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 $\mu$ such that conventional exponential mechanisms based on $\mu$ are capable of achieving arbitrarily good utility.

2 Definitions

We are given two metric spaces $(X, \rho)$ and $(Y, \sigma)$ and a continuous function $f : X \to Y$. We think of the input database as being an element $x \in X$, 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 $x, z \in X$. 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 $M : X \to \Delta(Y)$, where $\Delta(Y)$ denotes the set of all Borel probability measures on $Y$. For a point $x \in X$, we will often denote the probability measure $M(x)$ using the alternate notation $M_x$.

Definition 2.2. For $\varepsilon > 0$, we say that a mechanism $M$ achieves $\varepsilon$-differential privacy if the following relation holds for every $x, z \in X$ and every Borel set $T \subseteq Y$:
\[
M_x(T) \leq e^{\varepsilon \rho(x, z)} M_z(T) \tag{1}
\]

For $\gamma, \delta > 0$, we say that $M$ achieves $\gamma$-utility with probability at least $1 - \delta$ if the following relation holds for every $x \in X$:
\[
M_x(B_\sigma(f(x), \gamma)) \geq 1 - \delta. \tag{2}
\]

We abbreviate this relation by saying that $M$ achieves $(\gamma, \delta)$-utility.

Definition 2.3. Given a function $f : X \to Y$, the privacy-utility tradeoff of $f$ is the function
\[
\varepsilon^*(\gamma, \delta) = \inf\{\varepsilon > 0 | \exists \text{ a mechanism } M \text{ satisfying } \varepsilon\text{-differential privacy and } (\gamma, \delta)\text{-utility}\},
\]
where the right side is interpreted as $\infty$ 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 $(\varepsilon, \delta)$-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 $\rho$ 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 $\rho$ 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 $\varepsilon$-differential privacy and $(\gamma, \delta)$-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 $\Delta_f$ 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 $\varepsilon$-differential privacy and $(\gamma, \delta)$-utility.

Remark 2.5. Our definition of utility captures many prior formulations. For settings 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 $H$, and define the utility of a candidate output database $y$ on an input $x$ as $\max_{h \in H} |h(x) - h(y)|$. This setup can be viewed as mapping input databases $x$ to vectors $(h_1(x), h_2(x), \ldots)$ and taking the utility metric $\sigma$ to be the $L^\infty$ 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 $\sigma$ value exceeds $\gamma$.

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

$$C^{\mu, \beta}_x(T) = \frac{\int_T e^{-\beta \sigma(f(x), y)} d\mu(y)}{\int_Y e^{-\beta \sigma(f(x), y)} d\mu(y)},$$

(3)

provided that the denominator is finite. Otherwise $C^{\mu, \beta}_x$ is undefined.

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 $C^{\mu, \beta}$ is $(2C\beta)$-differentially private for every $\mu$.

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 $\gamma, \delta > 0$.

2We use the word “conventional” here to refer to the rich subclass of exponential mechanisms whose score function is $\sigma$; however, not all exponential mechanisms fall in this class.
Suppose that $f$ is Lipschitz continuous and that the metric space $(X, \rho)$ is bounded. We now prove that $f$ is privacy-compatible if and only if the completion of the metric space $f(X)$ is compact. Observe that rescaling the metrics $\rho, \sigma$ does not affect the question of whether $f$ is privacy-compatible nor whether $f(X)$ has compact completion, but it does rescale the Lipschitz constant of $f$ and the diameter of $X$. Accordingly, we may assume without loss of generality that the Lipschitz constant of $f$ and the diameter of $X$ are both bounded above by 1, i.e.

$$\sigma(f(x_1), f(x_2)) \leq \rho(x_1, x_2) \leq 1$$

for all $x_1, x_2 \in X$.

**Definition 3.2.** A probability measure $\mu$ on a metric space $(X, \sigma)$ 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 $(\gamma, \delta)$-utility;
3. There exists a uniformly positive measure on $(f(X), \sigma)$;
4. The completion of $(f(X), \sigma)$ is compact.

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

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

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

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

We have

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

and

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

Hence, for every $x \in X$,

$$M_x(B(f(x), \gamma)) = \frac{a}{a + b} = 1 - \frac{b}{a + b} > 1 - \frac{e^{-\beta \gamma}}{e^{-\beta \gamma} m} = 1 - \frac{1}{e^{\beta \gamma/2} m} = 1 - \delta.$$
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 $Y$. Now define a probability measure $\mu$ supported on the countable set $C = \cup_{i=1}^{\infty} C_i$, by specifying that for $y \in C$, $\mu(y) = \sum_{i:y \in C_i} \left( \frac{1}{2^{n(i)}} \right)$. Equivalently, one can describe $\mu$ by saying that a procedure for randomly sampling from $\mu$ 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 $\mu$ 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 $y \in Y$, 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 $\mu(B(y, r)) \geq \mu(y_{i,j}) \geq \frac{1}{2^{n(i)}}$. The right side depends only on $r$ (and not on $y$), hence $\inf_{y \in Y} \mu(B(y, r))$ is strictly positive, as desired.

We prove the contrapositive. Suppose that the completion of $Y$ 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 $Y = f(X)$, we may choose points $x_i$ such that $y_i = f(x_i)$ for all $i \geq 1$. Suppose we are given a mechanism $M$ that achieves $r$-utility with probability at least $1/2$. For every $\alpha > 0$ we must show that $M$ does not achieve $\alpha$-differential privacy. The relation $\sum_{i=1}^{\infty} M_{x_i}(B(y_i, r)) \leq 1$ implies that there exists some $i$ such that

$$M_{x_i}(B(y_i, r)) < e^{-\alpha}/2.$$  \hfill (5)

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

$$M_{x_i}(B(y_i, r)) > 1/2.$$  \hfill (6)

Combining (5) with (6) leads to

$$M_{x_i}(B(y_i, r)) > e^\alpha M_{x_1}(B(y_i, r)) \geq e^{\alpha \rho(x_i,x_1)} M_{x_1}(B(y_i, r)),$$  \hfill (7)

hence $M$ violates $\alpha$-differential privacy.

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 efficient privacy with multiple queries. In Proc. ACM Symposium on Theory of Computing (STOC), 2010. to appear.
A Appendix

Lemma A.1. If $f : X \rightarrow Y$ has Lipschitz constant 1, then the conventional exponential mechanism $M\mu;\beta$ achieves $(2\beta)$-differential privacy.

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

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

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

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

$$
\int_Y e^{-\beta \sigma(f(x), y)} \, d\mu(y) \geq \int_Y e^{\beta \sigma(f(x), y)} \, d\mu(y)
$$

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

$$
\geq e^{-\beta \rho(x, z)} \int_Y e^{\beta \sigma(f(x), y)} \, d\mu(y).
$$

The inequality $M_x(T) \leq e^{2\beta \rho(x, z)} M_z(T)$ follows upon taking the quotient of these two inequalities.

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

1. The completion of $X$ is a compact topological space.

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

3. For every $r > 0$, $X$ 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 $X$ 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 $i \geq k$; these two properties immediately imply that $z_1, z_2, \ldots$ is a Cauchy sequence as desired.

The construction begins by defining $S_0 = X$. 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 $X$. 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 $X$ is compact.

(1) $\Rightarrow$ (3) If $X$ 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 $X$, violating compactness.

(3) $\Rightarrow$ (2) Given $r > 0$, let $B(x_1, r/2), \ldots, B(x_n, r/2)$ be a maximal collection of disjoint balls of radius $r/2$. (Such a collection must be finite, by property (3).) The balls $B(x_1, r), \ldots, B(x_n, r)$
cover $X$, because if there were a point $y \in X$ 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. □