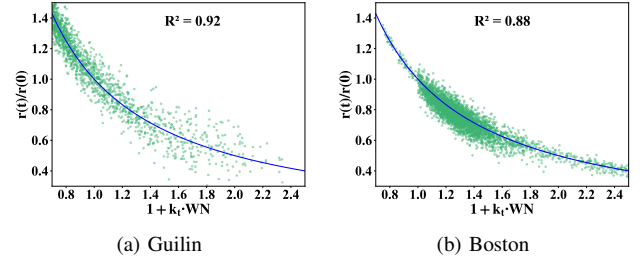


Fig. 4. Fitting performance of the spatiotemporal decay model across different cities.

V. EXPERIMENTS

A. Experimental Settings

To evaluate the performance of UniDisMob, we conduct extensive experiments on seven city datasets. A unified model is trained using data from four cities (Houston, Guilin, Phoenix, and Boston), and zero-shot testing is performed on the remaining three cities (Los Angeles, Worcester, and Sacramento). In this setting, the model generates trajectories for target cities without having seen any trajectory data under disaster scenarios, thereby assessing its cross-city generalization and transferability.

1) Dataset: The detailed information of the datasets is shown in Table I. We use seven real-world mobility datasets from different cities under various disaster types. We use four datasets (Houston, Guilin, Phoenix, and Boston) to jointly train our unified model, and conduct transfer experiments on three cities (Los Angeles, Worcester, and Sacramento) to evaluate the zero-shot capability of our model. For data preprocessing, we filter out users with fewer than five records per day. In the spatial dimension, raw GPS coordinates are converted to predefined grid IDs at a specified granularity. In the temporal dimension, timestamps are uniformly segmented into fixed-length intervals, using half-hour intervals as the unit.

As shown in Figure 4, we present the fitting performance of the spatiotemporal decay model, which was trained on multiple datasets, across different cities. The results indicate that the model effectively captures the decay patterns of human mobility over time and space under disaster scenarios, demonstrating its generalization ability across different types of disasters.

2) Metrics: We evaluate the generated trajectory using the following six metrics: The first four assess the behavioral characteristics of trajectories, and the last two measure mobility changes under disaster scenarios.

- **Distance:** It refers to the total length of a user's movement paths accumulated over a specific period and is evaluated using JSD.
- **Radius:** Radius of gyration measures the spatial extent of a user's daily activity range and is evaluated using JSD.
- **Duration:** It indicates the time a user spends at each visited location and is evaluated using JSD.
- **Daily-loc:** It represents the number of distinct places a user visits daily and is evaluated using JSD.
- **Decay Rate:** The degree of mobility decline compared to the normal state after a disaster is measured using MAPE.

TABLE I
BASIC STATISTICS OF MOBILITY DATASETS.

City	Disaster Type	Data Duration	Disaster Duration	Disaster Intensity	#Users
Houston	Hurricane	30 days	9 days	Rainfall	24162
Guilin	Rainstorm	17 days	6 days	Rainfall	12035
Phoenix	Extreme Heat	50 days	9 days	Temperature Rise	8637
Boston	Winter Storm	28 days	8 days	Temperature Drop	20191
Los Angeles	Wildfire	40 days	7 days	PM2.5	15408
Worcester	Winter Storm	31 days	7 days	Temperature Drop	6810
Sacramento	Extreme Heat	45 days	10 days	Temperature Rise	3784

- **Decay Speed:** The speed of recovery from the mobility disruption is represented by the parameter α in the function $k(t) = k_0 e^{-\alpha t}$, and is evaluated using MAPE.

MAPE is the mean absolute percentage error between the predicted and ground truth values. **JSD** is the Jensen-Shannon divergence, which is used to evaluate the similarity between the feature distributions of real trajectories and generated trajectories. The calculation formula is:

$$\text{JSD}(a, b) = \frac{1}{2} \text{KL}(a || \frac{a+b}{2}) + \frac{1}{2} \text{KL}(b || \frac{a+b}{2}), \quad (34)$$

where $\text{KL}(\cdot || \cdot)$ is the Shannon information, a and b are distributions.

3) Baselines:

a) *Trajectory Generation under Normal Scenario:* These models are designed for human mobility generation in normal scenarios.

- **TimeGeo** [28]: This method captures the temporal characteristics (travel rhythm, dwell rate, and burstiness) and the spatial characteristics (exploration and preferential return mechanisms) to generate individual trajectories.
- **MoveSim** [9]: This method is a GAN-based framework that leverages urban structure priors and mobility regularity-aware loss to generate more realistic trajectories.
- **PateGail** [10]: PateGail is a novel method that combines GAIL with differential privacy to generate high-quality trajectories while protecting user privacy.
- **DiffTraj** [7]: DiffTraj is a trajectory generation method based on a diffusion model, which is designed to simulate high-fidelity human mobility behavior.

b) *Trajectory Generation under Disaster Scenario:* These models are designed for human mobility modeling in disaster scenarios.

- **DeepMob** [11]: DeepMob leverages deep neural networks to extract rich representations of human mobility from large-scale data and models individual movement decision-making processes under emergency scenarios.
- **EmeMob** [29]: This method captures key patterns of human daily mobility and models how factors such as social relationships, disaster severity, government shelters, and news reports influence mobility behavior during disasters.

c) *Controllable Generation:* These baselines are controllable generation models that can use disaster information as conditions to guide the trajectory generation.

- **CSDI** [29]: This method is a conditional diffusion model that can reconstruct complete data by progressive denoising under the condition of observed values.

- **ControlTraj** [37]: This method is a controllable trajectory generation framework that leverages diffusion models combined with conditional constraints to generate high-fidelity trajectories.

d) *Cross-city Generation:* These baselines leverage knowledge from source cities to assist trajectory generation in target cities.

- **COLA** [38]: COLA is a cross-city human trajectory simulation framework that leverages model-agnostic meta-learning strategies to achieve knowledge transfer.
- **CHAML** [39]: This method is a cross-city prediction framework designed to leverage the experience of data-rich source cities to improve prediction performance in target cities.

B. Overall Performance

We evaluate the performance of various baselines and our approach across multiple mobility metrics in four cities affected by different types of disasters. We conduct multiple experiments and reported the average performance. Our method achieves an average improvement of over 13% across the four datasets. As shown in Table II, our approach outperforms all baseline methods across all six metrics on both the Guilin and Boston datasets. In particular, it shows notable advantages in modeling mobility changes caused by disasters. Specifically, our method achieves an average improvement of over 18% on the Decay Rate and Decay Speed metrics. As shown in Table III, our method also demonstrates strong performance on most metrics in the Phoenix and Houston datasets. These results indicate that our method excels in modeling human mobility across spatial and temporal dimensions and in capturing mobility perturbations caused by disasters. It more accurately replicates real-world human movement patterns under disaster scenarios, especially in quantifying the impact severity and recovery dynamics.

Notably, the two best-performing baselines, ControlTraj and COLA, represent two mainstream approaches. ControlTraj is based on diffusion models and adopts a condition-controlled generation strategy, while COLA is built on a Transformer architecture and incorporates meta-learning to enable cross-city transfer. Both methods enhance generalization by leveraging either conditional generation or transfer learning techniques. Our proposed UniDisMob integrates both strengths. On the one hand, it employs a diffusion transformer guided by physics-informed priors to guide the trajectory generation process, ensuring that the generated mobility patterns align

TABLE II
OVERALL PERFORMANCE ON GUILIN AND BOSTON DATASETS.

Dataset	Guilin						Boston					
	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed
TimeGEO	0.0628	0.4703	0.0912	0.4257	28.6%	23.2%	0.0438	0.3614	0.0795	0.3872	27.8%	22.3%
PateGail	0.0207	0.3361	0.0519	0.1841	20.6%	18.3%	0.0185	0.3149	0.0642	0.2256	19.7%	17.4%
MoveSim	0.0236	0.3875	0.0603	0.2740	22.4%	19.1%	0.0219	0.3017	0.0594	0.1908	21.6%	18.9%
DiffTraj	0.0156	0.2962	0.0480	0.1037	15.3%	12.8%	0.0152	0.2651	0.0487	0.1202	14.3%	12.2%
EmeMob	0.0149	0.2826	0.0458	0.0915	13.7%	11.8%	0.0151	0.2538	0.0471	0.1073	13.2%	11.3%
DeepMob	0.0142	0.2714	0.0439	0.0832	12.4%	10.4%	0.0146	0.2460	0.0453	0.0975	11.9%	9.7%
CSDI	0.0135	0.2517	0.0412	0.0694	11.6%	9.3%	0.0149	0.2385	0.0465	0.0864	10.7%	8.8%
ControlTraj	0.0119	0.2405	0.0381	0.0673	10.2%	8.2%	0.0138	0.2261	0.0448	0.0805	9.6%	7.7%
CHAML	0.0137	0.2612	0.0418	0.0723	10.7%	9.1%	0.0150	0.2351	0.0459	0.0815	10.2%	8.5%
COLA	0.0132	0.2526	0.0407	0.0690	9.3%	7.8%	0.0154	0.2298	0.0451	0.0772	8.8%	7.2%
Ours	0.0083	0.2059	0.0320	0.0638	7.8%	6.2%	0.0126	0.1842	0.0435	0.0714	7.4%	5.9%
Improvement	30.2%	14.4%	15.9%	5.2%	15.9%	18.1%	2.3%	18.5%	2.9%	7.5%	16.1%	20.5%

TABLE III
OVERALL PERFORMANCE ON PHOENIX AND HOUSTON DATASETS.

Dataset	Phoenix						Houston					
	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed
TimeGEO	0.0735	0.5184	0.0983	0.4526	32.4%	30.1%	0.0647	0.4861	0.1127	0.4076	30.7%	29.2%
PateGail	0.0418	0.3897	0.0773	0.2537	25.3%	22.8%	0.0375	0.3564	0.0895	0.2283	23.6%	22.1%
MoveSim	0.0462	0.4228	0.0835	0.2981	27.0%	24.1%	0.0451	0.3952	0.0943	0.2661	25.2%	23.3%
DiffTraj	0.0357	0.3426	0.0721	0.1934	18.2%	16.0%	0.0312	0.3117	0.0855	0.1826	17.4%	15.6%
EmeMob	0.0331	0.3290	0.0692	0.1785	16.5%	14.7%	0.0304	0.2989	0.0837	0.1704	15.9%	14.2%
DeepMob	0.0316	0.3175	0.0674	0.1692	14.9%	12.9%	0.0291	0.2865	0.0818	0.1612	14.1%	12.4%
CSDI	0.0305	0.3194	0.0688	0.1659	13.6%	11.6%	0.0287	0.2311	0.0831	0.1635	13.0%	11.0%
ControlTraj	0.0281	0.3078	0.0675	0.1513	11.3%	9.8%	0.0240	0.2180	0.0801	0.1285	11.0%	9.2%
CHAML	0.0270	0.2945	0.0660	0.1492	12.0%	10.5%	0.0235	0.2350	0.0830	0.1462	11.7%	10.0%
COLA	0.0276	0.3052	0.0661	0.1485	10.4%	9.0%	0.0225	0.2226	0.0817	0.1440	10.1%	8.6%
Ours	0.0247	0.2873	0.0624	0.1251	8.3%	7.5%	0.0201	0.1937	0.0795	<u>0.1363</u>	8.0%	7.3%
Improvement	10.5%	5.9%	7.5%	17.3%	20.2%	16.7%	10.6%	11.1%	0.8%	-	20.8%	15.1%

with real-world disaster-induced changes. On the other hand, it utilizes a meta-learning framework to extract universal mobility patterns across cities while adapting to city-specific characteristics, thereby achieving stronger transferability. As the first unified framework tailored for trajectory generation in disaster scenarios, UniDisMob significantly expands the frontier of research in human mobility modeling.

C. Zero-shot Performance

Figures 5, 6 and 7 compare the transferability of UniDisMob with several baseline models across three representative cities (Los Angeles, Worcester, and Sacramento), each corresponding to a different type of disaster: wildfire, winter storm, and extreme heat. The transfer setting assumes that no trajectory data under disaster scenarios is available for the target city. Instead, only a small amount of trajectories under normal scenarios is provided. We leverage these normal trajectories to perform lightweight fine-tuning of the general model, enabling effective adaptation to disaster scenarios in the target city. As shown in the figures, UniDisMob consistently achieves the best zero-shot performance across all metrics and datasets, outperforming even those baselines specifically designed for cross-city generalization. On average, it delivers over 8% improvement across the three datasets. These results demonstrate that UniDisMob, through its physics-informed design, can effectively capture the universal patterns of mobility changes under disaster scenarios. This enables it to generalize

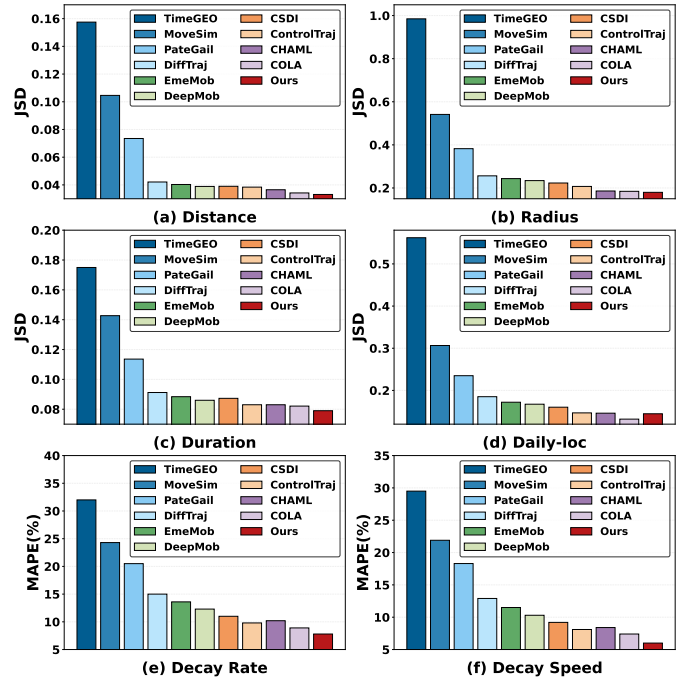


Fig. 5. Zero-shot performance on Los Angeles datasets.

well to entirely new cities or disaster types, showcasing strong adaptability and broad applicability.

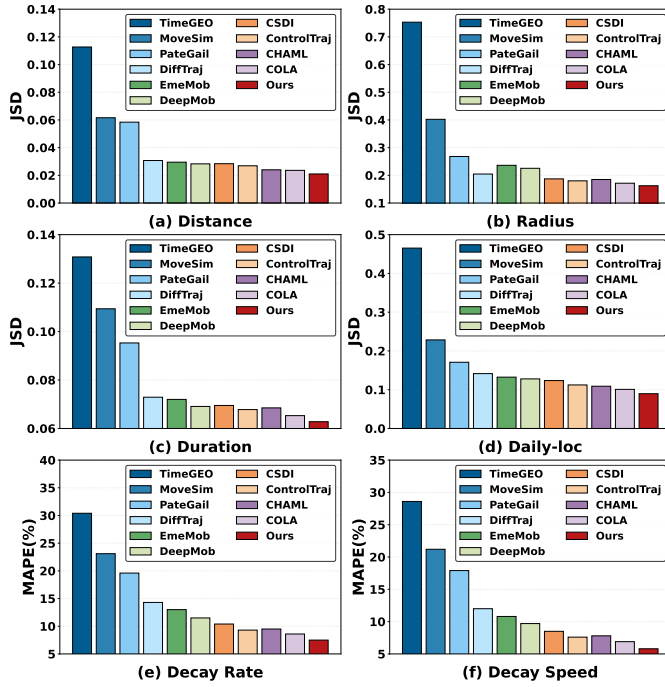


Fig. 6. Zero-shot performance on Worcester datasets.

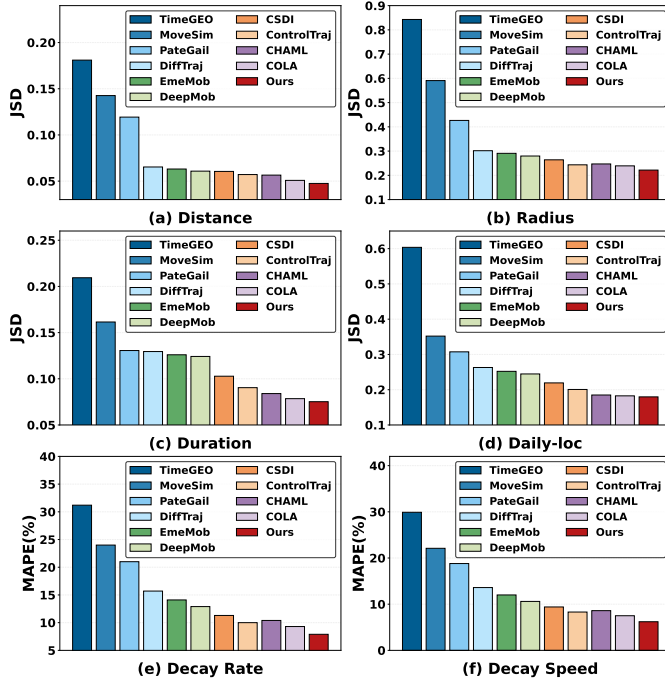


Fig. 7. Zero-shot performance on Sacramento datasets.

D. Ablation Study

To evaluate the contributions of different components in UniDisMob, we conduct comprehensive ablation studies across four cities—Guilin, Boston, Phoenix, and Houston. We successively ablated three modules: (1) w/o prompt: removes the physics-informed prompt module. (2) w/o alignment: removes the physics-guided alignment module. (3) w/o meta learning: removes the meta-learning mechanism and uses a unified model

without city-specific adaptation.

The results are shown in Tables II and V. Removing the physics-guided alignment module has the most significant impact on performance, especially on metrics that describe disaster-induced mobility disruptions, such as Decay Rate and Decay Speed. This highlights the critical role of explicitly aligning spatiotemporal decay patterns under disasters with individual trajectories to ensure physical consistency in the generated mobility patterns. Removing the physics-informed prompt module leads to moderate performance drops across multiple metrics, indicating that the physics-informed prompt effectively guides the model to perceive disaster conditions and learn mobility patterns that better reflect real-world disaster scenarios. Removing the meta-learning mechanism results in relatively smaller performance degradation, but consistently negative effects are observed across all metrics and cities. This confirms the importance of meta-learning in adapting to the heterogeneity across different cities.

VI. CONCLUSION

In this paper, we propose a unified model named UniDisMob for human mobility generation under disaster scenarios. The model demonstrates strong generalization across different cities and disaster types. We design a physics-informed neural network to capture the underlying common patterns of mobility disruptions caused by various disasters. Moreover, UniDisMob introduces a meta-learning mechanism to address the heterogeneity among different cities and enhance adaptability. Extensive experiments on multiple real-world datasets show the superiority of our model, particularly under zero-shot settings. In the future, we plan to broaden the unified model's applicability to various tasks and explore the integration of more data modalities, such as social network data, to enhance human mobility modeling.

REFERENCES

- [1] Z. Xu and S. S. Chopra, "Interconnectedness enhances network resilience of multimodal public transportation systems for safe-to-fail urban mobility," *Nature communications*, vol. 14, no. 1, p. 4291, 2023.
- [2] J. Tang, P. Zhao, Z. Gong, H. Zhao, F. Huang, J. Li, Z. Chen, L. Yu, and J. Chen, "Resilience patterns of human mobility in response to extreme urban floods," *National science review*, vol. 10, no. 8, p. nwad097, 2023.
- [3] Z. Zhang, H. Wang, Z. Fan, X. Song, and R. Shibasaki, "Human decision-making strategy analysis during large-scale disaster," 2022.
- [4] W. Li, Q. Wang, Y. Liu, M. L. Small, and J. Gao, "A spatiotemporal decay model of human mobility when facing large-scale crises," *Proceedings of the National Academy of Sciences*, vol. 119, no. 33, p. e2203042119, 2022.
- [5] Q. Long, C. Rong, H. Wang, and Y. Li, "One fits all: General mobility trajectory modeling via masked conditional diffusion," *arXiv preprint arXiv:2501.13347*, 2025.
- [6] M. Jin, S. Wang, L. Ma, Z. Chu, J. Y. Zhang, X. Shi, P.-Y. Chen, Y. Liang, Y.-F. Li, S. Pan *et al.*, "Time-llm: Time series forecasting by reprogramming large language models," *arXiv preprint arXiv:2310.01728*, 2023.
- [7] Y. Zhu, Y. Ye, S. Zhang, X. Zhao, and J. Yu, "DiffTraj: Generating gps trajectory with diffusion probabilistic model," *Advances in Neural Information Processing Systems*, vol. 36, pp. 65 168–65 188, 2023.
- [8] Q. Long, H. Wang, T. Li, L. Huang, K. Wang, Q. Wu, G. Li, Y. Liang, L. Yu, and Y. Li, "Practical synthetic human trajectories generation based on variational point processes," in *Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining*, 2023, pp. 4561–4571.

TABLE IV
ABLATION RESULTS ON GUILIN AND BOSTON DATASETS.

Dataset	Guilin						Boston					
	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed
Ours	0.0083	0.2059	0.0320	0.0638	7.8%	6.2%	0.0126	0.1842	0.0435	0.0714	7.4%	5.9%
w/o prompt	0.0095	0.2203	0.0351	0.0694	8.5%	6.8%	0.0136	0.1974	0.0482	0.0780	8.2%	6.6%
	(-14.5%)	(-7.0%)	(-9.7%)	(-8.8%)	(-9.0%)	(-9.7%)	(-7.9%)	(-7.2%)	(-10.8%)	(-9.2%)	(-10.8%)	(-11.9%)
w/o alignment	0.0102	0.2347	0.0376	0.0719	8.8%	7.1%	0.0145	0.2029	0.0501	0.0813	8.4%	6.8%
	(-22.9%)	(-14.0%)	(-17.5%)	(-12.7%)	(-12.8%)	(-14.5%)	(-15.1%)	(-10.2%)	(-15.2%)	(-13.9%)	(-13.5%)	(-15.3%)
w/o meta learning	0.0090	0.2138	0.0336	0.0659	7.9%	6.4%	0.0130	0.1895	0.0457	0.0745	7.6%	6.1%
	(-8.4%)	(-3.8%)	(-5.0%)	(-3.3%)	(-1.3%)	(-3.2%)	(-3.2%)	(-2.9%)	(-5.1%)	(-4.3%)	(-2.7%)	(-3.4%)

TABLE V
ABLATION STUDY ON PHOENIX AND HOUSTON DATASETS.

Dataset	Phoenix						Houston					
	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed	Distance	Radius	Duration	Daily-loc	Decay Rate	Decay Speed
Ours	0.0247	0.2873	0.0624	0.1251	8.3%	7.5%	0.0201	0.1937	0.0795	0.1363	8.0%	7.3%
w/o prompt	0.0261	0.2941	0.0657	0.1376	9.3%	8.5%	0.0217	0.2075	0.0806	0.1408	8.8%	7.8%
	(-5.7%)	(-2.4%)	(-5.3%)	(-10.0%)	(-12.0%)	(-13.3%)	(-8.0%)	(-7.1%)	(-1.4%)	(-3.3%)	(-10.0%)	(-6.8%)
w/o alignment	0.0279	0.3075	0.0680	0.1491	9.8%	8.6%	0.0228	0.2135	0.0823	0.1457	9.3%	8.2%
	(-13.0%)	(-7.0%)	(-9.0%)	(-19.2%)	(-18.1%)	(-14.7%)	(-13.4%)	(-10.2%)	(-3.5%)	(-6.9%)	(-16.3%)	(-12.3%)
w/o meta learning	0.0254	0.2902	0.0642	0.1309	8.6%	7.8%	0.0210	0.2018	0.0799	0.1385	8.2%	7.5%
	(-2.8%)	(-1.0%)	(-2.9%)	(-4.6%)	(-3.6%)	(-4.0%)	(-4.5%)	(-4.2%)	(-0.5%)	(-1.6%)	(-2.5%)	(-2.7%)

- [9] J. Feng, Z. Yang, F. Xu, H. Yu, M. Wang, and Y. Li, "Learning to simulate human mobility," in *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 2020, pp. 3426–3433.
- [10] H. Wang, C. Gao, Y. Wu, D. Jin, L. Yao, and Y. Li, "Pategail: A privacy-preserving mobility trajectory generator with imitation learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 12, 2023, pp. 14 539–14 547.
- [11] X. Song, R. Shibusaki, N. J. Yuan, X. Xie, T. Li, and R. Adachi, "Deepmob: learning deep knowledge of human emergency behavior and mobility from big and heterogeneous data," *ACM Transactions on Information Systems (TOIS)*, vol. 35, no. 4, pp. 1–19, 2017.
- [12] R. Jiang, Z. Wang, Y. Tao, C. Yang, X. Song, R. Shibusaki, S.-C. Chen, and M.-L. Shyu, "Learning social meta-knowledge for nowcasting human mobility in disaster," in *Proceedings of the ACM Web Conference 2023*, 2023, pp. 2655–2665.
- [13] J. Liu, Z. Shen, Y. He, X. Zhang, R. Xu, H. Yu, and P. Cui, "Towards out-of-distribution generalization: A survey," *arXiv preprint arXiv:2108.13624*, 2021.
- [14] K. Lakara, A. Bhandari, P. Seth, and U. Verma, "Evaluating predictive uncertainty and robustness to distributional shift using real world data," *arXiv preprint arXiv:2111.04665*, 2021.
- [15] H. Xue, T. Tang, A. Payani, and F. D. Salim, "Prompt mining for language models-based mobility flow forecasting," in *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*, 2024, pp. 113–122.
- [16] N. Gruver, M. Finzi, S. Qiu, and A. G. Wilson, "Large language models are zero-shot time series forecasters," *Advances in Neural Information Processing Systems*, vol. 36, pp. 19 622–19 635, 2023.
- [17] M. Jin, Q. Wen, Y. Liang, C. Zhang, S. Xue, X. Wang, J. Zhang, Y. Wang, H. Chen, X. Li *et al.*, "Large models for time series and spatio-temporal data: A survey and outlook," *arXiv preprint arXiv:2310.10196*, 2023.
- [18] H. Liu, Z. Zhao, J. Wang, H. Kamarthi, and B. A. Prakash, "Lstprompt: Large language models as zero-shot time series forecasters by long-short-term prompting," *arXiv preprint arXiv:2402.16132*, 2024.
- [19] S. Zhong, W. Ruan, M. Jin, H. Li, Q. Wen, and Y. Liang, "Time-vlm: Exploring multimodal vision-language models for augmented time series forecasting," *arXiv preprint arXiv:2502.04395*, 2025.
- [20] Z. Li, J. Xu, S. Wang, Y. Wu, and H. Li, "Streetviewllm: Extracting geographic information using a chain-of-thought multimodal large language model," *arXiv preprint arXiv:2411.14476*, 2024.
- [21] G. Mai, W. Huang, J. Sun, S. Song, D. Mishra, N. Liu, S. Gao, T. Liu, G. Cong, Y. Hu *et al.*, "On the opportunities and challenges of foundation models for geospatial artificial intelligence," *arXiv preprint arXiv:2304.06798*, 2023.
- [22] Z. Li, L. Xia, J. Tang, Y. Xu, L. Shi, L. Xia, D. Yin, and C. Huang, "Urbangpt: Spatio-temporal large language models," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 5351–5362.
- [23] D. Cao, W. Ye, Y. Zhang, and Y. Liu, "Timedit: General-purpose diffusion transformers for time series foundation model," *arXiv preprint arXiv:2409.02322*, 2024.
- [24] H. Wen, Y. Lin, Y. Xia, H. Wan, Q. Wen, R. Zimmermann, and Y. Liang, "Diffstg: Probabilistic spatio-temporal graph forecasting with denoising diffusion models," in *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*, 2023, pp. 1–12.
- [25] Y. Yuan, J. Ding, J. Feng, D. Jin, and Y. Li, "Unist: A prompt-empowered universal model for urban spatio-temporal prediction," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 4095–4106.
- [26] Y. Yuan, C. Han, J. Ding, D. Jin, and Y. Li, "Urbangpt: A foundation model for open-world urban spatio-temporal learning," *arXiv preprint arXiv:2411.12164*, 2024.
- [27] C. Song, T. Koren, P. Wang, and A.-L. Barabási, "Modelling the scaling properties of human mobility," *Nature Physics*, vol. 6, no. 10, pp. 818–823, 2010.
- [28] S. Jiang, Y. Yang, S. Gupta, D. Veneziano, S. Athavale, and M. C. González, "The timegeo modeling framework for urban mobility without travel surveys," *Proceedings of the National Academy of Sciences*, vol. 113, no. 37, pp. E5370–E5378, 2016.
- [29] X. Song, Q. Zhang, Y. Sekimoto, and R. Shibusaki, "Prediction of human emergency behavior and their mobility following large-scale disaster," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 5–14.
- [30] Y. Tashiro, J. Song, Y. Song, and S. Ermon, "Csd: Conditional score-based diffusion models for probabilistic time series imputation," *Advances in Neural Information Processing Systems*, vol. 34, pp. 24 804–24 816, 2021.
- [31] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.
- [32] Y. Yuan, J. Ding, C. Shao, D. Jin, and Y. Li, "Spatio-temporal diffusion point processes," *arXiv preprint arXiv:2305.12403*, 2023.
- [33] Z. Zhou, J. Ding, Y. Liu, D. Jin, and Y. Li, "Towards generative modeling of urban flow through knowledge-enhanced denoising diffusion," in *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*, 2023, pp. 1–12.
- [34] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [35] Z. Kong, W. Ping, J. Huang, K. Zhao, and B. Catanzaro, "Dif-fwave: A versatile diffusion model for audio synthesis," *arXiv preprint arXiv:2009.09761*, 2020.

- [36] J. Ho and T. Salimans, “Classifier-free diffusion guidance,” *arXiv preprint arXiv:2207.12598*, 2022.
- [37] Y. Zhu, J. J. Yu, X. Zhao, Q. Liu, Y. Ye, W. Chen, Z. Zhang, X. Wei, and Y. Liang, “Controltraj: Controllable trajectory generation with topology-constrained diffusion model,” in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 4676–4687.
- [38] Y. Wang, T. Zheng, Y. Liang, S. Liu, and M. Song, “Cola: Cross-city mobility transformer for human trajectory simulation,” in *Proceedings of the ACM Web Conference 2024*, 2024, pp. 3509–3520.
- [39] Y. Chen, X. Wang, M. Fan, J. Huang, S. Yang, and W. Zhu, “Curriculum meta-learning for next poi recommendation,” in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021, pp. 2692–2702.