Optimal Local Bayesian Differential Privacy over Markov Chains

Extended Abstract

Darshan Chakrabarti
Carnegie Mellon University
Pittsburgh, PA, USA
darshanc@alumni.cmu.edu

Jie Gao
Rutgers University
New Brunswick, NJ, USA
jg1555@rutgers.edu

Aditya Saraf
University of Washington
Seattle, WA, USA
sarafa@cs.washington.edu

Grant Schoenebeck
University of Michigan
Ann Arbor, MI, USA
schoeneb@umich.edu

Fang-Yi Yu
Harvard University
Cambridge, MA, USA
fangyiyu@seas.harvard.edu

ABSTRACT

In this paper, we focus on data generated from a Markov chain and provide optimal mechanisms for local Bayesian differential privacy (BDP) guarantees. Our main theoretical contribution is to provide a mechanism for achieving BDP when data is drawn from a binary Markov chain. We improve on the state-of-the-art BDP mechanism and show that our mechanism provides the *optimal* noise-privacy tradeoffs for any local mechanism up to negligible factors. We perform experiments on synthetic data to show that a correlation aware adversary can launch successful attacks on data that satisfies only the vanilla differential privacy guarantees. Finally, we perform experiments on real data to show that our privacy guarantees are robust to underlying distributions that are not simple Markov chains.

KEYWORDS

Pufferfish privacy, Markov Chains, Differential privacy

ACM Reference Format:
Darshan Chakrabarti, Jie Gao, Aditya Saraf, Grant Schoenebeck, and Fang-Yi Yu. 2022. Optimal Local Bayesian Differential Privacy over Markov Chains: Extended Abstract. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 3 pages.

1 INTRODUCTION

Differential privacy (DP) [2] is the most popular data privacy definition. An algorithm is differentially private if its outputs with and without any individual's data are (nearly) indistinguishable. However, DP does not offer meaningful privacy guarantees against an adversary who knows and exploits correlations in data [4]. Real world data typically exhibits natural correlation structures. Social networks mediate interactions and influence which often lead to strongly correlated personal attributes. Similarly, spatial and temporal proximity leads to strong correlations in data from sensors recorded as discretized time series. Examples include data from human mobility traces, power grids, health data from personal wearable devices, and US census data. In many applications the correlation structure can be learned from historical data and so should be assumed to be public knowledge.

The Pufferfish framework, proposed by Kifer and Machanavajjhala [5], allows for robust privacy guarantees on correlated data. Bayesian differential privacy (BDP), initially proposed by Yang et al. [9], is an instantiation of the Pufferfish privacy framework and a generalization of differential privacy. The distinction between DP and BDP is most salient when comparing which adversaries they protect against. As an example, suppose we are trying to protect Alice's time series data. Differential privacy guarantees that someone who knows Alice's location at all but one time step will not learn very much by analyzing the sanitized data set (their posterior after seeing the sanitized data will not change much) [9]. BDP guarantees that, regardless of what the adversary knows about Alice's mobility data, they will not learn much by analyzing the sanitized data set. Presumably, Alice is not very interested in protecting her data only against an adversary who knows almost all of it. Thus, we hope to achieve privacy guarantees that do not depend on how many data points the adversary knows.

1.1 Related Work

The limitations of differential privacy with correlated data has been pointed out in different contexts [1, 3, 4, 8–11]. As a result, there have been new correlation-aware privacy definitions in recent years. Zhao et al. [10] consider a definition equivalent to BDP that they term *dependent differential privacy*. Their proofs show that many attractive properties of DP, like post-processing and composition guarantees, also hold for BDP. Liu et al. [6] consider a different privacy definition that they also term dependent differential privacy; however, their definition is quite different, as it does not imply DP [10]. Naim et al. [7] consider a privacy definition rooted in information theory, that is more applicable to Internet privacy.

The original BDP paper considers a mechanism for the sum query, with data drawn from a Gaussian query model [9]. We model data generated from a Markov chain, so their mechanism does not apply. Song et al. [8] provide a very general mechanism along with guarantees for any privacy definition in the Pufferfish framework. However, their mechanism may be computationally heavy (e.g. it runs in $O(n^3)$ in our setting, while ours runs in time $O(n)$, for data of size n), and must be re-run (with additional privacy loss) for multiple queries. Zhao et al. [10] provides a very general reduction theorem which explains how ε-BDP is implied by a lower ε'-DP.

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

This result, while very powerful, does not produce optimal noise-privacy tradeoffs, as we show in the full paper.

2 WHY CONSIDER BDP?

In this section, we first explain the differences between the technical definitions of DP and BDP. We then provide one interpretation of the increased privacy that BDP ensures.

2.1 Differential Privacy

A randomized mechanism $\mathcal{M}$ is a function with domain $\mathcal{X}$ that consists of all possible input databases, and range S denoting all possible outputs.

Definition 1 (Dwork et al. [2]). *A randomized mechanism $\mathcal{M}$ is said to be* ε-differentially private *(ε-DP) if, for any databases x and y that differ in exactly one tuple (i.e. one data point),*

$$\sup_{s \in S} \frac{\Pr[\mathcal{M}(x) = s]}{\Pr[\mathcal{M}(y) = s]} \leq e^{\varepsilon}.$$

2.2 Bayesian Differential Privacy

For BDP, we now assume that the domain $\mathcal{X}$ of $\mathcal{M}$ is generated according to some probabilistic model. We also introduce adversaries $A = A(i, K)$, where $i \in [n]$ denotes the tuple A is trying to infer (attack) and $K \subseteq [n] \setminus \{i\}$ denotes the tuples A already knows.

Definition 2. *The Bayesian Differential Privacy Loss (BDPL) of $\mathcal{M}$ with respect to A is defined as*

$$\text{BDPL}(A; \mathcal{M}) = \sup_{x_i, x_i', \mathbf{x}_K, s} \frac{\Pr[\mathcal{M}(X) = s | x_i, \mathbf{x}_K]}{\Pr[\mathcal{M}(X) = s | x_i', \mathbf{x}_K]}.$$

The probability, taken over the mechanism and data generation process, should should be interpreted as the probability of the database being X, given x_i and $\mathbf{x}_K$, and then observing $\mathcal{M}(X) = s$.

Definition 3 (Yang et al. [9]). *A randomized mechanism $\mathcal{M}$ is said to satisfy ε-BDP, if $\sup_A \text{BDPL}(A; \mathcal{M}) \leq e^{\varepsilon}$, where the adversaries range over all possible i and K.*

ε-DP is equivalent to requiring $\text{BDPL}(A; \mathcal{M}) \leq e^{\varepsilon}$, for all adversaries $A = A(i, [n] \setminus \{i\})$, i.e all adversaries that know all but one tuple. It follows that ε-BDP is at least as strong as ε-DP.

2.3 Semantic Differences between DP and BDP

Differential privacy forms privacy guarantees *without* a model for how the data is generated. Thus, if an adversary has reasonable background knowledge regarding the data distributions, it may be the case that an ε-differentially private mechanism will produce an output that is disproportionately more likely (from the perspective of the adversary) given one of two neighboring datasets. This limitation of DP is well understood; Dwork et al. [2] reframes this limitation in terms of the *semantic* guarantee that DP provides. Suppose you are a medical researcher tasked with convincing users to divulge sensitive health data, which will then be published online. Differential privacy, according to the semantic interpretation, can be used as a *tool* to encourage participation. Individuals, who can only control their own participation in the study, know that they will receive minimal (privacy) harms *directly tied* to their participation in the study. Viewed this way, DP is an entirely end-user focused persuasive tool. However, an ethical researcher not only wants to persuade users to participate, but also to understand and limit the harms caused by the study itself. DP answers the question of whether a single user ought to participate; BDP answers the question of whether the study ought to be performed (in terms of the privacy "cost" of the study). Put another way, BDP persuades the researcher to publish a sanitized copy of their data, by more comprehensively limiting the harms to any study participant.

We stress that in this medical example, since the database consists of records that each belong to different individuals, DP can still be a reasonable choice (e.g. if the study could significantly help the participants, and the additional noise hinders utility). However, when considering data from a single individual, like mobility or heartbeat time series data, (standard) DP does not provide sufficient guarantees.[1] BDP provides much more meaningful guarantees than DP in these settings.

3 MAIN RESULTS

Our main theoretical contribution is providing a mechanism for achieving BDP when the data is drawn from a binary Markov chain. We focus on *non-interactive* mechanisms, which return "sanitized" (i.e. noisy) estimates of the real database that can be queried offline without further privacy loss. In contrast, query-based mechanisms must be rerun with additional privacy losses for each query. In order to achieve *local* privacy guarantees, our primary mechanism adds independent noise to each tuple. Local privacy means that no centralized curator needs to have access to the agents' true data. Instead, the owners of each tuple can privately sanitize their entry before submission, which would be ideal for IoT settings. When privacy guarantees are framed in terms of trust or persuasive power, this property is extremely attractive.

Our model is simple yet fundamental; we represent data correlations via Markov chains, and these preliminary results consider only binary state spaces. We assume that the data is *positively* correlated, as in the case of location data, where an individual is more likely to stay in the same location than leave (given a fine-enough time scale). We improve on the state-of-the-art BDP mechanism and show that our mechanism provides the *optimal* noise-privacy tradeoffs (among local/independent noise mechanisms) up to negligible factors. This is significant because the previous general results only provide a sufficient bound on the noise [6, 10]. The main challenge with finding an exact bound is describing how the privacy loss evolves through the Markov chain. We also consider a mechanism which adds *correlated* noise to the data, but find no additional improvement to noise-privacy tradeoffs. Lastly, we perform experiments on real and synthetic data. We first demonstrate that DP does not bound the correlated advantage, by providing a concrete, correlation aware attack that more than doubles the DP bound. Further, even on real data that is not entirely Markovian, a neural network adversary, using Long Short Term Memory (LSTM) models, cannot surpass our mechanism's privacy bounds, suggesting that our mechanism is *robust* to varying correlation structures in practice. [2]

[1] There are definitions such as group differential privacy [2] that apply to this setting, but they require significantly more noise. See [8] for a more thorough analysis.

[2] The authors would like to acknowledge support from NSF (OAC-1939459, CCF-2118953, CCF-2007256, and CAREER 1452915.)

REFERENCES

[1] Rui Chen, Benjamin CM Fung, S Yu Philip, and Bipin C Desai. 2014. Correlated network data publication via differential privacy. *The VLDB Journal* 23, 4 (2014), 653–676.

[2] Cynthia Dwork, Aaron Roth, et al. 2014. *The algorithmic foundations of differential privacy*. Vol. 9. Now Publishers, Inc. 211–407 pages.

[3] Xi He, Ashwin Machanavajjhala, and Bolin Ding. 2014. Blowfish privacy: Tuning privacy-utility trade-offs using policies. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*. 1447–1458.

[4] Daniel Kifer and Ashwin Machanavajjhala. 2011. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*. 193–204.

[5] Daniel Kifer and Ashwin Machanavajjhala. 2014. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)* 39, 1 (2014), 1–36.

[6] Changchang Liu, Supriyo Chakraborty, and Prateek Mittal. 2016. Dependence Makes You Vulnberable: Differential Privacy Under Dependent Tuples.. In *NDSS*, Vol. 16. 21–24.

[7] Carolina Naim, Fangwei Ye, and Salim El Rouayheb. 2019. ON-OFF privacy with correlated requests. In *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 817–821.

[8] Shuang Song, Yizhen Wang, and Kamalika Chaudhuri. 2017. Pufferfish privacy mechanisms for correlated data. In *Proceedings of the 2017 ACM International Conference on Management of Data*. 1291–1306.

[9] Bin Yang, Issei Sato, and Hiroshi Nakagawa. 2015. Bayesian differential privacy on correlated data. In *Proceedings of the 2015 ACM SIGMOD international conference on Management of Data*. 747–762.

[10] Jun Zhao, Junshan Zhang, and H Vincent Poor. 2017. Dependent differential privacy for correlated data. In *2017 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 1–7.

[11] Tianqing Zhu, Ping Xiong, Gang Li, and Wanlei Zhou. 2014. Correlated differential privacy: Hiding information in non-IID data set. *IEEE Transactions on Information Forensics and Security* 10, 2 (2014), 229–242.