Privacy-Compatibility For General Utility Metrics

Robert Kleinberg [*][†]   Katrina Ligett [*][‡]

Abstract

In this note, we present a complete characterization of the utility metrics that allow for non-trivial differential privacy guarantees.

1 Introduction

The field of data privacy is, at its heart, the study of tradeoffs between utility and privacy. The theoretical computer science community has embraced a strong and compelling definition of privacy — differential privacy [2, 3] — but utility definitions, quite naturally, depend on the application at hand. For a given function f, can we achieve arbitrarily close to perfect utility by relaxing the privacy parameter sufficiently? We show that this question has a satisfyingly simple answer: yes, if and only if the image of f has compact completion. Furthermore, in this case there exists a single base measure μ such that conventional exponential mechanisms based on μ are capable of achieving arbitrarily good utility.

2 Definitions

We are given two metric spaces $(\mathbf{X}, \rho)$ and $(\mathbf{Y}, \sigma)$ and a continuous function $f : \mathbf{X} \to \mathbf{Y}$. We think of the input database as being an element $x \in \mathbf{X}$, and our goal is to disclose an approximation to the value of $f(x)$ while preserving privacy. To allow for a cleaner exposition, we will assume throughout this paper that f has Lipschitz constant 1, i.e. $\sigma(f(x), f(z)) \leq \rho(x, z)$ for all $x, z \in \mathbf{X}$. All of our results generalize to arbitrary Lipschitz continuous functions, an issue that we return to in Remark 2.4.

Definition 2.1. A *mechanism* is a function $\mathcal{M} : \mathbf{X} \to \mathbf{\Delta}(\mathbf{Y})$, where $\mathbf{\Delta}(\mathbf{Y})$ denotes the set of all Borel probability measures on Y. For a point $x \in \mathbf{X}$, we will often denote the probability measure $\mathcal{M}(x)$ using the alternate notation $\mathcal{M}_x$.

Definition 2.2. For $\varepsilon > 0$, we say that a mechanism $\mathcal{M}$ *achieves ε-differential privacy* if the following relation holds for every $x, z \in X$ and every Borel set $T \subseteq \mathbf{Y}$:

$$\mathcal{M}_x(T) \leq e^{\varepsilon \rho(x,z)} \mathcal{M}_z(T).^1 \tag{1}$$

For $\gamma, \delta > 0$, we say that $\mathcal{M}$ *achieves γ-utility with probability at least $1 - \delta$* if the following relation holds for every $x \in X$:

$$\mathcal{M}_x(B_\sigma(f(x), \gamma)) \geq 1 - \delta. \tag{2}$$

We abbreviate this relation by saying that $\mathcal{M}$ achieves (γ, δ)-utility.

Definition 2.3. Given a function $f : \mathbf{X} \to \mathbf{Y}$, the *privacy-utility tradeoff* of f is the function

$$\varepsilon^*(\gamma, \delta) = \inf\{\varepsilon > 0 \,|\, \exists \text{ a mechanism } \mathcal{M} \text{ satisfying } \varepsilon\text{-differential privacy and } (\gamma, \delta)\text{-utility}\},$$

where the right side is interpreted as ∞ if the set in question is empty.

Remark 2.4. In prior work on differential privacy, it is more customary to express differential privacy guarantees in terms of an *adjacency relation* on inputs, rather than a metric space on the inputs. In this framework, the sensitivity of f (the maximum of $|f(a) - f(b)|$ over all adjacent pairs a, b) plays a pivotal role in determining the privacy achieved by a mechanism. The Lipschitz constant of f plays the equivalent role in our setting.

---

[1] A number of results in the literature, including recent work of Roth and Roughgarden [6] on mechanisms for predicate queries, achieve only a weakened definition of privacy known as (ε, δ)-differential privacy; such results do not fit in the framework presented here.

One could of course equate the two frameworks by defining the privacy metric ρ to be the shortest-path metric in the graph defined by the adjacency relation. This would equate the Lipschitz constant of f with its sensitivity. However, it is much more convenient to describe our mechanisms and their analysis under the assumption that f has Lipschitz constant 1; for any Lipschitz continuous f this can trivially be achieved by rescaling both ρ and the corresponding privacy bound by C, the Lipschitz constant of f.

Thus, for example, if one is given a function f and wishes to know whether there exists a mechanism achieving ε-differential privacy and (γ, δ)-utility, the answer is yes if and only if $\varepsilon/\varepsilon^*(\gamma, \delta)$ is greater than the Lipschitz constant (i.e., sensitivity) of f. In cases where the sensitivity Δ_f depends on the number of points in an input database, N, the relation $\varepsilon/\varepsilon^*(\gamma/\delta) \geq \Delta_f$ can be used to solve for N in terms of the parameters $\varepsilon, \gamma, \delta$. For example, in many papers (e.g. [1]) $\Delta_f = 1/N$ and then we find that $N = \varepsilon^*(\gamma, \delta)/\varepsilon$ is the minimum number of points in the input database necessary to achieve ε-differential privacy and (γ, δ)-utility.

Remark 2.5. Our definition of utility captures many prior formulations. For setings where the output space is simply $\mathbb{R}$, the traditional utility metric reflecting the difference between the given answer and the true answer is easily captured in our framework. A variety of prior work on problems involving more complex outputs can also be cast as measuring utility in a metric space. For example, Blum et al. [1] propose utility with respect to a concept class $\mathcal{H}$, and define the utility of a candidate output database y on an input x as $\max_{h \in \mathcal{H}} |h(x) - h(y)|$. This setup can be viewed as mapping input databases x to vectors $(h_1(x), h_2(x), \ldots)$ and taking the utility metric σ to be the L^∞ metric on output vectors. Hardt and Talwar [4] use L^2 as their utility metric, but whereas they compute the mean square (or p-th moment) of its distribution, we define disutility to be the probability that the σ value exceeds γ.

Definition 2.6. Given a measure μ on $\mathbf{X}$, and a scalar $\beta > 0$, the *(conventional) exponential mechanism* $\mathcal{C}^{\mu;\beta}$ is given by the formula:

$$\mathcal{C}_x^{\mu;\beta}(T) = \frac{\int_T e^{-\beta\sigma(f(x),y)}\, d\mu(y)}{\int_{\mathbf{Y}} e^{-\beta\sigma(f(x),y)}\, d\mu(y)}, \tag{3}$$

provided that the denominator is finite. Otherwise $\mathcal{C}_x^{\mu;\beta}$ is undefined.[2]

The differential privacy guarantee for exponential mechanisms is given by the following theorem, whose proof parallels the original proof of McSherry and Talwar [5] and is given in the Appendix.

Theorem 2.7. *If f has Lipschitz constant C then the conventional exponential mechanism $\mathcal{C}^{\mu;\beta}$ is $(2C\beta)$-differentially private for every μ.*

3 A topological criterion for privacy-compatibility

A surprising result of Blum et al. [1] shows that, in the natural setting of one-dimensional range queries over continuous domains, *no* mechanism can simultaneously achieve non-trivial privacy and utility guarantees. What is it about this application that makes privacy fundamentally impossible? In this section, we introduce a definition of *privacy-compatibility* and give a complete characterization of the applications that satisfy this definition.

Definition 3.1. We say that f is *privacy-compatible* if $\varepsilon^*(\gamma, \delta) < \infty$ for all $\gamma, \delta > 0$.

---

[2]We use the word "conventional" here to refer to the rich subclass of exponential mechanisms whose score function is σ; however, not all exponential mechanisms fall in this class.

Suppose that f is Lipschitz continuous and that the metric space $(\mathbf{X}, \rho)$ is bounded. We now prove that f is privacy-compatible if and only if the completion of the metric space $f(\mathbf{X})$ is compact. Observe that rescaling the metrics ρ, σ does not affect the question of whether f is privacy-compatible nor whether $f(\mathbf{X})$ has compact completion, but it does rescale the Lipschitz constant of f and the diameter of $\mathbf{X}$. Accordingly, we may assume without loss of generality that the Lipschitz constant of f and the diameter of $\mathbf{X}$ are both bounded above by 1, i.e.

$$\sigma(f(x_1), f(x_2)) \leq \rho(x_1, x_2) \leq 1 \tag{4}$$

for all $x_1, x_2 \in \mathbf{X}$.

Definition 3.2. A probability measure μ on a metric space $(\mathbf{X}, \sigma)$ is *uniformly positive* if it is the case that for all $r > 0$,

$$\inf_{x \in X} \mu(B_\sigma(x, r)) > 0.$$

Example 3.3. The uniform measure on $[0, 1]$ is uniformly positive. The Gaussian measure on $\mathbb{R}$ is not uniformly positive because one can find intervals of width $2r$ with arbitrarily small measure by taking the center of the interval to be sufficiently far from 0.

Theorem 3.4. *If the Lipschitz constant of f and the diameter of X are both bounded above by 1, then the following are equivalent:*

1. *f is privacy-compatible;*

2. *For every $\gamma, \delta > 0$, there is a conventional exponential mechanism that achieves (γ, δ)-utility;*

3. *There exists a uniformly positive measure on $(f(\mathbf{X}), \sigma)$;*

4. *The completion of $(f(\mathbf{X}), \sigma)$ is compact.*

Proof. For simplicity, throughout the proof we assume without loss of generality that $\mathbf{Y} = f(\mathbf{X})$. The notation $B(y, r)$ denotes the ball of radius r around y in the metric space $(\mathbf{Y}, \sigma)$.

$(\mathbf{2}) \Rightarrow (\mathbf{1})$ The exponential mechanism $\mathcal{M}^{\mu;\beta}$ achieves (2β)-differential privacy.

$(\mathbf{3}) \Rightarrow (\mathbf{2})$ For μ a uniformly positive measure on (Y, σ), and $\gamma, \delta > 0$, let $m = \inf_{y \in \mathbf{Y}} \mu(B(y, \gamma/2))$ and let $\beta = \frac{2}{\gamma} \ln \left(\frac{1}{\delta m} \right)$. We claim that the exponential mechanism $\mathcal{M} = \mathcal{M}^{\mu;\beta}$ achieves (γ, δ)-utility. To see this, let $x \in \mathbf{X}$ be an arbitrary point, let $z = f(x)$, and let

$$a = \int_{B(x,\gamma)} e^{-\beta\sigma(z,y)} \, d\mu(y) \qquad b = \int_{\mathbf{X} \setminus B(x,\gamma)} e^{-\beta\sigma(z,y)} \, d\mu(y).$$

We have

$$a \geq \int_{B(z,\gamma/2)} e^{-\beta\sigma(z,y)} \, d\mu(y) \geq \int_{B(z,\gamma/2)} e^{-\beta\gamma/2} \, d\mu(y) = e^{-\beta\gamma/2} \mu(B(z, \gamma/2)) \geq e^{-\beta\gamma/2} m$$

$$b < \int_Y e^{-\beta\gamma} \, d\mu(y) = e^{-\beta\gamma}.$$

Hence, for every $x \in \mathbf{X}$,

$$\mathcal{M}_x(B(f(x), \gamma)) = \frac{a}{a+b} = 1 - \frac{b}{a+b} > 1 - \frac{e^{-\beta\gamma}}{e^{-\beta\gamma/2}m} = 1 - \frac{1}{e^{\beta\gamma/2}m} = 1 - \delta.$$

3

(4) $\Rightarrow$ **(3)** We use the following fact from the topology of metric spaces: a complete metric space is compact if and only, for every r, if it has a finite covering by balls of radius r. (See Theorem A.2 in the Appendix.) For $i = 1, 2, \ldots$, let $C_i = \{y_{i,1}, \ldots, y_{i,n(i)}\}$ be a finite set of points such that the balls of radius 2^{-i} centered at the points of C_i cover $\mathbf{Y}$. Now define a probability measure μ supported on the countable set $C = \cup_{i=1}^{\infty} C_i$, by specifying that for $y \in C$, $\mu(y) = \sum_{i\,:\,y \in C_i} \left(\frac{1}{2^i n(i)} \right)$. Equivalently, one can describe μ by saying that a procedure for randomly sampling from μ is to flip a fair coin until heads comes up, let i be the number of coin flips, and sample a point of C_i uniformly at random. We claim that μ is uniformly positive. To see this, given any $r > 0$ let $i = \lceil \log_2(1/r) \rceil$, so that $2^{-i} \leq r$. For any point $y \in \mathbf{Y}$, there exists some j $(1 \leq j \leq n(i))$ such that $y \in B(y_{i,j}, 2^{-i})$. This implies that $B(y, r)$ contains $y_{i,j}$, hence $\mu(B(y, r)) \geq \mu(y_{i,j}) \geq \frac{1}{2^i n(i)}$. The right side depends only on r (and not on y), hence $\inf_{y \in \mathbf{Y}} \mu(B(y, r))$ is strictly positive, as desired.

(1) $\Rightarrow$ **(4)** We prove the contrapositive. Suppose that the completion of $\mathbf{Y}$ is not compact. Once again using point-set topology (Theorem A.2) this implies that there exists an infinite collection of pairwise disjoint balls of radius r, for some $r > 0$. Let $y_1, y_2, , \ldots,$ be the centers of these balls. By our assumption that $\mathbf{Y} = f(\mathbf{X})$, we may choose points x_i such that $y_i = f(x_i)$ for all $i \geq 1$. Suppose we are given a mechanism $\mathcal{M}$ that achieves r-utility with probability at least $1/2$. For every $\alpha > 0$ we must show that $\mathcal{M}$ does not achieve α-differential privacy. The relation $\sum_{i=1}^{\infty} \mathcal{M}_{x_1}(B(y_i, r)) \leq 1$ implies that there exists some i such that

$$\mathcal{M}_{x_1}(B(y_i, r)) < e^{-\alpha}/2. \tag{5}$$

The fact that $\mathcal{M}$ achieves r-utility with probability at least $1/2$ implies that

$$\mathcal{M}_{x_i}(B(y_i, r)) > 1/2. \tag{6}$$

Combining (5) with (6) leads to

$$\mathcal{M}_{x_i}(B(y_i, r)) > e^{\alpha} \mathcal{M}_{x_1}(B(y_i, r)) \geq e^{\alpha \rho(x_i, x_1)} \mathcal{M}_{x_1}(B(y_i, r)), \tag{7}$$

hence $\mathcal{M}$ violates α-differential privacy. $\square$

References

[1] A. Blum, K. Ligett, and A. Roth. A learning theory approach to non-interactive database privacy. In *Proc. ACM Symposium on Theory of Computing (STOC)*, pages 609–618, 2008.

[2] I. Dinur and K. Nissim. Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 202–210. ACM Press New York, NY, USA, 2003.

[3] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Proc. Theory of Cryptography Conference*, pages 265–284, 2006.

[4] M. Hardt and K. Talwar. On the geometry of differential privacy. In *Proc. ACM Symposium on Theory of Computing (STOC)*, 2010. to appear.

[5] F. McSherry and K. Talwar. Mechanism design via differential privacy. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 94–103, 2007.

[6] A. Roth and T. Roughgarden. The median mechanism: Interactive and efcient privacy with multiple queries. In *Proc. ACM Symposium on Theory of Computing (STOC)*, 2010. to appear.

A Appendix

Lemma A.1. *If $f : \mathbf{X} \to \mathbf{Y}$ has Lipschitz constant 1, then the conventional exponential mechanism $\mathcal{M}^{\mu;\beta}$ achieves (2β)-differential privacy.*

Proof. The proof follows the original proof of McSherry and Talwar [5]. The triangle inequality implies that for any x, z

$$\int_T e^{-\beta\sigma(f(x),y)\,d\mu(y)} \le \int_T e^{-\beta[\sigma(f(z),y)-\sigma(f(x),f(z))]}\,d\mu(y)$$

$$= e^{\beta\sigma(f(x),f(z))} \int_T e^{-\beta\sigma(f(x),y)}\,d\mu(y)$$

$$\le e^{\beta\rho(x,z)} \int_T e^{-\beta\sigma(f(z),y)}\,d\mu(y)$$

$$\int_{\mathbf{Y}} e^{-\beta\sigma(f(x),y)\,d\mu(y)} \ge \int_{\mathbf{Y}} e^{-\beta[\sigma(f(z),y)+\sigma(f(x),f(z))]}\,d\mu(y)$$

$$= e^{-\beta\sigma(f(x),f(z))} \int_{\mathbf{Y}} e^{-\beta\sigma(f(x),y)}\,d\mu(y)$$

$$\ge e^{-\beta\rho(x,z)} \int_{\mathbf{Y}} e^{-\beta\sigma(f(z),y)}\,d\mu(y).$$

The inequality $\mathcal{M}_x(T) \le e^{2\beta\rho(x,z)}\mathcal{M}_z(T)$ follows upon taking the quotient of these two inequalities. $\square$

Theorem A.2. *For a metric space $(\mathbf{X}, \sigma)$, the following are equivalent:*

1. *The completion of $\mathbf{X}$ is a compact topological space.*

2. *For every $r > 0$, $\mathbf{X}$ can be covered by a finite collection of balls of radius r.*

3. *For every $r > 0$, $\mathbf{X}$ does not contain an infinite collection of pairwise disjoint balls of radius r.*

Proof. **(2) $\Rightarrow$ (1)** Assume that property (2) holds. Recall that a metric space is compact if and only if every infinite sequence of points has a convergent subsequence, and it is complete if and only if every Cauchy sequence is convergent. Thus, we must prove that every infinite sequence $x_1, x_2, \ldots$ in $\mathbf{X}$ has a Cauchy subsequence. We can use a pigeonhole-principle argument to construct the Cauchy subsequence. In fact, the construction will yield a sequence of points $z_1, z_2, \ldots$ and sets $S_1, S_2, \ldots$ such that the diameter of S_k is at most $1/k$ and $z_i \in S_k$ for all $i \ge k$; these two properties immediately imply that $z_1, z_2, \ldots$ is a Cauchy sequence as desired.

The construction begins by defining $S_0 = \mathbf{X}$. Now, for any $k > 0$, assume inductively that we have a set S_{k-1} such that the relation $x_i \in S_{k-1}$ is satisfied by infinitely many i. Let $B_1, B_2, \ldots, B_{n(k)}$ be a finite collection of balls of radius $\frac{1}{2k}$ that covers $\mathbf{X}$. There must be at least one value of j such that the relation $x_i \in S_{k-1} \cap B_j$ is satisfied by infinitely many i. Let $S_k = S_{k-1} \cap B_j$ and let z_k be any point in the sequence $x_1, x_2, \ldots$ that belongs to S_k and occurs strictly later in the sequence than z_{k-1}. This completes the construction of the Cauchy subsequence and establishes that the completion of $\mathbf{X}$ is compact.

(1) $\Rightarrow$ (3) If $\mathbf{X}$ contains an infinite collection of pairwise disjoint balls of radius r, then the centers of these balls form an infinite set with no limit point in $\mathbf{X}$, violating compactness.

(3) $\Rightarrow$ (2) Given $r > 0$, let $B(x_1, r/2)$, ..., $B(x_n, r/2)$ be a maximal collection of disjoint balls of radius $r/2$. (Such a collection must be finite, by property (3).) The balls $B(x_1, r)$, ..., $B(x_n, r)$

cover $\mathbf{X}$, because if there were a point $y \in \mathbf{X}$ not covered by these balls, then $B(y, r/2)$ would be disjoint from $B(x_i, r/2)$ for $i = 1, \ldots, n$, contradicting the maximality of the collection. $\qquad \square$