VTSV: A Privacy-Preserving Vehicle Trajectory Simulation and Visualization Platform Using Deep Reinforcement Learning

Jinmeng Rao
Geospatial Data Science Lab,
University of Wisconsin-Madison
Madison, USA
jinmeng.rao@wisc.edu

Song Gao
Geospatial Data Science Lab,
University of Wisconsin-Madison
Madison, USA
song.gao@wisc.edu

Xiaojin Zhu
Department of Computer Sciences,
University of Wisconsin-Madison
Madison, USA
jerry.zhu@wisc.edu

ABSTRACT

Trajectory data is among the most sensitive data and the society increasingly raises privacy concerns. In this demo paper, we present a privacy-preserving Vehicle Trajectory Simulation and Visualization (VTSV) web platform (demo video: https://youtu.be/NY5L4bu2kTU), which automatically generates navigation routes between given pairs of origins and destinations and employs a deep reinforcement learning model to simulate vehicle trajectories with customized driving behaviors such as normal driving, overspeed, aggressive acceleration, and aggressive turning. The simulated vehicle trajectory data contain high-sample-rate of attributes including GPS location, speed, acceleration, and steering angle, and such data are visualized in VTSV using streetscape.gl, an autonomous driving data visualization framework. Location privacy protection methods such as origin-destination geomasking and trajectory k-anonymity are integrated into the platform to support privacy-preserving trajectory data generation and publication. We design two application scenarios to demonstrate how VTSV performs location privacy protection and customize driving behavior, respectively. The demonstration shows that VTSV is able to mitigate data privacy, sparsity, and imbalance sampling issues, which offers new insights into driving trajectory simulation and GeoAI-powered privacy-preserving data publication.

CCS CONCEPTS

• **Security and privacy** → *Privacy protections*; • **Computing methodologies** → *Artificial intelligence*; • **Human-centered computing** → Visualization systems and tools.

KEYWORDS

reinforcement learning, privacy protection, transportation, vehicle trajectory, data visualization

ACM Reference Format:
Jinmeng Rao, Song Gao, and Xiaojin Zhu. 2021. VTSV: A Privacy-Preserving Vehicle Trajectory Simulation and Visualization Platform Using Deep Reinforcement Learning. In *GeoAI'21, November 02, 2021, Beijing, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3486635.3491073

GeoAI'21, November 02, 2021, Beijing, China

ACM ISBN 978-1-4503-9120-7/21/11…$15.00
https://doi.org/10.1145/3486635.3491073

1 INTRODUCTION

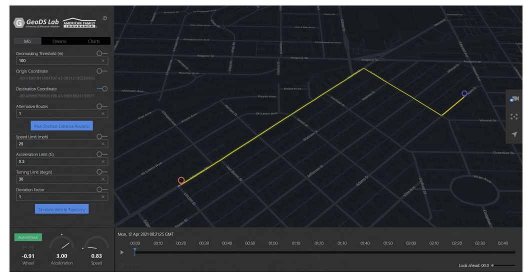

Figure 1: The web interface of VTSV.

The prevalence of nowadays location-aware devices and the ubiquitous Internet enables us to record and collect massive large-scale individual location trajectory data. Such big data bring novel opportunities for understanding and evaluating transportation, urban planning, human dynamics, and so forth [8], while raising privacy concerns [2]. As an important source of trajectory data, vehicle trajectory data can be recorded by either the sensors embedded in drivers' mobile phones or vehicle-mounted devices. A prominent property of vehicle trajectory data is that, unlike social media check-in data or mobility tracking data, they usually contain not only GPS locations but also many other attributes that can be used to better describe vehicle driving status, such as speed, acceleration, and steering angles [6]. Such enriched trajectory data support various application scenarios such as driving behavior profiling, Usage-Based Insurance (UBI), and traffic simulation. Although promising in these applications, vehicle trajectory data face three main challenges.

Challenge 1: Privacy concerns. Many people, especially private car drivers, may not agree to service providers for collecting or sharing their driving data since they worry that others may be able to identify their sensitive locations (e.g., home/work locations) from their trajectory data, such as through location clustering algorithms. Such concerns result in the fact that there are few private car trajectory datasets available.

Challenge 2: Data sparsity. Many existing trajectory datasets have relatively sparse sample rates (e.g., 2 ~ 10 seconds or even 2 minute per point), leading to data sparsity issues. Such sparse data are usually not sufficient to support some application scenarios such as travel time estimation, driving behavior extraction and lane change detection [1].

Challenge 3: Data imbalance. Most of the existing vehicle trajectory datasets are taxi trip datasets. Regardless of data quality, these data are not very representative for reflecting traffic characteristics or the driving behavior of private car drivers. Also, in the real world, trajectories with normal driving behavior account for the vast majority, while the trajectories with other driving behaviors (e.g., aggressive driving) may be more valuable for analyzing accidents or adjusting insurance premiums.

To this end, we present the VTSV, a privacy-preserving Vehicle Trajectory Simulation and Visualization platform that simulates vehicle trajectories using deep reinforcement learning with customized driving behaviors and supports privacy-preserving data generation. The web interface of VTSV is shown in Figure 1. Users can pick up pairs of origin and destination locations on the map, and the driving routes between them (shortest paths by default) along the road network will be planned. Privacy protection methods such as geomasking and trajectory k-anonymity are enabled here to protect location privacy of individual routes. A reinforcement learning model will then simulate high-sample-rate (e.g., 10 Hz) vehicle trajectory data (including GPS locations and sensor data such as speed, acceleration) along the planned routes, and several parameters can be set to simulate trajectory data with different driving behaviors (overspeed, aggressive acceleration, etc.). The simulated data are visualized on the platform as autonomous driving trips and can also be downloaded so that users can better understand and analyze driving behaviors comprehensively. To the best of our knowledge, our work is the first platform to simulate and visualize high-sample-rate privacy-preserving vehicle trajectory using deep reinforcement learning with customized driving behaviors, which offers new insights into driving trajectory simulation and GeoAI-powered privacy-preserving data publication.

2 OVERALL ARCHITECTURE

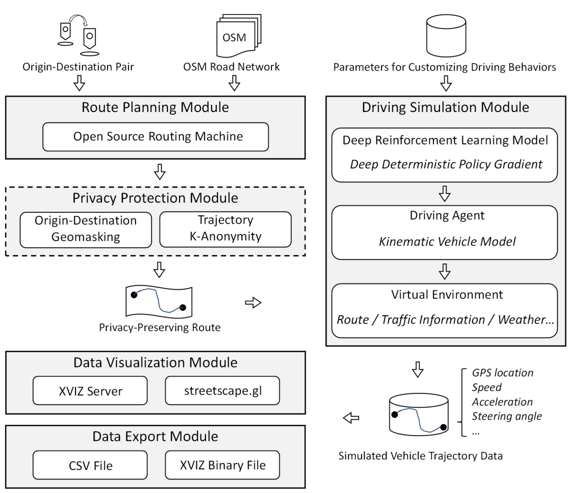

Figure 2: The overall architecture of VTSV.

Figure 2 shows the overall architecture of VTSV, which consists of the following five modules:

- **Route Planning Module**: given a pair of coordinates (origin and destination), this module will plan a route from the origin to the destination.
- **Privacy Protection Module**: this module integrates several location privacy protection techniques such as geomasking and trajectory k-anonymity to protect location privacy of vehicles.
- **Driving Simulation Module**: this module includes a pre-trained reinforcement learning model that can simulate customized driving behavior along a given route.
- **Data Visualization Module**: the planned route, simulated driving data, and the driving process are visualized using a driving data visualization framework.
- **Data Export Module**: the simulated data can be exported and downloaded from the module. Comma-separated values (CSV) file and the XVIZ binary file are supported.

2.1 Route Planning Module

The first step for vehicle trajectory simulation is to determine a driving route from an origin to a destination. A driving route R consists of a set of connected road segments $\mathbb{S}$ and a set of intersections $\mathbb{I}$. We record intersections and the centerlines of road segments using GPS points (lng, lat) and polylines $[(lng_1, lat_1), (lng_2, lat_2), ...]$. The goal of the Route Planning Module is to plan such a driving route between the origin and the destination picked up by users. Specifically, we use the Open Source Routing Machine (OSRM), an open-source route planning engine based on OpenStreetMap data, to plan driving routes. OSRM uses contraction hierarchies and the multilevel Dijkstra shortest path algorithm [3], which results in fast and flexible routing and thus suitable for web-based routing applications.

OSRM plans shortest-path routes by default, and it also accepts options for customizing route planning. For example, option "steps" allows returning route steps for each road segment; option "overview" defines the fineness of the route (e.g., complete or simplified).

2.2 Privacy Protection Module

Privacy Protection Module processes origins, destinations, and routes using location privacy protection methods and produces privacy-preserving routes that do not disclose location privacy or identification of mobile users or vehicles. The methods we currently use in VTSV are Origin-Destination Geomasking and Trajectory K-anonymity.

Origin-Destination Geomasking: Geomasking displaces locations in uncertain ways so that original locations are concealed, which has been broadly used in public health and spatial analysis [2]. In the VTSV, we implement random perturbation, one of the popular geomasking methods, to randomly blur origins and destinations based on a user-customizable threshold (e.g., 1,000 m. The best threshold value varies among different scenes). The planned routes based on blurred origins and destinations reduce the chance of disclosing actual departure and arrival locations of users.

Trajectory K-anonymity: The idea of trajectory k-anonymity is to mix k number of similar trajectories so that no one can distinguish one unique trajectory from the others. In VTSV, we integrated a simplified version of the generalization-based approach [5] that allows planning multiple alternative routes to achieve k-anonymity, where k can be customized by users.

2.3 Driving Simulation Module

We use the Driving Simulation Module to simulate high-sample-rate vehicle trajectory data along given routes. This module first

Table 1: Parameters in the kinematic vehicle model

Parameter	Description
Length	The length of the vehicle (4m by default)
Position	The position of the vehicle center
Orientation	The direction where the vehicle heads to
Steering Angle	The current steering angle of the vehicle
Velocity	The current velocity of the vehicle

Table 2: Parameters for customizing driving behaviors

Parameter	Example
Speed Limit	Higher -> overspeeding
Acceleration Limit	Higher -> sudden acceleration/brake
Turning Speed Limit	Higher -> aggressive turning
Deviation to Centerline	Higher -> dangerous lane changing

trains a deep reinforcement learning model to simulate the driving agent with customized driving behaviors, and then it puts the driving agent into a virtual environment built upon given routes. The driving agent observes the environment and simulates the driving process along the route from the origin to destination, and its real-time GPS locations, acceleration, speed, and steering angle are recorded at a customized sample rate (i.e., 10 Hz). There are four important concepts in this module, namely deep reinforcement learning, driving agent, driving behavior, and virtual environment.

Deep Reinforcement Learning: Reinforcement Learning (RL) is a subarea in machine learning that focuses on guiding an agent to take expected actions in an environment so as to maximize the cumulative reward. Deep Reinforcement Learning (DRL) incorporates deep learning models into RL so that the agent learns to make right decisions without manual feature engineering. In this work, we use the Deep Deterministic Policy Gradient (DDPG), a model-free DRL algorithm for learning continuous actions [4].

Driving Agent: A driving agent represents a simulated driver driving a vehicle – it observes the environment, makes decisions, takes actions, updates its state, and makes new observations. The decision-making process is supported by DDPG. To naturally sample vehicle trajectories, we use a simplified kinematic vehicle model to represent the vehicle. This model regards the vehicle as an object moving in space and uses kinematic variables (e.g., velocity, acceleration) to describe and control it without understanding the internal mechanism of the vehicle. Some key parameters used in the model are introduced in Table 1.

Driving Behavior: Driving behavior reflects a driver's driving habits and styles, which has a great impact on travel safety. Generally, driving behavior can be classified into aggressive behavior, normal behavior, etc. Specifically, driving behaviors can be quantitatively described by a series of driving-related indicators. Table 2 shows a series of indicators in the VTSV and how they can be customized to produce different driving behaviors.

Virtual Environment: A virtual environment is established based on the planned route and driving-related environmental factors such as traffic information (e.g., road speed limit), weather, etc. The goal of a virtual environment is to simulate real-world situations as close as possible. The driving agent interacts with the virtual environment and produces moving trajectories.

2.4 Data Visualization Module

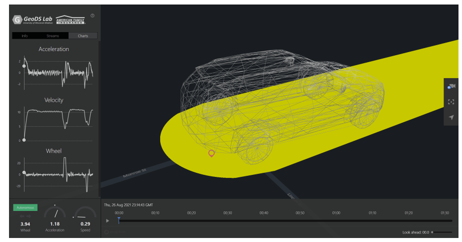

Figure 3: The visualization of driving data.

The VTSV provides a web-based visualization interface for simulated vehicle trajectory data using streetscape.gl, a toolkit for visualizing autonomous and robotics data, which can help users understand the vehicle trajectory data and the corresponding driving behaviors. As shown in Figure 1, the origin, destination, planned routes, and road networks are all marked and visualized on a 3D base map, and users can freely change the camera perspective to observe the map and route. The VTSV also puts a simplified vehicle 3D model on the map to indicate the current location and orientation of the driving agent. After the vehicle trajectory simulation is completed, users can click on the "play" button on the bottom timeline to play the 3D driving animation of the driving agent along the route. As shown in Figure 3, the real-time sensor information (e.g., speed, acceleration, and steering angle) is displayed on the bottom left panel, and their changes over time are visualized as line plots and displayed on the left side. We can clearly see from the plots when and how much the vehicle agent accelerates, decelerates, and makes turns, which helps users better interpret the vehicle trajectory data.

2.5 Data Export Module

This module allows users to download the simulated vehicle trajectory data to local storage for further analysis and downstream applications. By default, we support two file types: comma-separated values (CSV) file and the binary file following Uber XVIZ protocol. For the CSV file, the rows record trajectory location points at different timestamps, and the columns contain attributes including latitude, longitude, speed, acceleration, steering angle, etc. The XVIZ binary file contains the necessary information for real-time transfer and visualization in XVIZ-based applications.

3 IMPLEMENTATION AND APPLICATIONS

The implementation of VTSV follows classic Client/Server architecture. On the server side, we use *Flask* as backend web framework and deploy our pre-trained DDPG model (in PyTorch) and the OSRM routing engine. We also deploy an XVIZ server for converting the simulated data into XVIZ binary file for visualization. On the client side, we use *React* as frontend web framework and the base map from Mapbox, and utilize streetscape.gl for visualizing simulated vehicle trajectory data. Below we demonstrate two use cases of VTSV. One uses VTSV to simulate vehicle trajectories with customized driving behaviors, the other uses VTSV to generate privacy-preserving trajectories.

3.1 Customized Driving Behaviors

Here we refer to the thresholds from [7] to distinguish between normal and aggressive driving behaviors. Two vehicle trajectories produced by the VTSV using different parameters are shown in Figure 4. Figure 4A describes a trajectory with normal driving behavior (speed within the road speed limit 25mph, about 11.18 m/s; acceleration within 0.3G gravitational acceleration, about 2.94 m/s^2). Figure 4B, in contrast, describes a vehicle trajectory with aggressive driving behavior (speed reaches up to 35mph, about 15.65 m/s; acceleration reaches up to 0.4G, about 3.92 m/s^2).

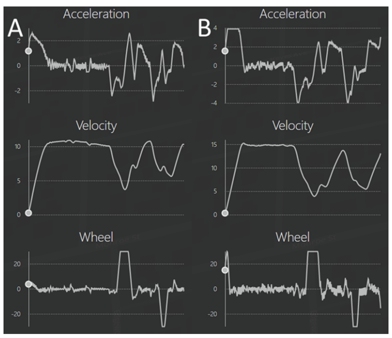

Figure 4: Customized driving behaviors. A: Normal; B: Aggressive. The units for acceleration, velocity, and wheel in the figure are m/s^2, m/s, and *degree*, respectively.

3.2 Privacy-Preserving Trajectory Generation

Figure 5 demonstrates the origin-destination geomasking methods to protect vehicles' location privacy. If enabled, the origin and destination will be randomly picked up within the circle (radius = r) around the location users click on the map, respectively. For example, a user may choose a neighborhood-scale radius or a city-scale radius to displace their true locations. By doing so, the actual origin and destination can be concealed so as to lower the risk of privacy leakage.

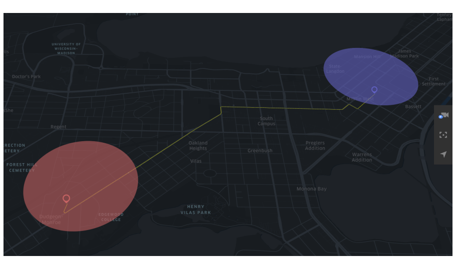

Figure 5: Illustration of origin-destination geomasking.

Figure 6 shows how we use trajectory k-anonymity to protect identity. If enabled, VTSV will generate multiple alternative routes between given origin and destination using the aforementioned generalization-based k-anonymity approach. Users can mix them with the real trajectory, making it harder to distinguish the real one from the others.

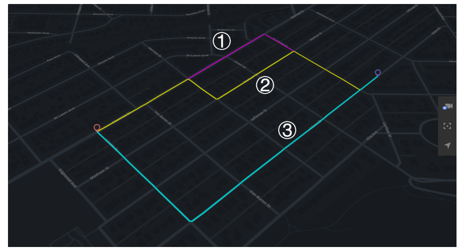

Figure 6: Illustration of trajectory k-anonymity.

4 CONCLUSIONS

In this demo paper, we present the VTSV, a web-based privacy-preserving vehicle trajectory simulation and visualization platform. It supports the planning of vehicle routes between origins and destinations and utilizes a deep reinforcement learning model to simulate vehicle trajectories with customized driving behaviors. Origin-destination geomasking and trajectory k-anonymity are incorporated to achieve privacy-preserving data generation. The results show that the VTSV helps mitigating data privacy, sparsity, and imbalance issues in vehicle trajectory data and brings novel insights into driving trajectory simulation and privacy-preserving data publication using GeoAI approaches.

ACKNOWLEDGMENTS

We acknowledge the funding support provided by the American Family Insurance Data Science Institute Funding Initiative at the University of Wisconsin-Madison.

REFERENCES

[1] Ove Andersen and Kristian Torp. 2017. Sampling Frequency Effects on Trajectory Routes and Road Network Travel Time. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 1–10.

[2] Song Gao, Jinmeng Rao, Xinyi Liu, Yuhao Kang, Qunying Huang, and Joseph App. 2019. Exploring the effectiveness of geomasking techniques for protecting the geoprivacy of Twitter users. *Journal of Spatial Information Science* 2019, 19 (2019), 105–129.

[3] Robert Geisberger, Peter Sanders, Dominik Schultes, and Daniel Delling. 2008. Contraction hierarchies: Faster and simpler hierarchical routing in road networks. In *International Workshop on Experimental and Efficient Algorithms*. Springer, 319–333.

[4] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).

[5] Mehmet Ercan Nergiz, Maurizio Atzori, and Yucel Saygin. 2008. Towards trajectory anonymization: a generalization-based approach. In *Proceedings of the SIGSPATIAL ACM GIS 2008 International Workshop on Security and Privacy in GIS and LBS*. 52–61.

[6] Eduardo Romera, Luis M Bergasa, and Roberto Arroyo. 2016. Need data for driver behaviour analysis? Presenting the public UAH-DriveSet. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 387–392.

[7] Chalermpol Saiprasert, Thunyasit Pholprasit, and Wasan Pattara-Atikom. 2013. Detecting driving events using smartphone. In *Proceedings of the 20th ITS World Congress*, Vol. 11. 11.

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