Location Trace Privacy Under Conditional Priors

Casey Meehan
cmeehan@eng.ucsd.edu
UC San Diego

Kamalika Chaudhuri
kamalika@eng.ucsd.edu
UC San Diego

Abstract

Providing meaningful privacy to users of location based services is particularly challenging when multiple locations are revealed in a short period of time. This is primarily due to the tremendous degree of dependence that can be anticipated between points. We propose a Rényi divergence based privacy framework for bounding expected privacy loss for conditionally dependent data. Additionally, we demonstrate an algorithm for achieving this privacy under Gaussian process conditional priors. This framework both exemplifies why conditionally dependent data is so challenging to protect and offers a strategy for preserving privacy to within a fixed radius for sensitive locations in a user's trace.

1 Introduction

Location data is acutely sensitive information, detailing where we live, work, eat, shop, worship, and often when, too. Yet increasingly, location data is being uploaded for smartphone services such as ride hailing and weather forecasting and then being brokered in a thriving user location aftermarket to advertisers and even investors (Valentino-DeVryes, 2018). Users share location 'traces' when they release a sequence of locations, often across a short period of time. These traces are then used by central servers to monitor traffic trends, track individual fitness, target marketing, and even to study the effectiveness of social-distancing ordinances (Fowler, 2020). Here, we aim to provide a *local* privacy guarantee, wherein traces are sanitized at the user level before being transmitted to a centralized service. Note that this requires different guarantees and mechanisms than in *aggregate* applications making queries on large location trace databases.

Specifically, we guarantee a radius r of privacy at any sensitive time point or combination of time points within a given trace. This is challenging due to the fact that the locations within traces are highly interdependent. Informally, traces tend to follow relatively smooth trajectories in time. If not sanitized carefully, that knowledge alone may be exploited to infer actual locations from the released version of the trace. This work centers on designing meaningful privacy definitions and corresponding mechanisms that takes this dependence into account.

Broadly speaking, the vast majority of prior work on rigorous data privacy can be divided into two classes that differ by the kind of guarantee offered: differential and inferential privacy. Differential privacy (DP) guarantees that the participation of a single person in a dataset does not change the probability of any outcome by much. In contrast, inferential privacy guarantees that an adversary who has a certain degree of prior knowledge cannot make certain sensitive inferences.

DP for releasing aggregate statistics of a spatiotemporal dataset has been well studied (Fan et al., 2013; Cao et al., 2017; Yang et al., 2015; dep). There, the idea is to add enough noise to released statistics such that the effect of any user's participation is obscured, even if their locations are highly correlated to each other or to those of other users. Here, such a guarantee does not apply since we aim to release a sanitized version of a single user's trace.

In this local case we cannot rule out the possibility that the data curator knows who each individual is and who participated. Instead, we want to guarantee that event level information *about* each trace remains private. In this work, at any sensitive time t we mask whether the user visited location A or location B for any A,B less than r apart. Without *ad hoc* modifications, standard DP tools are insufficient for achieving this for the primary reasons that 1) the domain of location is virtually unbounded and 2) locations are

Proceedings of the 24[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2021, San Diego, California, USA. PMLR: Volume 130.

highly dependent across a short period of time. To see this, consider the following instinctual approaches to achieving location trace privacy.

Approach A: apply Local Differential Privacy (LDP) to each trace. Imagine a dataset of traces, each from a separate individual. Applying LDP implies that every trace has nearly the same probability of releasing the same sanitized version. This would be robust to arbitrary side information about dependence between locations in any one trace. Unfortunately, the amount of additive noise needed to achieve this would destroy nearly all utility: sanitized traces from California would have almost the same probability of showing up in Connecticut as do those from New York. Even if we constrained the domain to just Manhattan, this definition would not permit enough utility to perform e.g. traffic monitoring.

Approach B: apply LDP to each location within a trace. To preserve some utility, imagine a single trace as a dataset of n locations, each of which enjoys ε-LDP guarantees. This alone is not robust to arbitrary dependence between locations. By the logic of group LDP, it does satisfy $k\varepsilon$-LDP regardless of the dependence between any k locations. This approach has two setbacks. First, how to set k is unclear. Technically, all points in the trace are correlated, so to ward off worst-case correlations one might set it to the length of the trace, which is identical to Approach A. Second, even if location is bounded to a single city or county, satisfying this definition would still destroy nearly all utility. We cannot use sanitized traces for traffic monitoring if locations from either side of town have about same probability of being sanitized to the same value.

Approach C: apply LDP guarantees to each location within a trace, but only within any region less than width r. This definition is known as Geo-Indistinguishability (GI) (Andrés et al., 2012). GI provides a substitute for restricting the domain of location allowing us to salvage some utility. Here, only locations within r of each other are required to have ε-LDP guarantees. In DP parlance, we might say that 'neighboring traces' have one location altered by $\leq r$ and are identical everywhere else. This gives us the guarantee we want for a trace with one location, but not with more than one location. To see why, compare with Approach B. Analogously, (ε, r)-GI along a trace provides $(k\varepsilon, r)$-GI to any subset of k locations. Like Approach B, setting k is unclear. Yet unlike Approach B, GI is not resistant to arbitrary dependence between any k locations. Any dependence where a change in one or more location(s) by r implies a change in some other location(s) by $\geq r$ breaks the GI guarantee. Even with the simplest models of dependence (e.g. if we know the true trace ought to move in a straight line) this is a problem.

To reiterate, applying LDP to traces or to locations within traces (Approaches A & B) does not provide a principled method for meaningful privacy with reasonable utility. GI adapts LDP by giving guarantees only within a radius r. But in relaxing LDP, GI compromises the standard DP tools for handling obvious dependences between data-points like group DP. In our eyes, this warrants an *inferentially private* approach. Here, we continue to provide privacy within a radius r, thus allowing for utility. Yet instead of providing resistance to arbitrary dependence across any k locations, we aim to provide resistance to natural models of dependence between all locations. One may view such models as an adversary's prior beliefs about what traces are likely, like the straight-line prior mentioned earlier.

In contrast with differential privacy, providing inferential privacy guarantees is more complex, and has been less studied. It is however appropriate for applications such as ours, where information must be released based on a single person's data, the features of which are private and dependent. Kifer & Machanavajjhala (2014) provide a formal inferential privacy framework called Pufferfish, and design mechanisms for specific Pufferfish instances. As these instances do not apply to our setting, we adapt the Pufferfish framework to location privacy and more broadly to releasing any sequence of real-valued private information.

Contributions: In this work, we propose an inferentially private approach to guaranteeing a radius r of privacy for sensitive points in location traces in three parts:

- First, we propose an adaptable privacy framework tailored to sequences of highly dependent data-points that adapts Pufferfish privacy (Kifer & Machanavajjhala, 2014) to use Rényi Differential Privacy (RDP) (Mironov, 2017). Given a model of dependence between points, this framework more appropriately estimates the risk of inference within radius r on points of interest than do vanilla LDP approaches.
- We then demonstrate how to implement our framework for the highly flexible and expressive setting of Gaussian process (GP) priors. These nonparametric models capture the spatiotemporal aspect of location data (Liang & Haas, 1999; Liu et al., 1998; Chen et al., 2015). GPs have a natural synergy with Rényi privacy enabling an interpretable upper bound on privacy loss for additive Gaussian

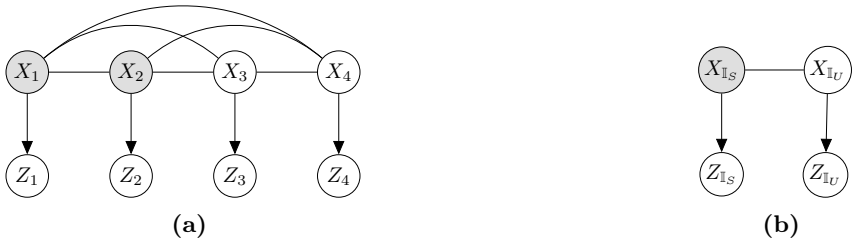

Figure 1: (a) An example graphical model of a four point trace X. (b) The more general grouped version of the model in (a), with the secret set $X_{\mathbb{I}_S} = \{X_1, X_2\}$ and the remaining set $X_{\mathbb{I}_U} = \{X_3, X_4\}$.

privacy mechanisms (that add Gaussian noise to each point). Using this, we design a semidefinite program (SDP) that optimizes the correlation of such mechanisms to minimize privacy loss without destroying utility, efficiently thwarting the inference of sensitive locations.

- Finally, we provide experiments on both location trace and home temperature data to demonstrate the advantage of these techniques over Approach C mechanisms like GI. We find that our mechanisms successfully obscure sensitive locations while respecting utility constraints, even when the prior model is misspecified.

Ultimately, by resisting only reasonable kinds of dependence in the data we are able to offer both meaningful privacy and utility. We show that our framework is robust to misspecification of this reasonable dependence and offers a privacy loss that is both tractable and interpretable.

2 Preliminaries and Problem Setting

A user transmits a sequence of N 2-dimensional locations along with their corresponding timestamps, collectively forming a 'trace'. We 'unroll' the trace into n real-valued random variables $X = \{X_1, X_2, \ldots, X_n\}$. A trace of 10 2d locations has $n = 2 \times 10 = 20$ random variables X_i. Instead of releasing the raw trace X, the user releases a private version $Z = \{Z_1, Z_2, \ldots, Z_n\}$, by way of an additive noise mechanism $Z = X + G$, where $G = \{G_1, G_2, \ldots, G_n\}$ is random noise produced by a privacy mechanism.

An adversary, receiving the obscured trace Z, then reasons about the true locations at some sensitive time(s). To reference the sensitive times, we use index set $\mathbb{I}_S$. If the sensitive indices are $\mathbb{I}_S = \{1, 2\}$, the corresponding location values are $X_{\mathbb{I}_S} = \{X_1, X_2\}$ (e.g. referring to the two coordinates of one location). When inferring the true value of $X_{\mathbb{I}_S}$, the adversary makes use of the remaining points in the trace at indices $\mathbb{I}_U = [n] \backslash \mathbb{I}_S$, denoted $X_{\mathbb{I}_U}$, with obscured values $Z_{\mathbb{I}_U}$. This separation of points into $X_{\mathbb{I}_S}$ and $X_{\mathbb{I}_U}$ is represented in **Figure 1**.

We use location as a guiding example, but such interdependent traces X could take the form of home temperature time series data or spatial data like 3D facial maps used for identification. Going forward, we will continue to denote $X = \{X_1, X_2, \ldots, X_n\}$ with the understanding that *any* subsequence of d points e.g. $X_{\mathbb{I}_S} = \{X_2, X_6, \ldots\}$ could represent a d-dimensional sensitive value, or Nd points could represent N d-dimensional sensitive values.

For the real-valued distributions considered here, $P_\times(\bullet)$ refers to a density of distribution $\times$ on r.v. $\bullet$ and $P_\times(\bullet|*)$ is its regular conditional density given $*$.

2.1 Background

GI limits what can be inferred about the sensitive $X_{\mathbb{I}_S}$ from its corresponding $Z_{\mathbb{I}_S}$, but not from the remaining locations $Z_{\mathbb{I}_U}$. To do so we need a privacy definition that specifies what events of random variable $X_{\mathbb{I}_S}$ we wish to obscure, which realistic priors of inter-dependence to protect against, and a privacy loss.

2.2 Basic and Compound Secrets

We borrow heavily from the Pufferfish framework (Kifer & Machanavajjhala, 2014), and specialize it for the setting of location traces. We define our own set of *secrets* — the collection of events we wish to obscure — and *discriminative pairs*, the pairs of secret events we do not want an adversary to tell between.

Basic Secrets & Pairs After releasing Z, we do not want an adversary with a reasonable prior on X, $\mathcal{P} \in \Theta$, to have sharp posterior beliefs about the user's location at some sensitive time (e.g. one of the sensitive times in **Figure 3** of Appendix 7.1). As such, the adversary cannot distinguish whether the user visited location A or some nearby location B at that time. Let $x_s \in \mathbb{R}^2$ represent a possible assignments to $X_{\mathbb{I}_S}$, hypothesizing the true sensitive location. Any such assignment is secret, $\mathcal{S} = \{X_{\mathbb{I}_S} = x_s : x_s \in \mathbb{R}^2\}$. Specifically, we want the posterior probability of any

two assignments to $X_{\mathbb{I}_S}$ within a radius r to be close: $\mathcal{S}_{\text{pairs}} = \{(x_s, x'_s) : \|x_s - x'_s\|_2 \leq r\}$. This protects a single time within a trace of locations. More generally, in the context of spatiotemporal data of any dimension, we call this a *basic secret*.

Compound Secrets & Pairs Suppose we have three sensitive times (again as in **Figure 3**). A mechanism that blocks inference on each of these separately does not prevent inference on the combination of them simultaneously. To obscure hypotheses on *all three* of these, we modify our set of secrets to any combination of assignments to each secret location:

$$\mathcal{S} = \big\{\{X_{\mathbb{I}_S 1} = x_{s1}\} \cap \{X_{\mathbb{I}_S 2} = x_{s2}\} \cap \{X_{\mathbb{I}_S 3} = x_{s3}\} : x_{si} \in \mathbb{R}^2, i \in [3]\big\} .$$

Now, the set of discriminative pairs is any two assignments to all three secret locations:

$$\mathcal{S}_{\text{pairs}} = \Big\{\big(\{x_{s1}, x_{s2}, x_{s3}\}, \{x'_{s1}, x'_{s2}, x'_{s3}\}\big) : \|x_{si} - x'_{si}\|_2 \leq r,\ i \in [3]\Big\}$$

This protects against compound hypotheses: if daycare and work are within r of each other, this keeps an adversary from inferring $X_{\mathbb{I}_S 1}$ = 'daycare' *and* $X_{\mathbb{I}_S 2}$ = 'work' versus $X_{\mathbb{I}_S 1}$ = 'work' *and* $X_{\mathbb{I}_S 2}$ = 'daycare'. More generally, in the context of spatiotemporal data of any dimension, we call this a *compound secret*. Intuitively, a mechanism that protects a compound secret of locations close together in time prevents a Bayesian adversary from leveraging the remainder of the trace to infer direction of motion at those sensitive times. Note that bounding the privacy loss of a compound secret does not bound the privacy loss of its constituent basic secrets.

Going forward, we refer to $\mathbb{I}_S$ as the 'secret set'.

2.2.1 Gaussian Processes

For the purpose of location privacy, it is important to choose a prior class Θ such that the conditional distribution $P_{\mathcal{P}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S})$ is simple to compute for any secret set $\mathbb{I}_S$ and any prior $\mathcal{P} \in \Theta$. Of course, it is also critical that the prior class naturally models the data, and thus consists of 'reasonable assumptions' for adversaries. GPs satisfy both these requirements. We model a full d-dimensional trace sampled at N times by 'unrolling' it into a $n = dN$ dimensional GP.

Definition 2.1. *Gaussian process* A trace X is a Gaussian process if $X_{\mathbb{I}_M}$ has a multivariate normal distribution for any set of indices $\mathbb{I}_M \subset [n]$. If X is a gaussian process, then the function $i \to \mathbb{E}[X_i]$ is called the mean function and the function $(i, j) \to \text{Cov}(X_i, X_j)$ is called the kernel function.

In this work, the kernel uses locations' time stamps to compute their covariance $(t_i, t_j) \to \text{Cov}(X_i, X_j)$, but generally could use any side information provided with each location.

GPs have simple, closed form conditional distributions. Let $X \sim \mathcal{N}(\mu, \Sigma)$, where $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$. Then, the random variable $X_{\mathbb{I}_U}|\{X_{\mathbb{I}_S} = x_s\} \sim \mathcal{N}(\mu_{u|s}, \Sigma_{u|s})$, where $\mu_{u|s} = \mu_u + \Sigma_{us}\Sigma_{ss}^{-1}(x_s - \mu_s)$ and $\Sigma_{u|s} = \Sigma_{uu} - \Sigma_{us}\Sigma_{ss}^{-1}\Sigma_{su}$. Here, μ_s denotes the mean vector μ accessed at indices $\mathbb{I}_S$ and Σ_{su} denotes the covariance matrix Σ accessed at rows $\mathbb{I}_S$ and columns $\mathbb{I}_U$.

For GP priors, we will use additive noise $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$. Thus $Z = X + G$, too, is multivariate normal. Furthermore, the distribution of any set of variables conditioned on any other set of variables in **Figure 1** belongs to some multivariate normal distribution.

GPs have been shown to successfully model mobility (Chen et al., 2015; Liang & Haas, 1999; Liu et al., 1998), even in the domain of surveillance video (Kim et al., 2011). Furthermore, although these non-parametric models are characterized by second order statistics, GPs are capable of complexity rivaling that of deep neural networks (Lee et al., 2018), allowing for scalability to more complex models and domains. Our proposed results and algorithms may be applied regardless of the complexity of the chosen GP.

2.2.2 Rényi Differential Privacy

In the following section, we propose a privacy definition that adapts Rényi Differential Privacy (RDP) (Mironov, 2017) to the Pufferfish framework. RDP resembles Differential Privacy (Dwork, 2006), except instead of bounding the maximum probability ratio or *max divergence* of the distribution on outputs for two neighboring databases, it bounds the *Rényi divergence* of order λ, defined in Equation (1) for distributions $\mathcal{P}_1$ and $\mathcal{P}_2$. The Rényi divergence bears a nice synergy with Gaussian processes. If $\mathcal{P}_1 = \mathcal{N}(\mu_1, \Sigma)$ and $\mathcal{P}_2 = \mathcal{N}(\mu_2, \Sigma)$ — two mean-shifted normal distributions — the Rényi divergence takes on a simple closed form shown in Equation (2).

$$D_\lambda\binom{\mathcal{P}_1}{\mathcal{P}_2} = \frac{1}{\lambda - 1}\log \mathbb{E}_{x \sim \mathcal{P}_2}\Big(\frac{P_{\mathcal{P}_1}(X = x)}{P_{\mathcal{P}_2}(X = x)}\Big)^\lambda \quad (1)$$

$$= \frac{\lambda}{2}(\mu_1 - \mu_2)^\intercal \Sigma^{-1}(\mu_1 - \mu_2) \quad (2)$$

We will make use of this in defining and bounding privacy loss in the next section.

3 Conditional Inferential Privacy

We now propose a privacy framework that is tailored to sequences of correlated data, Conditional Inferential Privacy (CIP). CIP guarantees a radius r of indistinguishability for the basic or compound secrets associated with any secret set $\mathbb{I}_S$. Specifically, CIP protects against any adversary with a specific prior on *the shape* of the trace, and is agnostic to their prior on the absolute location of the trace. We call the set of such prior distributions a Conditional Prior Class.

Definition 3.1. *Conditional Prior Class* For $X = \{X_1, \ldots, X_n\}$, prior distributions $\mathcal{P}_i, \mathcal{P}_j$ on X are said to belong to the same conditional prior class Θ if a constant shift in the conditioned x_s results in a constant shift on the distribution of $X_{\mathbb{I}_U}$. Formally, if conditional distributions $P_{\mathcal{P}_i}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s) = P_{\mathcal{P}_j}(X_{\mathbb{I}_U} + c^u_{ij\mathbb{I}_S}|X_{\mathbb{I}_S} = x_s + c^s_{ij\mathbb{I}_S})$ for all x_s.

For instance, prior $P_{\mathcal{P}_i}$ may concentrate probability on traces passing through Los Angeles, while $P_{\mathcal{P}_j}$ concentrates on traces passing through London. Conditioning on each secret in the pair (x_s, x'_s) in L.A. is analogous to conditioning on each secret in the pair $(x_s + c^s_{ij\mathbb{I}_S}, x'_s + c^s_{ij\mathbb{I}_S})$ in London. The corresponding pair of conditional distributions on $X_{\mathbb{I}_U}$ in London ($P_{\mathcal{P}_j}$) are copies of those in L.A. ($P_{\mathcal{P}_i}$) shifted by $c^u_{ij\mathbb{I}_S}$. What matters is that the set of all pairs of conditional distributions under $P_{\mathcal{P}_i}$ induced by secret pairs (x_s, x'_s) is identical to those under $P_{\mathcal{P}_j}$ up to a mean shift. See Appendix 7.5 for a more detailed discussion of conditional prior classes.

Definition 3.2. (ε, λ)-*Conditional Inferential Privacy* $(\mathcal{S}_{pairs}, r, \Theta)$ Given compound or basic discriminative pairs $\mathcal{S}_{\text{pairs}}$ associated with $\mathbb{I}_S$, a radius of privacy r, a conditional prior class, Θ, and a privacy parameter, $\varepsilon > 0$, a privacy mechanism $Z = \mathcal{A}(X)$ satisfies (ε, λ)-CIP$(\mathcal{S}_{\text{pairs}}, r, \Theta)$ if for all $(s_i, s_j) \in \mathcal{S}_{\text{pairs}}$, and all prior distributions $\mathcal{P} \in \Theta$, where $P_{\mathcal{P}}(s_i), P_{\mathcal{P}}(s_j) > 0$,

$$D_\lambda \begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_j) \end{pmatrix} \leq \varepsilon \tag{3}$$

CIP departs from DP type notions of privacy like Approaches A→C primarily by resisting only a restricted class of inter-dependence — the conditional prior class — as opposed to arbitrary dependence of any k locations. Unlike approaches A and B, we are able to preserve utility for tasks like traffic monitoring. Unlike approach C, CIP is still resistant to realistic models of location inter-dependence.

While this definition borrows heavily from the Pufferfish framework, it has a few key modifications. Pufferfish is generally described from a central, not local model. We specialize the kinds of secrets and discriminative pairs for the case of local location trace privacy. Additionally, we specialize the type of prior distribution class needed for this local setting: the conditional prior class. Finally, we relax the strict max divergence (max log odds) criterion of the Pufferfish definition to a Rényi divergence. This guarantees that — with high probability on draws of *realistic* traces $Z|X_{\mathbb{I}_S}$ — the log odds will be bounded by ε. As $\lambda \to \infty$, the log odds are bounded for all traces, i.e. the max divergence is bounded. We formalize this in Theorem 3.1.

The Rényi criterion of CIP greatly improves its flexibility. Unlike the standard DP Approaches A→C which only take probabilities over the mechanism, we do not have full control over the randomness at play: it is partially from $\mathcal{A}$ defined by us and from $\mathcal{P}$ intrinsic to the data. Unlike max divergence, Rényi divergence is available in closed form for many distributions, allowing for a more flexible privacy framework. The λ parameter helps us tune how strict a CIP definition is and how much noise we need to add. This allows us to design mechanisms that are resistant to natural models of dependence while preserving utility.

3.1 Properties

We now identify key properties that make the CIP guarantee interpretable and robust.

Interpretability: CIP guarantees that a Bayesian adversary with any prior distribution on traces $\mathcal{P}$ in the conditional prior class Θ does not learn much about basic or compound secrets from the released trace Z. For basic secrets, this means that the adversary's posterior beliefs regarding sensitive location $X_{\mathbb{I}_S}$ are not much sharper than their prior beliefs before witnessing Z.

Theorem 3.1. Prior-Posterior Gap: *An (ε, λ)-CIP mechanism with conditional prior class Θ guarantees that for any event O on sanitized trace Z*

$$\left| \log \frac{P_{\mathcal{P},\mathcal{A}}(s_i|Z \in O)}{P_{\mathcal{P},\mathcal{A}}(s_j|Z \in O)} - \log \frac{P_{\mathcal{P}}(s_i)}{P_{\mathcal{P}}(s_j)} \right| \leq \varepsilon'$$

for any $\mathcal{P} \in \Theta$ with probability $\geq 1 - \delta$ over draws of $Z|X_{\mathbb{I}_S} = s_i$ or $Z|X_{\mathbb{I}_S} = s_j$, where ε' and δ are related by

$$\varepsilon' = \varepsilon + \frac{\log 1/\delta}{\lambda - 1} \ .$$

This holds under the condition that $Z|X_{\mathbb{I}_S} = s_i$ and $Z|X_{\mathbb{I}_S} = s_j$ have identical support.

A CIP mechanism depends only on the conditional prior describing the data, not the data itself. Suppose an adversary's prior beliefs on $X_{\mathbb{I}_S}$ are uniform over some

region. For $\lambda = 5$ and $\varepsilon = 0.1$, there is only a $\approx 1\%$ chance that their posterior odds on s_i, s_j will be more than 3.5, and a $\approx 10\%$ chance that they will be more than 2. This 'chance' is over draws of likely remaining locations $X_{\mathbb{I}_U}$ and the additive noise G. Proofs of all results are in Appendix 7.2.

For additive noise mechanisms like $\mathcal{A}(X) = X + G = Z$, the CIP loss can be split into two terms: one accounting for the direct privacy loss of $Z_{\mathbb{I}_S}$ on $X_{\mathbb{I}_S}$ and a second accounting for the inferential privacy loss of $Z_{\mathbb{I}_U}$ on $X_{\mathbb{I}_S}$ via $X_{\mathbb{I}_U}$.

Lemma 3.2. Conditional Independence *For an additive noise mechanism, a fully dependent trace as in* ***Figure 1a****, and any prior $\mathcal{P}$ on X the CIP loss may be expressed as*

$$D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_j)\end{pmatrix} \tag{4}$$
$$= \sum_{i\in\mathbb{I}_S}\left[D_\lambda\begin{pmatrix} P_{\mathcal{A}}(Z_i|X_i = s_i) \\ P_{\mathcal{A}}(Z_i|X_i = s_j)\end{pmatrix}\right] + D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_j)\end{pmatrix}$$

One interpretation of GI is that it assumes all locations X_i are independent. In this case, the second term vanishes and the privacy loss only depends on randomness of the mechanism, not the prior.

Robustness: Kifer & Machanavajjhala (2011) show that it is impossible to achieve both utility and privacy resistant to all priors. CIP provides resistance to a reasonable class of priors $\mathcal{P} \in \Theta$, but it is possible that the true distribution $\mathcal{Q} \notin \Theta$. In this case, the privacy guarantees degrade gracefully as the divergence between $\mathcal{Q}$ and $\mathcal{P} \in \Theta$ grows.

Theorem 3.3. Robustness to Prior Misspecification *Mechanism $\mathcal{A}$ satisfies $\varepsilon(\lambda)$-CIP for prior class Θ. Suppose the finite mean true distribution $\mathcal{Q}$ is not in Θ. The CIP loss of $\mathcal{A}$ against prior $\mathcal{Q}$ is bounded by*

$$D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{Q}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{Q}}(Z|X_{\mathbb{I}_S} = s_j)\end{pmatrix} \le \varepsilon'(\lambda)$$

where

$$\varepsilon'(\lambda) = \frac{\lambda - \frac{1}{2}}{\lambda - 1}\,\Delta(2\lambda) + \Delta(4\lambda - 3) + \frac{2\lambda - \frac{3}{2}}{2\lambda - 2}\,\varepsilon(4\lambda - 2)$$

and where $\Delta(\lambda)$ is

$$\inf_{\mathcal{P}\in\Theta}\sup_{s_i\in\mathcal{S}}\max\left\{D_\lambda\begin{pmatrix} P_{\mathcal{P}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{Q}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_i)\end{pmatrix}, D_\lambda\begin{pmatrix} P_{\mathcal{Q}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{P}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_i)\end{pmatrix}\right\}$$

As long as the conditional distribution on $X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = s_i$ of prior $\mathcal{Q}$ is close to that of some $\mathcal{P} \in \Theta$, the privacy guarantees should change only marginally. This bound is tightest when $\varepsilon(\lambda)$ does not grow quickly with order λ.

3.2 CIP for Gaussian Process Priors

A *GP conditional prior class* is the set of all GP prior distributions with the same kernel function $(i, j) \to \mathrm{Cov}(X_i, X_j)$ and any mean function $i \to \mathbb{E}[X_i]$. With an additive Gaussian mechanism $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$, the CIP loss of Equation (4) can be bounded for any GP conditional prior class. See Appendix 7.5 for further discussion of the GP conditional prior class.

Theorem 3.4. CIP loss bound for GP conditional priors: *Let Θ be a GP conditional prior class. Let Σ be the covariance matrix for X produced by its kernel function. Let $\mathcal{S}$ be the basic or compound secret associated with $\mathbb{I}_S$, and S be the number of unique times in $\mathbb{I}_S$. The mechanism $\mathcal{A}(X) = X + G = Z$, where $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$, then satisfies (ε, λ)-Conditional Inferential Privacy $(\mathcal{S}_{pairs}, r, \Theta)$, where*

$$\varepsilon \le \frac{\lambda}{2} S r^2 \Big(\frac{1}{\sigma_s^2} + \alpha^*\Big) \tag{5}$$

where σ_s^2 is the variance of each $G_i \in G_{\mathbb{I}_S}$ (diagonal entries of $\Sigma_{ss}^{(g)}$) and α^* is the maximum eigenvalue of $\Sigma_{\mathit{eff}} = \big(\Sigma_{us}\Sigma_{ss}^{-1}\big)^\intercal\big(\Sigma_{u|s} + \Sigma_{uu}^{(g)}\big)^{-1}\big(\Sigma_{us}\Sigma_{ss}^{-1}\big)$.

The above bound is tight for basic secrets ($S = 1$). The two terms of Equation (5) represent the direct ($\frac{1}{\sigma_s^2}$) and inferential (α^*) loss terms of Equation (4). We assume that each diagonal entry of $\Sigma_{ss}^{(g)}$ equals some σ_s^2, so that each $X_i \in X_{\mathbb{I}_S}$ experiences identical direct privacy loss, which is optimal under utility constraints.

The above bound composes gracefully when multiple traces of an individual are released.

Corollary 3.4.1. Graceful Composition in Time *Suppose a user releases two traces X and $\hat{X}$ with additive noise $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$ and $\hat{G} \sim \mathcal{N}(\mathbf{0}, \hat{\Sigma}^{(g)})$, respectively. Then basic or compound secret $X_{\mathbb{I}_S}$ of X enjoys $(\bar{\varepsilon}, \lambda)$-CIP, where*

$$\bar{\varepsilon} \le \frac{\lambda}{2} S r^2 \Big(\frac{1}{\sigma_s^2} + \bar{\alpha}^*\Big)$$

and where $\bar{\alpha}^*$ is the maximum eigenvalue of $\bar{\Sigma}_{\mathit{eff}} = \big(\Sigma_{us}\Sigma_{ss}^{-1}\big)^\intercal\big(\Sigma_{u|s} + \bar{\Sigma}_{uu}^{(g)}\big)^{-1}\big(\Sigma_{us}\Sigma_{ss}^{-1}\big)$. Σ is the covariance matrix of the joint distribution on $X, \hat{X}$ and

$$\bar{\Sigma}^{(g)} = \begin{bmatrix}\Sigma^{(g)} & 0 \\ 0 & \hat{\Sigma}^{(g)}\end{bmatrix}.$$

This bound is identical to that of Theorem 3.4, only using the joint distribution over X, $\hat{X}$ and $G, \hat{G}$. This provides some insight to the fact that, unlike DP, even parallel composition guarantees are not automatic. Composition depends on the conditional prior. In the GP

setting, if the chosen kernel function decays over time, we can expect composition to have minimal effects on privacy for traces separated by long durations.

To reduce the upper bound of Theorem 3.4, we optimize the correlation (off-diagonal) of $\Sigma^{(g)}$ to minimize α^*, and optimize its variance (diagonal) to balance a noise budget between lowering inferential (α^*) and direct ($\frac{1}{\sigma_s^2}$) loss.

4 Optimized Privacy Mechanisms

Theorem 3.4 characterizes the privacy loss for GP conditional priors. We next show how to use this Theorem to design mechanisms that can strategically reduce CIP loss given a utility constraint. We measure 'utility loss' as the total mean squared error (MSE) between the released (Z) and true (X) traces: $\text{MSE}(\Sigma^{(g)}) = \sum_{i=1}^{n} \mathbb{E}[Z_i - X_i] = \mathbf{tr}(\Sigma^{(g)})$. We bound the utility loss by $\mathbf{tr}(\Sigma^{(g)}) \leq no_t$, where o_t is the average per-point utility loss.

It can be shown that optimizing the privacy loss under this utility constraint can be described by a semidefinite program (SDP) (formalization/derivation of SDPs in Appendix 7.3). For a given trace X, define its covariance matrix Σ using the the kernel of the GP conditional prior $\Sigma_{ij} = k(i, j)$. Then pass Σ, the secret set $\mathbb{I}_S$, and the utility constraint o_t to our first program, SDP_A, which returns noise covariance $\Sigma^{(g)}$. This defines an additive noise mechanism $G \sim \mathcal{N}(0, \Sigma^{(g)})$ that minimizes CIP loss to $\mathbb{I}_S$.

$$\Sigma^{(g)} = \text{SDP}_\text{A}(\Sigma, \mathbb{I}_S, o_t)$$

We can thus use a SDP to minimize the CIP loss to any single compound or basic secret. However, a trace may contain multiple locations or combinations thereof that one wishes to protect. It remains to produce a single mechanism $\Sigma^{(g)}$ that bounds the CIP loss to multiple basic and/or compound secrets in a single trace.

For this we propose SDP_B, which uses the fact that if ${\Sigma^{(g)}}' \succ \Sigma^{(g)}$ it will have lower CIP loss (see Appendix 7.3.2). SDP_B takes in a set of covariance matrices $\mathcal{F} = \{\Sigma_1^{(g)}, \ldots, \Sigma_m^{(g)}\}$, each designed to minimize CIP loss for a single compound or basic secret $\mathbb{I}_{Si}$. It then returns a single covariance matrix $\Sigma^{(g)} \succeq \Sigma_i^{(g)}, i \in [m]$ that maintains the privacy guarantee each $\Sigma_i^{(g)}$ offered its corresponding $\mathbb{I}_{Si}$, while minimizing utility loss.

In our experiments, we use Algorithm 1 to design a single mechanism that protects all locations in the trace — all basic secrets — while minimizing utility loss.

Algorithm 1: Multiple Secrets

Input: $\mathbb{I}_{S1}, \ldots, \mathbb{I}_{Sm}, o_t, \Sigma$
Output: $\Sigma^{(g)}$

```
F = ∅;
for i ∈ [m] do
    Σ_i^(g) = SDP_A(Σ, I_Si, o_t);
    F = F ∪ Σ_i^(g);
end
Σ^(g) = SDP_B(F);
return Σ^(g);
```

5 Experiments

Here, we aim to empirically answer: **1)** Do our SDP mechanisms maintain high posterior uncertainty of sensitive locations? How do they compare to Approach C baselines of equal MSE? **2)** How robust is the SDP_A mechanism when the prior covariance Σ is misspecified?

Methods To answer these questions, we look at the range of conditional prior classes that fit real-world data. For location trace data, we use the GeoLife GPS Trajectories dataset (Zheng et al., 2010) containing 10k human mobility traces after preprocessing (see Appendix 7.4 for details). We also consider the privacy risk of room temperature data (Nef et al., 2015), using the SML2010 dataset (Zamora-Martinez et al., 2014), which contains approximately 40 days of room temperature data sampled every 15 minutes.

For the location data, having observed that the correlation between latitude and longitude is low (≈ 0.06) we treat each dimension as independent. By way of Corollary 7.2.1, this allows us to bound privacy loss and design mechanisms for each dimension separately. Furthermore, having observed that each dimension fits nearly the same conditional prior, we treat our dataset of 10k 2-dimensional traces as a dataset of 20k 1-dimensional traces, where each trace represents one dimension of a 2d location trajectory.

We model the location trace data with a Radial Basis Function (RBF) kernel GP and the temperature series data with a periodic kernel GP:

$$k_{\text{RBF}}(t_i, t_j) = \sigma_x^2 \exp\Big(-\frac{(t_i - t_j)^2}{2l^2}\Big)$$
$$k_{\text{PER}}(t_i, t_j) = \sigma_x^2 \exp\Big(\frac{-2\sin^2(\pi|t_i - t_j|/p)}{l^2}\Big)$$

In both kernels, the intrinsic degree of dependence between points is captured by the lengthscale l. However, the fact that sampling rates vary significantly between traces means that traces with equal length scales can

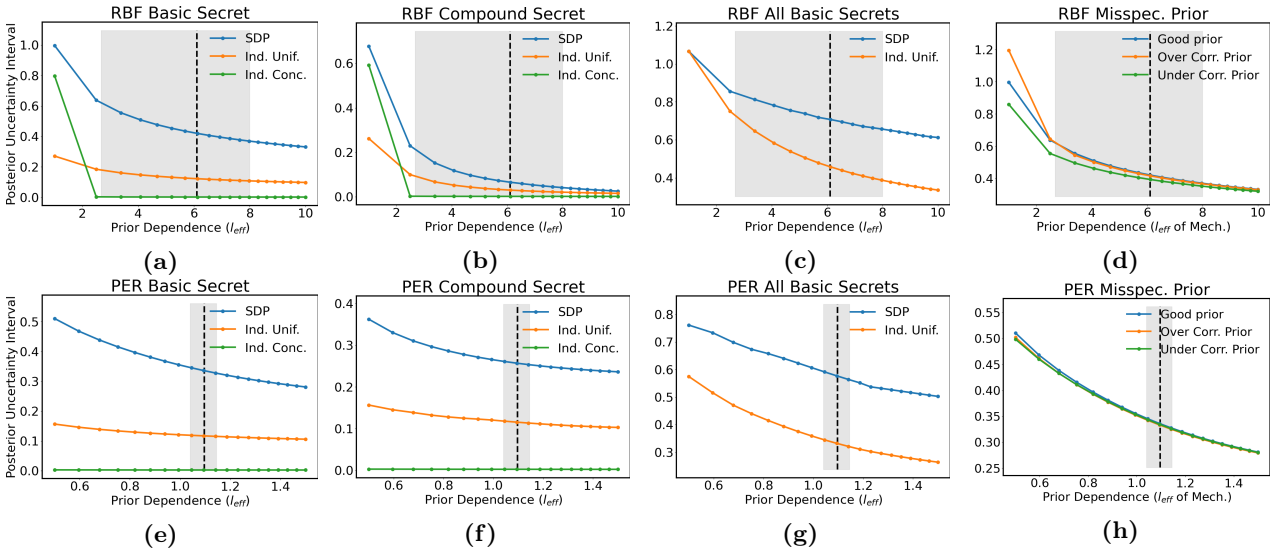

Figure 2: [1]Posterior uncertainty interval (higher=better privacy) on $X_{\mathbb{I}_S}$ of a GP Bayesian adversary. A larger l_{eff} corresponds to greater inter-dependence and reduces posterior uncertainty. The gray interval depicts the middle 50% of the MLE l_{eff} among traces in each dataset, and the black dotted line the median l_{eff}. **(a)→(c)**, **(e)→(g)** show SDP mechanisms (blue) maintaining relatively high uncertainty compared to two GI (Approach C) baselines of equal utility (MSE). **(d)**, **(h)** show the (minor) change in posterior uncertainty when the prior covariance Σ used in SDP_A is misspecified: when it is identical to the true covariance Σ^* known to the adversary (blue), is more correlated (orange), or is less correlated (green).

have very different degrees of correlation. To encapsulate both of these effects, we study the empirical distribution of *effective* length scale of each trace

$$l_{\text{eff},x} = \frac{l_x}{P} \quad l_{\text{eff},y} = \frac{l_y}{P}$$

where P is the trace's sampling period and l_x, l_y are the its optimal length scales for each dimension.

$l_{\text{eff},x}, l_{\text{eff},y}$ tell us the average number of neighboring locations that are highly correlated, instead of time period. For instance, a given trace with an optimal $l_{\text{eff},x} = 8$ tells us that every eight neighboring location samples in the x dimension have correlation > 0.8. The empirical distribution of effective length scales across all traces describes – over a range of logging devices (sampling rates), users, and movement patterns – how many neighboring points are highly correlated in location trace data. After this preprocessing, we are able to use the kernels that take indices (not time) as arguments:

$$k_{\text{RBF}}(i,j) = \exp\Big(-\frac{(i-j)^2}{2l_{\text{eff}}^2}\Big)$$

$$k_{\text{PER}}(i,j) = \exp\Big(\frac{-2\sin^2(\pi|i-j|/p)}{l_{\text{eff}}^2}\Big)$$

See Appendix 7.4 for a more detailed discussion of how the empirical distribution of l_{eff} across traces is measured.

To impart the range of realistic conditional priors the gray interval of each plot depicts the middle 50% of the empirical l_{eff} among traces in each dataset. The dashed vertical line reports the median l_{eff}.

Each figure increases the degree of dependence, l_{eff}, used by the kernel to compute the prior covariance $\Sigma(l_{\text{eff}})$. $\Sigma(l_{\text{eff}})$ is then used in one of the SDP routines of Section 4 to produce a mechanism $\Sigma^{(g)}(l_{\text{eff}})$ that protects a basic secret (SDP_A), a compound secret (SDP_A), or the union of all basic secrets (Multiple Secrets). We then observe the 68% confidence interval of the Gaussian posterior on sensitive points $X_{\mathbb{I}_S}$ (blue line). This is the 2σ uncertainty of a Bayesian adversary with a GP prior represented by $\Sigma(l_{\text{eff}})$ (see Appendix 7.4 for how this is computed). As l_{eff} increases, their posterior uncertainty will reduce. Our aim is to mitigate this as much as possible with the given utility constraint. For scale, recall that prior variance $\mathbf{diag}(\Sigma)$ is normalized to one. In the case of all basic secrets, we report the average posterior uncertainty over locations.

We compare the SDP mechanisms with two mechanisms using the logic of Approach C (all three of equal MSE utility loss): *independent/uniform* and *independent/concentrated*. The uniform approach adds independent Gaussian noise evenly along the whole trace regardless of $\mathbb{I}_S$, $\Sigma^{(g)} = o_t I$. The concentrated approach allocates the entire noise budget to the sensitive set $\mathbb{I}_S$.

Results For our first question, see **Figures 2a→2c, 2e→2g**. For both location and temperature data, our SDP mechanisms maintain higher posterior uncertainty

than the baselines with identical utility cost for a single basic secret, a compound secret, and all basic secrets. By actively considering the conditional prior class parametrized by Σ, the SDP mechanisms can strategize to both correlate noise samples and concentrate noise power such that posterior inference is thwarted at the sensitive set $\mathbb{I}_S$. For an intuitive illustration of the chosen $\Sigma^{(g)}$'s, see Appendix 7.1.2.

To answer our second question, see **Figures 2d** and **2h**. When the prior covariance Σ does not represent the true data distribution known to the adversary, a smaller posterior uncertainty may be achieved. The orange line indicates the uncertainty interval of an adversary who knows the data is *less* correlated than we believe i.e. the true $\Sigma^* = \Sigma(0.5l_{\text{eff}})$. The blue line represents an adversary who knows the data is *more* correlated than we believe i.e. the true $\Sigma^* = \Sigma(1.5l_{\text{eff}})$. Both plots confirm the robustness of our privacy guarantees stated by Theorem 3.3. Particularly around the median l_{eff} we see that the change in posterior uncertainty with this change in prior is indeed marginal.

6 Discussion

Related Work Few works have proposed solutions to the *local* guarantee when releasing individual traces. A mechanism offered in Bindschaedler & Shokri (2016) releases synthesized traces satisfying the notion of *plausible deniability* (Bindschaedler et al., 2017), but this is distinctly different from providing a radius of privacy to sensitive locations. Meanwhile, the frameworks proposed in Xiao & Xiong (2015) and Cao et al. (2019) nicely characterize the risk of inference in location traces, but use only first-order Markov models of correlation between points, do not offer a radius of indistinguishability as in this work, and are not suited to continuous-valued spatiotemporal traces.

Perhaps more technically similar to this work, Song et al. (2017) provide a general mechanism that applies to any Pufferfish framework, as well as a more computationally efficient mechanism that applies when the joint distribution of an individual's features can be described by a graphical model. The first is too computationally intensive. The second is for discrete settings, and cannot accommodate spatiotemporal effects.

Conclusion This work proposes a framework for both identifying and quantifying the *inferential* privacy risk for highly dependent sequences of spatiotemporal data. As a starting point, we have provided a simple bound on the privacy loss for Gaussian process priors, and an SDP-based privacy mechanism for minimizing this bound without destroying utility. We hope to extend this work to other data domains with different conditional priors, and different sets of secrets.

Acknowledgements

KC and CM would like to thank ONR under N00014-20-1-2334 and UC Lab Fees under LFR 18-548554 for research support. We would also like to thank our reviewers for their insightful feedback.

References

Dependence Makes You Vulnerable: Differential Privacy Under Dependent Tuples. San Diego, CA. ISBN 978-1-891562-41-9.

Andrés, M. E., Bordenabe, N. E., Chatzikokolakis, K., and Palamidessi, C. Geo-indistinguishability: Differential privacy for location-based systems. *arXiv preprint arXiv:1212.1984*, 2012.

Bindschaedler, V. and Shokri, R. Synthesizing Plausible Privacy-Preserving Location Traces. In *2016 IEEE Symposium on Security and Privacy (SP)*, pp. 546–563, May 2016. doi: 10.1109/SP.2016.39. ISSN: 2375-1207.

Bindschaedler, V., Shokri, R., and Gunter, C. A. Plausible deniability for privacy-preserving data synthesis. *Proceedings of the VLDB Endowment*, 10 (5):481–492, January 2017. ISSN 2150-8097. doi: 10.14778/3055540.3055542. URL https://doi.org/10.14778/3055540.3055542.

Cao, Y., Yoshikawa, M., Xiao, Y., and Xiong, L. Quantifying Differential Privacy under Temporal Correlations. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, pp. 821–832, April 2017. doi: 10.1109/ICDE.2017.132. ISSN: 2375-026X.

Cao, Y., Xiao, Y., Xiong, L., and Bai, L. PriSTE: From Location Privacy to Spatiotemporal Event Privacy. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, pp. 1606–1609, April 2019. doi: 10.1109/ICDE.2019.00153. ISSN: 2375-026X.

Chen, J., Low, K. H., Yao, Y., and Jaillet, P. Gaussian Process Decentralized Data Fusion and Active Sensing for Spatiotemporal Traffic Modeling and Prediction in Mobility-on-Demand Systems. *IEEE Transactions on Automation Science and Engineering*, 12(3):901–921, July 2015. ISSN 1558-3783. doi: 10.1109/TASE.2015.2422852. Conference Name: IEEE Transactions on Automation Science and Engineering.

Dwork, C. *Differential Privacy*, volume 4052. July 2006. ISBN 978-3-540-35907-4. URL https://www.microsoft.com/en-us/research/publication/differential-privacy/.

Fan, L., Xiong, L., and Sunderam, V. Differentially private multi-dimensional time series release for traffic monitoring. In *IFIP Annual Conference on Data and Applications Security and Privacy*, pp. 33–48. Springer, 2013.

Fowler, G. A. Perspective | Smartphone data reveal which Americans are social distancing (and not). *Washington Post*, 2020. ISSN 0190-8286. URL https://www.washingtonpost.com/technology/2020/03/24/social-distancing-maps-cellphone-location/.

Kifer, D. and Machanavajjhala, A. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, SIGMOD '11, pp. 193–204, Athens, Greece, June 2011. Association for Computing Machinery. ISBN 978-1-4503-0661-4. doi: 10.1145/1989323.1989345. URL https://doi.org/10.1145/1989323.1989345.

Kifer, D. and Machanavajjhala, A. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, 39(1):3, 2014.

Kim, K., Lee, D., and Essa, I. Gaussian process regression flow for analysis of motion trajectories. In *2011 International Conference on Computer Vision*, pp. 1164–1171, November 2011. doi: 10.1109/ICCV.2011.6126365. ISSN: 2380-7504.

Lee, J., Bahri, Y., Novak, R., Schoenholz, S. S., Pennington, J., and Sohl-Dickstein, J. Deep Neural Networks as Gaussian Processes. *arXiv:1711.00165 [cs, stat]*, March 2018. URL http://arxiv.org/abs/1711.00165. arXiv: 1711.00165.

Liang, B. and Haas, Z. Predictive distance-based mobility management for PCS networks. In *IEEE INFOCOM '99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, volume 3, pp. 1377–1384 vol.3, March 1999. doi: 10.1109/INFCOM.1999.752157. ISSN: 0743-166X.

Liu, T., Bahl, P., and Chlamtac, I. Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. *IEEE Journal on Selected Areas in Communications*, 16(6):922–936, August 1998. ISSN 1558-0008. doi: 10.1109/49.709453. Conference Name: IEEE Journal on Selected Areas in Communications.

Mironov, I. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275. IEEE, 2017.

Nef, T., Urwyler, P., Büchler, M., Tarnanas, I., Stucki, R., Cazzoli, D., Müri, R., and Mosimann, U. Evaluation of three state-of-the-art classifiers for recognition of activities of daily living from smart home ambient data. *Sensors*, 15(5):11725–11740, 2015.

Song, S., Wang, Y., and Chaudhuri, K. Pufferfish Privacy Mechanisms for Correlated Data. In *Proceedings of the 2017 ACM International Conference on Management of Data*, SIGMOD '17, pp. 1291–1306, Chicago, Illinois, USA, May 2017. Association for Computing Machinery. ISBN 978-1-

4503-4197-4. doi: 10.1145/3035918.3064025. URL https://doi.org/10.1145/3035918.3064025.

Valentino-DeVryes, Jennifer; Singer, N. K. M. K. A. Your apps know where you were last night, and they're not keeping it secret. *The New York Times*, 2018.

Vandenberghe, L. The cvxopt linear and quadratic cone program solvers. *Online: http://cvxopt. org/documentation/coneprog. pdf*, 2010.

Vandenberghe, L. and Boyd, S. Semidefinite Programming. *SIAM Review*, 38(1):49–95, March 1996. ISSN 0036-1445, 1095-7200. doi: 10.1137/1038003. URL http://epubs.siam.org/doi/10.1137/1038003.

Xiao, Y. and Xiong, L. Protecting locations with differential privacy under temporal correlations. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 1298–1309. ACM, 2015.

Yang, B., Sato, I., and Nakagawa, H. Bayesian Differential Privacy on Correlated Data. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, SIGMOD '15, pp. 747–762, Melbourne, Victoria, Australia, May 2015. Association for Computing Machinery. ISBN 978-1-4503-2758-9. doi: 10.1145/2723372.2747643. URL https://doi.org/10.1145/2723372.2747643.

Zamora-Martinez, F., Romeu, P., Botella-Rocamora, P., and Pardo, J. On-line learning of indoor temperature forecasting models towards energy efficiency. *Energy and Buildings*, 83:162–172, 2014.

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