An adversarial approach to quantitative analysis of trajectory data

Aparna Bhutani
Center for Urban Science and Progress
New York University
New York, NY
ab8473@nyu.edu

Shreeraman Arunachalam Karikalan
Center for Urban Science and Progress
New York University
New York, NY
sak869@nyu.edu

Siqi Huang
Center for Urban Science and Progress
New York University
New York, NY
sh5688@nyu.edu

Vivek Patel
Center for Urban Science and Progress
New York University
New York, NY
vp1338@nyu.edu

Abstract

Recent years have seen an alarming increase in the usage of digital devices with location capabilities in them. Despite being the most sensitive data of individuals, mobility/trajectory data are collected by various devices, such as, smartphones and cameras. With pervasive usage of these personal devices every user leaves a non-erasable digital trace and is prone to being exposed to privacy risks. This research tries to quantify the privacy leakage in a given mobility dataset after it has been anonymized with a location privacy protection mechanism (LPPM). By privacy, herein, it means identity of a particular user. The research has been conducted from two different roles, publisher and adversary. As a publisher we anonymize the dataset using different anonymization schemes, while from the adversarial approach we try to recover and reconstruct the anonymized dataset. The research tries to quantify the privacy leakage based on how much of original trace is being recovered by the adversary. In addition, the research also automates the trace generation process for different trajectory data sets. Our research will be useful to increase privacy of the mobility data being collected everyday and making it difficult for attackers to obtain sensitive information of the users.

1 Introduction

Smart cities are one of the key development areas around the globe, with city governments competing against each other to automate life around the city. People enjoy the convenience brought by smart cities, but they usually overlook the risks associated with the data-driven approaches used in the applications. As one of the most sensitive data of individuals, the mobility/trajectory data is collected by various devices, e.g., smartphones and cameras. These devices, although providing convince, leaves a digital trace of the user. The contextual information attached to a trace can be used to crack the habitual patterns, interests, and activities of users. The sensitive information in the trace can be subjected to identity theft wherein an adversary couples the given anonymized data or traces with additional information gained elsewhere such as social media to deanonymize the identity of the user. Privacy breach using the New York City taxi data is one of the classic examples of the privacy risk exposed by publishing mobility data.[8] Herein, it was a simple hack wherein the adversary utilized few pictures from the internet on the celebrities getting on or off a taxi and matched it with the trip released on the public portal. After a successful match, it was possible to retrieve not only where they resided but also their identity. In addition, the adversary was also able to recover the tipping patterns of certain celebrities. Nevertheless, Trajectory data is one of the vital resources for any urban planner/researcher to propose better plans and policies. Hence, it is very important to publish these datasets to drive better innovation and creative insights. It becomes a critical question for the data publisher on how to achieve the trade-off between usability and privacy, that is, publish these mobility datasets to preserve maximum usability while also preserving the privacy of individuals. This research focuses on the latter aspect of preserving the privacy and identity of individuals.

1.1 Literature Review

Location privacy is really important in this digital age where several websites and applications such as Google, Apple, Facebook offer multiple location based services can store and track user locations. Extensive research has been conducted in this field to promote better user location privacy.

The goal of the present work is to maximize the privacy of individuals in a published mobility dataset by using adversarial approach. This work fits into the larger effort of promoting maximum privacy for users in this digital era.

While Decker [1] gives us an overview of why location based services are important, the most relevant threats which occur in these scenarios of location information when using such services and the technical approaches which were used to avoid misuse of location information Duckham [2] proposes the seven key principles of location privacy in the future, which make this field of research different from other research topics in privacy. These principles majorly focuses on how to improve and distinguish the location privacy methods and techniques from other privacy methodologies and spatial research.

Shokri et al [5] explains the details of different location privacy preserving mechanisms and the various metrics that can be used for measuring location privacy such as uncertainty-based, error-based and k-anonymity. This work also gives an entire framework that provides a logical structure for classifying and organizing fundamental components and concepts of location privacy.

All this work has been a source of inspiration and our research leverages the tool developed by [7] to achieve the end goal.

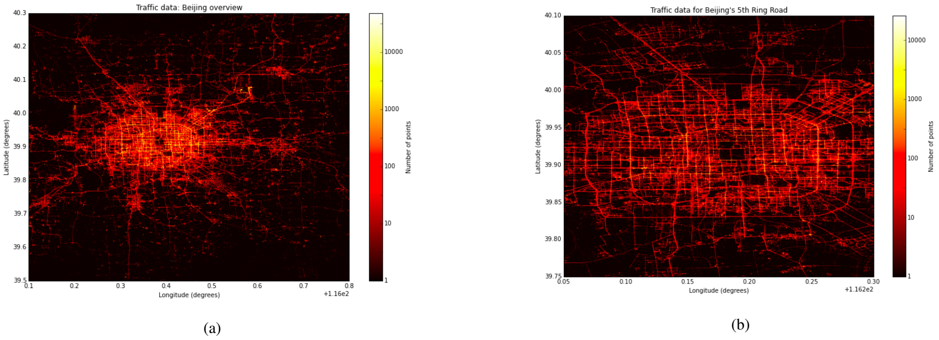
1.2 Problem Statement

We formulate our research question as: *How much user information is leaked after a trajectory dataset has been protected with a Location Privacy Preservation Mechanism.* The research tries to follow two approaches to understand how much privacy is leaked. The publisher's approach focuses on collecting the confidential dataset, that is, mobility/ trajectory dataset, processing it into a trace and using the LPPM tool [7] and its location preserving methods to protect identity of a user before publishing. In the adversarial approach, we will take the role of an attacker to recover the original trajectories given the anonymized trajectory dataset and mobility profiles of individuals. We then quantify the privacy leakage by calculating the similarity between the reconstructed traces and actual traces. At the end of the research, we hope to have a quantitative analysis of the level of privacy leakage after applying certain privacy preserving mechanisms. The scope of the research will be limited to urban trajectories. We hope to make the following contributions by the end of the research:

- Identify the right metric to measure privacy of users in an anonymized mobility dataset
- Quantify privacy for users in a published anonymized mobility dataset
- Identify the best set of parameters that can be used to promote maximum privacy for the mobility dataset

2 Data

In our work we are using the following two trajectory datasets:



(a)

(b)

Figure 1: T-Drive Trajectory data of all taxis a) Entire City b) Beijing's 5th Ring Road

1. **EPFL**[4]: This dataset provides mobility traces of cabs in San Francisco, USA. It contains GPS coordinates of 536 taxis collected over 30 days. GPS mobility traces of this dataset were collected and aggregated in May 2008. Trace format for this dataset is- [latitude, longitude, occupancy, time]. Time is in UNIX epoch format, latitude longitude in decimal degrees, and occupancy is either 0 or 1 denoting if the cab is occupied or not.

2. **T-Drive Trajectory**[11][9][10]: This is a trajectory dataset containing GPS traces of taxis for the city of Beijing provisioned by T-Drive and Microsoft Research License Agreement (MSR-LA). It contains week-long trajectories of 10,537 taxis with around 15 million data points recorded during Feb 2, 2008 to Feb 8, 2008.

3 Methodology

The project aims to quantify the user location privacy of a given mobility dataset through adversarial attacks against different standard Location Privacy Preserving Mechanisms(LPPM). The original datasets will be privacy protected using different LPPMs and then conduct attacks against the protected datasets using different statistical inference algorithms. The privacy of a mechanism is evaluated based on the correctness of the attack i.e based on how much of the original trace was reconstructed by the attack.

3.1 Location Privacy Preserving Mechanism

The mechanism that performs the modification in order to protect the user's location-privacy is called an LPPM[7]. A mobility dataset is modelled as a set of users U , each moving in an area partitioned into a set of regions R, and observed in a time interval T. An event is defined as triplet $\langle u, r, t \rangle$, where $u \epsilon U$, $r \epsilon R$ and $t \epsilon T$. A user's trace can be viewed as a vector of these events . A location-privacy mechanism receives a user trace and perturb based on two steps: (1) Location obfuscation and (2) Anonymization. The goal is to hide user information without destroying too much data utility.

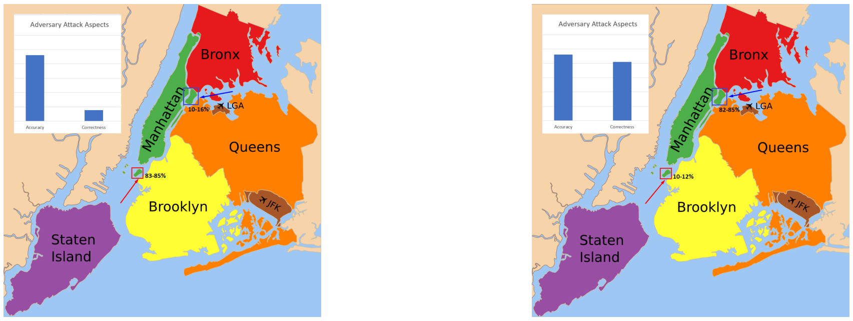
In the obfuscation process, location of each event is replaced with a location pseudonym constructed with an obfuscation function (f). LPPM includes methods like (i). perturbation (adding noise to the location), (ii) adding dummy regions, (iii) reducing location precision and (iv) location hiding. All of these methods can be viewed as the function (f) that maps each event $\langle u, r, t \rangle$ to $\langle u, r', t \rangle$ where $r' \epsilon R'$. For the anonymization process, the research will leverage on random permutation, the anonymization mechanism(g) provided by the LPPM tool, by which each user is represented with a pseudonym set $U' \epsilon \{1, 2, ..N\}$, a permutation of the users is randomly chosen from $N!$ permutations. Each user's pseudonym is his position in the permutation.

Any LPPM is a pair (f, g). Given an actual trace $a \epsilon \{\langle u, r, t \rangle\}$ the goal of an LPPM is to produce an observed trace $o \epsilon \{\langle u', r', t' \rangle\}$

3.2 Adversarial attack

An adversary is modelled by his knowledge and attacks. It is assumed that the adversary will have a knowledge about the obfuscation function and the anonymization function (f, g). Also, the adversary will have a knowledge about some part of the actual trace. An adversary's approach can be divided into three steps. (i) knowledge construction (ii) de-obfuscation (iii) tracing attack. The scope of the tracing attack in the research will be limited to maximum likelihood tracing i.e reconstructing the entire trajectory. The reconstruction will be carried out in two steps, deobfuscation and deanonymization. De-obfuscation will be done using Forward-Backward algorithm[6] coupled with Hungarian algorithm[3]. Deanonymization of the trace will be conducted using viterbi's algorithm[6]. The most likely actual trace is generated by keeping track of the region that maximizes the joint probability at each step of the algorithm.

3.3 Evaluation



(a)

(b)

Figure 2: Dilineating correctness as a metric. a) High accuracy – Low correctness. Correct location is Randall's island. Accuracy without correctness is insignificant. b) High accuracy – High correctness. Correct location is Randall's island

The evaluation of the reconstructed trace can be done using three evaluation metrics:

1. Anonymity: Anonymity measures how successful the adversary is in de-anonymizing the user
2. Entropy/ Certainty: Entropy is used to measure the certainty of the observed traces. The entropy shows how uniform and concentrated the estimated distribution is.
3. Distortion/ Correctness: Distortion measures the location privacy of the user for every time instance. [7] There are two types of distortion:
    (a) Location Distortion: It measures location privacy of the user for every time instance with respect to the most likely location of the user
    (b) Trace Distortion: It measures location privacy of the user at every time instance with respect to the most likely trace of the user

3.4 Experiments

(a) Generation of Actual Trace:

The location trace for 10 mobile users on each day was generated from the epfl/mobility dataset [4]. The actual trace was generated from the actual raw data using the trace generator tool for a particular date setting.

(b) Knowledge Construction: We constructed the knowledge or the mobility profiles for each user using LPPM tool. These mobility profiles represent the knowledge that adversary collects for each simulated user.

(c) Generation of Learning Trace:

The learning trace is the knowledge that the adversary has about the users. For our experiments to emulate stronger adversaries we sample the actual trace of users. The adversary extracts information from this learning trace to build mobility profiles and reconstruct the entire trace for each user. Learning trace was generated from the actual trace using the formula:

$$Learning\ Trace = Sampling\ Rate * Total\ Actual\ Trace \tag{1}$$

In the experiments, we used 5 different sampling rates (SR) levels: 20 % SR, 40 % SR, 60 % SR, 80 % SR, 100 % SR. Each levels represents what is the percentage of actual trace that is known to the adversary. For example, 20 % sampling rate implies the attacker has 20 % of actual trace of the user.

An ablation study was conducted to understand the impact of each hyper parameter of the privacy preserving mechanism. Different experiments were conducted to understand how to process the actual data into a trace:

i. Number of Regions: The number of regions of the entire research area is divided. For example, we can divide the entire area into 10 or 20 regions etc. Herein, we vary these regions from 10 to 100 (in steps of 10) to understand how it influences the privacy.

ii. Amount of prior knowledge of the user: Prior knowledge implies the learning trace that the adversary has or the sampling level of the actual trace. We varied the sampling rate of the user from 20 % to 100 % (in steps of 20) keeping the other parameters constant.

(d) Applying the LPPM Mechanism: The LPPM mechanism is used to convert the actual traces into observed traces. These observed traces are then published online.

(e) Reconstructing the Trace: Using the learning trace and prior knowledge of the adversary, we acted as an adversary and performed an adversarial attack to generate reconstructed trace for each user.

(f) Identifying metrics to measure privacy of mobility dataset: The metrics measure similarity between the most likely trace generated by adversary and the actual trace. We identified three metrics to evaluate the privacy of datasets: anonymity, entropy and distortion. Anonymity gives the anonymized data and entropy is used to quantify the certainty, that is, the unique values which the observed traces point to. These metrics do not tell us how similar the adversary's reconstructed trace is to the actual trace. Hence, for evaluation of similarity of privacy of our datasets, we use distortion/ correctness.

These experiments were conducted to understand the impact of the following LPPM parameters:

i. Number of regions to be obfuscated: Impact of changing the values of the number of regions that were obfuscated in our observed trace was observed. The values ranged from 0 to 2^m where m denotes the level of obfuscation.

ii. Probability of hiding a particular event in the observed trace: The impact of hiding the location of the events in our observed trace was studied. The values were varied from 0 to 1 keeping other lppm parameters constant.

iii. Probability of injecting noise into the observed trace: Impact of injecting noise into our observed trace was observed. The values were varied from 0 to 1 keeping other lppm parameters constant.

4 Results

1. Quantifying Privacy for Users:

Distortion/Correctness is the metric that quantifies the success of the attack. We use Correctness as the metric to evaluate the privacy of users. Correctness is quantified by measuring the distance between our outcome x and the actual outcome x_c. x_c is the value that we want to hide from the adversary. The distance d between the actual and likely outcome can be represented as :

$$\sum_{x} Pr(x|o)||x - x_c|| \qquad (2)$$

This is also known as the estimation error of the adversary. Here, $Pr(x|o)$ denotes the probability of error of adversary.

If distance $d = 0$ then $x = x_c$ implies that estimated error of adversary is very low , the correctness value is very high, and adversary has complete knowledge and is able to identify the actual trace of user.

Hence, we can say that correctness gives us how close the adversary is to getting to actual trace. For example, if $d = 0$ if and only if $x = x_c$ and $d = 1$ otherwise for each event in the trace; then the Incorrectness can be defines as $1 - Pr(x|o)$. The incorrectness of the user is directly proportional to the privacy. The more incorrect the adversary, the better is the privacy of the user.

2. Parameters to promote maximum privacy for the mobility dataset:

   (a) Number of Regions with respect to Privacy: As illustrated in Figure 3, as the number of regions increase the privacy of the users, that is, the incorrectness of the adversary also increases.

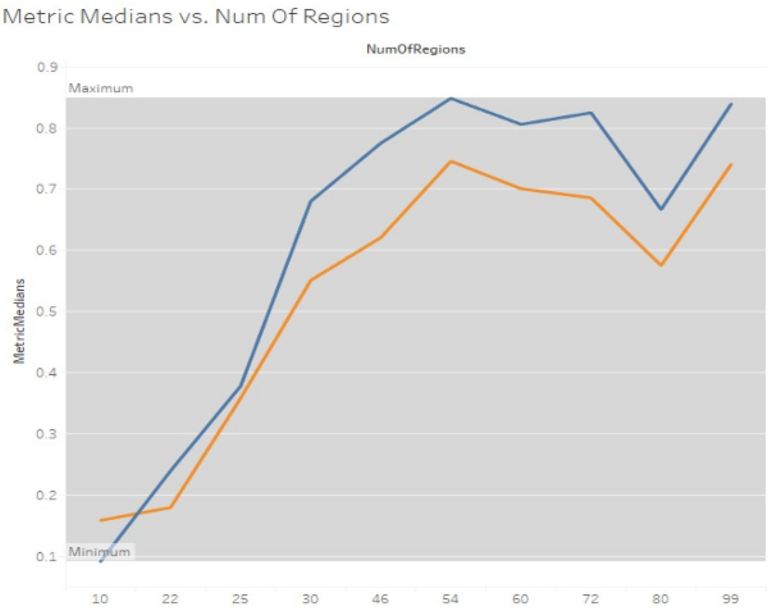


Figure 3: Number of Regions with respect to Privacy

   (b) Sampling level with respect to Privacy: With an increase in the amount of prior knowledge of the adversary, the privacy of the data set increases. Figure 4 illustrates, as sampling rate (knowledge of user) increases, there is a decrease in correctness/ distortion, that is, the adversary predicts the likelihood of a particular user at a particular time and region with more certainty.

   Various other parameters are studied, such as, obfuscation level with respect to Privacy as illustrated in Figure 5. Increase in obfuscation level causes increase in the privacy of users. Similarly, Fake injection probability vs Distortion. Fake injection can be in the form of adding noise to our events in observed trace. As the fake injection probability increases, the privacy of the user increases.

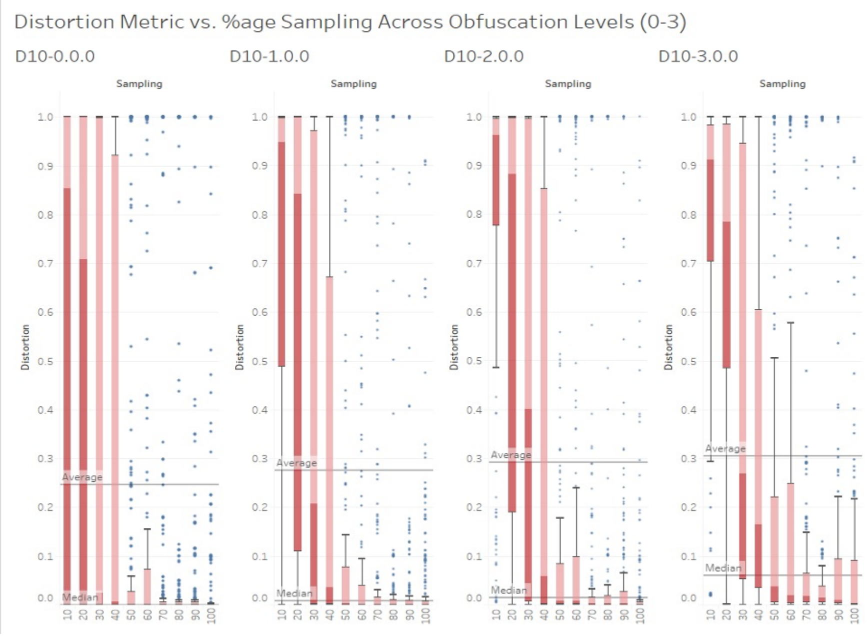


Figure 4: Sampling level with respect to Privacy

5 Conclusion

In this paper, we have quantified the privacy of users for mobility datasets. For quantifying, we have followed two approaches. In the publisher approach, LPPM mechanism was used to generate observed trace which are published online. In the adversarial approach, adversarial attack was performed to reconstruct actual trace of user. We automated the trace generation process, as a realization of our framework. For evaluation, we used correctness to measure privacy of users. We also tuned the parameters to achieve increasing privacy. From results, we concluded that with an increase in number of regions being obfuscated, hidden and increase in the noise being injected, the privacy of user increases. However, one question remains "How usable the dataset will be to any researcher if the actual information is being hidden and distorted". The scope of our project was limited to reconstruction of the adversary's trace. The usability of datasets after applying these LPPM mechanisms can be done as a part of future work.

References

[1] M Decker. “Location privacy-an overview”. In: *ICMB ’08: Proceedings of the 2008 7th International Conference on Mobile Business* 2 (2008), pp. 221–230.

[2] M Duckham. “Moving forward: location privacy and location awareness”. In: *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS* (2010), pp. 1–3.

[3] Harold W Kuhn. “The Hungarian method for the assignment problem”. In: *Naval research logistics quarterly* 2.1-2 (1955), pp. 83–97.

[4] Michal Piorkowski, Natasa Sarafijanovic-Djukic, and Matthias Grossglauser. *CRAWDAD dataset epfl/mobility (v. 2009-02-24)*. Downloaded from `https://crawdad.org/epfl/mobility/20090224`. Feb. 2009. DOI: `10.15783/C7J010`.

[5] J. Freudiger R. Shokri and J.-P. Hubaux. “A unified framework for location privacy”. In: *3rd Hot Topics in Privacy Enhancing Technologies (HotPETs)* (2010).

[6] Lawrence R Rabiner. “A tutorial on hidden Markov models and selected applications in speech recognition”. In: *Proceedings of the IEEE* 77.2 (1989), pp. 257–286.

[7] Reza Shokri et al. “Quantifying location privacy”. In: *2011 IEEE symposium on security and privacy*. IEEE. 2011, pp. 247–262.

[8] J. Trotter. “Public nyc taxicab database lets you see how celebrities tip”. en. In: (Oct. 2014). URL: `http://gawker.com/the-public-nyc-taxicab-%5C%5Cdatabase-that-accidentally-track-1646724546`.

[9] Jing Yuan et al. “Driving with knowledge from the physical world”. In: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2011, pp. 316–324.

[10] Jing Yuan et al. “T-drive: driving directions based on taxi trajectories”. In: *Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems*. 2010, pp. 99–108.

[11] Yu Zheng. *T-Drive trajectory data sample*. T-Drive sample dataset. Aug. 2011. URL: `https://www.microsoft.com/en-us/research/publication/t-drive-trajectory-data-sample/`.

Appendix A Team Contribution

- Literature review - Shreeraman AK, Aparna B
- Formulating problem statement - Shreeraman AK, Aparna B, Siqi H, Vivek P
- Dataset identification - Shreeraman AK, Vivek P
- Automating framework- Shreeraman AK, Vivek P
- Experiments - Shreeraman AK, Aparna B, Vivek P, Siqi H
- Evaluation - Vivek P, Shreeraman AK
- Data Visualization - Siqi H, Vivek P, Aparna B
- Trace Generation, Visualization - Siqi H
- Website, Github - Siqi H, Vivek P, Aparna B
- Final progress Report - Aparna B, Shreeraman AK

Appendix B Graphs

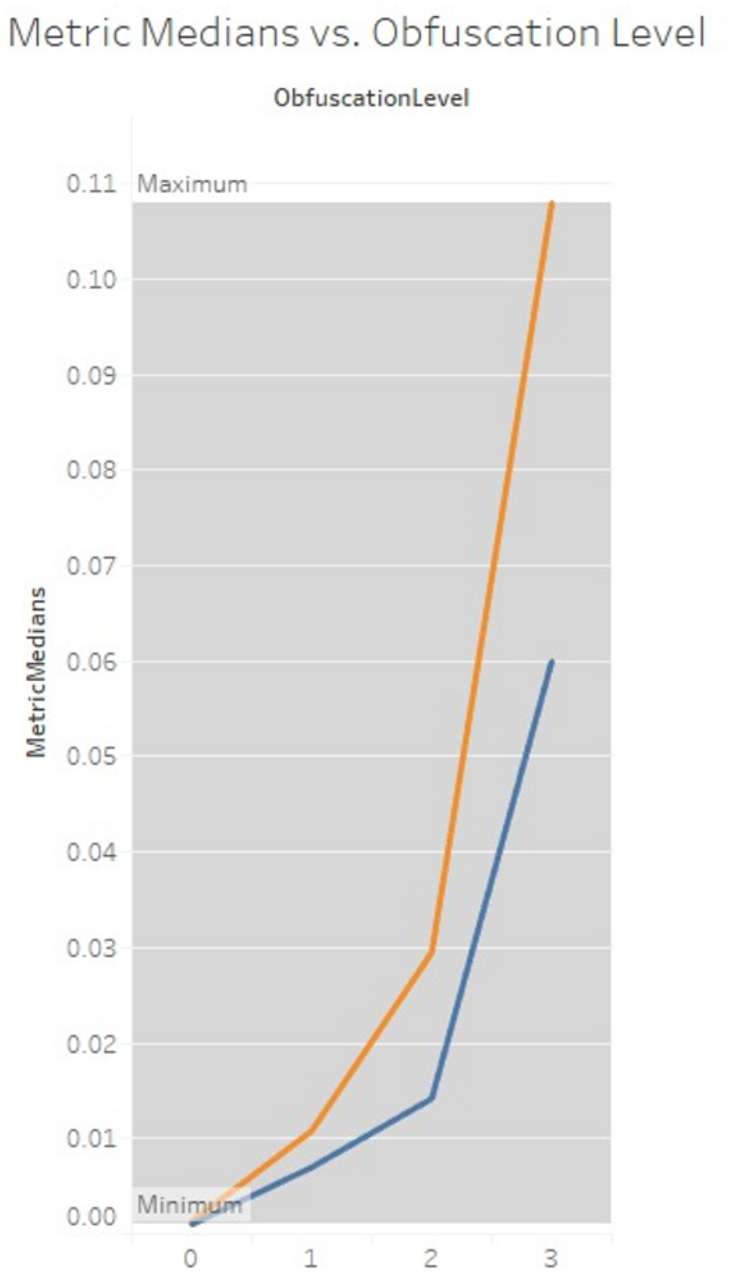


Figure 5: Obfuscation Level with respect to Privacy: Obfuscation level represents the number of regions that are combined and obfuscated in our observed trace. As the obfuscation level increases the privacy of the users, that is, the incorrectness of the adversary also increases

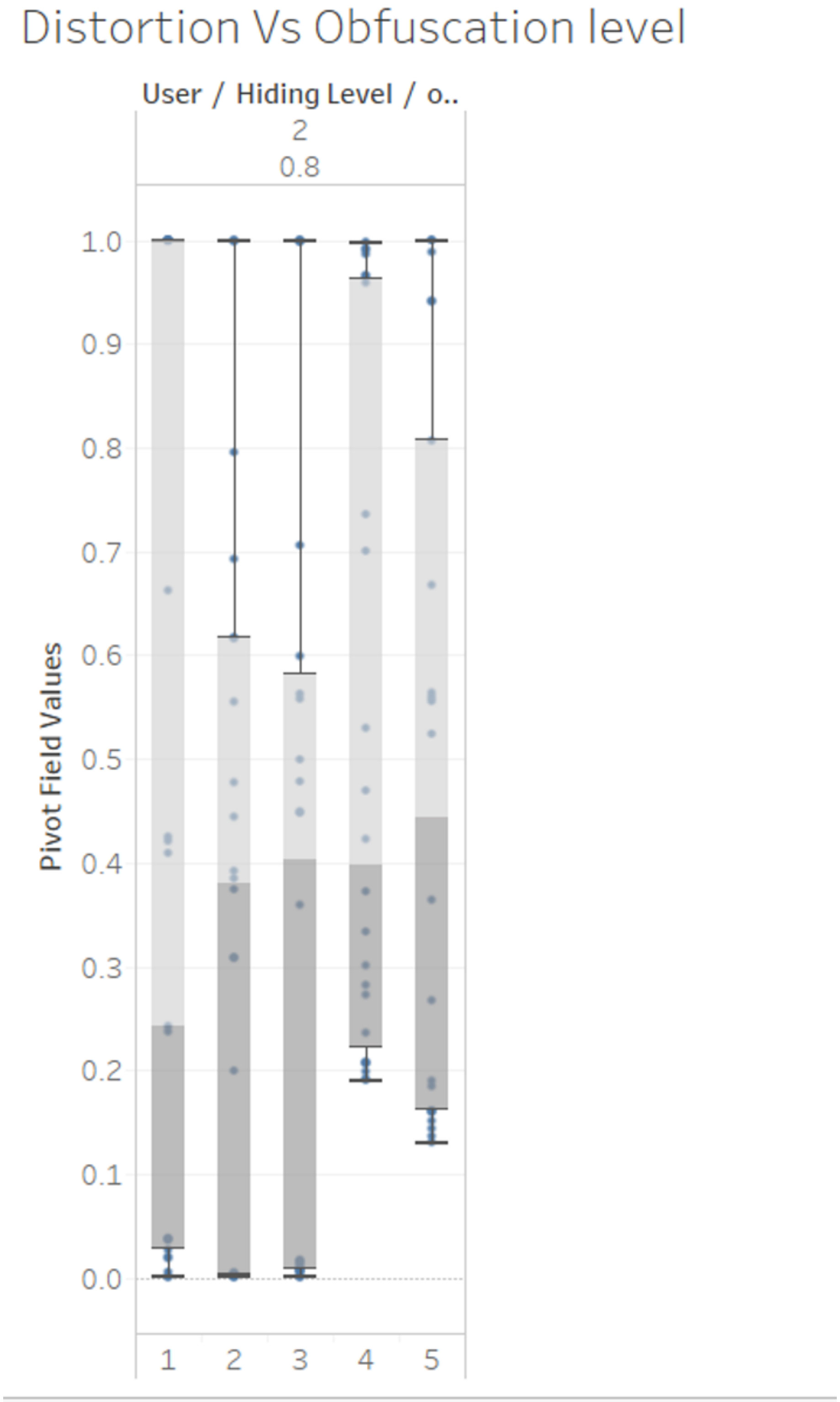


Figure 6: Distortion vs Obfuscation Level: Setting the parameters obfuscation level from 1-5, Fake injection probability=0.8 and hiding level probability=1. We observe that, as distortion increases the privacy also increases

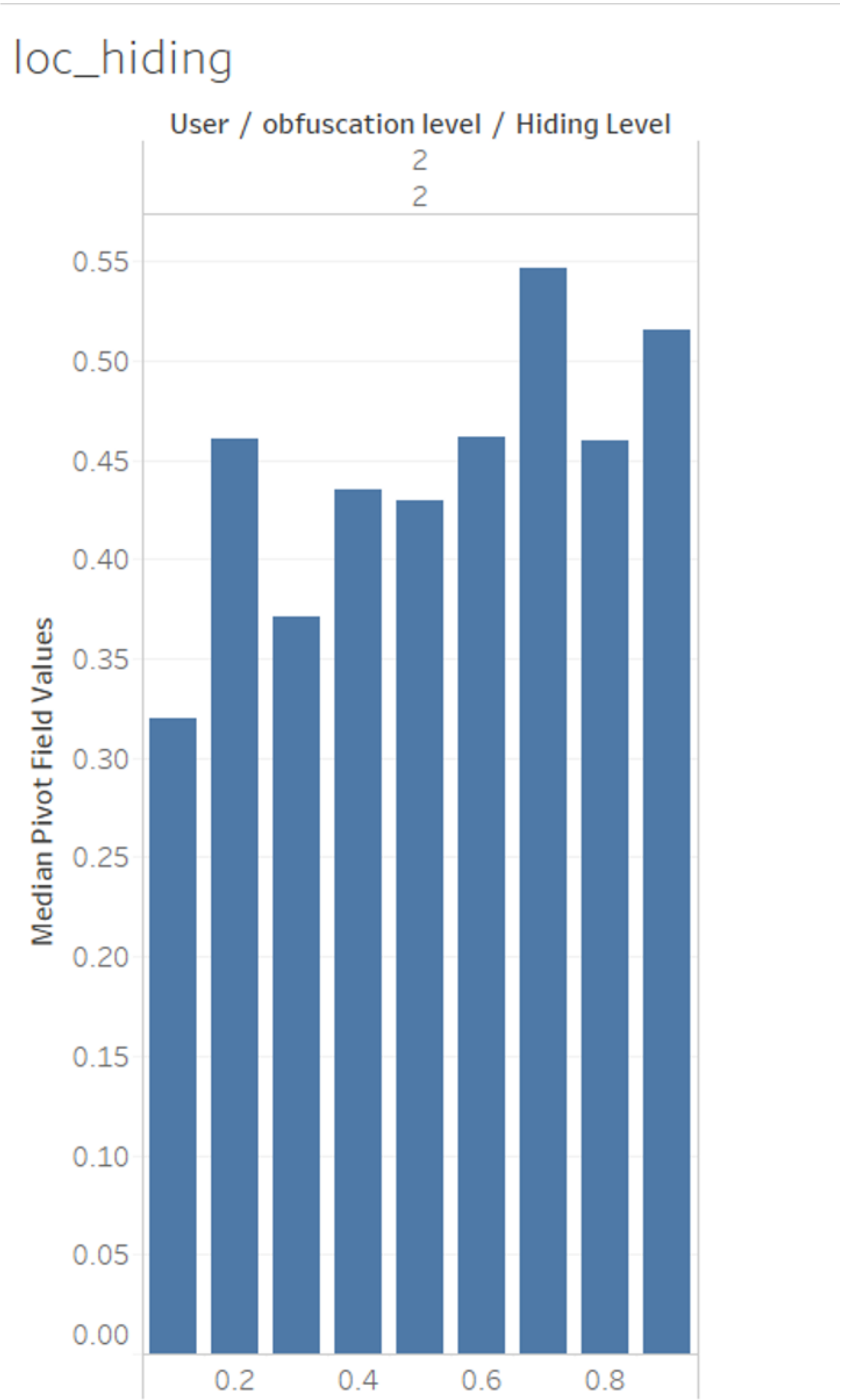


Figure 7: Location hiding level vs Privacy: Location hiding level represents hiding the location of the events in our observed trace. As we increase the location hiding level, the privacy of the user increases

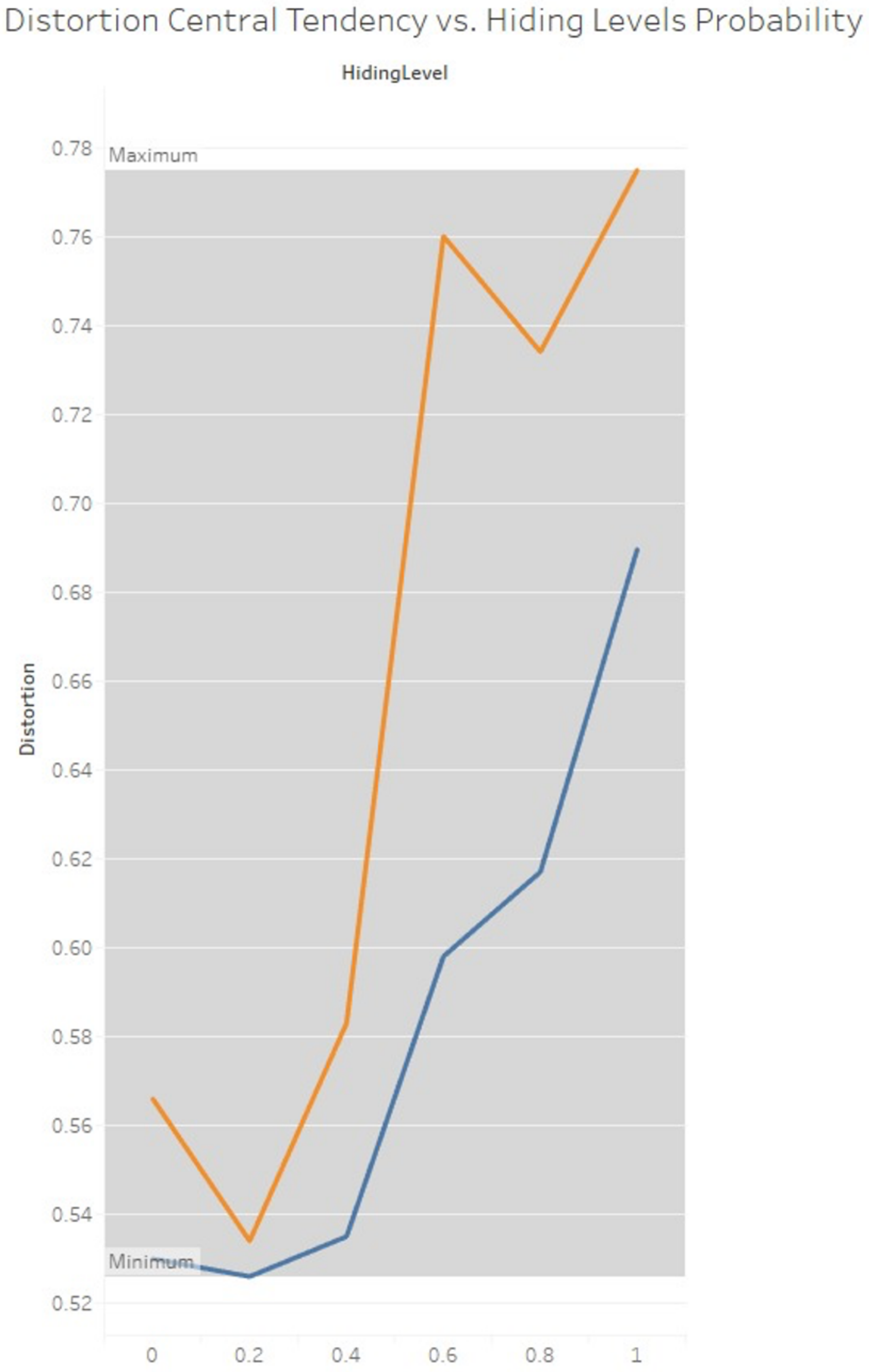


Figure 8: Hiding Level probability vs Central Tendency for Distortion: This gives us the hiding level probability with respect to the distortion. We can see with increase in hiding level probability, the distortion increases and the privacy of the user increases

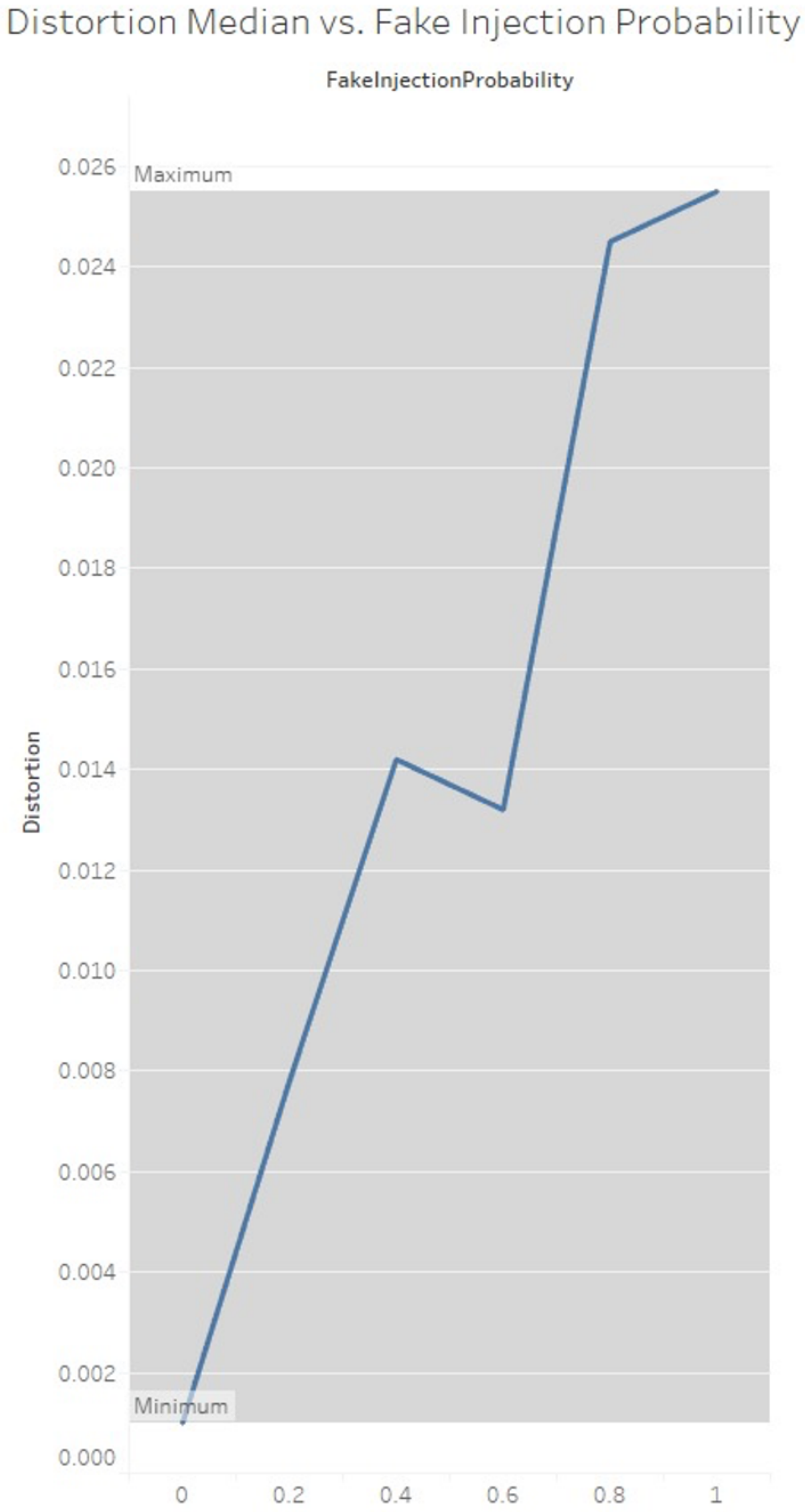


Figure 9: Fake injection probability vs median value for distortion: As the fake injection probability increases, the distortion increases and the privacy of the user increases

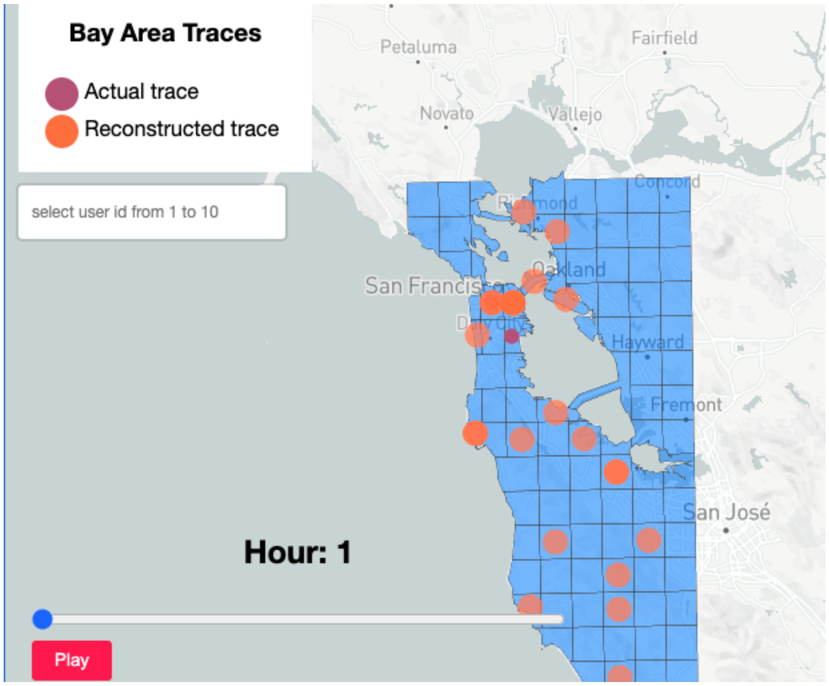


Figure 10: Traces for San Francisco Bay Area - 99 regions

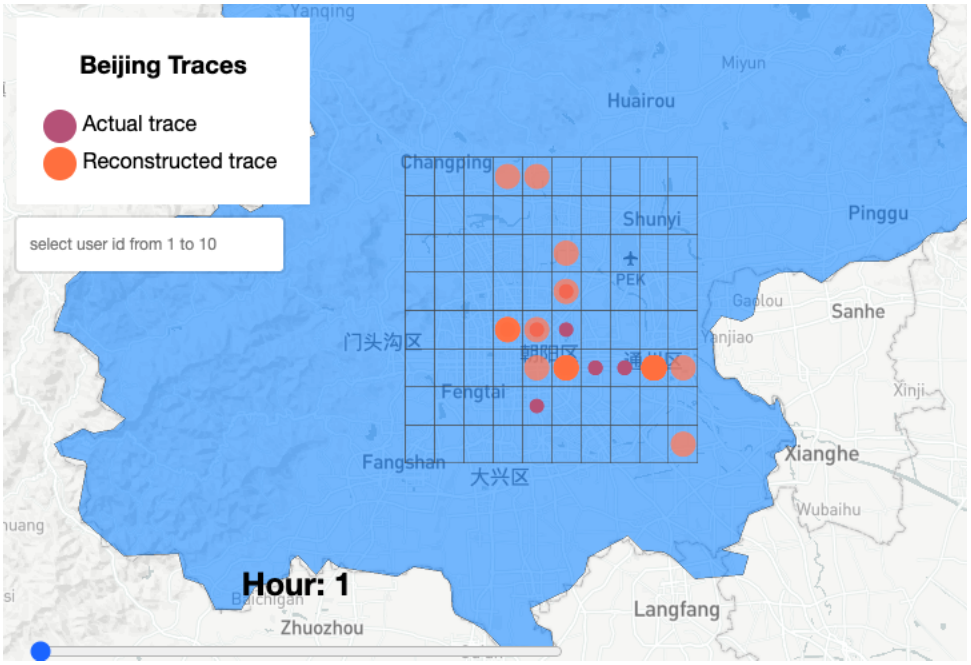


Figure 11: Traces for Beijing Area - 81 regions

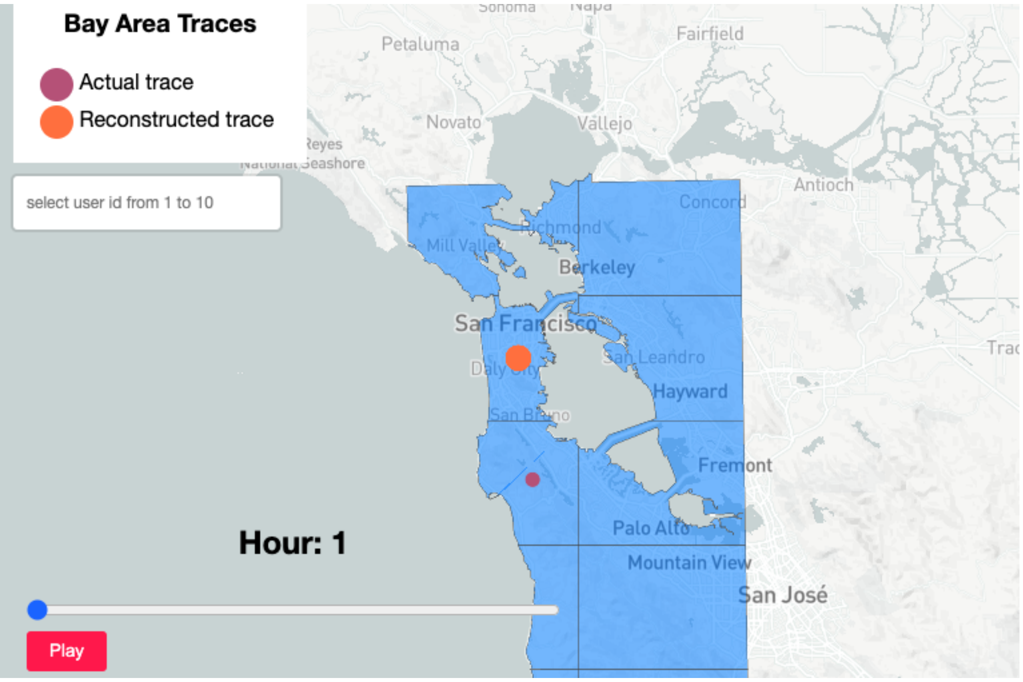


Figure 12: Traces for San Francisco Bay Area - 10 regions

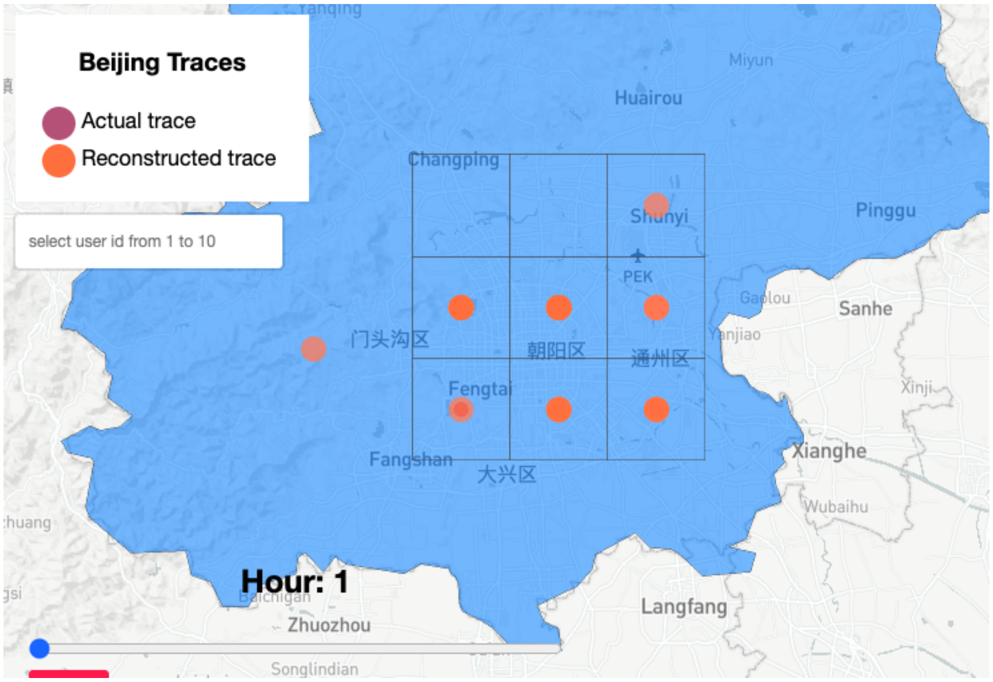


Figure 13: Traces for Beijing Area - 10 regions