PRIVIC: A privacy-preserving method for incremental collection of location data

Sayan Biswas, Catuscia Palamidessi

PRIVIC: A privacy-preserving method for incremental collection of location data

Sayan Biswas
INRIA and LIX, École Polytechnique
Palaiseau, France
sayan.biswas@inria.fr

Catuscia Palamidessi
INRIA and LIX, École Polytechnique
Palaiseau, France
catuscia@lix.polytechnique.fr

ABSTRACT

With recent advancements in technology, the threats of privacy violations of individuals' sensitive data are surging. Location data, in particular, have been shown to carry a substantial amount of sensitive information. A standard method to mitigate the privacy risks for location data consists in adding noise to the true values to achieve geo-indistinguishability (geo-ind). However, geo-ind alone is not sufficient to cover all privacy concerns. In particular, isolated locations are not sufficiently protected by the state-of-the-art Laplace mechanism (LAP) for geo-ind. In this paper, we focus on a mechanism based on the Blahut-Arimoto algorithm (BA) from the rate-distortion theory. We show that BA, in addition to providing geo-ind, enforces an elastic metric that mitigates the problem of isolation. Furthermore, BA provides an optimal trade-off between information leakage and quality of service. We then proceed to study the utility of BA in terms of the statistics that can be derived from the reported data, focusing on the inference of the original distribution. To this purpose, we de-noise the reported data by applying the iterative Bayesian update (IBU), an instance of the expectation-maximization method. It turns out that BA and IBU are dual to each other, and as a result, they work well together, in the sense that the statistical utility of BA is quite good and better than LAP for high privacy levels. Exploiting these properties of BA and IBU, we propose an iterative method, PRIVIC, for a privacy-friendly incremental collection of location data from users by service providers. We illustrate the soundness and functionality of our method both analytically and with experiments.

KEYWORDS

location privacy, geo-indistinguishability, rate-distortion theory, privacy-utility trade-off

1 INTRODUCTION

As the need and development of various kinds of research and analysis using personal data are becoming more and more significant, the risk of privacy violations of sensitive information of the data owners is also increasing manifold. One of the most successful proposals to address the issue of privacy protection is *differential privacy (DP)* [23, 24], a mathematical property that makes it difficult for an attacker to detect the presence of a record in a dataset. This is typically achieved by answering queries performed on the dataset in a (controlled) noisy fashion. Lately, the *local variant of differential privacy (LDP)* [22] has gained popularity due to the fact that the noise is applied at the data owner's end without needing a trusted curator. LDP is particularly suitable for situations where a data owner is a user who communicates her personal data in exchange for some service. One such scenario is the use of location-based services (LBS), where a user typically sends her location in order to obtain information like the shortest path to a destination, nearby points of interest, traffic information, etc. The security and the convenience of implementing the local model directly on a user's device (tablets, smartphones, etc.) make LDP very appealing.

Typically, in exchange for their service, providers incrementally collect their users' data and then make them available to other parties which process them to provide useful statistics to companies and institutions. Obviously, the statistical precision of the collected data is essential for the quality of the analytics performed (*statistical utility*). However, injecting noise locally into the data to protect the privacy of the users usually has a negative effect on the statistical utility. Additionally, the noise degrades the *quality of service* (QoS) as well, since, obviously, the service results from the elaboration of the information received.

Substantial research has been done to address the privacy-utility trade-off in the context of DP. In LDP, the primary focus has been to optimize the utility from the data collector's perspective, i.e., devising mechanisms and post-processing methods that would allow deriving the most accurate statistics from the collection of the noisy data [22, 57]. In contrast, in domains such as location privacy, the focus usually has been on optimizing the QoS, i.e., the utility from the point of view of the users. In particular, this is the case for the framework proposed by Shokri et al. [47, 49].

We argue that it is important to meet the interest of all parties involved, and hence to consider both kinds of utility at the same time. Hence, the first goal of this paper is to develop a *location-privacy preserving mechanism (LPPM)* that, in addition to providing formal location-privacy guarantees, preserves as much as possible *both* the statistical utility and the QoS.

Now, one may think that statistical utility and QoS are aligned since they both benefit from preserving as much original information as possible under the privacy constraint. However, this is not true in general: the optimization of statistical utility does not necessarily imply a significant improvement in the QoS, nor vice-versa. A counterexample is provided by Example 5.1 in Section 5. Hence, the preservation of both statistical utility and QoS is trickier than it may appear at first sight.

One of the approaches which have been proposed to protect location privacy is *geo-indistinguishability* (geo-ind) [4], which essentially obfuscates locations based on the distance between them. This idea works particularly well for protecting the precision of

Proceedings on Privacy Enhancing Technologies 2024(1), 582–596

https://doi.org/10.56553/popets-2024-0033

the location as it ensures that an attacker would not be able to differentiate between points that are close on the map by observing the reported noisy location. At the same time, it does not inject an enormous amount of noise that would be necessary to make far-away locations indistinguishable. Moreover, geo-ind has been shown to formally satisfy the basic sequential compositionality theorem [29], just like DP and its local variant. Although this approach of distance-based obfuscation seems enticing at a first glance, one of the issues it poses is that it may leave the geo-spatially isolated locations vulnerable, i.e., identifiable despite being formally geo-indistinguishable [17]. To improve the situation, [17] introduced the notion of *elastic distinguishability metrics*, which essentially leads to injecting more noise when the location to protect is isolated.

The *Blahut-Arimoto algorithm* (BA) [6, 9] from rate-distortion theory (a branch of information theory) Pareto-optimizes the trade-off between *mutual information* (MI) and average distortion. This property is appealing in the context of privacy because MI is often considered a measure of information leakage and average distortion is a commonly used metric for quantifying QoS. Moreover, BA was proven to satisfy geo-ind in [42] opening a door to study it as a potential LPPM. In this paper, we start off by exploring the privacy-preserving properties of BA and comparing them with those of the *Laplace mechanism* (LAP) [4] which is considered as the state-of-the-art mechanism for geo-ind. We show that, besides geo-ind, BA provides an elastic distinguishability metric and, hence, protects even the most isolated points in the map, unlike LAP. We then examine the statistical utility, focusing on the estimation of the most general statistical information, namely the distribution of the original location data (true distribution). The "best" estimation is known in statistics as the *maximum likelihood estimation* (MLE), and can be computed using the *iterative Bayesian update* (IBU) [2], an instance of the *expectation maximization* (EM) method. We discover a duality between BA and IBU, which in our opinion is quite intriguing, because BA and IBU were developed in different contexts, using different concepts and metrics, and for completely different purposes. We prove experimentally that the statistical utility of BA is very good, i.e., the MLE is very close to the true distribution. We conjecture that this is probably due to the duality between the mechanism that injects the noise (BA) and the one that de-noises the noisy data (IBU). In any case, the experiments show that the statistical utility of BA outperforms that of LAP for high levels of privacy, eventually becoming comparable as the level of privacy decreases.

One important point to note is that BA requires the knowledge of the original distribution to provide the optimal mechanism. When it is fed with only an approximation of the distribution, it only provides an approximated result. We acknowledge that the distribution of the original data is usually off-limits and, even when available, it typically gets outdated over time. In any case, we can soundly assume that it is not available because it is essentially the reason for collecting the data. Hence we have a vicious circle: we want to collect data in a privacy-friendly fashion to estimate the original distribution while wanting to use a privacy mechanism that requires knowing a good approximation of the original distribution. Motivated by this dilemma, we propose PRIVIC, an incremental data collection method providing extensive privacy protection for the users of LBS, while retaining a high utility for both them and the service providers, and ensuring that both parties, acting in their best interest, would benefit from the end mechanism.

Finally, we prove formally the convergence of PRIVIC to the true distribution and illustrate empirically the privacy-utility trade-off of our method. The experiments also demonstrate the efficacy of combining BA and IBU, in that the estimation of the original distribution is very accurate, especially when measured using a notion of distance between distributions compatible with the ground distance used to measure the QoS (e.g., the Earth Mover's distance). All the experiments were performed using real location data from the Gowalla dataset for Paris and San Francisco.

Contributions. The key contributions of this paper are:

(1) We show, analytically and with experiments on real datasets, that the BA mechanism, in addition to geo-ind, provides an elastic distinguishability metric. As such, it protects the privacy of isolated locations, which the standard LAP for geo-ind fails at.
(2) We prove that BA produces an invertible mechanism, which means that the MLE is unique. This is crucial to prove that the IBU always converges to the true distribution and that, therefore, we can get a good statistical utility.
(3) We establish a duality between BA and IBU, thus demonstrating a connection between rate-distortion theory and the expectation-maximization method from statistics.
(4) We show experimentally that BA provides a better statistical utility than LAP for high levels of privacy, eventually becoming comparable as the level of privacy decreases.
(5) Since the construction of the optimal BA requires precise knowledge of the true distribution, we propose an iterative method (PRIVIC) that alternates between BA and IBU, thus getting a better and better estimation of the true distribution as more (noisy) data get collected. We show, both formally and with experiments on real location datasets, that PRIVIC converges to the true distribution. In summary, PRIVIC produces a geo-indistinguishable LPPM with an elastic distinguishability metric, which optimizes the trade-off with the QoS and provides high statistical utility.
(6) We investigate the effect on the privacy guarantees of our method by considering adversarial users who report their locations falsely to compromise the privacy of the isolated locations in the map.

Related Work. The trade-off between privacy and utility has been widely studied in the literature [11, 36]. Optimization techniques for DP and utility for statistical databases have been analyzed by the community from various perspectives [30, 31, 40]. There have been works focusing on devising privacy mechanisms that are optimal to limit the privacy risk against Bayesian inference attacks while maximizing the utility [47, 49]. In [42], Oya et al. examine an optimal LPPM w.r.t. various privacy and utility metrics for the user.

In [43], Oya et al. consider the optimal LPPM proposed by Shokri et al. in [49] which maximizes a notion of privacy (the *adversarial error*) under some bound on the QoS. The construction of the optimal LPPM requires the knowledge of the original distribution, and [43] uses the EM method to estimate it and design *blank-slate models* empirically shown to outperform the traditional hardwired models. However, a problem with their approach is that there may

exist LPPMs that are optimal in the sense of [49], but with no statistical utility, see Example 5.1 in Section 5. Furthermore, for the mechanisms considered in [43] the EM method may fail to converge to the true distribution. Indeed, [25] points out various mistakes in the results of [2], on which [43] intrinsically relies to prove the convergence of their method.

[44] proposed a method for generating privacy mechanisms that tend to minimize mutual information using an ML-based approach. However, this work assumes the knowledge of the exact prior from the beginning, unlike ours. Moreover, [44] does not provide formal guarantees for location privacy (e.g., geo-ind) which is one of the main aspects captured by our work. In [55], Zhang et al. consider the Blahut-Arimoto algorithm in the context of location privacy. However, their proposed method also requires the knowledge of the prior distribution to construct the LPPM. Additionally, [55] focuses on measuring privacy for the trace of a single user. On the contrary, our notion of privacy assumes the collection of single check-ins (or check-ins separated in time) by a set of users.

The Laplace mechanism has been rigorously studied in the literature in various scenarios as the cutting-edge standard to achieve geo-ind [4, 7, 29] and has been proven to be optimal for one-dimensional data w.r.t. Bayesian utility [27]. Despite its wide popularity, it has been recently criticized due to its limitation to protect geo-spatially isolated points from being identified by adversaries [17]. The authors of [17] addressed this concern by proposing the idea of *elastic distinguishability metrics*.

Our paper also considers mutual information (MI) as an additional privacy guarantee. MI and its closely related variants (e.g. conditional entropy) have been shown to nurture a compatible relationship with DP [21]. MI measures the correlation between observations and secrets, and its use as a privacy metric is widespread in the literature. Some key examples are: gauging anonymity [16, 56], estimating privacy in training ML models with a typical cross-entropy loss function [1, 35, 44, 51], and assessing location-privacy [42].

A popular choice of utility metric for the users is the *average distortion*, which quantifies the expected quality loss of the service due to the noise induced by the mechanism. Such a metric has gained the spotlight in the community [4, 10, 15, 18, 49] due to its intuitive and simple nature. On the other hand, a standard notion of statistical utility for the data consumer is the precision of the estimation of the distribution on the original data from that of the noisy data. Iterative Bayesian update [2, 3] provides one of the most flexible and powerful estimation techniques and has recently become in the focus of the community [25, 26].

Incremental and privacy-friendly data collection has been explored both in the context of k-anonymity [5, 12, 13] and DP [32, 52]. However, to the best of our knowledge, the problem of providing a rather robust privacy guarantee while preserving utility for both data owners and data consumers has not been addressed by the community so far.

Plan of the paper. Section 2 introduces preliminary ideas from the literature relevant to this work. Section 3 highlights BA as an LPPM because of its extensive privacy-preserving properties. Section 4 establishes the duality between BA and IBU. Section 5 explains our proposed method (PRIVIC). Section 6 exhibits the working of PRIVIC with experiments using real locations from the Gowalla dataset illustrating the convergence of our method. Section 7 discusses and illustrates with experiments the vulnerability of PRIVIC under adversarial data submission and Section 8 concludes. Appendices A and B contain the proofs of the theorems derived in the paper and the relevant tables supporting the experimental analysis of PRIVIC, respectively.

The code used for implementing our mechanism for experiments is available at `https://anonymous.4open.science/r/PRIVIC`.

2 PRELIMINARIES

2.1 Standards of privacy

Definition 2.1 (*d-privacy*, a.k.a. *metric privacy* [14])**.** For any space $\mathcal{X}$ equipped with a metric $d : \mathcal{X}^2 \mapsto \mathbb{R}_{\geq 0}$ and an output space $\mathcal{Y}$, a mechanism $\mathcal{R} : \mathcal{X} \mapsto \mathcal{Y}$ is *ϵ-d-private* if $\mathbb{P}[\mathcal{R}(x) = y] \leq e^{\epsilon d(x,x')} \mathbb{P}[\mathcal{R}(x') = y]$ for every $x, x' \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Note that:

- Setting d as the *discrete metric* on any $\mathcal{X}$, we obtain the definition of *local differential privacy (LDP)* [22].
- Setting $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$ and d as the *Euclidean metric*, we get the definition of *geo-ind* [4].

Definition 2.2 (Mutual information[46])**.** Let (X, Y) be a pair of random variables defined over the discrete space $\mathcal{X} \times \mathcal{Y}$ such that μ is the joint *probability mass function* (PMF) of X and Y, and p_X and p_Y are the marginal PMFs of X and Y, respectively, and $p_{X|Y}$ is the conditional probability of X given Y. Then the (Shannon) *entropy* of X, $H(X)$, is defined as $H(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x)$. The *residual entropy* of X given Y is defined as $H(X|Y) = \sum_{y \in \mathcal{Y}} p_Y(y) H(X|Y = y) = -\sum_{y \in \mathcal{Y}} p_Y(y) \sum_{x \in \mathcal{X}} p_{X|Y} p(x|y) \log p_{X|Y}(x|y)$, and, finally, the *mutual information (MI)* is given by:

$$I(X|Y) = H(X) - H(X|Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \mu(x, y) \log \frac{\mu(x, y)}{p_X(x) p_Y(y)}$$

Remark 1. MI has often been used as a notion of privacy (and security) in the literature. In particular, [38] has provided an operational interpretation of MI in terms of an attacker model. On the other hand, other researchers have strongly criticized the use of Shannon entropy and MI as measures of privacy, see for example [50].

We do not take sides in this controversy: for us, MI is only a means to construct a mechanism that provides geo-ind under an elastic metric, which is our reference privacy notion.

2.2 Notions of utility

Definition 2.3 (Quality of service)**.** For discrete spaces $\mathcal{X}$ and $\mathcal{Y}$, let $d: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_{\geq 0}$ be any distortion metric (a generalization of the notion of distance). Let X be a random variable on $\mathcal{X}$ with PMF p_X and C be any randomizing mechanism where C_{xy} is the probability of x being mapped by C into y. We define the *quality of service (QoS)* of X for C as the *average distortion w.r.t. d*, given as:

$$AvgD(X, C, d) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_X(x) C_{xy} d(x, y)$$

Definition 2.4 (Full-support probability distribution)**.** Let θ be a probability distribution defined on the space $\mathcal{X}$. θ is a *full-support* distribution on $\mathcal{X}$ if $\theta(x) > 0$ for every $x \in \mathcal{X}$.

Definition 2.5 (Iterative Bayesian update [2])**.** Let C be a privacy mechanism that locally obfuscates points from the discrete space $\mathcal{X}$ to $\mathcal{Y}$ such that $C_{xy} = \mathbb{P}[y|x]$ for all $x, y \in \mathcal{X}, \mathcal{Y}$. Let $X_1, \ldots, X_n$ be i.i.d. random variables on $\mathcal{X}$ following some PMF $\pi_{\mathcal{X}}$. Let Y_i denote the random variable of the output when X_i is obfuscated with C.

Let $\overline{y} = \{y_1, \ldots, y_n\}$ be a realisation of $\{Y_1, \ldots, Y_n\}$ and $\boldsymbol{q}$ be the empirical distribution obtained by counting the frequencies of each y in $\overline{y}$. The *iterative Bayesian update (IBU)* estimates $\pi_{\mathcal{X}}$ by converging to its maximum likelihood estimate (MLE) with the knowledge of $\boldsymbol{q}$ and C. IBU works as follows:

(1) Start with any full-support PMF θ_0 on $\mathcal{X}$.
(2) Iterate $\theta_{r+1}(x) = \sum_{y \in \mathcal{Y}} \boldsymbol{q}(y) \frac{\theta_r(x) C_{xy}}{\sum_{z \in \mathcal{X}} \theta_r(z) C_{zy}}$ for all $x \in \mathcal{X}$.

The convergence of IBU has been studied in [2, 25]. For a given set of observed locations, the limiting estimate $\hat{\pi}_{\mathcal{X}} = \lim_{r \to \infty} \theta_r$ is well-defined by the privacy mechanism in use, C, and the empirical distribution of the noisy locations, $\boldsymbol{q}$. We will functionally denote $\hat{\pi}_{\mathcal{X}}$ as $\texttt{IBU}(\boldsymbol{q}, C)$.

Next, we recall a generalization of IBU from the literature that we use in this work. Generalized IBU (GIBU) [26] applies IBU in parallel to several empirical distributions derived from the application of (possibly different) obfuscation mechanisms to various sets of samples from the same distribution.

Definition 2.6 (Generalized iterative Bayesian update [26])**.**

Let $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}$, with $\boldsymbol{x}^{(t)} = (x_1^{(t)}, \ldots, x_n^{(t)})$ for every $t \in \{1, \ldots, N\}$, be N datasets s.t. the entries $x_i^{(t)}$ for each $i \in \{1, \ldots, n\}$ are i.i.d. samples from the discrete space $\mathcal{X}$ following the probability distribution $\pi_{\mathcal{X}}$. Let $C^{(1)}, \ldots, C^{(N)}$ be N privacy mechanisms that locally obfuscate points from $\mathcal{X}$ to $\mathcal{Y}$ such that the mechanism $C^{(t)}$ is applied to the dataset $\boldsymbol{x}^{(t)}$ and $C_{xy}^{(t)} = \mathbb{P}[y|x]$ for all $x, y \in \mathcal{X}, \mathcal{Y}$ and $i \in \{1, \ldots, N\}$. Denoting the random variable of the output when $x_i^{(t)}$ is obfuscated with $C^{(t)}$ as $Y_i^{(t)}$, let $(y_1^{(t)}, \ldots, y_n^{(t)})$ be a realisation of $(Y_1^{(t)}, \ldots, Y_n^{(t)})$ for every $t \in \{1, \ldots, N\}$.

Let $\mathcal{G} = \begin{pmatrix} C^{(1)} & \ldots & C^{(N)} \end{pmatrix}$ be referred to as the *combined mechanism* a.k.a. the *output probability matrix* satisfying:

$$\mathcal{G}\left(x, y_i^{(t)}\right) = \mathbb{P}\left[y_i^{(t)} \middle| x\right] = C^{(t)}_{x, y_i^{(t)}}$$
$$\forall x \in \mathcal{X},\ i \in \{1, \ldots, n\}.$$

GIBU estimates $\pi_{\mathcal{X}}$ by converging to the maximum likelihood estimate (MLE) of $\pi_{\mathcal{X}}$ with the knowledge of the noisy data and the obfuscating channels. GIBU works as follows:

(1) Start with any full-support PMF θ_0 on $\mathcal{X}$.
(2) Iterate $\theta_{r+1}(x) = \frac{1}{Nn} \sum_{t=1}^{N} \sum_{i=1}^{n} \frac{\theta_r(x) \mathcal{G}\left(x, y_i^{(t)}\right)}{\sum_{z \in \mathcal{X}} \theta_r(z) \mathcal{G}\left(z, y_i^{(t)}\right)}$ for all $x \in \mathcal{X}$.

Setting $\hat{\pi}_{\mathcal{X}} = \lim_{r \to \infty} \theta_r$ and $\boldsymbol{y}^t = (y_1^{(t)}, \ldots, y_n^{(t)})$, let $\hat{\pi}_{\mathcal{X}}$ (the MLE of the prior obtained with GIBU) be functionally denoted by:

$$\texttt{GIBU}\left(\left(C^{(1)}, \boldsymbol{y}^{(1)}\right), \ldots, \left(C^{(N)}, \boldsymbol{y}^{(N)}\right)\right).$$

Definition 2.7 (Earth mover's distance [37])**.** Let π_1 and π_2 be PMFs defined over a discrete space of locations $\mathcal{X}$. For a metric $d: \mathcal{X}^2 \mapsto \mathbb{R}_{\geq 0}$, the *earth mover's distance (EMD)* (aka the *Kantorovich–Rubinstein metric*) is defined as

$$EMD(\pi_1, \pi_2) = \min_{\mu \in \Pi(\pi_1, \pi_2)} \mu(x, y) d(x, y)$$

where $\Pi(\pi_1, \pi_2)$ is the set of all joint distributions over $\mathcal{X}^2$ such that for any $\eta \in \Pi(\pi_1, \pi_2)$, $\sum_{x \in \mathcal{X}} \eta(x_0, x) = \pi_1(x_0)$ and $\sum_{x \in \mathcal{X}} \eta(x, x_0) = \pi_2(x)$ for every $x_0 \in \mathcal{X}$.

EMD is considered a canonical way to lift a distance on a certain domain to a distance between distributions on the same domain.

Definition 2.8 (Statistical utility)**.** Let C be a privacy mechanism that obfuscates data on the discrete space $\mathcal{X}$. Let $\pi_{\mathcal{X}}$ be the PMF of the original locations and let $\hat{\pi}_{\mathcal{X}}$ be its estimate by IBU. Then we define the *statistical utility* of the mechanism C as $EMD(\hat{\pi}_{\mathcal{X}}, \pi_{\mathcal{X}})$.

2.3 Optimization of MI and QoS

Definition 2.9 (Blahut-Arimoto algorithm [6, 9])**.** Let X be a random variable on the discrete space $\mathcal{X}$ with PMF $\pi_{\mathcal{X}}$ and $C(\mathcal{X}, \mathcal{Y})$ be the space of all mechanisms encoding $\mathcal{X}$ to $\mathcal{Y}$. For a distortion $d: \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}_{\geq 0}$ and fixed $d^* \in \mathbb{R}^+$, we wish to find the mechanism $\hat{C} \in C(\mathcal{X}, \mathcal{Y})$ that minimizes MI given the bound d^* on distortion:

$$\hat{C} = \underset{\substack{C \in C(\mathcal{X}, \mathcal{Y}) \\ AvgD(X, C, d) \leq d^*}}{\arg\min} I(X|Y_{X,C})$$

where, for any $C \in C(\mathcal{X}, \mathcal{Y})$, $Y_{X,C}$ is the random variable on $\mathcal{Y}$ denoting the output of the encoding of X. The *Blahut-Arimoto algorithm (BA)* provides an iterative method to construct $\hat{C}$ as follows:

(1) Start with any full-support PMF c_0 on $\mathcal{X}$ and any $C^{(0)}$.
(2) Iterate:

$$C_{xy}^{(t+1)} = \frac{c_t(y) \exp\{-\beta d(x, y)\}}{\sum_{z \in \mathcal{Y}} c_t(z) \exp\{-\beta d(x, z)\}} \tag{1}$$

$$c_{t+1}(y) = \sum_{x \in \mathcal{X}} \pi_{\mathcal{X}}(x) C_{xy}^{(t+1)} \tag{2}$$

where $\beta > 0$ is the negative of the slope of the *rate-distortion* function $RD(X, d^*) = \min_{C \in C(\mathcal{X}, \mathcal{Y})} I(X|Y_{X,C})$ under $AvgD(X, C, d) \leq d^*$. We call β the *loss parameter*, capturing the role of d^* in BA.

Remark 2. The equations (1) and (2) above define two transformations $\mathcal{F} : \boldsymbol{D}(\mathcal{X}) \to \boldsymbol{C}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{G} : \boldsymbol{C}(\mathcal{X}, \mathcal{Y}) \to \boldsymbol{D}(\mathcal{X})$, where $\boldsymbol{D}(\mathcal{X})$ is the space of distributions on $\mathcal{X}$, so that $C^{(t+1)} = \mathcal{F}(c_t)$ and $c_{t+1} = \mathcal{G}(C^{(t+1)})$.

Remark 3. In [20], Csiszár proved the convergence of BA when $\mathcal{X}$ is finite. The limit $\lim_{n \to \infty} (\mathcal{F} \circ \mathcal{G})^n (C^{(0)})$ is the optimal mechanism $\hat{C}$ (parametrized by β), and it is uniquely determined by the prior $\pi_{\mathcal{X}}$ and by the initial PMF c_0. Note that $\hat{C}$ is a fixpoint of $\mathcal{F} \circ \mathcal{G}$, i.e. $\hat{C} = (\mathcal{F} \circ \mathcal{G})(\hat{C})$, and that $\hat{c} = \mathcal{G}(\hat{C})$ is a fixpoint of $\mathcal{G} \circ \mathcal{F}$.

Remark 4. In [42], Oya et al. proved that, when d is the Euclidean metric, the mechanism $\hat{C}$ obtained from BA with loss parameter β satisfies 2β-geo-ind.

In the context of the location-privacy, as addressed in this work, we obfuscate the original locations to points in the same space and, hence, for the rest of the paper we consider the spaces of the secrets and the noisy locations to be the same, i.e., $\mathcal{X} = \mathcal{Y}$.

Table 1: Key notations

Notation	Meaning
$\mathcal{X}$	Finite space of locations
$\mathcal{C}(\mathcal{X}, \mathcal{Y})$	Space of all mechanisms encoding $\mathcal{X}$ to $\mathcal{Y}$
d^*	Maximum average distortion
β	Loss parameter of RD function
$\pi_{\mathcal{X}}$	distribution of the original locations (true prior)
$\hat{\pi}_{\mathcal{X}}$	Estimation of the true prior
BA	Blahut-Arimoto algorithm
IBU	Iterative Bayesian Update
δ_{BA}	Precision parameter for BA to converge
δ_{IBU}	Precision parameter for IBU to converge
N	Number of iterations of PRIVIC
$\hat{C}_{\text{BA}}(\theta, N)$	Mechanism produced by PRIVIC

3 LOCATION-PRIVACY WITH THE BLAHUT-ARIMOTO ALGORITHM

Definition 2.9 shows that the BA mechanism optimizes between MI and average distortion, which is a standard choice for measuring QoS. Furthermore, Remark 4 formally links the mechanism produced by BA with geo-ind, which is our reference privacy notion.

In this section, we investigate the privacy protection offered by BA beyond geo-ind, study the statistical utility it renders, and compare it with LAP, the canonical mechanism for geo-ind.

3.1 Elastic location-privacy with BA

One of the concerns harboured by geo-ind is that it treats the space in a uniform way, thus making isolated locations vulnerable to an attacker that knows the prior distribution. This issue has been raised and addressed by Chatzikokolakis et al. in [17] where the authors introduce a variant of LAP based on an *elastic distinguishability metrics*, which they refer to as *elastic mechanisms*. Such mechanisms obfuscate locations not only by considering the Euclidean distance between them but also by taking into account an abstract attribute of the reported location, called *mass*, which is a parameter of the definition.

Formally, if $\mathcal{R}_{\text{elas}}$ is an elastic mechanism with privacy parameter ϵ defined on $\mathcal{X}$, then, for all $x, y \in \mathcal{X}$, $\mathcal{R}_{\text{elas}}$ must satisfy:

$$\mathbb{P}[\mathcal{R}_{\text{elas}}(x) = y] \propto \exp\{-\epsilon d_{\text{E}}(x, y)\} \tag{3}$$

$$\mathbb{P}[\mathcal{R}_{\text{elas}}(x) = y] \propto q(y) \tag{4}$$

where q is the probability distribution of the reported locations.

Note that Equations 3 and 4 characterize the properties of an elastic mechanism $\mathcal{R}_{\text{elas}}$, but they *do not define what $\mathcal{R}_{elas}$ exactly is, as a function.* In fact, as a definition, Equation 4 would be circular, since it uses the probability mass q generated by $\mathcal{R}_{\text{elas}}$ without knowing what $\mathcal{R}_{\text{elas}}$ is. As we will see, BA solves this problem by constructing the mechanism $\mathcal{R}_{\text{elas}}$ as a fixpoint of a recursive process starting from a uniform output distribution q. (To be precise the process is mutually recursive, alternating the generation of a new mechanism and a new output distribution, that, in turn, is fed into BA to generate the mechanism at the next step.)

$\mathcal{R}_{\text{elas}}$, unlike LAP, protects a point in a densely populated area (e.g. city) and a geo-spatially isolated point (e.g. island) differently by considering not only the ground distance between the true and the reported locations but also the mass of the reported location. The exact mechanism depends of course on how we define the notion of mass. A natural way, and the most meaningful from the privacy point of view, is to set the mass of y to be the probability to be reported (from any true location x). Under this definition, the interpretation of (4) is in the spirit of obtaining privacy by ensuring that the set of possible true locations (given the reported one) is large. In other words, given a true location x, we tend to report with higher probability those locations y that are reported with high probability from other locations as well so that it becomes harder to re-identify x as the original one. Note that this property is not incompatible with the geo-ind guarantee. However, LAP does not provide it.

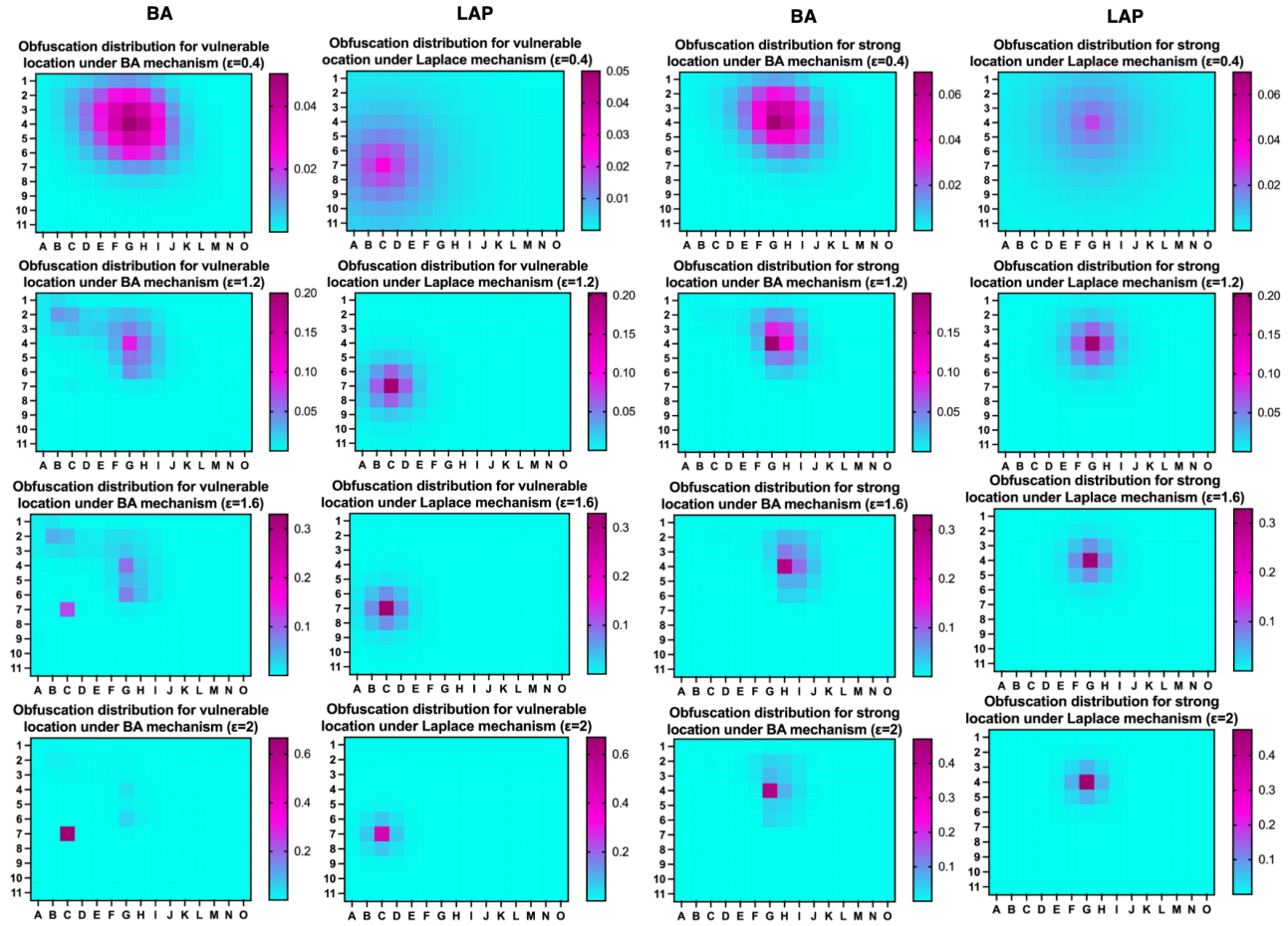
Obviously, the definition of mass as the probability to be reported would be circular, because it would depend on the mechanism, which in turn is defined in terms of the mass. The authors of [17] do not explain how this mechanism could be constructed. Fortunately, the following theorem shows that an elastic mechanism of this kind can be constructed using BA. The proof is provided in Appendix A.

Theorem 3.1. The privacy mechanism generated by BA produces an elastic location-privacy mechanism.

Note also that there can be many mechanisms satisfying (3) and (4) (also with the mass interpreted as probability). The one produced by the BA is the mechanism that offers the best trade-off between QoS and MI among these. Finally, a consequence of the connection with BA is that it provides an understanding of the elastic mechanism in terms of information theory and of the attacker illustrated in the previous section.

Experimental validation. Having furnished the theoretical foundation, we now enable ourselves to empirically validate that BA, indeed, satisfies the properties of the elastic mechanism unlike LAP, its state-of-the-art geo-indistinguishable counterpart. We perform experiments using real location data from the Gowalla dataset [19, 39]. We consider 10,078 Gowalla check-ins from a central part of Paris bounded by latitudes (48.8286, 48.8798) and longitudes (2.2855, 2.3909) covering an area of 8Km×6Km discretized with a 16×12 grid.

In order to demonstrate the property of an elastic mechanism, we artificially introduced an "island" amidst the locations in Paris by choosing a grid A in a low-density area of the dataset (in the south-west region), assigning the probability mass of the grids around A to 0, and dumping this cumulative mass from the surrounding region to A, ensuring that the sum of the probability masses of all the grids remains to be 1. We call A as a *vulnerable location* in the map as it is isolated from the crowded area. To visualize the elastic behaviour of the mechanisms for locations in crowded regions, we consider another grid B in the central part of the map which has a high probability mass and has a highly populated surrounding – we refer to such a grid B as a *strong location* in the map. Figure 1 illustrates the selection of vulnerable and strong locations in the Paris dataset.



(a) Reporting distribution of the vulnerable location A **(b) Reporting distribution of the strong location B**

Figure 2: Distribution of privatizing the vulnerable and the strong locations for different levels of privacy. Top-down, the rows illustrate the results for $\epsilon = 0.4, 1.2, 1.6, 2$, respectively.

For the mechanism derived from BA with a loss parameter β, we know, by Remark 4, that the privacy parameter ϵ is 2β, which we use to tune the privacy level of LAP in order to compare the two mechanisms under the same level of geo-ind. Figure 2 illustrates the probability distribution of reporting a privatized point on the map by obfuscating the vulnerable and the strong locations with different levels of geo-ind – we vary the value of ϵ to be 0.4, 1.2, 1.6, 2.

By comparing with the distribution of the true locations in Paris given by Figure 1, we observe that when the value of ϵ is low (privacy is high), the reported location with BA is likely to be mapped to a nearby densely populated place. For example, with $\epsilon = 0.2$, the highest level of privacy considered in the experiments, the location reported by BA will most probably be around the most crowded region of Paris. As ϵ increases, the location most likely to be reported by BA systematically moves to a densely populated region closer and closer to the true vulnerable location. LAP, on the other hand, always obfuscates every location around its true position in the map – varying the value of ϵ changes the spread of the distribution around the true location. As explained in the introduction, this might be problematic as the vulnerable location is known to be isolated and, hence, even being reported somewhere nearby would potentially result in its re-identification.

For example, we would like to highlight the setting of $\epsilon = 1.6$ for the vulnerable location to show that the distribution of the location reported by LAP is almost completely around the true vulnerable point covering an area that is deserted, i.e., there is no realistic chance of someone being located in that region. Thus, despite providing formal 1.6-geo-ind, LAP fails to protect such a vulnerable location from being potentially identified. BA, on the other hand, does the job quite well, adhering to the principles of the elastic mechanism – it distributes the reported location in the crowded areas nearby providing a sense of camouflage amidst the many possibilities, in addition to 1.6-geo-ind.

In the case of privatizing the strong location, Figure 2b shows that both BA and LAP behave similarly by concealing the point around its true position. This would not give rise to a similar issue as for the vulnerable location because, by definition, the strong location B is already positioned in a highly dense region of the map

and, hence, being privatized, it will still remain among the crowd with a high probability.

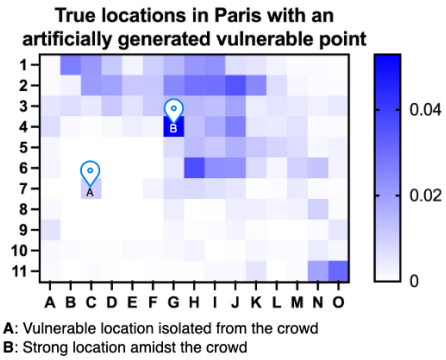


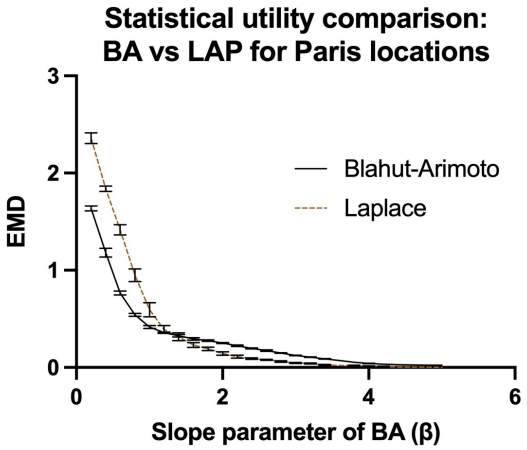
Figure 1: Gowalla check-in locations in Paris with an artificially planted vulnerable point, A, in isolation, and a strong point, B, in a crowded area.

Focusing on the utility of individual users, we note that due to theories from Nash equilibrium [41] and Hotelling's spatial competition [28], a huge fraction of the typical points of interest (POIs) like cinemas, theatres, restaurants, retails, etc. lie in crowded areas syncing with the distribution of population. Therefore, for an isolated point in the map that is located in some extremely unpopulated area (e.g. some forest or island far from the city), the closest POI is usually going to be in the nearest urban region, i.e., a region on the map with a high density of population. Suppose A is one such isolated location and let A_{BA} and A_{LAP} be the reported locations for A obfuscated with BA and LAP, respectively. Due to the elastic property of BA, A_{BA} is likely to be at a nearby crowded location to A, while A_{LAP} is likely to be around the true location A. Let P_{BA} and P_{LAP} be the nearest POIs from the reported locations A_{BA} and A_{LAP}, respectively. The most likely scenario is that P_{BA} and P_{LAP} are almost at a similar place under the assumption that typical POIs follow the distribution of the crowd and, therefore, a vulnerable user has to travel a similar distance from their true position in both the cases, except that under LAP, the privacy of A will be compromised much more than that under BA.

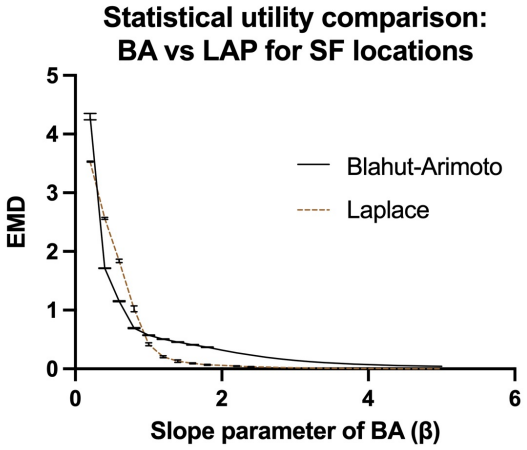
3.2 Statistical utility: BA vs LAP

Now we proceed to empirically compare the statistical utility of BA and LAP by performing experiments on the locations obtained from the Gowalla dataset for two different cities: Paris and San Francisco. In addition to the same setting for the Gowalla check-ins in Paris as considered in the experiments of Section 3.1, here we also test for 123,025 check-in locations from the Gowalla dataset in a northern part of San Francisco bounded by latitudes (37.7228, 37.7946) and longitudes (-122.5153, -122.3789) covering an area of 12Km×8Km discretized with a 24×17 grid. The locations were privatized with BA and LAP under varying levels of privacy – the loss parameter, β, for BA ranged from 0.2 to 5.0, which implies that the value of the geo-ind parameter, ϵ, ranged from 0.4 (very high level of privacy) to 10.0 (almost no privacy). To account for the randomness in the process of generating the sanitized locations, 5 simulations were run for each value of the privacy parameter for obfuscating every location in both datasets.

Figure 3 reveals that BA possesses a significantly better statistical utility than LAP for a high level of privacy (for $\beta \in (0.4, 1.4]$ and $\beta \in (0, 1)$, i.e., ϵ up to 2.8 and 2, in Paris and San Francisco



(a) Statistical utility for BA and Laplace on locations in Paris



(b) Statistical utility for BA and Laplace on locations in SF

Figure 3: Statistical utility in terms of *earth mover's distance* (EMD) between the true and the estimated distributions for locations in Paris and San Francisco, under BA and LAP.

datasets, respectively). As the level of privacy decreases, the EMD of BA becomes worse than that of LAP. We conjecture that this is the price to pay for the added privacy provided by the elasticity of the mechanism. Eventually, the EMD between the true and the estimated PMFs converge to 0 in both mechanisms, as we would expect, fostering the maximum possible statistical utility with, practically, no privacy guarantee.

Summarizing the results from Sections 3.1 and 3.2, we can establish that:

- in addition to providing a formal geo-ind guarantee, BA also gives an LPPM with an elastic distinguishability metric to enhance the privacy of vulnerable locations.
- BA optimizes the trade-off between QoS and MI.
- the statistical utility for high levels of privacy is significantly better for BA than LAP.

Therefore, we conclude that BA is a key contender for providing a comprehensive notion of location privacy while preserving the utility of the data for both the users and the service providers.

4 DUALITY BETWEEN IBU AND BA

We now explore a relationship between BA and IBU which we found rather intriguing. For a metric space $(\mathcal{X}, d)$, let X be a random variable on $\mathcal{X}$ with PMF $\pi_{\mathcal{X}}$. Recalling the iteration of BA from (2) and (1):

$$c_t(y) = \sum_{x\in\mathcal{X}} \pi_{\mathcal{X}}(x) C_{xy}^{(t)} \text{ and } C_{xy}^{t+1} = \frac{c_t(y)\exp\{-\beta d(x,y)\}}{\sum_{z\in\mathcal{X}} c_t(z)\exp\{-\beta d(x,z)\}}$$

Hence, we obtain:

$$c_{t+1}(y) = \sum_{x\in\mathcal{X}} \pi_{\mathcal{X}}(x) C_{xy}^{(t+1)} = \sum_{x\in\mathcal{X}} \pi_{\mathcal{X}}(x) \frac{c_t(y)\exp\{-\beta d(x,y)\}}{\sum_{z\in\mathcal{X}} c_t(z)\exp\{-\beta d(x,z)\}} \tag{5}$$

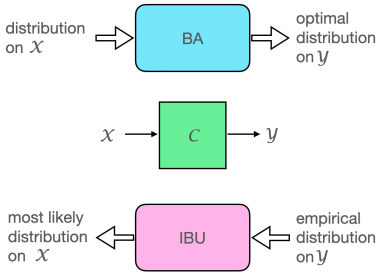


Figure 4: Illustration of the duality between BA and IBU.

Comparing it with the iteration of IBU as in Definition 2.5, we observe that (5) BA is dual to IBU. Indeed, consider an exponential mechanism of the form $C = c\exp\{-\beta d(x,y)\}$. Flipping the roles of x and y in (5), and replacing the input distribution $\pi_{\mathcal{X}}$ with the empirical distribution in output to C, we obtain the iterative step of IBU.

Due to this duality between BA and IBU (illustrated in Figure 4) and taking advantage of the fact that BA converges [20], i.e., $\lim_{t\to\infty} c_t$ exists, we obtain that also IBU converges.

5 PRIVIC: A PRIVACY-PRESERVING METHOD FOR INCREMENTAL DATA COLLECTION

To ensure that the produced mechanism is truly optimal, BA needs a good approximation of the prior distribution. In the beginning, we cannot assume to have such knowledge, but as the service providers incrementally collect data from their users, we can use these data to refine the estimation of the prior and get a better mechanism. These data, however, are obfuscated by the privacy mechanism and, hence, it is not obvious that the estimation of the prior really improves in the process. We show that this is the case, and, summarizing all results obtained for BA so far, we propose a method that facilitates the service providers to incrementally collect data and gradually achieve a high statistical utility with respect to the QoS. We shall refer to our proposed method for **PRIV**acy-preserving **I**ncremental **C**ollection of location data as *PRIVIC*.

The goal of PRIVIC is to construct an obfuscation mechanism that guarantees formal geo-ind, acts as an elastic mechanism, and eventually optimizes between MI and QoS, while producing, at the same time, a good estimation of the distribution on the data.

We shall consider locations sampled from a finite space $\mathcal{X} = \{x_1, \ldots, x_m\}$. Let the *true distribution* or *true PMF* on $\mathcal{X}$ (from which the users' locations are sampled) be $\pi_{\mathcal{X}}$. Note that *we do not assume the knowledge of* $\pi_{\mathcal{X}}$ *in our method*. We assume that the new locations are sampled independently from the previous ones. This hypothesis is reasonable if the collection of the new data is enough separated in time from the previous one, otherwise, we would have a potential correlation between samplings due to the possibility that a user sends repeated check-ins from spatially closed locations. In any case, geo-ind, like DP, satisfies the property of sequential compositionality [29], which means that privacy degradation is under control.

In this work, to achieve geo-ind, we shall adhere to the Euclidean metric d_{E} to measure the ground distance between locations.

PRIVIC proceeds as follows (cf. also Figure 5):

1. Set θ_0, c_0 to be the uniform distributions on $\mathcal{X}$, i.e., $\theta_0(x) = c_0(x) = 1/|\mathcal{X}|$ for all $x \in \mathcal{X}$.
2. In step $t \geq 1$:
   i) For a fixed the maximum average distortion, set $\hat{C}^{(t)} = \mathsf{BA}(\theta_{t-1}, c_0)$.
   ii) Sample a new set of locations $\boldsymbol{x}^{(t)}$ from the (unknown) true distribution and obfuscate them locally by the mechanism $\hat{C}^{(t)}$ to get $\boldsymbol{y}^{(t)}$, thus obtaining the empirical distribution of the reported locations $\boldsymbol{q}_t = \{\boldsymbol{q}_t(x) : x \in \mathcal{X}\}$.
   iii) $\mu_t = \mathsf{IBU}\left(\hat{C}^{(t)}, \theta_{t-1}, \boldsymbol{q}_t\right)$.
   iv) if $t = 1$ then $\theta_t = \mu_t$ else $\theta_t = \mu_t \oplus \theta_{t-1}$ (combination of previous and new estimation proportional to the respective number of samples).
3. $\hat{\pi}_{\mathcal{X}} = \mathsf{GIBU}\left(\left(\hat{C}^{(1)}, \boldsymbol{y}^{(1)}\right), \ldots, \left(\hat{C}^{(N)}, \boldsymbol{y}^{(N)}\right)\right)$.

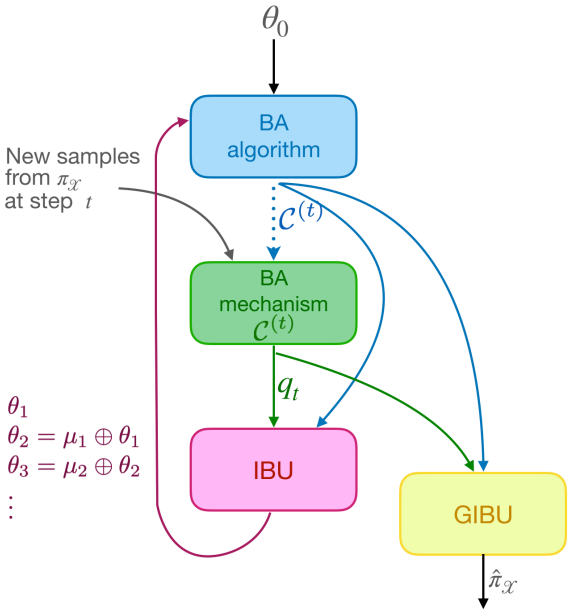


Figure 5: Illustration of the iterative process of PRIVIC

Algorithm 1: PRIVIC

Input: Loss parameter: β, No. of iterations: N, precision of BA: δ_{BA}, precision of IBU: δ_{IBU}, precision of GIBU: δ_{GIBU};
Output: Optimal channel: $\hat{C}$, Estimation of true PMF:$\hat{\pi}_{\mathcal{X}}$;
$\theta_0(x) \leftarrow 1/|\mathcal{X}|$;
$c_0 \leftarrow 1/|\mathcal{X}|$;
$t \leftarrow 0$;
while $t \leq N$ **do**
  $\hat{C}^{(t+1)} = \texttt{BA}(\theta_t, c_0, \beta, \delta_{\text{BA}})$;
  $\boldsymbol{y}^{(t)} \leftarrow \left(y_1^{(t)}, \ldots, y_n^{(t)}\right)$: New noisy locations reported by users after obfuscating their newly sampled true locations with $\hat{C}^{(t)}$;
  $\boldsymbol{q} \leftarrow \{q(x): x \in \mathcal{X}\}$: Empirical PMF obtained from $\mathcal{L}$ by the service provider;
  $\mu \leftarrow \texttt{IBU}(\hat{C}^{(t+1)}, \theta_t, \boldsymbol{q}, \delta_{\text{IBU}})$;
  if $t = 0$ then $\theta_{t+1} \leftarrow \mu$ else $\theta_{t+1} \leftarrow \mu \oplus \theta_t$;
  $t \leftarrow t + 1$;
$\hat{\pi}_{\mathcal{X}} \leftarrow \texttt{GIBU}\left(\left(\hat{C}^{(1)}, \boldsymbol{y}^{(1)}\right), \ldots, \left(\hat{C}^{(N)}, \boldsymbol{y}^{(N)}\right), \delta_{\text{GIBU}}\right)$;
$\hat{C} \leftarrow \text{BA}(\hat{\pi}_{\mathcal{X}}, c_0, \beta, \delta_{\text{BA}})$;
Return: $\hat{C}$, $\hat{\pi}_{\mathcal{X}}$

Algorithm 2: Blahut-Arimoto algorithm (BA)

Input: PMF: π, initial mechanism: $C^{(0)}$, loss parameter: β, precision: δ_{BA};
Output: mechanism giving minimum mutual information for maximum avg. distortion encapsulated by β: $\hat{C}$;
Function $\texttt{BA}(\pi, c_0, \beta, \delta_{BA})$**:**
  $t \leftarrow 0$;
  while $\delta_{BA} \leq |C^{(t)} - C^{(t-1)}|$ **do**
    $C_{xy}^{(t+1)} \leftarrow \frac{c_t(y)\exp\{-\beta d_{\text{E}}(x,y)\}}{\sum_{z\in\mathcal{X}} c_t(z)\exp\{-\beta d_{\text{E}}(z,y)\}}$;
    $c_{t+1}(y) \leftarrow \sum_{x\in\mathcal{X}} \pi(x) C_{xy}^{(t+1)}$;
    $t \leftarrow t + 1$
  $\hat{C} \leftarrow C^{(t)}$;
  Return: $\hat{C}$

Remark 5. The initial distribution θ_0 does not need to be a uniform distribution, any fully-supported distribution would suffice for the process to eventually converge to an optimal mechanism. However, starting with a uniform distribution allows us to avoid any bias in the mechanisms produced in the intermediate steps.

Remark 6. We believe that the last step (3) is not really necessary: The combination of all estimations should already be the MLE of the true distribution, and this is also what we have witnessed in the experiments. However, applying this last step allows us to *formally prove* the converge to the MLE, using the results for GIBU in [26].

In the practical implementation of PRIVIC, we use the precision parameters δ_{BA}, δ_{IBU}, and δ_{GIBU} to set the threshold of empirical convergence of BA, IBU, and GIBU, respectively. Let the privacy mechanism generated this way after N iterations, for fixed parameters c_0, β, δ_{BA}, δ_{IBU}, and δ_{GIBU}, be functionally represented as $\hat{C}_{\text{BA}}(\theta_0, N)$.

Algorithm 3: iterative Bayesian update (IBU)

Input: Privacy mechanism: C, Full-support PMF: ϑ_0, empirical PMF from observed data: $\boldsymbol{q}$, precision: δ_{IBU};
Output: MLE of true PMF: θ;
Function $\texttt{IBU}(C, \vartheta_0, \boldsymbol{q}, \delta_{IBU})$**:**
  Set $t \leftarrow 0$;
  while $\delta_{IBU} < |\vartheta_t - \vartheta_{t-1}|$ **do**
    $\vartheta_{t+1}(x) \leftarrow \sum_{y\in\mathcal{X}} \boldsymbol{q}(y) \frac{C_{xy}\vartheta_t(x)}{\sum_{z\in\mathcal{X}} C_{zy}\vartheta_t(z)}$;
    $t \leftarrow t + 1$
  $\theta \leftarrow \vartheta_t$;
  Return: θ;

Concerning statistical utility, it is important to ensure that IBU converges to the true distribution. As a matter of fact, IBU always converges to an MLE but the MLE may not be unique [26]. More precisely, there can be more than one distribution that is the most likely input to the obfuscation mechanism, for a given empirical distribution on the noisy data. Thus, even though IBU converges, it may converge to a distribution different from the true one. This is a problem in the method by Oya et al. in [43] which computes the obfuscation mechanism via the algorithm of Shokri et al. [48]. The resulting mechanism optimizes the trade-off between distortion and a Bayesian notion of privacy, but may not have a unique MLE, as illustrated in the example below. They probably did not realize the problem, because they relied on the flawed results by [2] according to which every mechanism would have a unique MLE.

The following example is a simplified version of the example given in [26] (Sections 3.1 and 3.2) which was aimed at showing the non-uniqueness of the MLE, and consequent convergence to the wrong distribution, in a more general setting. However, for the scope of our work, a simpler variant suffices.

Example 5.1. Consider three collinear locations, a, b and c, where b lies in between a and c at a unit distance from each of them. Assume that the prior distribution on these three locations is uniform and that the constraint on the utility is that it should not exceed 2/3. Then a mechanism that optimizes the QoS in the sense of [48] is the one that maps all locations to b. However, this mechanism has no statistical utility, as the b's do not provide any information about the original distribution. Indeed, given n obfuscated locations (i.e., n b's) all distributions on a, b and c of the form $k_a/n, k_b/n, k_c/n$ with $k_a + k_b + k_c = n$, have the same likelihood to be the original one.

Fortunately, our method does not have this problem, because the BA produces an invertible mechanism, and invertibility implies the uniqueness of the MLE [26]. In particular, we are now able to show the convergence of PRIVIC as a whole using the results of [26].

Theorem 5.1. For any $t \geq 1$, the mechanism generated by BA over $\mathcal{X}$ at the t'th iteration, seen as a stochastic matrix, is invertible.

Proof. In Appendix A. □

Theorem 5.2. PRIVIC converges to the unique MLE of the true distribution.

Proof. In Appendix A. □

To evaluate the statistical utility of $\hat{C}_{\mathrm{BA}}(\theta_0, N)$ (cf. Section 6), we will measure the EMD between the true and the estimated PMFs at the end of N iterations of PRIVIC. Thus, the quantity $EMD(\hat{\pi}_X, \pi_X)$ parameterizes the utility of $\hat{C}_{\mathrm{BA}}(\theta_0, N)$ for the service providers. We use the same Euclidean distance as the underlying metric for computing, both, the EMD and the average distortion – this consistency threads together and complements the notion of *utility* of the service providers and that from the sense of the QoS of the users.

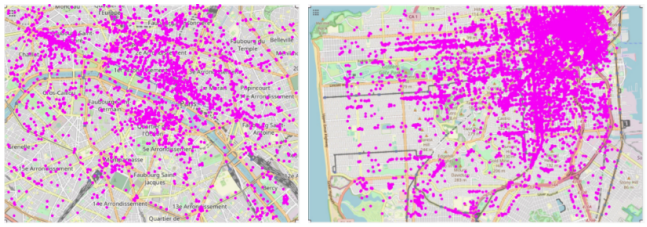
6 EXPERIMENTAL ANALYSIS OF PRIVIC

In this section, we describe the empirical results obtained by carrying out experiments to illustrate and validate the working of our proposed method. Standard Python packages were used to run the experiments in a MacOS Ventura 13.2.1 environment with an Intel core i9 processor and 32 GB of RAM. Like in the previous experiments to compare the statistical utilities of BA and LAP, as elaborated in Section 3.2, we use real locations from the same regions in Paris and San Francisco from the Gowalla dataset [19, 39]. In particular, we consider Gowalla check-ins from (i) a northern part of San Francisco bounded by latitudes (37.7228, 37.7946) and longitudes (-122.5153, -122.3789) covering an area of 12Km×8Km discretized with a 24×17 grid; (ii) a central part of Paris bounded by latitudes (48.8286, 48.8798) and longitudes (2.2855, 2.3909) covering an area of 8Km×6Km discretized with a 16×12 grid. In this setting, we work with 123,108 check-in locations in San Francisco and 10,260 check-in locations in Paris. Figure 6a shows the particular points of check-in from Paris and San Francisco and Figure 6b highlights their distribution.

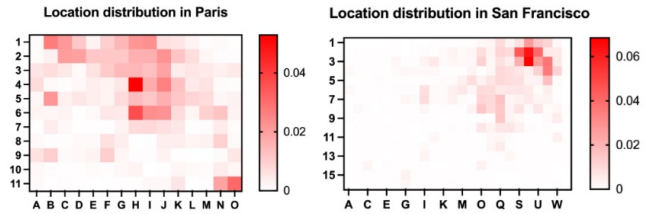
Table 2: Run-time and complexity of BA and IBU in each cycle of PRIVIC

Dataset	BA		IBU	
	Mean run-time (sec.)	Complexity	Mean run-time (sec.)	Complexity
Paris	3.256	$O(n^2)$	1.30	$O(n^2)$
San Francisco	16.805	$O(n^2)$	128.192	$O(n^2)$
Framework: MacOS Ventura 13.2.1 with Intel core i9 processor and 32 GB RAM				

We implemented PRIVIC on the locations from Paris and San Francisco separately to judge its performance on real data with very different priors. In both cases, we ran our mechanism until it empirically converged. 15 cycles of PRIVIC were required for the Paris dataset where each cycle comprised 8 iterations of BA until it converged to generate the privacy mechanism and 10 iterations of IBU until it converged to the MLE of the prior. For the San Francisco dataset, PRIVIC needed 8 cycles to converge with 5 iterations of BA and IBU each to converge in every cycle. The complexities and the run-times of BA and IBU are summarised in Table 2. In both cases, we assigned the value of the loss parameter signifying the QoS of the users, β, to be 0.5 and 1. This was done to test the performance of PRIVIC in estimating the true PMF under two different levels of privacy. Each experiment was run for 5 rounds of simulation to calibrate the randomness of the sampling and obfuscation. In each



(a) Check-in locations in Paris and San Francisco



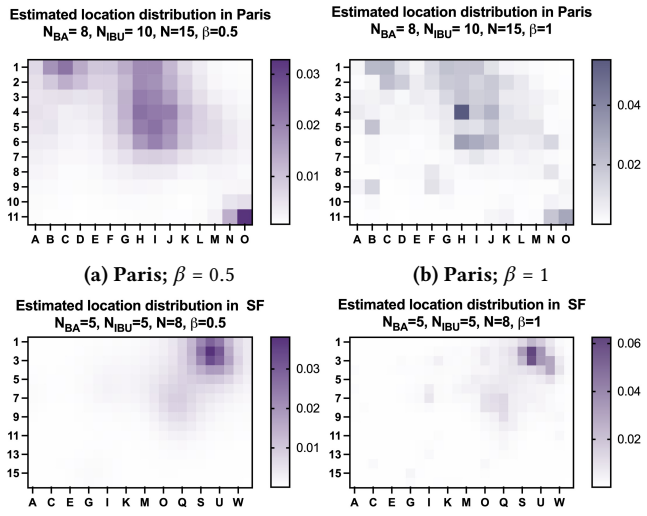
(b) Density of the original locations from Paris and San Francisco

Figure 6: (a) visualizes the original locations from Gowalla dataset from Paris and San Francisco. (b) illustrates a heatmap representation of the locations in the two cities to capture the distribution of the data.

cycle of PRIVIC, across all the settings, BA was initiated with the uniform marginal c_0 and a uniform distribution over the space of locations as the "starting guess" of the true distribution.

With $\beta = 1$, BA produces a geo-indistinguishable mechanism that injects less local noise than that obtained with $\beta = 0.5$. As a result, PRIVIC obtains a more accurate estimate of the true PMF for $\beta = 1$ than for $\beta = 0.5$. However, in both cases, the EMD between the true and the estimated distributions is very low, indicating that the PRIVIC mechanism is able to preserve a good level of statistical utility. Moreover, for both Paris and San Francisco, PRIVIC seems to significantly improve its estimation of the true PMF with every iteration until it converges to the MLE. Comparing Figures 7 and 6b, we see that the estimations of the true distributions of the locations in Paris and San Francisco by IBU under PRIVIC for both the settings of the loss parameter are fairly accurate. However, as we would anticipate, the statistical utility for $\beta = 1$ is better than that for $\beta = 0.5$.

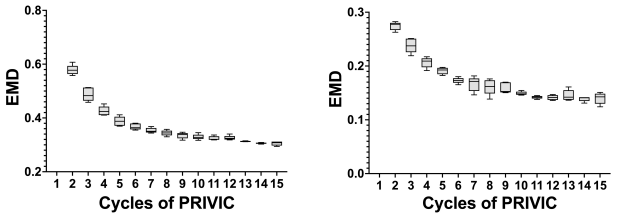
Now we shift our attention to analyze the performance of PRIVIC in preserving the statistical utility and its long-term behaviour of the two datasets. Figure 8 shows us the EMD between the true distribution of the locations in Paris and its estimate by IBU under PRIVIC in each of its 15 cycles under the two settings of privacy ($\beta = 0.5, 1$). One of the most crucial observations here is that the EMD between the true and the estimated PMFs seems to decrease with the number of iterations and it finally converges, implying that the estimation of PMFs given by PRIVIC seems to improve at the end of each cycle and, eventually, it converges to the MLE of the prior of the noisy locations, giving the estimate of the true PMF. This, empirically, suggests the convergence of the entire method. This is a major difference from the work of [43] which, as we pointed out before, has the potential of encountering an LPPM which is optimal according to the standards set by Shokri et al. in [48] but



(a) Paris; $\beta = 0.5$ **(b) Paris; $\beta = 1$**

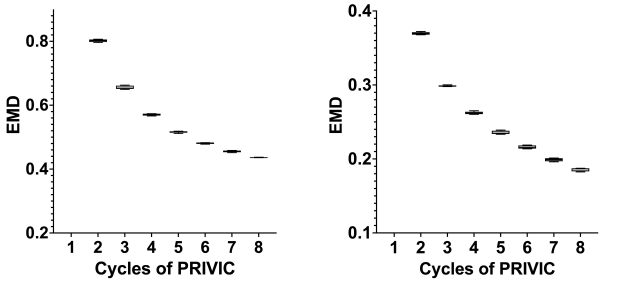
(c) San Francisco; $\beta = 0.5$ **(d) San Francisco; $\beta = 1$**

Figure 7: Visualization of the estimated true distribution of the locations in Paris ((a) and (b)) and San Francisco ((c) and (d)) by PRIVIC after its convergence; the first column is for $\beta = 0.5$ and the second column is for $\beta = 1$.



(a) $\beta = 0.5$ **(b) $\beta = 1$**

Figure 8: (a) and (b) show the EMD between the true PMF of the Paris locations and its estimation by PRIVIC in each of its cycle for $\beta = 0.5$ and $\beta = 1$, respectively.

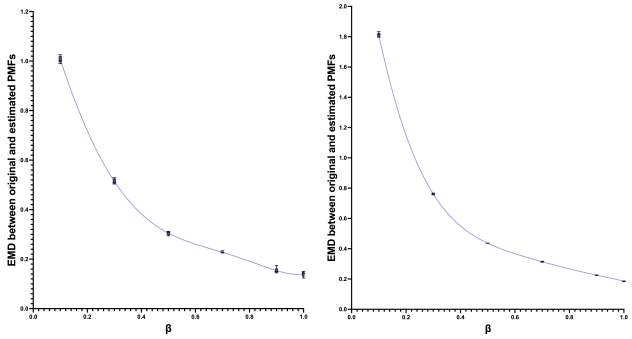


(a) $\beta = 0.5$ **(b) $\beta = 1$**

Figure 9: (a) and (b) show the EMD between the true PMF of the San Francisco locations and its estimation by PRIVIC in each of its cycles for $\beta = 0.5$ and $\beta = 1$, respectively.

the EM method used to estimate the true distribution would fail to converge for that mechanism as illustrated in Example 5.1. We observe a very similar trend for the San Francisco dataset. Figure 9 shows the statistical utility of the mechanism generated by PRIVIC under each of its 8 cycles for $\beta = 0.5$ and $\beta = 1$. The explicit values of the EMD between the true and the estimated PMFs on the location data from Paris and San Francisco for both the settings of the loss parameter can be found in Tables 3 and 4 in Appendix B.

In the next part of the experiments, we set ourselves to dissect the trend of the statistical utility harboured by PRIVIC w.r.t. the level of geo-ind it guarantees. We recall that the higher the value of β, the lesser the local noise that is injected into the data, and, hence, the worse will be the statistical utility, staying consistent with our observations in Figure 7. We continue working with the location data from Paris and San Francisco obtained from the Gowalla dataset in the same framework as described before. We consider β taking the values 0.1, 0.3, 0.5, 0.7, 0.9, 1, and for each value of the loss parameter, we run PRIVIC on both datasets using the same number of iterations as in the previous experiments. We adhere to 5 rounds of simulation for each β to account for the randomness generated in the obfuscation process.



(a) Paris **(b) San Francisco**

Figure 10: (a) and (b) illustrate that EMD between the true and the estimated distributions of the locations in Paris and San Francisco, respectively, after the empirical convergence of PRIVIC for the different values of the loss parameters β.

Figure 10 shows us that the difference between the true and the estimated PMFs under PRIVIC starts by sharply decreasing and then eventually stabilizes with an increase in the value of the loss parameter. In other words, as the intensity of the local noise decreases, we will end up estimating the unique MLE of the original distribution while optimizing MI and the users' QoS. Both the location datasets result in a Pareto curve showing a similar trend. This depicts an improvement of the estimated PMF until it converges to the true distribution. This observation complements the Pareto-optimality of MI with the maximum average distortion as studied in *rate-distortion theory* [46], and thus, we empirically weave together the two ends of utility with the information theoretical notion of privacy under PRIVIC.

Discussion. As a justification for the applicability and the working of our method, in a setting where the service providers periodically collect location data from clients, it is reasonable to assume

that, over time, they would like to maximize their utility by accurately approximating the true distribution of the population for improving their service in various aspects (crowd management, security enhancement, WLAN hotspot positioning, etc.). BA, in addition to guaranteeing geo-ind, acts as an elastic location-privacy mechanism and optimizes between MI and the data owners' QoS when it initiates with the true prior. Therefore, as every iteration of PRIVIC improves the estimation of the original distribution, as seen in Figures 3a and 3b, which is used as the starting distribution in its next cycle, the overall privacy protection and its trade-off with QoS of the users will also improve, motivating the users and the service providers comply with PRIVIC to act in their best interests and, in turn, engaging them in a positive feedback loop to maximize the corresponding privacy and utility goals.

7 VULNERABILITY OF PRIVIC

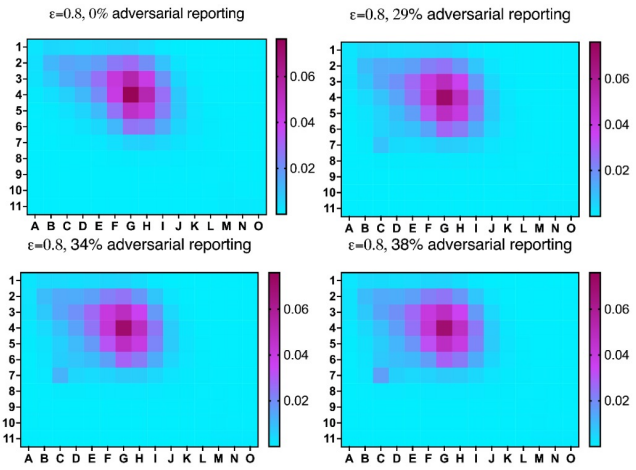
In this section, we illustrate a potential vulnerability of PRIVIC when a subset of colluding users (*adversarial users*) intentionally deviate from the correct use of the protocol.

The attack consists in falsely reporting their location in order to alter the estimation of the true distribution and, consequently, the obfuscation mechanism produced by BA. Specifically, we study two cases: (i) adversaries reporting a crowded location (*strong location*) and adversaries reporting an isolated location (*vulnerable location*).

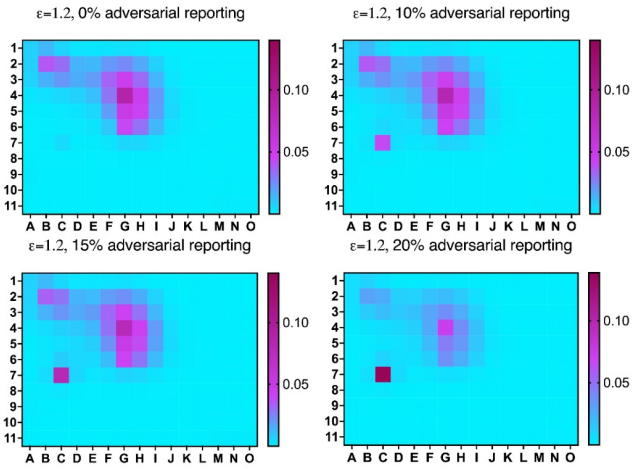
We used the real locations from the Paris dataset with the geo-spatially isolated "island" (as illustrated by Figure 1 presented in Section 3) representing a strong and a vulnerable location in the map denoted by points A and B in the figure, respectively. We performed the experiments with two different levels of formal geo-indistinguishability ($\epsilon = 0.8$ and $\epsilon = 1.2$) considering different fractions of "adversarial data submissions" (i.e., adversarial users reporting their locations falsely to compromise the privacy of other users) in each case.

Although, in both cases (i) and (ii), we observed that with an increase in the fraction of adversaries, the probability mass assigned by BA (used to obfuscate the corresponding points locally) becomes higher in and around the corresponding reported points, the impact of privacy differs across both settings. For (i), the obfuscation distribution happens to be weighed heavily around the true crowded location (point B) by both BA and LAP (as illustrated by Figures 2b and 2a) even without any adversarial users. This trend was seen to continue even when we assumed different levels of adversaries.

However, case (ii) represents a much more serious attack. As the number of adversarial users increases, BA and, in turn, PRIVIC become less potent to be able to protect the privacy of honest users who are genuinely located in an isolated location on the map. We also observe that BA and PRIVIC start behaving more like LAP. In particular, Figures 11a and 11b illustrate that the obfuscation distribution generated by BA satisfying geo-ind with $\epsilon = 0.8$ and $\epsilon = 1.2$, respectively, of the (non-adversarial) users located in point A assigns more and more weight to and around point A which, as a result, makes them more and more identifiable. This evaluation of the vulnerability of PRIVIC under adversarial data submission



(a) Obfuscation distribution of the vulnerable location in the map using BA satisfying 0.8-geo-ind with (clockwise) 0%, 29%, 34%, and 39% adversarial users, respectively.



(b) Obfuscation distribution of the vulnerable location in the map using BA satisfying 1.2-geo-ind with (clockwise) 0%, 10%, 15%, and 20% adversarial users, respectively.

Figure 11: Effect on the privacy provided by BA to obfuscate the geo-spatially isolated location (Location A as in Figure 1) for different fractions of adversarial users who intentionally report their locations falsely under two different levels of formal geo-ind guarantees by BA.

essentially exposes a weakness of the elastic distinguishability metric. We plan to address this aspect and aim to make PRIVIC more robust against adversarial users in our future works.

8 CONCLUSION

We have bridged some ideas from information theory and statistics to develop a method allowing an incremental collection of location data while protecting the privacy of the data owners, upholding their quality of service, and preserving the statistical utility

for the data consumers. Specifically, we have proposed the Blahut-Arimoto algorithm as a location-privacy mechanism, showing its extensive privacy-preserving properties and its other advantages over the state-of-the-art Laplace mechanism for geo-ind. Further, we have exhibited its duality with the iterative Bayesian update and explored this connection to present an iterative method (PRIVIC) for incremental collection of location data with formal guarantees of geo-ind and an elastic distinguishability metric, while optimizing the QoS of the users and their privacy from an information theoretical perspective. Moreover, PRIVIC efficiently estimates the MLE of the distribution of the original data and, thus, yields a high statistical utility for the service providers. Finally, we have illustrated the convergence and the general functioning of PRIVIC with experiments on real location datasets. We believe that our results can be extended easily to other kinds of data, including those with high dimensions, and to other notions of distortion measures since the analysis carried out in this paper does not depend on the notion of distance used. ‘

ACKNOWLEDGMENTS

This work was supported by the European Research Council (ERC) project HYPATIA under the European Union’s Horizon 2020 research and innovation programme. Grant agreement n. 835294.

REFERENCES

[1] Martín Abadi and David G. Andersen. 2016. Learning to Protect Communications with Adversarial Neural Cryptography. *CoRR* abs/1610.06918 (2016). arXiv:1610.06918 http://arxiv.org/abs/1610.06918

[2] Dakshi Agrawal and Charu C. Aggarwal. 2001. On the Design and Quantification of Privacy Preserving Data Mining Algorithms. In *Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (Santa Barbara, California, USA) *(PODS ’01)*. Association for Computing Machinery, New York, NY, USA, 247–255. https://doi.org/10.1145/375551.375602

[3] Rakesh Agrawal, Ramakrishnan Srikant, and Dilys Thomas. 2005. Privacy preserving OLAP. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*. 251–262.

[4] Miguel E. Andrés, Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2013. Geo-Indistinguishability: Differential Privacy for Location-Based Systems. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer &; Communications Security* (Berlin, Germany) *(CCS ’13)*. Association for Computing Machinery, New York, NY, USA, 901–914. https://doi.org/10.1145/2508859.2516735

[5] Adeel Anjum, Guillaume Raschia, Marc Gelgon, Abid Khan, Naveed Ahmad, Mansoor Ahmed, Sabah Suhail, M Masoom Alam, et al. 2017. τ-safety: A privacy model for sequential publication with arbitrary updates. *computers & security* 66 (2017), 20–39.

[6] Suguru Arimoto. 1972. An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Trans. Inf. Theory* 18 (1972), 14–20.

[7] Ugur Ilker Atmaca, Sayan Biswas, Carsten Maple, and Catuscia Palamidessi. 2022. A privacy preserving querying mechanism with high utility for electric vehicles. *arXiv preprint arXiv:2206.02060* (2022).

[8] Serge Bernstein. 1929. Sur les fonctions absolument monotones. *Acta Mathematica* 52, none (1929), 1 – 66. https://doi.org/10.1007/BF02592679

[9] Richard E. Blahut. 1972. Computation of channel capacity and rate-distortion functions. *IEEE Trans. Inform. Theory* 18 (1972), 460–473.

[10] Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2014. Optimal Geo-Indistinguishable Mechanisms for Location Privacy. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (Scottsdale, Arizona, USA) *(CCS ’14)*. Association for Computing Machinery, New York, NY, USA, 251–262. https://doi.org/10.1145/2660267.2660345

[11] Justin Brickell and Vitaly Shmatikov. 2008. The Cost of Privacy: Destruction of Data-Mining Utility in Anonymized Data Publishing. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Las Vegas, Nevada, USA) *(KDD ’08)*. Association for Computing Machinery, New York, NY, USA, 70–78. https://doi.org/10.1145/1401890.1401904

[12] Ji-Won Byun, Tiancheng Li, Elisa Bertino, Ninghui Li, and Yonglak Sohn. 2009. Privacy-Preserving Incremental Data Dissemination. *J. Comput. Secur.* 17, 1 (jan 2009), 43–68.

[13] Ji-Won Byun, Yonglak Sohn, Elisa Bertino, and Ninghui Li. 2006. Secure Anonymization for Incremental Datasets. In *Secure Data Management*, Willem Jonker and Milan Petković (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 48–63.

[14] Konstantinos Chatzikokolakis, Miguel E. Andrés, Nicolás E. Bordenabe, and Catuscia Palamidessi. 2013. Broadening the Scope of Differential Privacy Using Metrics. In *The 13th Privacy Enhancing Technologies Symposium (Lecture Notes in Computer Science, Vol. 7981)*, De Cristofaro, Emiliano, Wright, and Matthew (Eds.). Springer, Bloomington, Indiana, United States, 82–102. https://doi.org/10.1007/978-3-642-39077-7

[15] Konstantinos Chatzikokolakis, Ehab ElSalamouny, and Catuscia Palamidessi. 2017. Efficient Utility Improvement for Location Privacy. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 308–328. https://doi.org/doi:10.1515/popets-2017-0051

[16] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. 2008. Anonymity Protocols as Noisy Channels. *Information and Computation* 206, 2–4 (2008), 378–401. https://doi.org/10.1016/j.ic.2007.07.003

[17] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Marco Stronati. 2015. Constructing elastic distinguishability metrics for location privacy. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 156–170. https://doi.org/doi:10.1515/popets-2015-0023

[18] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Marco Stronati. 2015. Constructing elastic distinguishability metrics for location privacy. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 156–170. https://doi.org/doi:10.1515/popets-2015-0023

[19] Eunjoon Cho, Seth A Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1082–1090.

[20] Imre Csiszar. 1974. On the computation of rate-distortion functions (Corresp.). *Information Theory, IEEE Transactions on* 20 (02 1974), 122 – 124. https://doi.org/10.1109/TIT.1974.1055146

[21] Paul Cuff and Lanqing Yu. 2016. Differential Privacy As a Mutual Information Constraint. In *Proceedings of the 23rd ACM SIGSAC Conference on Computer and Communications Security (CCS)* (Vienna, Austria) *(CCS ’16)*. ACM, New York, NY, USA, 43–54. https://doi.org/10.1145/2976749.2978308

[22] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. 2013. Local Privacy and Statistical Minimax Rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. 429–438. https://doi.org/10.1109/FOCS.2013.53

[23] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. 2006. Our Data, Ourselves: Privacy Via Distributed Noise Generation. In *Advances in Cryptology - EUROCRYPT 2006*, Serge Vaudenay (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 486–503.

[24] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography*, Shai Halevi and Tal Rabin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 265–284.

[25] Ehab ElSalamouny and Catuscia Palamidessi. 2019. Full Convergence of the Iterative Bayesian Update and Applications to Mechanisms for Privacy Protection. *CoRR* abs/1909.02961 (2019). arXiv:1909.02961 http://arxiv.org/abs/1909.02961

[26] Ehab ElSalamouny and Catuscia Palamidessi. 2020. Generalized Iterative Bayesian Update and Applications to Mechanisms for Privacy Protection. In *2020 IEEE European Symposium on Security and Privacy (EuroS P)*. 490–507. https://doi.org/10.1109/EuroSP48549.2020.00038

[27] Natasha Fernandes, Annabelle McIver, and Carroll Morgan. 2021. The Laplace Mechanism has optimal utility for differential privacy over continuous queries. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021*. IEEE, 1–12. https://doi.org/10.1109/LICS52264.2021.9470718

[28] Esther Gal-or. 1982. Hotelling’s spatial competition as a model of sales. *Economics Letters* 9, 1 (1982), 1–6. https://doi.org/10.1016/0165-1765(82)90089-1

[29] Filippo Galli., Sayan Biswas., Kangsoo Jung., Tommaso Cucinotta., and Catuscia Palamidessi. 2023. Group Privacy for Personalized Federated Learning. In *Proceedings of the 9th International Conference on Information Systems Security and Privacy - ICISSP*. INSTICC, SciTePress, 252–263. https://doi.org/10.5220/0011885000003405

[30] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. 2012. Universally Utility-maximizing Privacy Mechanisms. *SIAM J. Comput.* 41, 6 (2012), 1673–1693. https://doi.org/10.1137/09076828X arXiv:https://doi.org/10.1137/09076828X

[31] Mangesh Gupte and Mukund Sundararajan. 2010. Universally Optimal Privacy Mechanisms for Minimax Agents. In *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (Indianapolis, Indiana, USA) *(PODS ’10)*. Association for Computing Machinery, New York, NY, USA, 135–146. https://doi.org/10.1145/1807085.1807105

[32] Mehmet Emre Gursoy, Acar Tamersoy, Stacey Truex, Wenqi Wei, and Ling Liu. 2021. Secure and Utility-Aware Data Collection with Condensed Local Differential Privacy. *IEEE Transactions on Dependable and Secure Computing* 18, 5 (2021), 2365–2378. https://doi.org/10.1109/TDSC.2019.2949041

[33] David Haussler. 1999. Convolution Kernels on Discrete Structures UCSC CRL.

[34] Thomas Hofmann, Bernhard Schölkopf, and Alexander J. Smola. 2008. Kernel methods in machine learning. *The Annals of Statistics* 36, 3 (jun 2008). https://doi.org/10.1214/009053607000000677
[35] Chong Huang, Peter Kairouz, Xiao Chen, Lalitha Sankar, and Ram Rajagopal. 2017. Context-aware generative adversarial privacy. *Entropy* 19, 12 (1 12 2017). https://doi.org/10.3390/e19120656
[36] Stratis Ioannidis, Andrea Montanari, Udi Weinsberg, Smriti Bhagat, Nadia Fawaz, and Nina Taft. 2014. Privacy Tradeoffs in Predictive Analytics. In *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems* (Austin, Texas, USA) *(SIGMETRICS '14)*. Association for Computing Machinery, New York, NY, USA, 57–69. https://doi.org/10.1145/2591971.2592011
[37] L. V. Kantorovich. 1960. *Mathematical Methods of Organizing and Planning Production*. Vol. 6. INFORMS. Issue 4. https://doi.org/10.1287/mnsc.6.4.366
[38] Boris Köpf and David A. Basin. 2007. An information-theoretic model for adaptive side-channel attacks. In *Proceedings of the 2007 ACM Conference on Computer and Communications Security (CCS 2007)*, Peng Ning, Sabrina De Capitani di Vimercati, and Paul F. Syverson (Eds.). ACM, 286–296. https://doi.org/10.1145/1315245.1315282
[39] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data/loc-Gowalla.html.
[40] Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. 2010. Optimizing Linear Counting Queries under Differential Privacy. In *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (Indianapolis, Indiana, USA) *(PODS '10)*. Association for Computing Machinery, New York, NY, USA, 123–134. https://doi.org/10.1145/1807085.1807104
[41] John Nash. 1951. Non-Cooperative Games. *Annals of Mathematics* 54, 2 (1951), 286–295. http://www.jstor.org/stable/1969529
[42] Simon Oya, Carmela Troncoso, and Fernando Pérez-González. 2017. Back to the Drawing Board: Revisiting the Design of Optimal Location Privacy-preserving Mechanisms. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA). ACM, 1959–1972. https://doi.org/10.1145/3133956.3134004
[43] Simon Oya, Carmela Troncoso, and Fernando Pérez-González. 2019. Rethinking Location Privacy for Unknown Mobility Behaviors. In *2019 IEEE European Symposium on Security and Privacy (EuroS P)*. 416–431. https://doi.org/10.1109/EuroSP.2019.00038
[44] Marco Romanelli, Kostantinos Chatzikokolakis, and Catuscia Palamidessi. 2020. Optimal Obfuscation Mechanisms via Machine Learning. In *2020 IEEE 33rd Computer Security Foundations Symposium (CSF)*. 153–168. https://doi.org/10.1109/CSF49147.2020.00019
[45] I. J. Schoenberg. 1938. Metric Spaces and Completely Monotone Functions. *Annals of Mathematics* 39, 4 (1938), 811–841. http://www.jstor.org/stable/1968466
[46] C. E. Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
[47] Reza Shokri. 2015. Privacy Games: Optimal User-Centric Data Obfuscation. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 299–315. https://doi.org/doi:10.1515/popets-2015-0024
[48] Reza Shokri, George Theodorakopoulos, George Danezis, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. 2011. Quantifying Location Privacy: The Case of Sporadic Location Exposure. In *Privacy Enhancing Technologies*, Simone Fischer-Hübner and Nicholas Hopper (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 57–76.
[49] Reza Shokri, George Theodorakopoulos, Carmela Troncoso, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. 2012. Protecting Location Privacy: Optimal Strategy against Localization Attacks. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security* (Raleigh, North Carolina, USA) *(CCS '12)*. Association for Computing Machinery, New York, NY, USA, 617–627. https://doi.org/10.1145/2382196.2382261
[50] Paul Syverson. 2013. Why I'm Not an Entropist. In *Security Protocols XVII*, Bruce Christianson, James A. Malcolm, Vashek Matyáš, and Michael Roe (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 213–230.
[51] Ardhendu Tripathy, Ye Wang, and Prakash Ishwar. 2019. Privacy-Preserving Adversarial Networks. In *57th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2019, Monticello, IL, USA, September 24-27, 2019*. IEEE, 495–505. https://doi.org/10.1109/ALLERTON.2019.8919758
[52] Yue Wang, Xintao Wu, and Donghui Hu. 2016. Using Randomized Response for Differential Privacy Preserving Data Collection.. In *EDBT/ICDT Workshops*, Vol. 1558. 0090–6778.
[53] Holger Wendland. 2004. *Scattered Data Approximation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511617539
[54] DV Widder. 1941. The Laplace transform, vol. 6 of. *Princeton Mathematical Series* (1941).
[55] Wenjing Zhang, Ming Li, Ravi Tandon, and Hui Li. 2019. Online Location Trace Privacy: An Information Theoretic Approach. *IEEE Transactions on Information Forensics and Security* 14, 1 (2019), 235–250. https://doi.org/10.1109/TIFS.2018.2848659
[56] Ye Zhu and Riccardo Bettati. 2005. Anonymity vs. Information Leakage in Anonymity Systems. In *Proc. of ICDCS*. IEEE Computer Society, 514–524.
[57] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In *Proceedings of the 21st ACM Conference on Computer and Communications Security*. Scottsdale, Arizona. https://arxiv.org/abs/1407.6981

A PROOFS

Theorem 3.1. The privacy mechanism generated by BA produces an elastic location-privacy mechanism.

Proof. Let $x, y \in \mathcal{X}$ be any true and reported location, respectively. Letting $\hat{C}_{\mathrm{BA}}$ to be the limiting mechanism generated by BA, to show that $\hat{C}_{\mathrm{BA}}$ possesses an elastic distinguishability metric, we need to ensure that:

(1) The probability of reporting y to obfuscate x given by $\hat{C}_{\mathrm{BA}}$ should be exponentially reducing w.r.t. the Euclidean distance between x and y, staying consistent with the essence of geo-ind (the property captured by (3)).
(2) Under $\hat{C}_{\mathrm{BA}}$, the probability of reporting y to obfuscate x should be taking into account the mass of reported points around y, i.e., the more geo-spatially isolated (from other reported points) y is in the space, the less likely it should be to report it, as, ideally, we would like to have x being reported as a location amidst a crowd of other reported locations (the property captured by (4)).

Let's simplify the notation and denote $\mathbb{P}[\hat{C}_{\mathrm{BA}}(x) = y]$ as $\mathbb{P}_{\mathrm{BA}}[y|x]$ and let $q(y)$ be the probability mass of the observed location y. Hence, for being an elastic location-privacy mechanism, $\hat{C}_{\mathrm{BA}}$ should satisfy (3) and (4), i.e., we must have:

$$\mathbb{P}_{\mathrm{BA}}[y|x] \propto \exp\{-\beta d_{\mathrm{E}}(x, y)\} \tag{6}$$

$$\mathbb{P}_{\mathrm{BA}}[y|x] \propto q(y) \tag{7}$$

Therefore, in order to satisfy (6) and (7), it is sufficient to have:

$$\begin{aligned} &\mathbb{P}_{\mathrm{BA}}[y|x] \propto \exp\{-\beta d_{\mathrm{E}}(x, y) + \ln q(y)\} \\ \implies &\mathbb{P}_{\mathrm{BA}}[y|x] \propto q(y)\exp\{-\beta d_{\mathrm{E}}(x, y)\} \\ &= \frac{q(y)\exp\{\beta d_{\mathrm{E}}(x, y)\}}{\sum\limits_{z\in\mathcal{X}} q(z)\exp\{-\beta d_{\mathrm{E}}(x, z)\}} \end{aligned} \tag{8}$$

Now, it's sufficient to note that, if we interpret the mass of y as the probability of being reported by the mechanism, (8) is exactly the fixpoint of $\mathcal{G} \circ \mathcal{F}$, cf. Remarks 2 and 3. □

Theorem 5.1. For any $t \geq 1$, the mechanism generated by BA over $\mathcal{X}$ at the t'th iteration, seen as a stochastic matrix, is invertible.

Proof. For notational convenience, in this proof, we shall denote the Euclidean distance $d_{\mathrm{E}}(.)$ as $d(.)$. For any $t \geq 1$, let $C^{(t)}$ be the channel generated at the t'th iteration of BA. Hence, we have:

$$C^{(t)}_{x,y} = \frac{c_{t-1}(y)\exp\{-\beta d(x, y)\}}{\sum_{z\in\mathcal{Y}} c_t(z)\exp\{-\beta d(x, z)\}} \tag{9}$$

Let $C^{'} \in C(\mathcal{X}, \mathcal{X})$ such that $C^{'}_{x,y} = \exp\{-\beta d(x, y)\}$. Correspondingly, let us define $C^{''(t)}, C^{'''(t)} \in C(\mathcal{X}, \mathcal{X})$ s.t. $C^{''(t)}_{x,y} = c_{t-1}(y)C^{'}_{x,y}$ and $C^{'''(t)}_{x,y} = K_x C^{''(t)}_{x,y}$ where $K_x = (\sum_{z\in\mathcal{Y}} c_t(z)\exp\{-\beta d(x, z)\})^{-1}$. Therefore, we have $C^{'''(t)} = C^{(t)}$.

Exploiting the fact that scaling of rows and columns of matrices by real numbers (elementary operations on rows and columns) preserves their linear independence, we ensure that if $C^{'}$ is invertible, then so is $C^{''(t)}$ (elementary column operation on $C^{'}$) which, in turn, implies that $C^{'''(t)} = C^{(t)}$ is invertible (elementary row operation on $C^{''(t)}$). Therefore, in order to show $C^{(t)}$ is invertible, it is sufficient to prove that $C^{'}$ is invertible.

Note that $\exp\{-\beta d(x,y)^2\} = \exp\{-\beta||x-y||_2^2\}$ is the *Gaussian kernel* for any $\beta > 0$ and is positive definite [33, 34]. Furthermore, Schoenberg [45] observed that for any *completely monotone function* $g: \mathbb{R}_{\geq 0} \mapsto \mathbb{R}$, we can use *Hausdorff–Bernstein–Widder theorem* [8, 54] to deduce that *radial basis function (RBF)* kernels such as $\exp\{-\beta g\left(||x-y||_2^2\right)\}$ are also positive definite. Moreover, in addition to being positive definite, it was also shown that Gaussian kernels are *strictly* positive definite [34, 53].

Let $f: \mathbb{R}_{\geq 0} \mapsto \mathbb{R}$ be the *square-root* function, i.e., $f(x) = \sqrt{x}$ for all $x \in \mathbb{R}_{\geq 0}$. Therefore, observing that f is completely monotone and recalling that Gaussian kernels are strictly positive definite, i.e., $\boldsymbol{z}^{\mathrm{T}} \exp\{-\beta||x_i - x_j||_2^2\}\boldsymbol{z} \geq 0$ for every $\boldsymbol{z} \in \mathbb{R}^m$ with equality holding iff $\boldsymbol{z} = \boldsymbol{0}$, we can use *Schoenberg theorem* [45] to show the strict positive definiteness of $\exp\{-\beta f\left(d(x,y)^2\right)\} = \exp\{-\beta d(x,y)\}$. Hence, noting that C' is the Gram matrix of the RBF kernel $\exp\{-\beta d(x,y)\}$, we must have $\boldsymbol{z}^{\mathrm{T}} C' \boldsymbol{z} \geq 0$ for every $\boldsymbol{z} \in \mathbb{R}^m$ with equality holding iff $\boldsymbol{z} = \boldsymbol{0}$. This implies that C' is positive definite and, hence, invertible. Therefore, in turn, $C^{(t)}$ is invertible.

□

Theorem 5.2. PRIVIC converges to the unique MLE of the true distribution.

Proof. For $1 \leq t \leq N$ and $1 \leq i \leq n$, in the t'th round of PRIVIC, the i'th sampled user locally sanitize their location $x_i^{(t)}$ with $\hat{C}^{(t)}$ and reports the noisy location $y_i^{(t)}$. Therefore, the *combined mechanism* (referred to as *output probability matrix* in [26]) for implementing GIBU is $\mathcal{G} = \left(\hat{C}^{(1)} \quad \dots \quad \hat{C}^{(N)}\right)$ s.t.

$$\mathcal{G}\left(x, y_i^{(t)}\right) = \mathbb{P}\left[y_i^{(t)} \middle| x\right] = \hat{C}^{(t)}\left(x, y_i^{(t)}\right)$$
$$\forall x \in \mathcal{X},\ i \in \{1, \dots, n\}.$$

By Theorem 5.1, $\hat{C}^{(t)}$ is invertible for every $t \geq 1$ and let $\hat{C}^{(t)-1}$ denote the corresponding inverse of $\hat{C}^{(t)}$. Therefore, defining $\mathcal{G}'$ s.t $\mathcal{G}' = \frac{1}{N}\left(\hat{C}^{(1)-1} \quad \dots \quad \hat{C}^{(N)-1}\right)^{\mathrm{T}}$ ensures that $\mathcal{G} \cdot \mathcal{G}' = \mathbb{I}_m$ where $m = |\mathcal{X}|$. Therefore, $\mathcal{G}$ is right-invertible.

Hence, combining the right-invertibility of $\mathcal{G}$ with Theorem 3 (GIBU converges to MLEs) and Corollary 1 (right-invertibility of the combined channel of GIBU implies unique MLE) of [26], we can conclude that $\texttt{GIBU}\left(\left(\hat{C}^{(1)}, \boldsymbol{y}^{(1)}\right), \dots, \left(\hat{C}^{(N)}, \boldsymbol{y}^{(N)}\right)\right)$ estimates the unique MLE of the prior $\pi_{\mathcal{X}}$, implying that PRIVIC converges.

□

B TABLES

Table 3: EMD between the true and the estimated PMFs by PRIVIC on the Paris locations.

N	$\beta = 1$					$\beta = 0.5$				
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1	Round 2	Round 3	Round 4	Round 5
1	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262
2	0.27104	0.27796	0.28247	0.27758	0.26276	0.57738	0.57994	0.55791	0.60717	0.56880
3	0.21916	0.23750	0.25035	0.25116	0.23241	0.51324	0.48295	0.47043	0.51285	0.45826
4	0.19156	0.21115	0.21726	0.20913	0.20408	0.43184	0.42398	0.41040	0.45230	0.41119
5	0.18241	0.19264	0.19570	0.19747	0.18728	0.39741	0.38771	0.37284	0.41176	0.36953
6	0.16526	0.18020	0.174578	0.17310	0.17268	0.37818	0.35482	0.36039	0.36375	0.38045
7	0.14643	0.18159	0.16092	0.17222	0.17139	0.35044	0.34383	0.34818	0.35760	0.36804
8	0.13860	0.17605	0.15938	0.17078	0.16192	0.34086	0.32983	0.35769	0.34430	0.34780
9	0.15047	0.16926	0.15153	0.17005	0.15266	0.33137	0.31749	0.34690	0.33840	0.34028
10	0.14734	0.14825	0.14585	0.15001	0.15459	0.3170	0.32975	0.32772	0.34529	0.32670
11	0.14227	0.14135	0.14326	0.13797	0.14507	0.31917	0.31851	0.32689	0.33667	0.32043
12	0.13818	0.14448	0.14703	0.13589	0.14142	0.32137	0.32451	0.31843	0.32556	0.34014
13	0.16111	0.13641	0.14893	0.14208	0.13808	0.31224	0.31219	0.31380	0.31515	0.31159
14	0.13894	0.13111	0.14094	0.14199	0.14192	0.30883	0.30496	0.30578	0.30726	0.30282
15	0.15106	0.14271	0.14601	0.13584	0.12413	0.29405	0.31198	0.30786	0.31167	0.30100

Table 4: EMD between the true and the estimated PMFs by PRIVIC on the San Francisco locations.

N	$\beta = 1$					$\beta = 0.5$				
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1	Round 2	Round 3	Round 4	Round 5
1	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595
2	0.37229	0.37038	0.36784	0.36949	0.36816	0.79621	0.80474	0.80219	0.80670	0.79940
3	0.29828	0.298370	0.30017	0.29784	0.29859	0.64931	0.66292	0.65362	0.65950	0.65285
4	0.26091	0.26029	0.26231	0.26180	0.26518	0.56896	0.57378	0.57338	0.57125	0.56618
5	0.23472	0.23419	0.23337	0.23740	0.23897	0.51672	0.51710	0.51735	0.51880	0.51138
6	0.21367	0.21432	0.21537	0.21777	0.21881	0.48194	0.48299	0.47992	0.48267	0.47791
7	0.19612	0.19761	0.19904	0.20120	0.20067	0.45531	0.45861	0.45122	0.45732	0.45450
8	0.18244	0.18412	0.18741	0.18724	0.18674	0.43588	0.43745	0.43558	0.43704	0.43584