PateGail: A Privacy-Preserving Mobility Trajectory Generator with Imitation Learning

Huandong Wang[1], Changzheng Gao[1], Yuchen Wu[2], Depeng Jin[1], Lina Yao[3], Yong Li[1]
[1]Beijing National Research Center for Information Science and Technology (BNRist),
Department of Electronic Engineering, Tsinghua University, China
[2]Carnegie Mellon University, USA
[3]CSIRO's Data61 and University of New South Wales, USA
{wanghuandong,liyong07}@tsinghua.edu.cn

Abstract

Generating human mobility trajectories is of great importance to solve the lack of large-scale trajectory data in numerous applications, which is caused by privacy concerns. However, existing mobility trajectory generation methods still require real-world human trajectories centrally collected as the training data, where there exists an inescapable risk of privacy leakage. To overcome this limitation, in this paper, we propose PateGail, a privacy-preserving imitation learning model to generate mobility trajectories, which utilizes the powerful generative adversary imitation learning model to simulate the decision-making process of humans. Further, in order to protect user privacy, we train this model collectively based on decentralized mobility data stored in user devices, where personal discriminators are trained locally to distinguish and reward the real and generated human trajectories. In the training process, only the generated trajectories and their rewards obtained based on personal discriminators are shared between the server and devices, whose privacy is further preserved by our proposed perturbation mechanisms with theoretical proof to satisfy differential privacy. Further, to better model the human decision-making process, we propose a novel aggregation mechanism of the rewards obtained from personal discriminators. We theoretically prove that under the reward obtained based on the aggregation mechanism, our proposed model maximizes the lower bound of the discounted total rewards of users. Extensive experiments show that the trajectories generated by our model are able to resemble real-world trajectories in terms of five key statistical metrics, outperforming state-of-the-art algorithms by over 48.03%. Furthermore, we demonstrate that the synthetic trajectories are able to efficiently support practical applications, including mobility prediction and location recommendation.

Introduction

Human mobility trajectory data is instrumental for a large number of applications. For example, for the mobile Internet service providers, based on mobility trajectories, the movement and communication process of mobile users can be simulated to implement a reliable and efficient performance evaluation of the mobile networks (Hess et al. 2016). For the government, mobility trajectories can characterize the travel demand of the population and the traffic condition of the city, and thus provide important guidance to the transportation system planning (Feng, Bai, and Xu 2019). However, the utilization of real-world mobility trajectories leads to a growing privacy concern, since sensitive information of users can be leaked from their trajectories, e.g., which places they have visited and who they have met. Thus, it is hard to obtain a large-scale human mobility trajectory dataset to support numerous downstream applications. Under these circumstances, simulating human mobility behavior to produce realistic and high-quality mobility trajectory data becomes an important task for downstream applications, and has drawn much attention from both academia and industry.

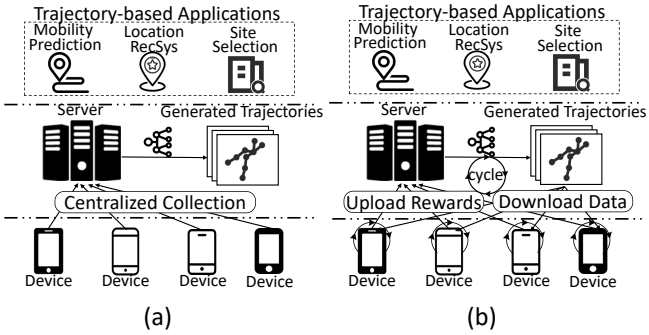


Figure 1: Illustration of (a) existing trajectory generator and our proposed (b) federated trajectory generator.

Numerous existing approaches have been proposed to generate mobility trajectories by utilizing powerful deep learning techniques including variational autoencoder (VAE) (Huang et al. 2019), and generative adversarial network (GAN) (Feng et al. 2018; Ouyang et al. 2018; Kulkarni et al. 2018; Liu, Chen, and Andris 2018), etc. However, as shown in Figure 1(a), these methods still require a number of real-world human trajectories centrally collected as the training data, where there exists the risk of privacy leakage. The rising paradigm of federated learning has provided a promising solution to this problem, which is a distributed machine learning framework with the goal of training machine learning models based on data distributed across multiple devices and protecting users' privacy at the same time. Federated learning has shown success in a num-

ber of practical applications, including personalized recommendation (Chen et al. 2018), keyboard prediction (Hard et al. 2018), etc. Thus, we seek to train the mobility trajectory generator in the manner of federated learning. As shown in Figure 1(b), in the federated mobility trajectory generation system, each user device keeps the private mobility trajectory data belonging to its owner (user). Only aggregated intermediate results proceeded by privacy protection mechanisms are shared between devices, while this system does not take any piece of the user trajectory data away from the device. In this way, we can train the mobility trajectory generator without privacy leakage.

However, training an efficient trajectory generator based on federated learning is not an easy task with the following challenges. First, mobility trajectories are with high dimensions and complicated interactions with both spatial venues and timestamps. How to develop a trajectory generator that accurately models human mobility behavior is the first challenge. Second, users' privacy is still possible to be leaked from the transmitted intermediate results in the training process, while most existing solutions do not provide privacy-preserving guarantees of the training process (Feng et al. 2020; Ouyang et al. 2018). How to provide the privacy-preserving guarantees of the training process is the second challenge.

To overcome these challenges, in this paper, we propose PateGail, a privacy-preserving imitation learning based mobility trajectory generator. Specifically, this model utilizes the powerful technique of generative adversary imitation learning (GAIL) to extract the hidden human movement decision process correlated with both spatial venues and timestamps, and thus is able to produce plausible mobility trajectories with preserved utility. Further, in order to provide privacy-preserving guarantees, we locally train a separate personal discriminator on each user device to distinguish and reward the generated and real-world decision-making sequences of human mobility, and then only share the generated trajectories and the rewards of the generated trajectories obtained based on personal discriminators between the devices and the server in the training process. Further, we propose a perturbation mechanism to prevent privacy leakage from the rewards of personal discriminators, which is theoretically proven to satisfy the differential privacy criterion. Finally, we propose a novel aggregation mechanism based on the mean and variance of reward obtained from different personal discriminators, which is able to model the dynamics of reward function across users. Furthermore, we theoretically prove that under the reward obtained based on our proposed aggregation mechanism, our proposed model maximizes the lower bound of the discounted total rewards of users. Our contributions can be summarized as follows:

- We propose a powerful mobile trajectory generator based on GAIL and federated learning, which is able to extract the hidden human decision process to generate plausible mobile trajectories and preserve user privacy with differential privacy guarantees at the same time.
- We propose a novel reward aggregation mechanism of reward obtained from personal discriminators of different users, which is able to model the dynamics of reward function across users. Furthermore, we theoretically prove that under our proposed reward aggregation mechanism, the obtained model maximizes the lower bound of the discounted total rewards of users.
- Extensive experiments show that the synthetic trajectories of our proposed model are able to preserve the statistical properties of the original dataset, and are able to efficiently support downstream applications by augmenting their training data. We release the code of our proposed algorithm as well as the datasets to better reproduce the experimental results[1].

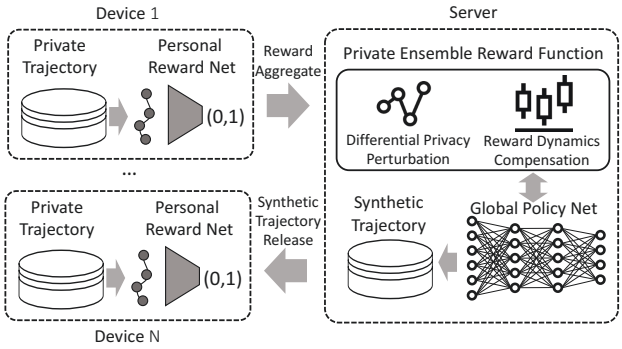


Figure 2: The framework of our system.

Mathematical Model and System Overview

Mathematical Model

For the sake of convenience, we summarize the notations used in this paper in Table A1 of the Appendix. Specifically, we consider the scenario where there are multiple users with their own mobile devices. Each device has recorded the historical mobility trajectory of the corresponding user. We define the set of users as $\mathcal{U}$. Further, for each user $u \in \mathcal{U}$, we define the mobility trajectory of u as a sequence of spatio-temporal points, i.e., $T_u = \{(t_1, l_1), (t_2, l_2), ..., (t_N, l_N)\}$, where l_i is the identifier of the visited location and t_i is the corresponding timestamp. Then, the human mobility generation problem can be defined as follows.

Definition 1 (Privacy-Preserving Federated Mobility Trajectory Generation Problem) *Given a set of user $\mathcal{U}$ and their historical mobility trajectory $\{T_u\}_{u\in\mathcal{U}}$, the goal of this problem is to train a mobility trajectory generator distributedly, which is able to generate mobility trajectory with the preserved utility. In addition, the privacy information involved in the trajectory of each user should be preserved.*

Specifically, the preserved utility indicates that the generated trajectories should statistically resemble the real-world trajectories. Furthermore, they should be able to effectively support the downstream applications relying on trajectory data. On the other hand, the preserved privacy indicates that any pieces of the users' private trajectories should neither be

[1] https://github.com/tsinghua-fib-lab/PateGail

taken away from their own devices, nor be inferred from the uploaded intermediate results of the user devices.

System Overview

We propose a privacy-preserving federated imitation learning system to solve this problem, of which the framework is shown in Figure 2.

As we can observe, each device keeps the private trajectory data belonging to its owner (user), and it trains a personal discriminator to measure to what degree arbitrary state-action pair (s, a) resembles its owner. Thus, this discriminator is trained based on the positive samples obtained from the real-world user trajectory belonging to the user, while the negative samples come from the trajectory generator in our system. Then, only the reward obtained from users' personal discriminators is uploaded to the server.

Although the trajectories are not uploaded, user privacy is also possible to be leaked from the uploaded intermediate result through membership inference attacks (Shokri et al. 2017) or reconstruction attack (Geiping et al. 2020). Thus, a private aggregation mechanism is utilized to aggregate the reward obtained from users' personal discriminators for arbitrary state-action pair (s, a). Specifically, a differential privacy perturbation mechanism is utilized in this process to protect users' privacy. Furthermore, a reward dynamics compensation based on the variance of the rewards is utilized to eliminate the potential noise introduced by the ensemble learning and model the dynamics of reward function across users. Finally, based on the obtained overall reward, the global policy network is able to be trained with the target of finding a policy to maximize the obtained rewards.

Method

In order to model the substantial human decision-making process to simulate the human mobility behavior, we utilize the powerful technique of the generative adversary imitation learning (GAIL) under the model of Markov decision processes (MDPs). Specifically, the MDP model is defined by a 4-tuple $< \mathcal{S}, \mathcal{A}, P, R >$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}+$ represents the state transition probability, and $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ represents the reward function. Specifically, we define the state of users as the set of their historical spatio-temporal points, i.e., $s_t = \{(t_\tau, l_\tau)\}_{\tau \leq t}$, and the action space is defined based on the widely-adopted *exploration and preferential return (EPR)* model (Jiang et al. 2016; Song et al. 2010), which includes four actions composed of *stay, home return, preferential return, and explore*. Then, the state transition probability, which defines the probability distribution of the next state given the current state and action, is also defined based on the EPR model (see Appendix for details). Then, each user is regarded as an "agent" who dynamically determines the action to be executed based on its current state through its policy function, and the goal of the agent is to maximize the discounted total rewards $\sum_t \gamma^t R(s_t, a_t)$, where $\gamma \leq 1$ is the discount factor. In the imitation learning problem, the reward function R as well as the policy function are unknown and thus need to be learned from the real-world data. Thus, in the following part of this section, we first introduce the utilized policy function and reward function. Then, we introduce how to train our proposed system based on GAIL. Finally, we analyze the system in terms of its theoretical privacy-preserving performance.

Policy Function

The policy function defines the decision strategy taken by users, which takes the current state s_t as the input and then outputs the action a_t to be executed. We follow the common settings adopted in most imitation learning problems, and consider the stochastic policy rather than the deterministic policy. In this case, the policy function actually gives the probabilistic distribution of executing arbitrary action $a_t \in \mathcal{A}$. Specifically, we utilize a self-attention transformer (Vaswani et al. 2017) parameterized by θ to model the policy function, which is denoted by $\pi_\theta(a_t|s_t)$.

Combining the policy function $\pi(a_t|s_t)$ and the state transition probability function $P(s_{t+1}|s_t, a_t)$, each agent is able to dynamically determine which action to be executed and then change the current state from s_t to s_{t+1} via interaction with $P(s_{t+1}|s_t, a_t)$. By repeating this process, the agent is able to sample synthetic trajectories corresponding to the policy net π_θ. Thus, the policy function acts as the trajectory generator in our system.

In our system, we only utilize a global policy network, which is trained on the server. Further, in the training process of our system, the policy network only interacts with the private user trajectories through the reward function. By designing a privacy-preserving reward function, we are able to obtain a policy network without privacy leakage. Then, the server is able to send the parameter of the global policy network to the devices, and thus the devices also have the ability to generate synthetic human trajectories. Specifically, the policy function is optimized to imitate real-world human trajectories, of which the degree is measured by the reward function introduced in the following section.

Reward Function

The reward function measures to what degree arbitrary given trajectories imitate real trajectories. Specifically, it takes a state-action pair (s_t, a_t) as the input, and outputs a real number, where higher values indicate that the state-action pair better imitates real-world human decisions.

In the standard GAIL, a discriminator network is utilized to model the reward function, which is trained based on the positive samples of real-world human state-action pairs and the negative samples of synthetic state-action pairs. However, in order to protect user privacy, the private trajectory data stored on mobile devices cannot be gathered together to train the discriminator. Thus, in our system, we replace it with a number of separate personal discriminators of users and a private aggregation mechanism, of which the idea is inspired by the techniques of Private Aggregation of Teacher Ensembles (PATE) (Papernot et al. 2016; Jordon, Yoon, and Van Der Schaar 2018). Specifically, each user device trains its personal discriminator based on its private trajectory data, and the final utilized reward function comes from the private aggregation of their ensemble. Note that the personal

discriminators play a similar role with the teacher models in the standard PATE and PATE-GAN model (Papernot et al. 2016; Jordon, Yoon, and Van Der Schaar 2018). However, different from them, there is no student discriminator trained in our designed system. Instead, the aggregated reward is utilized to directly update the quality function or the advantage function in reinforcement learning algorithms such as A2C (Mnih et al. 2016) and PPO (Schulman et al. 2017).

In the following part of this section, we first introduce the personal discriminator. Then, we introduce how to implement a private aggregation to obtain the reward function based on ensemble learning. Finally, we propose a reward dynamics compensation mechanism to eliminate the potential noise introduced by ensemble learning and model the dynamics of reward function across users.

Personal Discriminators: The discriminator takes the state-action pair as input and then outputs its plausibility. The personal discriminator plays a similar role, but it can only access the trajectory data belonging to its corresponding user and the synthetic trajectories generated based on the global policy network. Specifically, we denote the personal discriminator belonging to user u as D_{ϕ_u}, which is parameterized by ϕ_u. Then, D_{ϕ_u} is optimized based on the following loss function:

$$\mathcal{L}_D^u(\phi_u) = -\mathbb{E}_{\pi_{T_u}}[\log D_{\phi_u}(s,a)] - \mathbb{E}_{\pi}[\log(1 - D_{\phi_u}(s,a))], \tag{1}$$

where $\mathbb{E}_\pi$ represents the expectation with respect to the trajectories sampled based on the policy π. Specifically, for an arbitrary function $f : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$, we have $\mathbb{E}_\pi[f(s,a)] = \mathbb{E}[\sum_{i=1}^{N} f(s_i, a_i)]$, where $a_i \sim \pi(\cdot|s_i)$ and $s_{i+1} \sim P(\cdot|s_i, a_i)$. In addition, $\mathbb{E}_{\pi_{T_u}}$ represents the expectation in terms of the state-action pairs obtained from the real-world trajectory of user u.

Private Aggregation Mechanism: The trajectory data on each user device is insufficient and largely influenced by the user personality, which prevents the discriminator from capturing the principle plausibility of state-action pairs. Thus, it is necessary to incorporate the plausibility estimated by all personal discriminators.

Formally, for arbitrary state-action pair (s, a), each mobile device u estimates its plausibility based on the local personal discriminator $D_{\phi_u}(s, a)$, which is then uploaded to the server. The server computes the average value of the obtained personal rewards and then adds a perturbation to it, of which the process can be expressed as follows:

$$R(s,a) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} D_{\phi_u}(s,a) + \text{Laplace}(0, \lambda). \tag{2}$$

Note that this process from uploading $D_{\phi_u}(s, a)$ to calculating (2) can be protected by homomorphic encryption techniques. Specifically, the server is in charge of generating public keys and sending them to all devices. Each device u then encrypts $D_{\phi_u}(s, a)$ and sends it to another third-party server, which is in charge of implementing the computation of (2) and sending the results to the server in charge of training the trajectory generator.

The third-party server can only influence the performance of the trained trajectory generator and no user privacy can be leaked from it. To order to further prevent malicious third-party servers, we can randomly select a client to be the third-party server in each communication round. It is also possible for the malicious third-party server to incorrectly calculate (2), which can be solved by introducing verification information to the uploaded message $D_{\phi_u}(s, a)$ of the clients.

Another difference of the aggregation mechanism (2) from standard PATE is that we compute the average value of the uploaded outputs of the personal discriminators, while the aggregation results of the standard PATE are discrete. The reason is that utilizing a discrete aggregation mechanism may lead to a sparse reward function, which reduces the performance of our proposed system. Due to this difference, the required perturbation scale λ to achieve differential privacy is also different.

Reward Dynamics Compensation Mechanism: The above ensemble learning based method will introduce extra noise to the obtained reward function. Specifically, users' mobility behavior has intrinsic stochasticity. Furthermore, there also exist personal differences between the reward functions of different users. Since each personal discriminator can only observe the trajectory of a single user, it is more affected by stochasticity and personal differences, which might reduce the performance.

In order to eliminate the influence introduced by the distributed training method, we propose a reward dynamics compensation mechanism. Specifically, it models the reward dynamics actively based on the variance of the obtained rewards from the personal discriminators, which is further incorporated into the reward function to derive the lower bound of obtained reward. The process can be formally expressed by the following equations:

$$\begin{cases} \xi(s,a) = \sqrt{\text{var}(D_{\phi_u}(s,a)) + \text{Laplace}(0, \lambda_c)}, \\ \hat{R}(s,a) = R(s,a) - \beta \xi(s,a), \end{cases} \tag{3}$$

where $\text{var}(X)$ is the variance of the stochastic variable X, and β is a hyper-parameter to adjust the influence of the reward dynamics compensation mechanism. Intuitively, $\hat{R}$ can be regarded as the lower bound of the personal rewards of users, which is formally described in the following theorem.

Theorem 1 *Denote the discounted total reward based on the policy function π and reward function R as $\boldsymbol{J}(\pi, R) = \sum_i \gamma^i R(s_i, a_i)$, where $a_i \sim \pi(\cdot|s_i)$ and $s_{i+1} \sim P(\cdot|s_i, a_i)$. Let R_u denote the personal reward function of user u, i.e., $R_u = D_{\phi_u}$. Then, for a randomly selected user u and policy function π, we have $Pr\left(\boldsymbol{J}(\pi, R_u) \geq \boldsymbol{J}(\pi, \hat{R})\right) \geq 1 - 1/\beta^2$.*

Proof. See Appendix for proof. □

This theorem tells us that under arbitrary policy function π, the discounted total reward of user u obtained based on the reward function $\hat{R}$ is the lower bound of that obtained based on R with probability $1 - \frac{1}{\beta^2}$. By setting a sufficiently large β, we can obtain a probability very close to one. Thus, by utilizing the reward function (3) to replace the original reward function (2), our proposed model is able to maximize the lower bound discounted total reward of the majority

Dataset	ISP					GeoLife				
Metrics(JSD)	*Radius*	*DailyLoc*	*Distance*	*G-rank*	*I-rank*	*Radius*	*DailyLoc*	*Distance*	*G-rank*	*I-rank*
IO-HMM	0.1443	0.3929	0.0596	0.0635	0.1005	0.6146	0.6928	0.5108	0.1679	0.0529
TimeGeo	0.1609	0.6912	0.0337	0.0875	0.1125	0.0737	0.5349	0.0473	0.0553	0.0584
DeepMove	0.6425	0.6934	0.4483	0.1947	0.2310	0.6754	0.4914	0.0512	0.1302	0.0934
GAN	0.6267	0.6936	0.4421	0.1022	0.2485	0.6143	0.6932	0.5157	0.0550	0.3000
SeqGAN	0.6297	0.6931	0.4388	0.1537	0.2474	0.6146	0.6927	0.5068	0.0535	0.2867
MoveSim	0.3606	0.1130	0.0245	0.0578	0.0816	0.2845	0.2467	0.0138	0.0492	0.0408
Ours	**0.0556**	**0.0381**	**0.0051**	**0.0510**	**0.0096**	**0.0699**	**0.1046**	**0.0130**	**0.0256**	**0.0176**
Percentages	**61.47%**	**66.28%**	**79.18%**	**11.76%**	**88.24%**	**5.16%**	**57.60%**	**5.80%**	**47.97%**	**56.86%**

Table 1: Performance comparison of our model and baselines on two mobility datasets, where lower results are better. Bold denotes best (lowest) results and underline denotes the second-best results.

of users, which helps to eliminate the potential noise introduced by the ensemble learning and model the dynamics of reward function across users.

Model Training

The training process of our proposed model is summarized as follows. Firstly, a batch of synthetic trajectories is sampled to train the personal discriminators belonging to different users. Note that this generation process can be implemented on the server as well as the user devices by sending the parameters of the global generator to each device. Then, each user device samples a batch of positive samples from its private real-world trajectory data. Combined with the negative samples obtained from the synthetic trajectories, each device is able to optimize its own discriminator for a number of iterations. Next, another batch of synthetic trajectories sampled from the server is sent to all devices. Each device calculates its rewards based on its personal discriminator and uploads the results to the server. After receiving the rewards of synthetic trajectories obtained from different devices, the server aggregates them based on (2) (3), which is used as the reward function of the imitation learning. Finally, based on the obtained overall reward, the server is able to optimize the global policy function based on reinforcement learning algorithms, e.g., PPO. We give the pseudocode of the training process in the Appendix.

Privacy Analysis

Before we analyze our proposed system in terms of its theoretical privacy-preserving performance, we first provide preliminaries about differential privacy, which is the most widely-used privacy preserving criterion.

Definition 2 ((ϵ, δ)-differential privacy) *A randomized mechanism $\mathcal{M} : \mathcal{D} \to \mathcal{O}$ satisfies (ϵ, δ)-differential privacy if and only if, for arbitrary adjacent datasets D_1 and D_2, and subset $O \in \mathcal{O}$, we have $Pr(\mathcal{M}(D_1) \in O) \leq e^{\epsilon} Pr(\mathcal{M}(D_2) \in O) + \delta$.*

Then, based on the above definition, we will further examine the theoretical privacy-preserving performance of our system by investigating how to set the value of the parameters of the adding noise in (2) and (3) to achieve (ϵ, δ)-differential privacy. Overall, our obtained results can be summarized in the following theorems.

Theorem 2 *By setting $\lambda = \frac{1}{\epsilon|\mathcal{U}|}$, the random mechanism (2) satisfy $(\epsilon, 0)$-differential private.*

Proof. See Appendix for proof. □

Theorem 2 provides how to achieve (ϵ, δ)-differential privacy by using the reward function based on the private aggregation mechanism. If further utilizing the reward dynamics compensation mechanism, we can achieve (ϵ, δ)-differential privacy based on the following theorem.

Theorem 3 *By setting $\lambda = \frac{\kappa}{\epsilon|\mathcal{U}|}$ and $\lambda_c = \frac{3\kappa}{\epsilon(\kappa-1)|\mathcal{U}|}$ for all $\kappa > 1$, the random mechanism (3) satisfy $(\epsilon, 0)$-differential private.*

Proof. See Appendix for proof. □

From these theorems, we can observe that the minimum scales of the added perturbation to achieve $(\epsilon, 0)$-differential privacy are all inversely proportional to the number of users participating in the learning process. Thus, we can only add a small perturbation with sufficient users, indicating our proposed system is feasible in practice.

Experiments

Datasets

We utilize two trajectory datasets to evaluate the performance of our proposed algorithm, which includes a publicly available dataset from previous work and a large-scale dataset obtained from an Internet service provider (ISP).

ISP Dataset. This dataset is provided by an Internet service provider (ISP), which records over 100,000 mobile users' access logs to different cellular base stations covering the duration of one week. Users' locations are obtained based on their accessed cellular base stations, while the timestamps of the access logs are also recorded, together composing the spatio-temporal mobility trajectories of the users.

GeoLife Dataset. This dataset is collected by (Zheng et al. 2010), which contains the mobility trajectories of 178 users, of which the duration is from April 2007 to October 2011. Users' locations are obtained from the GPS logs of their mobile phones, with each record containing the latitude, longitude, and timestamp.

Experimental Settings

Compared Algorithms. In order to have a reliable evaluation of our proposed algorithm, we select the following

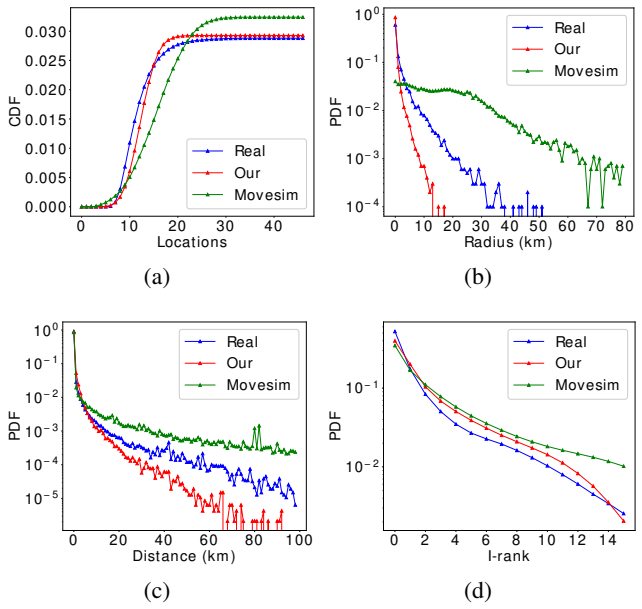


Figure 3: Visualization of the distribution of the selected statistical metrics on the ISP dataset.

state-of-the-art trajectory generation algorithms to be compared with: (1) **IO-HMM** (Yin et al. 2017) modifies the hidden Markov model to incorporate external context information, where the home and work locations of users are used as context information to generate synthetic trajectories. (2) **TimeGeo** (Jiang et al. 2016) is a rule-based probabilistic model based on the classical exploration and preferential return (EPR) model. (4) **GAN** (Goodfellow et al. 2014) utilizes the GAN model to directly generate trajectories, which trains the generator and the discriminator in an adversarial manner. (5) **SeqGAN** (Yu et al. 2017) modifies the standard GAN by enabling the generator to synthesize trajectories step by step, which is trained based on the policy gradient algorithm. (6) **MoveSim** (Feng et al. 2020) is a state-of-the-art trajectory generation algorithm based on GAN. Specifically, the domain knowledge of human mobility regularity is utilized to improve performance.

In addition, all baselines are implemented in the centralized setting, while only our proposed method is implemented in the distributed setting, of which the detailed parameter settings are given in the Appendix.

Statistical Evaluation Metrics. In order to comprehensively evaluate the quality of the generated mobility trajectories, we first evaluate the similarity of the generated trajectories in terms of statistical characteristics. Specifically, we select the following indicators to characterize the characteristics of the mobility trajectories at the record-level or trajectory-level: (1) **Radius** represents *radius of gyration* (Gonzalez, Hidalgo, and Barabasi 2008), i.e., a trajectory-level metric defined by the root mean square of each point of an arbitrary trajectory to its center of mass. (2) **DailyLoc** is a trajectory-level metric defined by the number of distinctive locations visited by each trajectory. (3) **Distance** measures the traveling distance between the adjacent records in users' trajectories, and is a metric at the record-level. (4) **G-rank** is a record-level metric defined by the normalized visited frequency of all users to the top visited locations. (5) **I-rank** is also a record-level metric. Different from G-rank, this metric is computed based on the normalized visited frequency of each user to its top visited locations and then takes the average between different users.

Specifically, these metrics are all expressed by probability distributions with each mobility record or each trajectory as an example. In order to compare the generated trajectories and real trajectories in a more intuitive way, we utilize the Jensen-Shannon divergence (JSD) to measure their difference. Specifically, for two distributions $\boldsymbol{p}$ and $\boldsymbol{q}$, the JSD between them can be defined as:

$$\mathrm{JSD}(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}\mathrm{KL}(\boldsymbol{p}||\frac{\boldsymbol{p}+\boldsymbol{q}}{2}) + \frac{1}{2}\mathrm{KL}(\boldsymbol{q}||\frac{\boldsymbol{p}+\boldsymbol{q}}{2}), \quad (4)$$

where $\mathrm{KL}(\cdot||\cdot)$ is the Kullback-Leibler divergence (Thomas and Cover 2006).

Statistical Evaluation Results

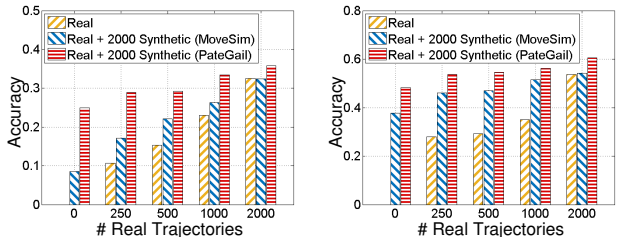
Statistical Metrics. We evaluate the performance of different algorithms in terms of statistical metrics. For fairness, in this group of experiments, no perturbation is added to our proposed algorithm. As the results shown in Table **??**, our model beats the baselines in most situations. Compared to the best baseline, our method can obtain a significant performance gap in most metrics. In addition, we can observe that the JSD of most metrics of all algorithms on the GeoLife data is larger than those on the ISP dataset, indicating worse performance. The reason is that the GeoLife dataset is sparser and with a smaller scale than the ISP dataset, leading to the difficulty of capturing the complex temporal and spatial features by only relying on the limited number of trajectories.

Statistical Distribution Visualization. In Figure 3, we compare the distribution of trajectory data generated from our proposed model and MoveSim on the ISP dataset in terms of the selected statistics metrics. We can observe that the distributions of trajectory data generated by our model are closer to the distribution of real data compared with MoveSim. We also present the corresponding experimental results of the GeoLife dataset in the Appendix.

Practical Demonstrations

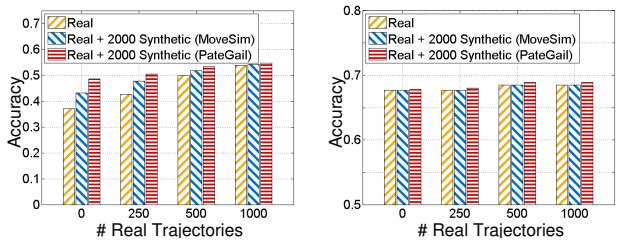
Due to privacy concerns and the collection cost, the available real trajectories are usually limited, which has become the bottleneck of performance of the machine learning models of the downstream applications. In this group of experiments, we examine whether the synthetic trajectory data can help to solve the problem of data scarcity in terms of realistic downstream applications, which include mobility prediction and location recommendation.

Mobility Prediction. In this experiment, we examine whether the synthetic trajectories can help to train a better mobility prediction model. Specifically, the real trajectories combined with synthetic trajectories are used to train an LSTM mobility prediction model (Fattore et al. 2020),



(a) ISP (b) Geolife

Figure 4: Performance of the individual mobility prediction based on trajectory data augmented by different models.



(a) ISP (b) Geolife

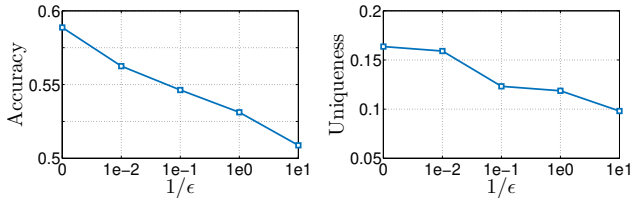
Figure 5: Location recommendation on different datasets.

which is then validated on another group of real trajectories as the test set. We consider three different scenarios which only use real-world trajectory data, use real-world trajectory data and synthetic data generated by MoveSim, and use real-world trajectory data and synthetic data generated by our proposed algorithms as the training set, respectively. We select MoveSim since it is the best baseline in the previous analysis. As we can observe from Figure 4, the mobility prediction performance has a significant improvement when adding the generated trajectories, and the improvement of adding trajectories generated by our proposed algorithm is significantly greater than MoveSim. In addition, we also evaluate the performance when only using generated data, and our performance is approximately three times of MoveSim. These experiment results prove the usability of our proposed model.

Location Recommendation. Further, we utilize synthetic trajectories to train a collaborative filtering algorithm (He et al. 2016), which is then utilized to recommend locations to real users. As we can observe from Figure 5, although the performance gap is not significant on the GeoLife dataset due to its sparsity, on the ISP dataset, the performance is greatly improved by adding generated data. In addition, the large performance gap of the location recommendation algorithm based on trajectories generated by our proposed algorithm compared with MoveSim indicates the superiority of our proposed algorithm.

Privacy Risk Analysis

To evaluate the privacy-preserving performance of our proposed algorithms, we further conduct two experiments in terms of actual privacy attacks. The first attack is the membership inference attack (MIA) (Shokri et al. 2017; Nasr, Shokri, and Houmansadr 2019). Specifically, we consider the white-box inference attack. Given a set of trajectories $\mathcal{T}_A$ as the target of attacks, for each trajectory $T_u \in \mathcal{T}_A$, the adversary calculates the reward of each state-action pair in T_u, and then uses them as the feature of a Random Forest classifier to infer whether T_u is included in the training dataset of the obtained trajectory generation model. The same number of real-world trajectories used for training and not used for training the trajectory generation model are sampled as the positive samples and negative samples, respectively. Then, a five-fold cross validation is implemented to evaluate the privacy risk in terms of MIA. The second attack is to examine the uniqueness of real trajectories with respect to generated trajectories. Specifically, for each real trajectory T_u, the adversary finds the generated trajectory T'_v with the highest overlapping rate with T'_v, where the highest overlapping rate is utilized as the uniqueness metric. Here, the overlapping rate between two trajectories is defined as the ratio between their identical locations at the same time slots and the total trajectory length. A lower accuracy of MIA and a smaller uniqueness indicate better privacy-preserving performance. The experimental results are shown in Figure 6. As we can observe, it is not surprising that both MIA accuracy and uniqueness decrease with $1/\epsilon$, indicating less privacy risk for smaller privacy budget ϵ. In addition, we can observe that our algorithm has low MIA accuracy and uniqueness, indicating that it is robust to actual privacy attacks.



(a) MIA (b) Uniqueness test

Figure 6: Performance in terms of privacy risk.

Related Work

Human Trajectory Generators. Human trajectory generation models have been investigated for decades. Early approaches mainly utilize classical probabilistic methods or rule-based methods to model and generate human trajectories (Jiang et al. 2016; Isaacman et al. 2012; Yin et al. 2017; Bindschaedler and Shokri 2016; Zhao et al. 2019; Pappalardo and Simini 2017). However, these methods are derived from strong assumptions of human mobility, and whether these assumptions hold true in reality is questionable. In addition, these assumptions also limit a small number of parameters describing the human mobility process, leading to their weakness in terms of modeling the complicated relationship of high-dimensional mobility trajectories. In recent years, more deep learning based human trajectory generators have been proposed by utilizing variational autoencoder (VAE) (Huang et al. 2019), genera-

tive adversarial network (GAN) (Feng et al. 2018; Ouyang et al. 2018; Kulkarni et al. 2018; Liu, Chen, and Andris 2018; Agrim Gupta 2018), imitation learning (Pan et al. 2020; Zhang et al. 2019; Wu et al. 2020; Zhang 2020; Seongjin Choi 2020; Menghai Pan 2020). However, none of them considers the critical privacy problem, indicating that there exist privacy leakage risks of the trajectory data utilized to train these models. Different from them, in this paper, we address the privacy-preserving issue, and propose a federated mobility generator by utilizing the techniques of imitation learning.

Imitation Learning. The goal of imitation learning is to learn the policy function, which gives the action to be executed based on the current state (Bain and Sammut 1995; Boularias, Kober, and Peters 2011; Ziebart et al. 2008; Ziebart, Bagnell, and Dey 2010; Ziebart et al. 2008; Ho and Ermon 2016). The most successful imitation learning method is generative adversarial imitation learning (GAIL), which utilizes the non-linear neural network to model the reward function and policy function. It has been adopted in numerous practical applications, including dynamic treatment regimes (Wang et al. 2020), traffic signal control (Xiong et al. 2019), and human drive behavior analysis (Pan et al. 2020; Zhang et al. 2019; Wu et al. 2020; Zhang 2020; Seongjin Choi 2020; Menghai Pan 2020), etc. In this paper, we utilize imitation learning techniques to solve the human mobility trajectory generation problem, which is able to model the crucial human decision-making process to generate human trajectories with preserved utility.

Conclusion

In this paper, we propose a privacy-preserving federated mobility trajectory generator based on imitation learning techniques, which is able to generate plausible synthetic mobility trajectories with the preserved utility to be utilized in downstream applications and preserve users' privacy at the same time. Extensive experiments validate the effectiveness of our proposed model. Specifically, the generated trajectories based on our proposed algorithm are able to preserve the statistical properties of the original dataset in terms of a number of key statistical metrics. Furthermore, the synthetic trajectories are able to efficiently support practical applications, including mobility prediction and location recommendation, demonstrating its effectiveness.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under U21B2036, 62171260, 61971267, 61972223, U1936217, U20B2060, the National Key Research and Development Program of China under 2020YFA0711403, and Young Elite Scientists Sponsorship Program by CIC (Grant No. 2021QNRC001).

References

Agrim Gupta, L. F.-F. S. S. A. A., Justin Johnson. 2018. Social GAN: Socially Acceptable Trajectories With Generative Adversarial Networks. In *Proc. CVPR*.

Bain, M.; and Sammut, C. 1995. A Framework for Behavioural Cloning. In *Machine Intelligence 15*, 103–129.

Bindschaedler, V.; and Shokri, R. 2016. Synthesizing plausible privacy-preserving location traces. In *Proc. IEEE SP*.

Boularias, A.; Kober, J.; and Peters, J. 2011. Relative entropy inverse reinforcement learning. In *Proc. AISTATS*, 182–189.

Chen, F.; Luo, M.; Dong, Z.; Li, Z.; and He, X. 2018. Federated meta-learning with fast convergence and efficient communication. *arXiv preprint arXiv:1802.07876*.

Fattore, U.; Liebsch, M.; Brik, B.; and Ksentini, A. 2020. AutoMEC: LSTM-based user mobility prediction for service management in distributed MEC resources. In *Proc. MSWiM*.

Feng, H.; Bai, F.; and Xu, Y. 2019. Identification of critical roads in urban transportation network based on GPS trajectory data. *Physica A: Statistical Mechanics and its Applications*, 535: 122337.

Feng, J.; Li, Y.; Zhang, C.; Sun, F.; Meng, F.; Guo, A.; and Jin, D. 2018. Deepmove: Predicting human mobility with attentional recurrent networks. In *Proc. WWW*.

Feng, J.; Yang, Z.; Xu, F.; Yu, H.; Wang, M.; and Li, Y. 2020. Learning to Simulate Human Mobility. *Proc. KDD*.

Geiping, J.; Bauermeister, H.; Dröge, H.; and Moeller, M. 2020. Inverting gradients-how easy is it to break privacy in federated learning? *Proc. NeurIPS*.

Gonzalez, M. C.; Hidalgo, C.; and Barabasi, A. L. 2008. Understanding individual human mobility patterns. *Nature*, 453(7196): p.779–782.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *Proc. NeurIPS*.

Hard, A.; Rao, K.; Mathews, R.; Ramaswamy, S.; Beaufays, F.; Augenstein, S.; Eichner, H.; Kiddon, C.; and Ramage, D. 2018. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.

He, X.; Zhang, H.; Kan, M.-Y.; and Chua, T.-S. 2016. Fast matrix factorization for online recommendation with implicit feedback. In *Proc. SIGIR*.

Hess, A.; Hummel, K. A.; Gansterer, W. N.; and Haring, G. 2016. Data-driven Human Mobility Modeling: A Survey and Engineering Guidance for Mobile Networking. *ACM Computing Surveys (CSUR)*.

Ho, J.; and Ermon, S. 2016. Generative adversarial imitation learning. In *Proc. NeurIPS*.

Huang, D.; Song, X.; Fan, Z.; Jiang, R.; Shibasaki, R.; Zhang, Y.; Wang, H.; and Kato, Y. 2019. A variational autoencoder based generative model of urban human mobility. In *Proc. MIPR*.

Isaacman, S.; Becker, R.; Caceres, R.; Martonosi, M.; Rowland, J.; Varshavsky, A.; and Willinger, W. 2012. Human mobility modeling at metropolitan scales. 239–252.

Jiang, S.; Yang, Y.; Gupta, S.; Veneziano, D.; Athavale, S.; and González, M. C. 2016. The TimeGeo modeling framework for urban mobility without travel surveys. *PNAS*, 113(37).

Jordon, J.; Yoon, J.; and Van Der Schaar, M. 2018. PATE-GAN: Generating synthetic data with differential privacy guarantees. In *Proc. ICLR*.

Kulkarni, V.; Tagasovska, N.; Vatter, T.; and Garbinato, B. 2018. Generative models for simulating mobility trajectories. *arXiv preprint arXiv:1811.12801*.

Liu, X.; Chen, H.; and Andris, C. 2018. trajGANs : Using generative adversarial networks for geo-privacy protection of trajectory data (Vision paper). In *Location Privacy and Security Workshop*.

Menghai Pan, Y. L.-X. Z. Z. L. J. B. Y. Z. J. L., Weixiao Huang. 2020. "Is Reinforcement Learning the Choice of Human Learners?: A Case Study of Taxi Drivers". In *Proc. SIGSPATIAL*.

Mnih, V.; Badia, A. P.; Mirza, M.; Graves, A.; Lillicrap, T.; Harley, T.; Silver, D.; and Kavukcuoglu, K. 2016. Asynchronous methods for deep reinforcement learning. In *Proc. ICML*.

Nasr, M.; Shokri, R.; and Houmansadr, A. 2019. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *Proc. IEEE SP*, 739–753.

Ouyang, K.; Shokri, R.; Rosenblum, D. S.; and Yang, W. 2018. A Non-Parametric Generative Model for Human Trajectories. In *IJCAI*, 3812–3817.

Pan, M.; Huang, W.; Li, Y.; Zhou, X.; and Luo, J. 2020. xGAIL: Explainable Generative Adversarial Imitation Learning for Explainable Human Decision Analysis. In *Proc. KDD*.

Papernot, N.; Abadi, M.; Erlingsson, U.; Goodfellow, I.; and Talwar, K. 2016. Semi-supervised knowledge transfer for deep learning from private training data. *arXiv preprint arXiv:1610.05755*.

Pappalardo, L.; and Simini, F. 2017. Data-driven generation of spatio-temporal routines in human mobility. *Data Mining and Knowledge Discovery*, 32(1).

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Seongjin Choi, H. Y., Jiwon Kim. 2020. "TrajGAIL: Generating Urban Vehicle Trajectories using Generative Adversarial Imitation Learning". In *arXiv preprint*.

Shokri, R.; Stronati, M.; Song, C.; and Shmatikov, V. 2017. Membership Inference Attacks Against Machine Learning Models. 3–18.

Song, C.; Qu, Z.; Blumm, N.; and Barabasi, A. 2010. Limits of Predictability in Human Mobility. *Science*, 327(5968): 1018–1021.

Thomas, J. A.; and Cover, T. M. 2006. *Elements of information theory*. John Wiley & Sons.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. In *Proc. NeurIPS*.

Wang, L.; Yu, W.; He, X.; Cheng, W.; and Zha, H. 2020. Adversarial Cooperative Imitation Learning for Dynamic Treatment Regimes. In *Proc. WWW*.

Wu, G.; Li, Y.; Luo, S.; Song, G.; Wang, Q.; He, J.; Ye, J.; Qie, X.; and Zhu, H. 2020. A Joint Inverse Reinforcement Learning and Deep Learning Model for Drivers' Behavioral Prediction. In *Proc. CIKM*.

Xiong, Y.; Zheng, G.; Xu, K.; and Li, Z. 2019. Learning Traffic Signal Control from Demonstrations. In *Proc. CIKM*.

Yin, M.; Sheehan, M.; Feygin, S.; Paiement, J.-F.; and Pozdnoukhov, A. 2017. A generative model of urban activities from cellular data. *IEEE Transactions on Intelligent Transportation Systems*, 19(6).

Yu, L.; Zhang, W.; Wang, J.; and Yu, Y. 2017. Seqgan: Sequence generative adversarial nets with policy gradient. In *Proc. AAAI*.

Zhang, L. Y. Z. X.-Z. Z. L. J., X. 2020. TrajGAIL: Trajectory Generative Adversarial Imitation Learning for Long-term Decision Analysis. In *Proc. ICDM*.

Zhang, X.; Li, Y.; Zhou, X.; and Luo, J. 2019. Unveiling taxi drivers' strategies via cgail: Conditional generative adversarial imitation learning. In *Proc. ICDM*.

Zhao, P.; Jiang, H.; Li, J.; Zeng, F.; and Zhang, G. 2019. Synthesizing Privacy Preserving Traces: Enhancing Plausibility With Social Networks. *IEEE/ACM Transactions on Networking*, PP(99): 1–14.

Zheng, Y.; Xie, X.; Ma, W.-Y.; et al. 2010. Geolife: A collaborative social networking service among user, location and trajectory. *IEEE Data Eng. Bull.*, 33(2): 32–39.

Ziebart, B. D.; Bagnell, J. A.; and Dey, A. K. 2010. Modeling interaction via the principle of maximum causal entropy. In *Proc. ICML*.

Ziebart, B. D.; Maas, A. L.; Bagnell, J. A.; Dey, A. K.; et al. 2008. Maximum entropy inverse reinforcement learning. In *Proc. AAAI*.