Practical Synthetic Human Trajectories Generation Based on Variational Point Processes

Qingyue Long
Department of Electronic Engineering
BNRist, Tsinghua University
Beijing, China

Huandong Wang*
Department of Electronic Engineering
BNRist, Tsinghua University
Beijing, China

Tong Li
Department of Electronic Engineering
BNRist, Tsinghua University
Beijing, China

Lisi Huang
China Mobile Research Institute,
Beijing, China

Kun Wang
China Mobile Research Institute,
Beijing, China

Qiong Wu
China Mobile Research Institute,
Beijing, China

Guangyu Li
China Mobile Research Institute,
Beijing, China

Yanping Liang
China Mobile Research Institute,
Beijing, China

Li Yu
China Mobile Research Institute,
Beijing, China

Yong Li
Department of Electronic Engineering
BNRist, Tsinghua University
Beijing, China

ABSTRACT

Human trajectories, reflecting people's travel patterns and the range of activities, are crucial for the applications like urban planning and epidemic control. However, the real-world human trajectory data tends to be limited by user privacy or device acquisition issues, leading to its insufficient quality to support the above applications. Hence, generating human trajectory data is a crucial but challenging task, which suffers from the following two critical challenges: 1) how to capture the user distribution in human trajectories (group view), and 2) how to model the complex mobility patterns of each user trajectory (individual view). In this paper, we propose a novel human trajectories generator (named **VOLUNTEER**), consisting of a user VAE and a trajectory VAE, to address the above challenges. Specifically, in the user VAE, we propose to learn the user distribution with all human trajectories from a group view. In the trajectory VAE, from the individual view, we model the complex mobility patterns by decoupling travel time and dwell time to accurately simulate individual trajectories. Extensive experiments on two real-world datasets show the superiority of our model over the state-of-the-art baselines. Further application analysis in the industrial system also demonstrates the effectiveness of our model.

*Huandong Wang is the corresponding author (wanghuandong@tsinghua.edu.cn).

KDD '23, August 6–10, 2023, Long Beach, CA, USA

ACM ISBN 979-8-4007-0103-0/23/08…$15.00
https://doi.org/10.1145/3580305.3599888

CCS CONCEPTS

• **Information systems** → **Spatial-temporal systems**; • **Networks** → *Network simulations*; • **Computing methodologies** → *Simulation tools.*

KEYWORDS

Mobility Trajectory, Generation Model, Variational Auto-Encoder, Temporal Point Process

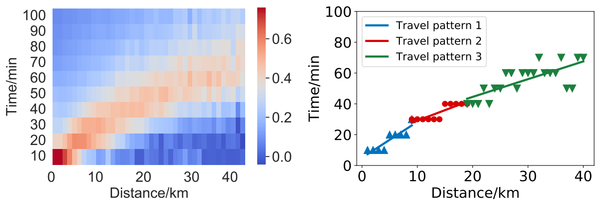
ACM Reference Format:

Qingyue Long, Huandong Wang, Tong Li, Lisi Huang, Kun Wang, Qiong Wu, Guangyu Li, Yanping Liang, Li Yu, and Yong Li. 2023. Practical Synthetic Human Trajectories Generation Based on Variational Point Processes. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23), August 6–10, 2023, Long Beach, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3580305.3599888

1 INTRODUCTION

Human trajectory data reveals human mobility patterns and has various downstream applications, such as urban planning [40], migration flow prediction [19], epidemic control [5], and environmental protection [14], which are highly dependent on users' high-quality trajectories. Specifically, for Internet service providers (ISP), based on the human trajectory data, the dynamics of the spatial distribution of mobile users can be predicted to enable numerous quality of service (QoS) optimization techniques, including bandwidth reservation [30] and content caching [42]. For the location-based services, human trajectories can also help to provide customized services by enabling intelligent spatiotemporal-aware recommendation [22–24, 26], friend recommendation [32], etc.

However, due to data privacy issues and collection expenses, we are unable to get enough high-quality human trajectory data in practice to support the aforementioned applications. Notably, the trajectory data provides fine-grained user activity traces, containing a lot of sensitive information such as users' home addresses,



(a) Heat map of distribution. (b) Scatter of the peak time conditioned on travel distance.

Figure 1: Illustration of the complex distribution (a) and travel patterns (b) of human trajectories in MME dataset.

working places, and social connections. To protect user privacy and comply with the General Data Protection Regulation (GDPR), the data provider typically adds perturbations to the trajectories to disguise users' precise spatiotemporal positions, which will seriously degrade the quality of the data and make it less useful for the downstream applications [2, 3, 10]. Fortunately, the emerging artificial intelligence (AI)-based simulation methods have given us a promising solution to this issue. These methods have demonstrated great success in a variety of applications, including weather prediction [33], fluid dynamics simulation [34], and driving behavior simulation [37]. In this situation, researchers seek to generate synthetic trajectories by simulating human mobility that retains useful information to support downstream applications, which simultaneously ensures the user privacy [11, 20, 36].

Despite the above methods [6, 15, 33, 34, 37] attempted to generate trajectory data in specific tasks, synthesizing practical human trajectories is still an open problem with the following unsolved critical challenges:

- **User distribution in human trajectories is critical but hard to capture (*group view*).** Different users tend to have totally different ranges of activities (*e.g.* home or workplace), which means their trajectories are typically quite different. For example, colleagues in the same company have totally different trajectories because of their unique home locations. Hence, how to capture the user distribution characteristics in the group view and generate the trajectory data matching the user distribution is a critical but challenging task. However, existing studies [6, 15] tend to focus on modeling at the trajectory level and totally neglect the user distribution characteristics in the group view.
- **Mobility patterns in each user trajectory are complex (*individual view*).** As Figure 1(b) shows, there are multiple speed (*line slope*) types in the real-world datasets, which means the travel patterns (*e.g.* driving, biking, and walking) in each user trajectory are varied and complex. Hence, how to capture the mobility patterns and perform trajectory generation with multiple travel patterns is challenging. Hence, in the individual view, generating the trajectory data with multiple travel patterns for a specific user is critical but challenging. However, existing works [17] only perform trajectory generation with a single travel mode.

To address the above challenges, we propose a VAE-based trajectory generation model named **VOLUNTEER** (VariatiOnaL hUmaN Trajectories gEnERator), which performs practical synthetic human trajectories generation based on variational point processes. Specifically, VOLUNTEER proposes a two-layer VAE structure, consisting of a user VAE and a trajectory VAE. In the user VAE, we first capture user-level features with a transformer (based on all human trajectories) to learn the user distribution. Then, we can obtain user attributes (such as residence and workplace) for different users to obtain better trajectory generation. In this way, we solve the challenge of how to model the user distribution from the group view. In the trajectory VAE, we generate the trajectory data for specific users with complex mobility patterns from the individual view. The trajectory generation is essential to generate a sequence of spatio-temporal points (location and time). Hence, we need to learn the location distribution (for location generation) and time distribution (for time interval generation) in the trajectory VAE. For the location distribution, we combine the user attribute sampled from user VAE and the trajectory feature obtained from trajectory data to learn the distribution parameters. For the time distribution, we proposed to decouple time intervals into dwell time and travel time. The dwell time is generated using a probability distribution with explicit physical meaning, and the parameters of the probability distribution are fitted by a neural network. As for the travel time, we implicitly model the travel modes (*e.g* driving or walking) with different distributions and sample the travel time according to the travel mode-related distributions. In this way, we generate the trajectory data with complex mobility patterns and solve challenge two. Finally, our model tightly combines the classical temporal point process with deep neural networks based on a new variational inference framework, leveraging the strong interpretability of the classical temporal point process model to play the role of connecting uninterpretable neurons in neural networks with real-world human trajectories. Such a framework can serve to model trajectories of non-fixed length with a continuous temporal distribution. Overall, the contributions of our paper can be summarized as follows:

- We propose a two-layer VAE model, including a user-level VAE and a trajectory-level VAE, from both group and individual views. User VAE is concerned with modeling user distributions, thus providing users' information to assist trajectory VAE in completing the trajectory generation. Such a two-layer structure, especially the user VAE, is designed to effectively generate a large number of users with a small number of users, as well as generate a large number of trajectory patterns with a small number of trajectories patterns.
- We propose a unique temporal modeling module for trajectory generation. Our model not only utilizes the temporal point process as a bridge between deep neural networks based on a variational inference framework and real-world mobility behavior for the purpose of modeling non-fixed-length trajectories with continuous time distributions. It also decouples dwell time and travel time, especially the travel time module utilizes a self-supervised framework to implicitly model different travel modes, solving the dilemma that the dataset is not labeled by travel mode and accurately models the time in trajectory generation by accurately capturing travel time and dwell time.
- Our model has been extensively evaluated on two real-world mobility datasets with a mix of multiple travel modes. The results show that our model performs well on a number of important

trajectory statistics compared to several representative baselines. Further application analysis in the industrial system also demonstrates the effectiveness of our model.

This paper is structured as follows. We begin by presenting a preliminary study and overview of our problem. Then, we propose our synthetic human trajectories generation method based on variational point processes. Following our methodology, we present the evaluation and validation of our method. Finally, after discussing important related work, we conclude our paper.

2 PRELIMINARIES

2.1 Problem Definition

In this section, we first introduce the problem definition of trajectory generation, as well as related background knowledge.

Mobility Trajectory. The mobility trajectory of user u consists of a set of spatial-temporal points $S = \{s_1, ..., s_n\}$, where each spatial-temporal point s_i can be expressed as (l_i, t_i), t_i is the timestamp of i-th visit, and l_i denotes the location information, either in the form of latitude and longitude or in the form of region ID.

Mobility Trajectory Generation. Given a real-world mobility trajectory dataset, generating new trajectories that can retain the key characteristics and important utility of the original trajectory data.

2.2 Backgroud

In this section, we briefly introduce Variational Auto-Encoder and temporal point process model.

Variational Auto-Encoder (VAE). Variational auto-encoder is a wildly-adopted deep generative model. Specifically, VAE is composed of an encoder and a decoder, which are both neural networks. The decoder is utilized to model the generative process of the observable data x based on the latent vector z, *i.e.*, $p_\phi(x|z)$, where ϕ is the trainable parameters of the neural network. In addition, z follows a pre-defined prior distribution, which is normally set to be the multi-dimensional diagonal Gaussian distribution. The encoder is utilized to model the variational distribution to approximate the posterior distribution derived based on the generative process and observation x, which is denoted by $q_\psi(z|x)$ with parameter ψ. Then, VAE is optimized by minimizing the similarity between the variational distribution $q_\psi(z|x)$ and the posterior distribution derived based on the generative process and observation x, *i.e.*, $p(z|x) = \frac{p(x|z)p(z)}{p(x)}$. By using KL divergence as the matrix of the similarity, the optimization target can be represented by $KL(q(z|x)||\frac{p(x|z)p(z)}{p(x)})$. Though the evidence $p(x)$ is unknown, by transforming the above optimization target, we can obtain:

$$KL(q(z|x)||\frac{p(x|z)p(z)}{p(x)}) = \log p(x) - \int q(z|x)\log\frac{p(x|z)p(z)}{q(z|x)}dz, \tag{1}$$

Thus, minimizing the KL divergence can be converted to maximizing the right item in (1), which is referred to as the evidence lower bound (ELBO). ELBO can be further transformed into:

$$ELBO = \mathbb{E}_{q(z|x)}p(x|z) - KL(q(z|x)||p(z)), \tag{2}$$

where the first term represents the probability of generating the original sample x based on the latent vector obtained based on the encoder, *i.e.*, the reconstruction probability. The second term is the KL divergence between the variational distribution and the prior distribution, which can be regarded as a regularizer [38].

Compared with GAN, which trains an additional neural discriminator to distinguish real and synthetic samples, VAE is able to explicitly model the probability of the observable data and generate samples with high stability and diversity [25]. Therefore, VAE is often used to model time-series data [7, 8, 21].

Temporal Point Process. Temporal point processes (TPP) are a powerful tool to model the generative process of time series, which are defined as the stochastic process of time-tagged event sequences in the continuous time domain. In addition, each event can also be associated with a specific event category. Thus, we can represent a time series generated by TPP as $\{x_i\}_{i=1}^N$, where each point x_i is composed of an event type k_i and a timestamp t_i. Normally, the temporal distribution of TPP is modeled based on the intensity function $\lambda(\cdot) \leq 0$. Specifically, conditioned on the previous $n-1$ points, the happening time t_n of the next point x_n can be calculated as follows:

$$p(t_n|x_{1:n-1}) = \lambda(t_n|x_{1:n-1})e^{-\int_{t_{n-1}}^{t_n}\lambda(t|x_{1:n-1})dt}. \tag{3}$$

In addition, the event category k_n for the next point x_n usually follows a categorical distribution dependent on the previous $n-1$ points and t_n.

The core of VAE-based trajectory generation is modeling the probability distributions of movement behavior. Constructing these probability distributions using the temporal point process framework enables us to effectively incorporate expert knowledge and enhance the interpretability of the model.

3 VARIATIONAL HUMAN TRAJECTORIES GENERATOR (VOLUNTEER)

3.1 Overview

The model's framework is shown in Figure 2. Our model is composed of two components: a user VAE and a trajectory VAE. To address the challenge of capturing user distributions in human trajectories, we design the user VAE to model the distribution of users' residences and workplaces in order to capture user distribution properties. To solve the problem of how to model the complex mobility patterns of each user's trajectory, we build the trajectory VAE to generate trajectories with periodicity and regularity to match the user distribution.

3.2 User VAE for User Distribution

To better model the user distribution and obtain specific user attributes such as residence and workplace for different users, we develop a user VAE. There are two phases to the user VAE: inference and generation. The objective of the inference phase is to extract the latent variable g_i, which contains information on the user's whole trajectory sequence and personal characteristics, such as the user's residence and workplace. In particular, the latent variable g_i is learned by designing appropriate embedding and encoding methods [1] to extract information from the entire spatio-temporal

[1]The encoding details of user VAE and trajectory VAE are similar, please refer to the encoding methods of trajectory VAE.

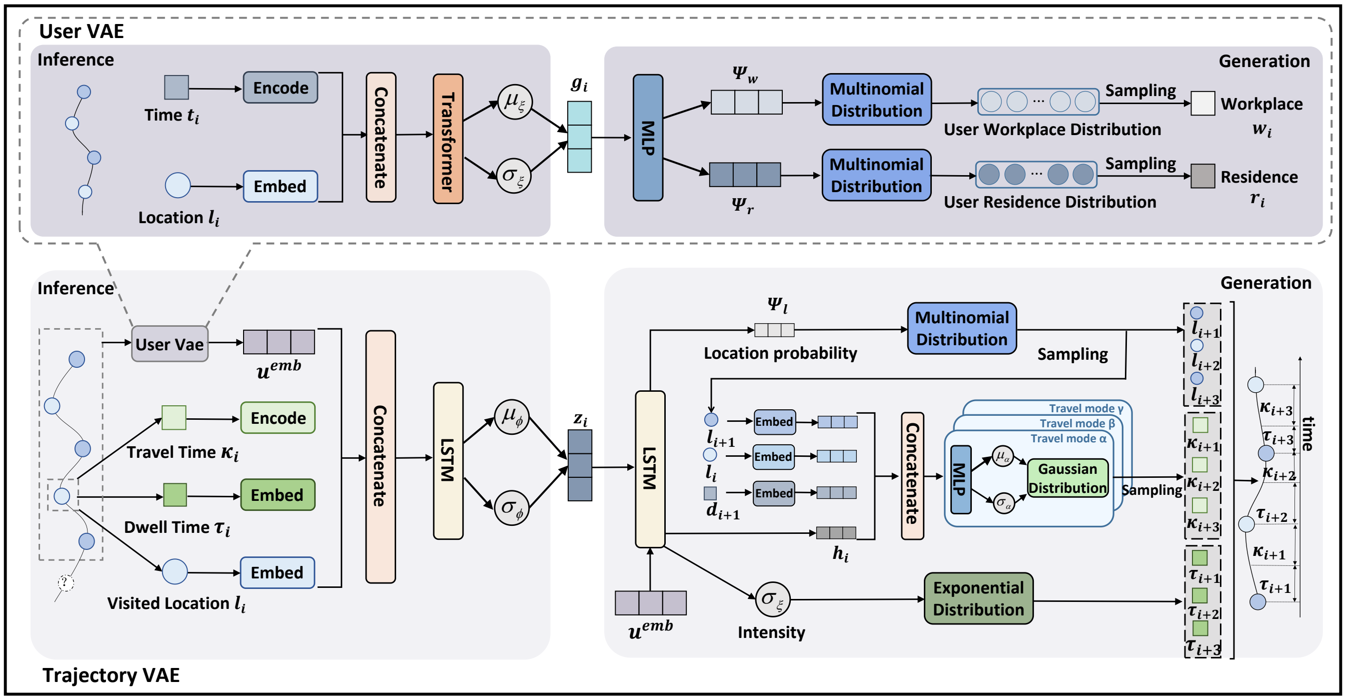


Figure 2: Framework of the proposed variational trajectory synthesizing model.

trajectory data. With the latent variable g_i as input, the generation phase aims to generate the user's residence and workplace.

3.2.1 Inference. The specific procedure of the inference process is to learn an approximate distribution $q_\xi(g_i|s_{1:i})$ with parameter ξ and estimate the posterior distribution of g_i directly from the observed movement records to obtain the probability distribution of the latent variable g_i containing the mobility characteristics.

Specifically, g_i is modeled to follow a Gaussian distribution with parameters μ_ξ and σ_ξ, which is fitted by a neural network. This process is represented as follows:

$$\begin{cases} [\mu_\xi; \sigma_\xi] = f_\xi(s_{1:i}), \\ q_\xi(g_i|s_{1:i}) = \mathcal{N}(\mu_\xi, \sigma_\xi^2), \end{cases} \tag{4}$$

where $s_{1:i}$ is the historical trajectory sequence of user u. To enhance the model capability, we utilize transformer [39] as the parameter generation function $f_\xi(\cdot)$ for the generation of user distribution parameters μ_ξ and σ_ξ.

3.2.2 Generation. We use a neural network to fit the parameters of a multinomial distribution and then generate the residence and workplace from the multinomial distribution. Inspired by the temporal point process (*e.g.*, Poisson process, Hawkes process), we model the probability distribution of residence r_i and workplace w_i as follows:

$$\begin{cases} p(r_i|g_{1:i}, s_{1:i}) = \Psi_r(z_{1:i}, s_{1:i}), \\ p(w_i|g_{1:i}, s_{1:i}) = \Psi_w(z_{1:i}, s_{1:i}). \end{cases} \tag{5}$$

where $\Psi_r(g_{1:i}, s_{1:i})$ indicates the probability of residence, and $\Psi_w(g_{1:i}, s_{1:i})$ indicates the probability of workplace.

Specifically, the latent variable g_i is mapped to the relevant parameters of the multinomial distribution as follows:

$$[\Psi_r, \Psi_w] = \mathrm{MLP}_\lambda(g_i), \tag{6}$$

where $g_i \in N(0, 1)$ is a random variable describing the user's mobility characteristics. Finally, the user's residence and workplace are selected from the multinomial distribution with the parameter Ψ_r and Ψ_w, respectively, which is able to model the randomness of the data generation process.

3.3 Trajectory VAE for Mobility Patterns

In order to capture the complex mobility patterns of each user trajectory and generate trajectories that match the actual mobility law, we propose the trajectory VAE. The trajectory VAE is divided into two phases: inference and generation. The purpose of the inference phase of trajectory VAE is to mine the user's decision propensity from his historical trajectory, which means his intention to choose which location to go to in which time period. Specifically, we can input historical trajectories, embed mobility records in spatial and temporal dimensions, and learn the latent variable z_i by extracting various information from the spatial and temporal mobility records through the encoder. The key information between inference and generation is encoded in the latent variable z_i, which represents the user's decision propensity. The generating step of the trajectory VAE is intended to generate new trajectories. One of the decoder's inputs is the latent variable z_i, which characterize mobility features of users and replicate the randomness of human mobility. In connection with the user's embedding and the encoding of previous mobility trajectories, the latent variable z_i models the evolution of the mobility, which is then mapped into parameters characterizing the probability distribution of subsequent mobility behaviors to build new trajectories.

3.3.1 Inference. Unlike traditional neural networks, our proposed temporal point process based model generates data through probability distributions, which are inherently non-derivable and therefore cannot be trained directly to obtain model parameters using the backpropagation algorithm. To solve this problem, we use the variational inference method to optimize the model. Based on the variational inference technique, we use an additional neural network to estimate the inverse process from the generative network, *i.e.*, to invert the latent features from the generative results, which we call the approximate distribution. The approximate distribution is used to derive a measure of good or bad generative results by combining Bayesian formulas, and ultimately to train the parameters of the entire variational temporal point process model. Specifically, we estimate the user's current latent variable z_i based on the historical mobility trajectory $s_{1:i}$, of which the corresponding neural network is denoted as $q_\phi(z_i|s_{1:i})$.

First, the dwell time τ_i, the visited location l_i, and the user embedding u^e obtained based on User VAE are used as features of the model. Then, the visited location l_i is embedded into representative vector l_i^e based on the embedding module. Next, we make some improvements to the classical positional encoding technique and propose a Fourier-based positional encoding mechanism with the goal of being able to capture the fine-grained periodic behavior of human mobility. The Fourier-based positional encoding mechanism is described as follows:

$$\begin{cases} \mathrm{PE}_{2i}(t) = \sin(2\pi it/\delta), \\ \mathrm{PE}_{2i+1}(t) = \cos(2\pi it/\delta), \end{cases} \tag{7}$$

where $\mathrm{PE}_i(t)$ represents the i-th element of the positional encoding of the absolute time or time difference t. δ is the fundamental frequency, which can be set to any desired periodic time to encode, either a day, a week or a month. In this way, the same times of different periods are encoded as similar vectors, reflecting the mobility periodicity. Based on the Fourier-based positional encoding mechanism, we encode the dwell time τ_i as τ_i^{emb}.

Then, the latent variable of user VAE g_i, the embedding of the output residence of user VAE r_i^{emb}, and the embedding of the output workplace of user VAE w_i^{emb} are concatenated to obtain the user embedding u^{emb}.

$$u^{emb} = [g_i; r_i^{emb}; w_i^{emb}]. \tag{8}$$

Finally, the encoding of dwell time τ_i^{emb}, the embedding of location information l_i^{emb}, and the embedding of user characteristics u^{emb} are concatenated to obtain the embedding of mobility record s_i denoted by s_i^{emb}, which is then fed into an LSTM as follows:

$$\begin{cases} s_i^{emb} = [\tau_i^{emb}; l_i^{emb}; u^{emb}], \\ h_i = \mathrm{LSTM}_\phi(h_{i-1}, s_i^{emb}). \end{cases} \tag{9}$$

Based on the output of this LSTM network, the final distribution of z_i is modeled as a Gaussian distribution of functions whose parameters are h_i. This process is represented as follows:

$$\begin{cases} [\mu_\phi; \sigma_\phi] = \mathrm{MLP}_\phi(h_i), \\ q_\phi(z_i|r_{1:i}) = \mathcal{N}(\mu_\phi, \sigma_\phi^2). \end{cases} \tag{10}$$

3.3.2 Generation. The generation process is simply to generate the next mobility record s_i based on the historical mobility record $s_{1:i-1}$ and the historical latent variable $z_{1:i}$. Specifically, we utilize neural networks to model the intensity function of user movement and fit the user selection of the different locations.

The next mobility record is mainly composed of visit location l_i, dwell time τ_i, and travel time T_i, and the detailed design of modeling the three parts are described below.

Dwell Time. Briefly, we mainly use the neural network LSTM to fit the parameters of the exponential distribution and then generate the dwell time from the exponential distribution. Drawing on the Poisson process, we model the probability distribution of the dwell time τ_i as follows:

$$p(\tau_i|z_{1:i}, s_{1:i-1}) = \lambda(z_{1:i}, s_{1:i-1}) \cdot e^{-\lambda(z_{1:i}, s_{1:i-1})\tau_i}, \tag{11}$$

where $\lambda(z_{1:i}, s_{1:i-1})$ denotes the intensity of the next movement and the probability distribution of τ_i is modeled by an exponential distribution.

Further, we use an LSTM network to model the correlation between the latent variable z and the embedding of historical mobility records, and the process is represented as follows:

$$\begin{cases} s_{i-1}^{emb} = [t_{i-1}^{emb}; l_{i-1}^{emb}; u^{emb}], \\ h_i = \mathrm{LSTM}_\theta(h_{i-1}, z_i, s_{i-1}^{emb}), \end{cases} \tag{12}$$

where h_i is the hidden state variable of the LSTM and $z_i \in \mathcal{N}(0,1)$ is a random variable describing the user's mobility state characteristics, which models the randomness of the data generation process. s_{i-1}^{emb} is composed of the embedding of location l_{i-1}^{emb}, dwell time τ_{i-1}^{emb}, and user characteristics u^{emb}, where u^{emb} denotes the attribute characteristics of the user that do not change with movement.

Further, the hidden states are mapped by a multi-layer perceptron (MLP) to the relevant parameters of the intensity function, expressed as follows:

$$\eta = \mathrm{MLP}_\theta(h_i), \tag{13}$$

where η is the value of the intensity function. Thus, we have:

$$\lambda(z_{1:i}, s_{1:i-1}) = \eta. \tag{14}$$

Then, the dwell time τ is drawn from (11).

Visited Location. To summarize, we choose neural network LSTM to fit the parameters of the multinomial distribution and then generate the visited location from the multinomial distribution.

Similar to generating the probability distribution of the dwell time, we model the probability distribution of the visited locations l_i as follows:

$$p(l_i|z_{1:i}, s_{1:i-1}) = \Psi_l(z_{1:i}, s_{1:i-1}), \tag{15}$$

where $\Psi_l(z_{1:i}, s_{1:i-1})$ indicates the visiting probability of each location l.

Specifically, the hidden state h_i obtained in (12) is mapped by an MLP to the relevant parameters of the multinomial distribution, expressed as follows,

$$\Psi_l = \mathrm{MLP}_\theta(h_i). \tag{16}$$

The final obtained vector $\Psi_l(z_{1:i}, s_{1:i-1})$ represents the propensity of users to visit these locations, and the next location is selected from the multinomial distribution with $\Psi_l(z_{1:i}, s_{1:i-1})$ as the parameter.

Travel Time. In Section 3.2 we conclude that there are three travel modes in the data by analyzing the density of travel distance and travel time in the dataset. Based on this prior knowledge, our travel time module is designed with three Gaussian distributions corresponding to the three travel modes.

To model the correlation between the travel time κ_i and the origin l_i, destination l_{i+1}, and distance d_i, we use a neural network to fit the parameters of each Gaussian distribution to generate the travel time. Specifically, the embeddings of origin location l_i, destination location l_{i+1}, and distance d_i as well as the hidden vector h_i of LSTM in (12) are connected as follows:

$$L_i^{emb} = [l_i^{emb}; l_{i+1}^{emb}; d_i^{emb}; h_i]. \tag{17}$$

The networks are then fed into different MLP networks as follows:

$$\begin{cases} [\mu_\alpha; \sigma_\alpha] = \mathrm{MLP}_\alpha(L_i^{emb}), \\ [\mu_\beta; \sigma_\beta] = \mathrm{MLP}_\beta(L_i^{emb}), \\ [\mu_\gamma; \sigma_\gamma] = \mathrm{MLP}_\gamma(L_i^{emb}), \\ [\alpha; \beta; \gamma] = \mathrm{MLP}_\xi(L_i^{emb}), \end{cases} \tag{18}$$

where the first three MLP networks are utilized to fit parameters corresponding to the travel time distribution of the three travel modes, and the last MLP network is used to fit the weights $[\alpha; \beta; \gamma]$ corresponding to the three travel modes.

Therefore, the distribution of travel time is given by the following equation:

$$p(\kappa_i | L_i^{emb}) = \alpha \cdot \mathcal{N}(\kappa_i | \mu_\alpha, \sigma_\alpha^2) + \beta \cdot \mathcal{N}(\kappa_i | \mu_\beta, \sigma_\beta^2) + \gamma \cdot \mathcal{N}(\kappa_i | \mu_\gamma, \sigma_\gamma^2), \tag{19}$$

At this point, all of our generation modules have been introduced. The generation module consists of three parts: dwell time, travel time, and visited location, and we iterate in a loop based on the generation network to generate the user's movement sequence.

3.3.3 Optimization. The optimization target $\mathcal{L}_u$ of User Vae is expressed as follows:

$$\begin{aligned} \mathcal{L}_u &= \sum_u \mathcal{L}_{\lambda,\xi}(s_{1:N_u}^u) \\ &= \sum_u \sum_{i=1}^{N_u} E_{q_\xi(z_i^u | s_{1:i}^u)} [\log p_\lambda(s_i^u | z_{1:i}^u)] \\ &\quad - \mathrm{KL}(q_\xi(z_i^u | s_{1:i}) || p_\lambda(z)). \end{aligned} \tag{20}$$

where N_u denotes the number of records for user u and s_i^u represents the i-th records for user u. z_i^u is the latent variable describing the decision tendency of user u corresponding to s_i^u.

The optimization target $\mathcal{L}_t$ of Trajectory Vae is expressed as follows:

$$\begin{aligned} \mathcal{L}_t &= \sum_u \mathcal{L}_{\theta,\phi}(s_{1:N_u}^u) \\ &= \sum_u \sum_{i=1}^{N_u} E_{q_\phi(z_i^u | s_{1:i}^u)} [\log p_\theta(s_i^u | z_{1:i}^u)] \\ &\quad - \mathrm{KL}(q_\phi(z_i^u | s_{1:i}) || p_\theta(z)). \end{aligned} \tag{21}$$

Given that joint optimizing two different VAE modules is challenging and may lead to worse performance, we first train the user VAE with $\mathcal{L}_u$, and then we fix the parameters of user VAE and further optimize the trajectory VAE with $\mathcal{L}_t$. Moreover, calculating the integration in (20) and (21) is intractable in practice. In order to solve this problem, following the common solution of the variational inference framework [18], we train the entire model using the technique of reparameterization trick.

4 OFFLINE EVALUATION OF TRAJECTORY GENERATION

4.1 Experimental Settings

4.1.1 Dataset. We conducted extensive experiments on two real-world mobility datasets, which are the MME and ISP datasets.

- **MME**: The MME dataset is provided by the China Mobile Research Institute. The mobility data in it contains a subscriber volume of 10,000, with a primary spatial extent of Nanchang and a time span of one week from May 18 to May 24, 2022. Each mobility record in the MME dataset contains an anonymous user id, timestamp, and cellular base station.
- **ISP**: The ISP dataset is collected through a partnership with a major Internet Service Provider (ISP) in China. The mobility data in it contains a volume of more than two million users, with a primary spatial scope of Shanghai and a time span from April 19 to April 26, 2016. Each mobility record in the ISP dataset contains an anonymous user id, timestamp, and cellular base station.

We preprocess the trajectory data of two datasets by filtering out users who have less than 10 records per day. The preprocessing for location involved mapping GPS points to pre-defined grid IDs with a granularity of 30m × 30m. For the preprocessing of time, we identified the dwell time and travel time using the Timegeo algorithm [17]. This algorithm identifies stationary points from GPS data points, allowing us to obtain the dwell time and travel time in the dataset.

4.1.2 Metrics. We evaluate the extent and effectiveness of the generated trajectories with the following five Metrics. *i.e.* **Distance**, **G-rank**, **Duration**, **Move**, and **Stay**. We introduce the details of metrics in Section A.1 of the Appendix.

4.1.3 Baselines. We compare the performance of our model with five state-of-the-art baselines, *i.e.* **TimeGEO** [17], **Semi-Markov** [27], **Hawkes** [4], **LSTM** [1], and **MoveSim** [12]. We introduce the details of baselines in Section A.2 of the Appendix.

4.1.4 Hyper-parameters Settings. As for the hyper-parameters settings of our model, we set the embedding size of time and location as 256 and set the user embedding size as 128. As for the encoder and decoder, we set the latent size of RNN as 512. For the hidden

Table 1: Performance comparisons on two mobility datasets, where bold denotes best (lowest) results and underline denotes the second best results.

	ISP					MME				
	Distance (JSD)	G-Rank (JSD)	Duration (JSD)	Move (JSD)	Stay (MSE×10^2)	Distance (JSD)	G-Rank (JSD)	Duration (JSD)	Move (JSD)	Stay (MSE×10^2)
Semi-Markov	0.016	**0.197**	0.026	0.062	0.028	0.012	0.213	0.018	0.053	0.030
TimeGEO	0.013	0.685	0.023	0.055	0.030	0.017	0.691	0.035	0.049	0.038
Hawkes	0.159	0.241	0.037	0.157	0.041	0.136	**0.189**	0.032	0.051	0.040
LSTM	0.125	0.269	0.018	0.059	0.031	0.011	0.245	0.013	0.048	0.032
MoveSim	0.028	0.238	0.312	0.121	0.056	0.042	0.314	0.298	0.107	0.031
VOLUNTEER	**0.010**	0.221	**0.012**	**0.048**	**0.024**	**0.008**	0.217	**0.009**	**0.041**	**0.023**
Improv.	**23.1%**	-	**33.3%**	**12.7%**	**14.3%**	**27.2%**	-	**30.7%**	**16.3%**	**23.3%**

size of z, we set both the mean and variance sizes as 256. We set dropout as 0.2, learning rate as 1e-5, and batch size as 8.

4.2 Experimental performance

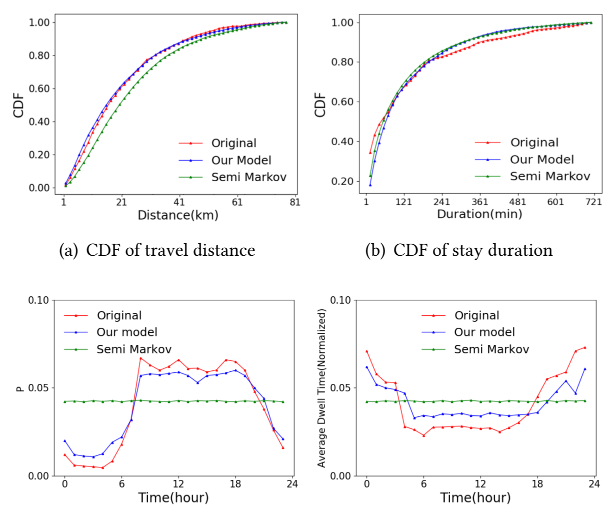
4.2.1 Overall Performance. The performance of our models on the ISP and MME datasets is shown in Table 1. Specifically, all our experiments on the MME dataset are performed on the Jiutian Artificial Intelligence Platform. We can observe that our model achieves the best performance in most of the usability metrics. Each of these baselines has its own advantages. LSTM is able to restore the Duration property well because it predicts time accurately enough, while Movesim does not perform well in the metric of Duration because Movesim requires a discrete-time input, while we finally evaluate the performance of Duration when compared to a continuous time distribution. Semi-Markov performs best on the metric of Stay because this model is designed specifically to characterize the distribution of dwell time, which utilizes prior knowledge to model the strength of dwell time for Bayesian inference. The reason why TimeGeo performs well in the metric of Move is that the design mechanism of the model specifically considers the possibility of exploration and explicitly simulates the circadian rhythm of human mobility. The above describes the best-performing model in several baselines, and our model is far superior to the best baseline in most metrics.

The average performance difference between our proposed algorithm and each baseline is around 20%, with our improvement in the metric of Duration being particularly significant. The significant improvement in Duration also demonstrates the effectiveness of our model in the well-designed time module. The decoupling of travel time and dwell time, as well as the ability to generate continuous time distributions based on the variational temporal point process, have helped us to model time efficiently. Overall, our proposed model outperforms existing algorithms in most cases, which demonstrates the superiority of our approach in synthesizing human trajectories.

4.2.2 Ablation experiments. To assess the importance of the different components in our model, we perform an ablation study on two mobility datasets, specifically by removing each component from the full model. Without loss of generality, we only show the results on the MME dataset. No-User means removing the user VAE module and No-Travel means removing the travel time module. We can observe from Table 2 that the performance of the knowledge

Table 2: Results of the ablation study in terms of different metrics. Bold denotes the best (lowest) results.

	Distance (JSD)	G-Rank (JSD)	Duration (JSD)	Move (JSD)	Stay (MSE×10^2)
No-User	0.009	0.246	0.011	0.054	0.030
No-Travel	0.010	0.223	0.014	0.065	0.035
Our Model	**0.008**	**0.217**	**0.009**	**0.041**	**0.023**



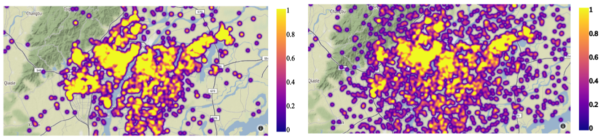
(a) CDF of travel distance (b) CDF of stay duration

(c) Move (d) Stay

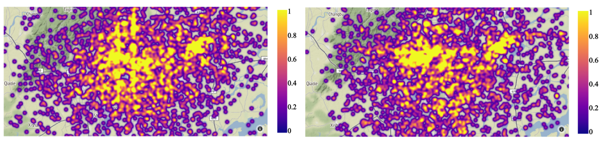
Figure 3: Performance visualization on ISP Dataset.

system becomes worse when an arbitrary module is removed. In addition, we also find that the travel time module plays a key role in improving the statistical metrics of time.

4.2.3 Visualization Analysis. In Figure 3, we have chosen the probability density function (PDF), the cumulative distribution function (CDF), and the time curves for different metrics to visualize the availability of our generated trajectories. From Figure 3(a) and Figure 3(b), we can observe that the CDF curves of our model and the real trajectory for two metrics are closer. From Figure 3(c) and Figure 3(d), we can observe that the trajectories generated by our model have higher travel frequency during the daytime and longer

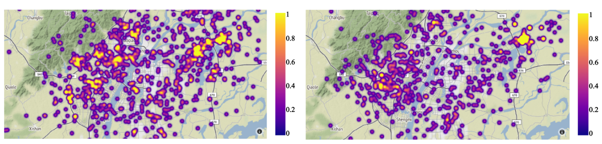


(a) Heatmap of real residence (b) Heatmap of generated residence based on our model

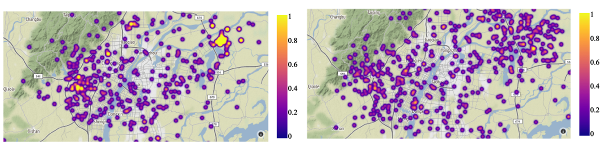


(c) Heatmap of generated residence based on Semi-Markov (d) Heatmap of generated residence based on MoveSim

Figure 4: Heatmap of the residence on MME dataset.

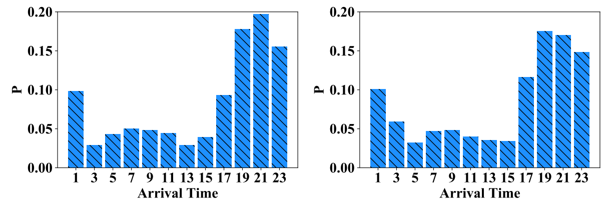


(a) Heatmap of real workplace (b) Heatmap of generated workplace based on our model

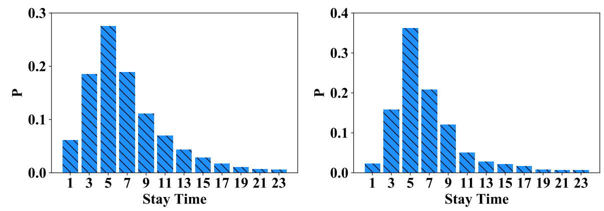


(c) Heatmap of generated workplace based on Semi-Markov (d) Heatmap of generated workplace based on MoveSim

Figure 5: Heatmap of the workplace on MME dataset.

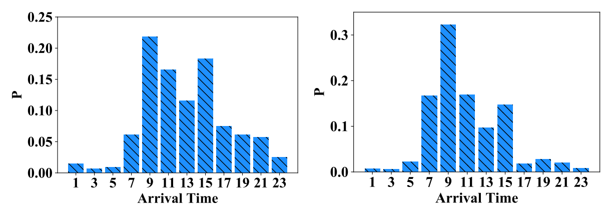


(a) Arrival time of real trajectories (b) Arrival time of generated trajectories

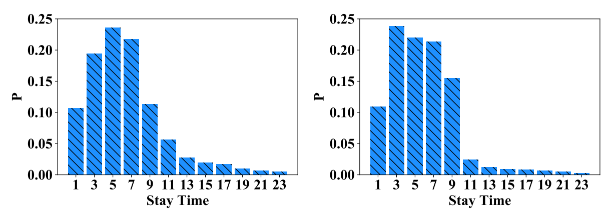


(c) Stay time of real trajectories (d) Stay time of generated trajectories

Figure 6: Distribution of arrival and stay time at the residence.



(a) Arrival time of real trajectories (b) Arrival time of generated trajectories



(c) Stay time of real trajectories (d) Stay time of generated trajectories

Figure 7: Distribution of arrival and stay time at the workplace.

average dwell time at night, which are in line with people's daily mobility patterns.

Figures 4 and 5 show the results of the generated visualization of residence and workplace. In Figure 4, by comparing the differences in the heatmap between the residence generated by our model, Semi-Markov, Movsim, and the real residence, it can be observed that the distribution of the residence generated by our model is closest to the real ones. In particular, the residence generated by our model satisfies the feature of less distribution of residence in mountainous areas and the distribution of residence along rivers on both sides, which is very close to the distribution in the real world. Similarly, the heatmap of the working place also reflects the same situation. This proves the effectiveness of our User VAE for modeling residences and workplaces.

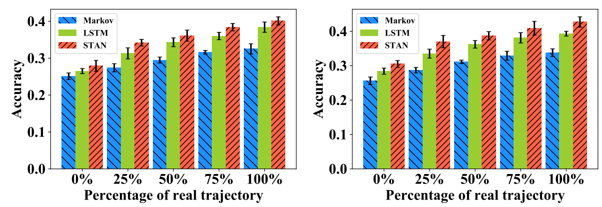
In Figures 6 and 7, we analyze the distribution of arrival and stay time at the residence and workplace. It can find that the time distribution of the generated trajectories is close to that of the real trajectories, and both have the pattern of staying at the residence for a long time in the morning and in the evening while staying at the workplace during the day. This proves that our model captures the regularity of human mobility and circadian rhythms well.

5 APPLICATION EVALUATION WITH INDUSTRIAL SYSTEM

In this section, we introduce the deployment and practical applications of our model on the industrial system, Jiutian Artificial Intelligence (AI) Platform.

5.1 System deployment

China Mobile is the largest wireless carrier in China, with more than 950 million subscribers. Jiutian Artificial Intelligence (AI) Platform is China Mobile's self-developed AI innovation platform, providing



(a) City A (b) City B

Figure 8: Mobility prediction based on generated trajectories from two cities of China.

open AI services from infrastructure to core capabilities. Jiutian Artificial Intelligence platform provides high-performance computing power. It supports over a hundred AI capability services such as vision, speech, natural language processing, and network intelligence, etc., which can meet the innovation needs of AI applications in various fields. Our model has been deployed on the Jiutian Artificial Intelligence platform to support network optimization and digital twin applications, which realize the generation models for practical downstream applications.

5.2 Applications of mobility prediction

We deployed our model on the Jiutian Artificial Intelligence platform and further conduct the application of mobility prediction to verify the effectiveness of our model. Specifically, we first collect trajectory data from two cities (*e.g.* City A and City B) in China via the Jiutian Artificial Intelligence platform. Then, we train the model with the collected data and generate trajectory data. Next, we train the prediction model of the Jiutian platform with both the generated data and real-world data. Finally, we evaluate the performance of mobility prediction on real trajectories. As we can observe from Figure 8, our model clearly outperforms the other two baselines for the online prediction task.

6 RELATED WORK

Mobility trajectory generation. Existing mobility trajectory generation methods can be divided into two categories. The first category of methods generates a new trajectory based on an existing trajectory [6, 50]. That is, each generated trajectory corresponds to a real-world human trajectory, which is regarded as the seed trajectory. For example, Bindschaedler et al. [6] first project a real-world trajectory into the semantic domain by transforming it to a semantic trace, where each record represents a semantic class containing a set of locations. Then, synthetic trajectories are generated by replacing the semantic class of each record with the location sampled from its location set. The other category of methods directly generated trajectories from sampled noise without seed trajectories [12, 36, 48, 49]. For example, He et al. [13] mainly break down the trajectory generation problem in a new city into three steps: 1) transfer of travel intention, 2) origin-destination (OD) generation, and 3) route generation. Jiang et al. [16] transfers urban human mobility across cities via POI embedding. However, although these methods can learn to generate more trajectories from a limited number of trajectories as the training set, they cannot control user-level attributes, such as users' homes and workplaces. Different from them, our proposed method utilizes a two-layer VAE structure to model the distribution of user-level attributes and the distribution of mobility patterns of a specific user simultaneously.

Trajectory-based Applications. There are many important applications that are based on human mobility data. Wu et al. [43] use user trajectories obtained from geotagged content posted on social networks as a way to understand traffic dynamics. Xie et al. [45] predict an individual's next location and recommend appropriate points of interest based on historical trajectories. Wang et al. [41] recommend spatial items by modeling and fusing sequential influences, cyclical patterns, and personal interests. Yuan et al. [46] design a context-aware location recommendation system that can consider user, spatial, temporal, and activity aspects simultaneously. These numerous trajectory-based applications show us the powerful potential of trajectory data, further demonstrating the need to develop trajectory synthesizing models to generate higher-quality trajectories.

Neural Temporal Point Process. In recent years, point-in-time processes have been combined with new deep neural network techniques [9, 29, 44, 47] to show outstanding performance in predicting time series data. Du et al. [9] propose the recursively labeled time point process (RMTPP), which uses a recurrent neural network (RNN) to model the intensity function of the time point process. Mehrasa et al. [28] propose Action Point Process VAE (APP-VAE), which uses VAE combined with RNN to model action sequences in videos. Pan et al. [31] combine VAE with point-in-time processes to model sequence data. Instead, we focus on trajectory data, which is different from general time-series data and has unique characteristics, such as periodicity, regularity, or spatiotemporal correlation.

7 CONCLUSION

In this paper, we propose VOLUNTEER, which is a two-layer VAE model that accurately captures the mobility characteristics of users by using a variational temporal point process framework. VOLUNTEER can decouple dwell time and travel time, excels in modeling more complex mobility patterns in trajectories, and is also generalizable to datasets that mix multiple modes of travel. Extensive experiments on real datasets show that the trajectories generated by our model retain the statistical characteristics and usability of real trajectories.

In the future, we intend to extend our framework to incorporate semantic information such as functions of location or points of interest visited by users in order to better understand the underlying motivation of users' movement and achieve semantic-aware trajectory generation.

ACKNOWLEDGMENT

This work is supported in part by the National Key Research and Development Program of China under 2022ZD0116402, the National Natural Science Foundation of China under 62171260, U22B2057, and U20B2060, the Young Elite Scientists Sponsorship Program by CIC (Grant No.2021QNRC001), and the joint project of China Mobile Research Institute and Tsinghua University. We thank China Mobile Research Institute for technical support, including data information, and the Jiutian Artificial Intelligence Platform for experiments.

REFERENCES

[1] Yazan Abu Farha, Alexander Richard, and Juergen Gall. 2018. When will you do what?-anticipating temporal occurrences of activities. In *Proc. CVPR.*

[2] Gergely Acs, Gergely Biczók, and Claude Castelluccia. 2018. Privacy-preserving release of spatio-temporal density. In *Handbook of Mobile Data Privacy.* 307–335.

[3] Miguel E Andrés, Nicolás E Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2013. Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security.* 901–914.

[4] Emmanuel Bacry, Thibault Jaisson, and Jean–Franois Muzy. 2016. Estimation of slowly decreasing Hawkes kernels: application to high-frequency order book dynamics. *Quantitative Finance* (2016), 1–23.

[5] Duygu Balcan and Alessandro Vespignani. 2011. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature physics* 7, 7 (2011), 581–586.

[6] Vincent Bindschaedler and Reza Shokri. 2016. Synthesizing plausible privacy-preserving location traces. In *Proc. IEEE SP.*

[7] Xinyu Chen, Jiajie Xu, Rui Zhou, Wei Chen, Junhua Fang, and Chengfei Liu. 2021. TrajVAE: A Variational AutoEncoder model for trajectory generation. *Neurocomputing* 428 (2021), 332–339.

[8] Abhyuday Desai, Cynthia Freeman, Zuhui Wang, and Ian Beaver. 2021. Timevae: A variational auto-encoder for multivariate time series generation. *arXiv preprint arXiv:2111.08095* (2021).

[9] Nan Du, Hanjun Dai, Rakshit Trivedi, Utkarsh Upadhyay, Manuel Gomez-Rodriguez, and Le Song. 2016. Recurrent marked temporal point processes: Embedding event history to vector. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* 1555–1564.

[10] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security.* 1054–1067.

[11] Jie Feng, Zeyu Yang, Fengli Xu, Haisu Yu, Mudan Wang, and Yong Li. 2020. Learning to Simulate Human Mobility. *Proc. KDD* (2020).

[12] Jie Feng, Zeyu Yang, Fengli Xu, Haisu Yu, Mudan Wang, and Yong Li. 2020. Learning to Simulate Human Mobility. In *Proc. KDD.*

[13] Tianfu He, Jie Bao, Ruiyuan Li, Sijie Ruan, Yanhua Li, Li Song, Hui He, and Yu Zheng. 2020. What is the human mobility in a new city: Transfer mobility knowledge across cities. In *Proceedings of The Web Conference 2020.* 1355–1365.

[14] Hsun-Ping Hsieh, Shou-De Lin, and Yu Zheng. 2015. Inferring air quality for station location recommendation based on urban big data. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining.* 437–446.

[15] Dou Huang, Xuan Song, Zipei Fan, Renhe Jiang, and Yugo Kato. 2019. A Variational Autoencoder Based Generative Model of Urban Human Mobility. In *Proc. MIPR.*

[16] Renhe Jiang, Xuan Song, Zipei Fan, Tianqi Xia, Zhaonan Wang, Quanjun Chen, Zekun Cai, and Ryosuke Shibasaki. 2021. Transfer urban human mobility via POI embedding over multiple cities. *ACM Transactions on Data Science* 2, 1 (2021), 1–26.

[17] Shan Jiang, Yingxiang Yang, Siddharth Gupta, Daniele Veneziano, Shounak Athavale, and Marta C González. 2016. The TimeGeo modeling framework for urban mobility without travel surveys. *PNAS* 113, 37 (2016).

[18] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

[19] Xiangjie Kong, Zhenzhen Xu, Guojiang Shen, Jinzhong Wang, Qiuyuan Yang, and Benshi Zhang. 2016. Urban traffic congestion estimation and prediction based on floating car trajectory data. *Future Generation Computer Systems* 61 (2016), 97–107.

[20] Vaibhav Kulkarni, Natasa Tagasovska, Thibault Vatter, and Benoit Garbinato. 2018. Generative models for simulating mobility trajectories. *arXiv preprint arXiv:1811.12801* (2018).

[21] Mihee Lee, Samuel S Sohn, Seonghyeon Moon, Sejong Yoon, Mubbasir Kapadia, and Vladimir Pavlovic. 2022. Muse-VAE: multi-scale VAE for environment-aware long term trajectory prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 2221–2230.

[22] Yinfeng Li, Chen Gao, Xiaoyi Du, Huazhou Wei, Hengliang Luo, Depeng Jin, and Yong Li. 2022. Automatically Discovering User Consumption Intents in Meituan. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.* 3259–3269.

[23] Yinfeng Li, Chen Gao, Xiaoyi Du, Huazhou Wei, Hengliang Luo, Depeng Jin, and Yong Li. 2022. Spatiotemporal-aware Session-based Recommendation with Graph Neural Networks. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management.* 1209–1218.

[24] Yinfeng Li, Chen Gao, Quanming Yao, Tong Li, Depeng Jin, and Yong Li. 2022. DisenHCN: Disentangled Hypergraph Convolutional Networks for Spatiotemporal Activity Prediction. *arXiv preprint arXiv:2208.06794* (2022).

[25] Kaikai Liu, Renjun Shuai, Li Ma, et al. 2021. Cells image generation method based on VAE-SGAN. *Procedia Computer Science* 183 (2021), 589–595.

[26] Yingtao Luo, Qiang Liu, and Zhaocheng Liu. 2021. Stan: Spatio-temporal attention network for next location recommendation. In *Proceedings of the Web Conference 2021.* 2177–2185.

[27] Leandros A Maglaras and Dimitrios Katsaros. 2016. Social Clustering of Vehicles Based on Semi-Markov Processes. *TVT* 65, 1 (2016), 318–332.

[28] Nazanin Mehrasa, Akash Abdu Jyothi, Thibaut Durand, Jiawei He, Leonid Sigal, and Greg Mori. 2019. A variational auto-encoder model for stochastic point processes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 3165–3174.

[29] Hongyuan Mei and Jason M Eisner. 2017. The neural hawkes process: A neurally self-modulating multivariate point process. *Advances in neural information processing systems* 30 (2017).

[30] Apollinaire Nadembega, Abdelhakim Hafid, and Tarik Taleb. 2014. Mobility-prediction-aware bandwidth reservation scheme for mobile networks. *IEEE Transactions on Vehicular Technology* 64, 6 (2014), 2561–2576.

[31] Zhen Pan, Zhenya Huang, Defu Lian, and Enhong Chen. 2020. A variational point process model for social event sequences. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 173–180.

[32] Dimitrios Rafailidis and Fabio Crestani. 2018. Friend recommendation in location-based social networks via deep pairwise learning. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).* 421–428.

[33] Suman Ravuri, Karel Lenc, Matthew Willson, Dmitry Kangin, Remi Lam, Piotr Mirowski, Megan Fitzsimons, Maria Athanassiadou, Sheleem Kashem, Sam Madge, et al. 2021. Skilful precipitation nowcasting using deep generative models of radar. *Nature* 597, 7878 (2021), 672–677.

[34] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. 2020. Learning to simulate complex physics with graph networks. In *International Conference on Machine Learning.* 8459–8468.

[35] Chaoming Song, Tal Koren, Pu Wang, and Albert-László Barabási. 2010. Modelling the scaling properties of human mobility. *Nature Physics* 6, 10 (2010), 818–823.

[36] Ha Yoon Song, Moo Sang Baek, and Minsuk Sung. 2019. Generating Human Mobility Route Based on Generative Adversarial Network. In *Proc. FedCSIS.*

[37] Simon Suo, Sebastian Regalado, Sergio Casas, and Raquel Urtasun. 2021. Trafficsim: Learning to simulate realistic multi-agent behaviors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 10400–10409.

[38] Jakub M Tomczak. 2022. Why Deep Generative Modeling? In *Deep Generative Modeling.* 1–12.

[39] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proc. NeurISP.*

[40] Jinzhong Wang, Xiangjie Kong, Azizur Rahim, Feng Xia, Amr Tolba, and Zafer Al-Makhadmeh. 2017. IS2Fun: Identification of subway station functions using massive urban data. *IEEE Access* 5 (2017), 27103–27113.

[41] Weiqing Wang, Hongzhi Yin, Xingzhong Du, Quoc Viet Hung Nguyen, and Xiaofang Zhou. 2018. TPM: A Temporal Personalized Model for Spatial Item Recommendation. *ACM Trans. Intell. Syst. Technol.* 9, 6 (2018).

[42] Xiaofei Wang, Min Chen, Tarik Taleb, Adlen Ksentini, and Victor CM Leung. 2014. Cache in the air: Exploiting content caching and delivery techniques for 5G systems. *IEEE Communications Magazine* 52, 2 (2014), 131–139.

[43] Fei Wu, Hongjian Wang, and Zhenhui Li. 2016. Interpreting traffic dynamics using ubiquitous urban data. *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (2016), 69.

[44] Weichang Wu, Junchi Yan, Xiaokang Yang, and Hongyuan Zha. 2018. Decoupled learning for factorial marked temporal point processes. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.* 2516–2525.

[45] Dongbo Xi, Fuzhen Zhuang, Yanchi Liu, Jingjing Gu, Hui Xiong, and Qing He. 2019. Modelling of bi-directional spatio-temporal dependence and users' dynamic preferences for missing poi check-in identification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5458–5465.

[46] Quan Yuan, Gao Cong, Kaiqi Zhao, Zongyang Ma, and Aixin Sun. 2015. Who, where, when, and what: A nonparametric bayesian approach to context-aware recommendation and search for twitter users. *ACM Transactions on Information Systems (TOIS)* 33, 1 (2015).

[47] Yuan Yuan, Jingtao Ding, Chenyang Shao, Depeng Jin, and Yong Li. 2023. Spatio-temporal Diffusion Point Processes. *arXiv preprint arXiv:2305.12403* (2023).

[48] Yuan Yuan, Jingtao Ding, Huandong Wang, Depeng Jin, and Yong Li. 2022. Activity Trajectory Generation via Modeling Spatiotemporal Dynamics. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.* 4752–4762.

[49] Yuan Yuan, Huandong Wang, Jingtao Ding, Depeng Jin, and Yong Li. 2023. Learning to Simulate Daily Activities via Modeling Dynamic Human Needs. In *Proceedings of the ACM Web Conference 2023.* 906–916.

[50] Paul A Zandbergen. 2014. Ensuring confidentiality of geocoded health data: assessing geographic masking strategies for individual-level data. *Advances in medicine* 2014 (2014).

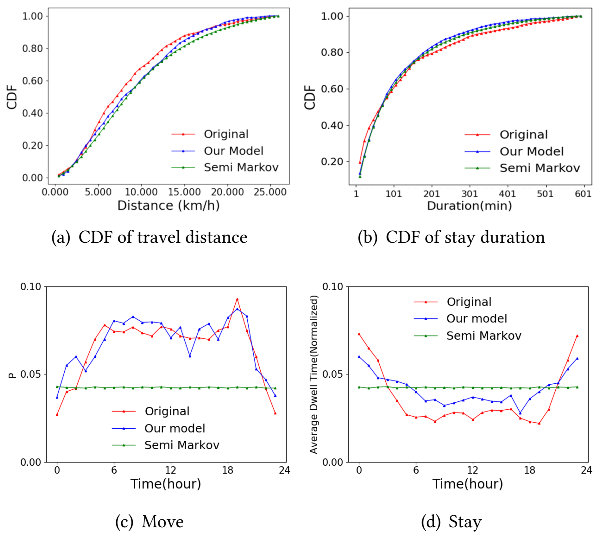


Figure 9: Performance visualization on MME Dataset.

A APPENDIX FOR REPRODUCIBILITY

A.1 Metrics

We need to evaluate the extent and effectiveness of the generated trajectories in retaining the statistical characteristics of real trajectories. The five evaluation metrics that we have chosen take into account the measurement of temporal statistical characteristics as well as spatial statistical characteristics and are described in detail as follows:

- **Distance**: This is a metric of spatial statistical characteristics to measure the distance between adjacent mobility records in a trajectory.
- **G-rank**: This is a metric of spatial statistical characteristics to measure the top visited frequency to different locations with respect to all users.
- **Duration**: This is a metric of temporal statistical characteristics to measure the time spent by users in different locations.
- **Move**: This is a metric of temporal statistical characteristics to measure the visiting time of each mobility record.
- **Stay**: This is a metric of temporal statistical characteristics to measure the correlation between average dwell time and visiting time.

Specifically, the metrics of Distance, Location, G-rank, Duration, and Move are expressed in terms of probability distributions. To visually measure the similarity between the generated sequences and the true sequences, we use the Jensen-Shannon scatter (JSD) to measure their differences. Specifically, for two distribution $\boldsymbol{p}$ and $\boldsymbol{q}$, the JSD between them can defined as:

$$\mathrm{JSD}(\boldsymbol{p},\boldsymbol{q}) = \frac{1}{2}\mathrm{KL}(\boldsymbol{p}||\frac{\boldsymbol{p}+\boldsymbol{q}}{2}) + \frac{1}{2}\mathrm{KL}(\boldsymbol{q}||\frac{\boldsymbol{p}+\boldsymbol{q}}{2}), \tag{22}$$

where $\mathrm{KL}(\cdot||\cdot)$ is the Kullback-Leibler divergence. On the other hand, the metric of Stay is not a probability distribution function but comes from the average dwell time as a function of access time in time units of minutes. Therefore, we use the mean squared error (MSE) to measure their difference.

A.2 Baselines

We compare the performance of our model with five state-of-the-art baselines, *i.e.* **TimeGEO** [17], **Semi-Markov** [27], **Hawkes** [4], **LSTM** [1], and **MoveSim** [12].

- **TimeGEO** [17]: As a model-based trajectory synthesizing method, TimeGEO defines dwell rate, burst rate, and the weekly home-based tour number to model the temporal choices and utilizes the explore and preferential return (EPR) model [35] to model the spatial choices.
- **Semi-Markov** [27]: In the Semi-Markov process, the dwell time is modeled by the exponential distribution. Dirichlet prior and gamma prior is used to model the transition matrix and the intensity of the dwell time to implement a Bayesian inference.
- **Hawkes** [4]: Hawkes process is a widely used classical temporal point process, where an occurred data point will influence the intensity function of future points.
- **LSTM** [1]: This is a model that treats the predicted results as a generated trajectory, specifically using an LSTM network to predict the next location and time.
- **MoveSim** [12]: This is a generative adversarial framework that incorporates the domain knowledge of human mobility regularities.

A.3 Additional Experiments

In Figure 9, for the MME dataset, we also choose the probability density function (PDF), the cumulative distribution function (CDF), and the time curves for different metrics to visualize the availability of our generated trajectories.