Randomness in Cryptography January 31, 2013

Lecture 4: Introduction to Differential Privacy

Lecturer: Yevgeniy Dodis *Scribe: Umut Orhan*

1 Last Time

Last time we talked about,

- Impossibility of privacy with weak, block and SV sources
- MACs with (enchanced) block/SV sources

2 Differential Privacy (DP)

Given a database containing confidential information, we would like to allow learning of statistical information about the contents of database without violating the privacy of any of its individual entries. The traditional notion of privacy is not suitable, because it only allows negligible information to be revealed from the database. Therefore a new notion of privacy is needed to allow a better trade-off between privacy and utility.

Informal definition of DP: Compared to traditional privacy, differential entropy has

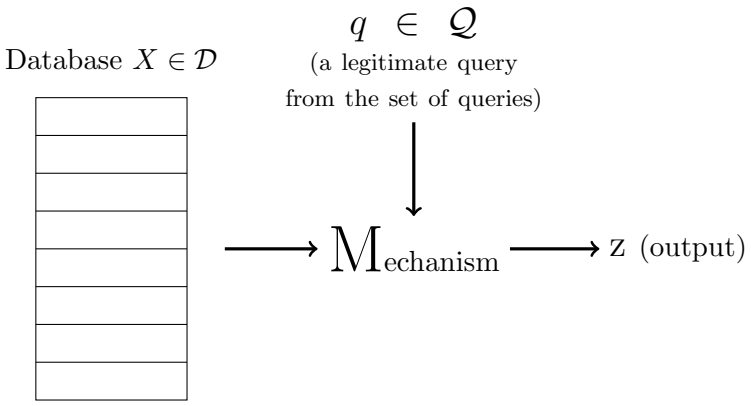
- "weaker" correctness/utility/accuracy
- "stronger" security (but with non-negligible ε)

Setting: We have the following setting.

- A sensitive database $x \in \mathcal{X}$, where $\mathcal{X}$ is the space of databases
- A discrete distance function between databases $\Delta : \mathcal{X} \times \mathcal{X} \to \mathbb{N} = \{0, 1, \cdots\}$
- Neighboring databases: $x, x' \in \mathcal{X}$ s.t. $\Delta(x, x') = 1$

Example:

- $\mathcal{D}$: universe of databases
- $\mathcal{X}$: the subsets of $\mathcal{D}$
- n : (maximum) number of records
- $\mathcal{D}_n$: all subsets of size $\leq n$, often make $\mathcal{X} = \mathcal{D}_n$
- $\delta(x, x')$: size of symmetric difference



Goal: The goal is to answer the query $q(X)$, while conserving the privacy. In essence we wish to approximate the true answer $q(X)$ with z, without revealing too much information. $\mathcal{Q}$ is defined the set of query functions,

$$\mathcal{Q} = \{q : \mathcal{X} \to \mathbb{Z}\}.$$

For simplicity, the codomain is selected to be the set of integers. For now, only one (arbitrary) query function will be considered.

To achieve the goal of privacy, the mechanism, M, should not be deterministic, and it should be randomized. Consequently, let $\mathcal{R}$ be the source of randomness with a sample r. Then the output of the query, generated by the mechanism, depends on the randomness source in the mechanism.

$$z = M(X, q; r), \text{ or equivalently } z \leftarrow M(X, q).$$

DEFINITION 1 A mechanism M is $(\mathcal{R}, \varepsilon)$-*differentially private* for $\mathcal{Q}$ if for any neighboring pair $X, X' \in \mathcal{X}$ and $\forall q \in \mathcal{Q}$, $\mathsf{RD}(M(x, q; \mathcal{R}), M(x', q; \mathcal{R})) \leq \varepsilon$, i.e. $\forall z$ (outcome)

$$\Pr_R(M(X, q; R) = z) \leq e^{\varepsilon} \Pr_R(M(X', q; R) = z), \text{ or equivalently}$$

$$\left(\iff \Pr_R(\mathsf{Eve}(M(X, q; R)) = 1) \leq e^{\varepsilon} \Pr_R(\mathsf{Eve}(M(X', q; R)) = 1) \right).$$

$\diamondsuit$ Equivalency of the definitions can be justified by noting that $e^{\varepsilon} \approx 1 + \varepsilon$ when $\varepsilon \ll 1$ and also $\mathsf{RD}(A, B) \leq \varepsilon \Rightarrow \mathsf{SD}(A, B) \leq e^{\varepsilon} - 1 \approx \varepsilon$.

Properties:

- Using the triangle inequality for RD, we obtain,

$$\forall X, X' \in \mathcal{X}, \mathsf{RD}(M(X, q; R), M(X', q; R)) \leq \varepsilon \Delta(X, X').$$

- If further $\Delta(X, X') \leq n$, then $\mathsf{RD}(M(X, q; R), M(X', q; R)) \leq \varepsilon \cdot n$. Let's consider the following case, n is polynomial over a secondary parameter, k, and ε is negligible over k. Consequently, $\varepsilon \cdot n$ is negligible over k and $\forall X, X', q \Rightarrow M(X, q) \approx M(X', q)$. This means if ε is negligible there is no "public utility", and answers to the queries are same for all databases.

> Negligible function: A function $\varepsilon(t) : \mathbb{N} \to \mathbb{R}$ is negligible if for every positive polynomial $p(t)$, $\exists N \in \mathbb{N}$, s.t. $\forall t > N, |\varepsilon(t)| < \frac{1}{p(t)}$.

- In contrast to traditional privacy, no external secret key, or other secrets are available to get utility/correctness.
- We cannot use a negligible ε, however a tiny non-negligible constant is utilizable.
- $(\mathcal{R}, \varepsilon)$ differential privacy is constructed using RD. For large ε, SD becomes too weak.

DEFINITION 2 A mechanism M is (R, ρ)-accurate w.r.t. $\mathcal{Q}$, if $\forall X \in \mathcal{X}, \forall q \in \mathcal{Q}$,

$$\mathbb{E}_R[|M(X, q; \mathcal{R}) - q(X)|] \leq \rho.$$

♢

It might be useful to note that, there is no trivial trade-off between ε and ρ. Some extreme examples,

- $M(X, q; \mathcal{R}) = q(X) \Rightarrow \varepsilon = \infty, \rho = 0$. In other words, if the mechanism just returns the original answer to the query no privacy, but perfect accuracy.
- $M(X, q; \mathcal{R})$ is constant $\Rightarrow \varepsilon = 0$, and a large $\rho = 0$. In essence, if the mechanism is not trying to answer the query according to the database privacy is protected perfectly, however answer to the query becomes very inaccurate.

DEFINITION 3 $\mathcal{Q}$ admits nontrivial differential privacy w.r.t. $\mathcal{R}$ if there exists a function $\rho(\varepsilon)$ s.t. $\forall \varepsilon > 0$ there exists a mechanism M_ε that is $(\mathcal{R}, \varepsilon)$-differentially private and $(\mathcal{R}, \rho(\varepsilon))$-accurate. Then we call $\mathcal{M} = \{M_\varepsilon\}$ a class of accurate and private mechanisms for $\mathcal{Q}$ w.r.t. $\mathcal{R}$. ♢

It is important to note that ρ does not depend on n, the number of records.

Counting Queries Example: Let $\mathcal{X} = \mathcal{D}_n$, and $p : \mathcal{D} \to \{0, 1\}$ be a given a property. Consider $q_p(X)$ to be the number of databases $a \in \mathcal{X}$ s.t. $p(a)$ is true. Correspondingly $\forall p$, the codomain of q_p is $\{0, \cdots, n\}$. A useful example of $\mathcal{Q}$ would be the set of counting queries, i.e. $\mathbb{C} = \{q_p | \forall p : \mathcal{D} \to \{0, 1\}\}$.

$\forall M$ which is (R, ε)-DP and (R, ρ)-accurate, $\exists M'$ such that M' is (R, ε)-DP and (R, ρ)-accurate and $\mathsf{Range}(M')$ is the maximum range of Q (e.g. $[0, n]$ for $\mathbb{C}$). M' is called the truncation of M, i.e. $M' = \mathsf{Trunc}(M)$. As expected, any deterministic operation on the mechanism can only improve security.

Lemma 1 *For M is $(\mathcal{R}, \varepsilon)$ secure, (R, ρ)-accurate and non-negative, $\forall X, X' \in \mathcal{X}$ and $q \in \mathcal{Q}$,*

$$\frac{q(X') - \rho}{q(X) + \rho} \leq e^{\varepsilon \cdot \Delta(X, X')}.$$

Proof: Let,

$$\alpha = \frac{\mathbb{E}[M(X', q; R)]}{\mathbb{E}[M(X, q; R)]}.$$

Using Definition 2 (accuracy),

$$\alpha \geq \frac{q(X') - \rho}{q(X) + \rho}.$$

By Definition 1 (differential privacy),

$$\alpha = \frac{\sum_z z \Pr(M(X', q; R)}{\sum_z z \Pr(M(X, q; R)} \leq \frac{\sum_z z e^{\varepsilon \cdot \Delta(X, X')} \Pr(M(X, q; R)}{\sum_z z \Pr(M(X, q; R)} = e^{\varepsilon \cdot \Delta(X, X')}.$$

Therefore,

$$\frac{q(X') - \rho}{q(X) + \rho} \leq e^{\varepsilon \cdot \Delta(X, X')}.$$

□

Corollary 2 *$\mathcal{Q} = \mathbb{C}$ and $\rho \leq \frac{n}{4} \Rightarrow \varepsilon \geq \frac{1}{n}$.*

Proof: Take q, X, X' s.t. $q(X) = 0, q(X') = n$ and $\Delta(X, X') \leq n$.

$$e < 3 = \frac{n - \frac{n}{4}}{\frac{n}{4}} \leq \frac{n - \rho}{n} \leq e^{\varepsilon \cdot n} \Rightarrow \varepsilon \cdot n > 1$$

□

Corollary 3 *If $q(X) = 0 \Rightarrow \rho \geq \Omega(\max_{X' \in Ball(X, \frac{1}{\varepsilon})} q(X')$, e.g. if $\mathcal{Q} = \mathbb{C} \Rightarrow \rho \geq \Omega(1/\varepsilon)$.*

Proof:

$$\forall X' \in Ball(X, \frac{1}{\varepsilon}) \Rightarrow \quad \frac{q(X') - \rho}{\rho} \leq e^{\varepsilon \cdot \Delta(X, X')} \leq e^{\varepsilon \cdot \frac{1}{\varepsilon}} \leq e \Rightarrow \quad \rho \geq \frac{1}{e + 1} q(X')$$

□

This corollary motivates putting a restriction on the query s.t. it doesn't change much on the neighboring databases.

DEFINITION 4 Sensitivity of $q \in \mathcal{Q}$ is defined as

$$\mathsf{Sen}(q) = \max_{X, X' : \Delta(X, X') = 1} |q(X) - q(X')|.$$

Let $\mathcal{Q}_S$ to be defined as $\mathcal{Q}_S = \{q : \mathsf{Sen}(q) \leq S\}$. ♢ Note $\mathbb{C} \subseteq \mathcal{Q}_1$, and the focus of the remainder of the notes will be $\mathcal{Q} = \mathcal{Q}_1$. In particular, we know $\rho = \Omega(1/\varepsilon)$ for $\mathcal{Q}_1$.

Question 1 *Can we match $\rho = \Omega(1/\varepsilon)$ with $\mathcal{R} = \mathcal{U}$? (Yes.)*

DEFINITION 5 M is additive noise (AN) if $\exists$ noise function $e(r)$ s.t. $M(X, q; r) = q(X) + e(r)$, where $e(r)$ is independent of q and X. ♢

Note 1: $(\mathcal{R}, \rho)$ utility for AN $\iff \mathbb{E}[|e(R)|] \leq \rho$.

Note 2: What about $(\mathcal{R}, \varepsilon)$-DP?
Take any X, X', q s.t. $\Delta(X, X') = 1, q(X') = q(X) - 1$. To satisfy differential privacy condition, $\forall z = q(X) + \alpha$,

$$\left.\begin{array}{c} \Pr(M(X,q;R) = z) = \Pr(e(R) = \alpha) \\ \Pr(M(X',q;R) = z) = \Pr(e(R) = \alpha + 1) \end{array}\right\} \Rightarrow \forall \alpha,\ \frac{\Pr(e(R) = \alpha)}{\Pr(e(R) = \alpha + 1)} \in [e^{-\varepsilon}, e^{\varepsilon}].$$

This can be achieved using the following distribution, i.e. discrete Laplace distribution,

$$\Pr(E = \alpha) = \frac{1 - e^{-\varepsilon}}{1 + e^{-\varepsilon}} e^{-\varepsilon|\alpha|}.$$

In this distribution, for every interval $[\alpha, \alpha + \frac{1}{\varepsilon}]$, $\Pr(E = \alpha), \cdots, \Pr(E = \alpha + \frac{1}{\varepsilon})$ are with the factor of e from each other.

How to sample from discrete-Laplacian distribution?

Let $B_1, B_2, \cdots$ be a family of independent Beurnoulli coins where $\Pr(B_i = 1) = 1 - e^{-\varepsilon} \approx \varepsilon$ (e.g. if $\varepsilon = 2^{-}i$, we can sample using unbiased coins). Let E^+ be the smallest i s.t. $B_i = 1$. In other words, let E^+ be the number of times we need to wait until we get the first one. Let B_f be a fair Beurnoulli coin and B_0 be a Beurnoulli coin with $\Pr(B_0 = 1) = \frac{1-e^{-\varepsilon}}{1+e^{-\varepsilon}}$. The variable corresponding to the discrete Laplace dist. (E) can be sampled as

$$E = \begin{cases} 0 & \text{, if } B_0 = 1 \\ E^+ & \text{, if } B_0 = 0 \text{ and } B_f = 1 \\ -E^+ & \text{, if } B_0 = 0 \text{ and } B_f = 0 \end{cases} \tag{1}$$

As expected $\forall \alpha \geq 0$,

$$\frac{\Pr(E^+ = \alpha)}{\Pr(E^+ = (\alpha + 1))} = \frac{(e^{-eps})^{a-1}(1 - e^{-\varepsilon})}{(e^{-eps})^{a}(1 - e^{-\varepsilon})} = e^{\varepsilon}$$

Let's consider $\rho = \mathbb{E}(|E|) = E(E^+) = \Omega(\frac{1}{\varepsilon})$, which corresponds to optimal ρ in $\mathcal{Q}_1$.

Theorem 1 *$M(X, q) \to q(X) + \mathsf{DLap}_{(0,1/\varepsilon)}$ is AN mechanism (w.r.t. $\mathcal{U}$) for $\mathcal{Q}_1$, which is $(\mathcal{U}, \varepsilon)$-DP and $(\mathcal{U}, \varepsilon)$-accurate.*

Question 2 *How much entropy do we need? Can we show,*

$$\mathbf{H}_{\infty}(R) = \Omega\left(\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}\right)?$$

Now, let's consider more realistic sources than $\mathcal{U}$. Letting $\mathfrak{R} = \{R\}$, we can define $(\mathfrak{R}, \varepsilon)$-DP and $(\mathfrak{R}, \rho)$-accuracy and non-triviality w.r.t. $\mathfrak{R}$ to mean it holds $\forall \mathcal{R} in \mathfrak{R}$.

Lemma 4 *If M is an AN for $\mathcal{Q}_1$, with error function e, then extractor $\mathsf{Ext}(r) = e(r)(\mod r)$ is $(\mathfrak{R}, \varepsilon)$-secure bit extractor for $\mathfrak{R}$.*

Proof: Take $\forall q$ and $\forall$ neighboring X, X' s.t. $q(X') = q(X) + 1$ and $\mathsf{Eve}(z) = z \ (\mod 2)$.

$$\Pr(\mathsf{Eve}(q(X) + e(R)) = 1) = \Pr(e(R) = (1 + q(X)) \ (\mod 2))$$

$$\Pr(\mathsf{Eve}(q(X') + e(R)) = 1) = \Pr(e(R) = q(X)) \ (\mod 2))$$

$$\Rightarrow \frac{\Pr(e(R) = 0 \ (\mod 2))}{\Pr(e(R) = 1 \ (\mod 2))} \in [e^{-\varepsilon}, e^{\varepsilon}] \Rightarrow \text{ bias}(e(R) \ (\mod 2) \leq e^{\varepsilon} - 1 \approx \varepsilon$$

□

Corollary 5 *There is no AN, non-trivial mechanism for weak, block & even, SV source.*

Theorem 2 *If $k < m - log(\rho\varepsilon) - \Omega(1)$, then no $(Weak_k(m), \varepsilon)$-DP and $(Weak_k(m), \rho)$-accurate mechanism exists.*

Proof: Assume there exists such a mechanism. Start with $\mathfrak{R} = \text{Weak}_m(k) = \{\mathcal{R} \in \{0,1\}^m | \mathbf{H}_\infty(\mathcal{R}) \geq k\}$.
$\forall \mathcal{R} \in \mathfrak{R}$, take $\forall q, X, X'$ s.t. $1 \leq \Delta(X, X') \leq \frac{1}{2\varepsilon}$. By ε-DP,

$$\mathsf{RD}(M(X, q; R), M(X', q; R)) \leq \varepsilon \cdot \Delta(X, X') \leq \frac{1}{2}$$

$$\Rightarrow \mathsf{SD}(M(X, q; R), M(X', q; R)) \leq 1.$$

Letting $f(R) = M(X, q; R)$ and $g(R) = M(X', q; R)$, from Theorem 2 of Lecture 3,

$$\mathsf{SD}(f(R), g(R)) < 1, \forall R \in \text{Weak}_m(k) \Rightarrow \Pr_{r \xleftarrow{} \mathcal{U}}[f(r) \neq g(r)] < 2^{k-m+2} = 4 \cdot 2^{k-m} \quad (2)$$

This means functions f and g are distinguishable with a very low probability.
Let's consider $q(X) = \text{size}(X) = |X|$, and

$$X_0 = \emptyset \subset X1 \subset \cdots \subset X_{8\rho\varepsilon} \text{ s.t. } |X_i| = \frac{i}{2\varepsilon}.$$

Letting $f_i(r) = M(X_i, q; r)$, from (2) for f_i and f_{i+1}, we obtain

$$\Pr_{r \xleftarrow{} \mathcal{U}_m} (f_i(r) \neq f_{i+1}) \leq 4 \cdot 2^{k-m}$$

$$\Rightarrow \Pr_{r \xleftarrow{} \mathcal{U}_m} (f_0(r) \neq f_{8\rho\varepsilon}(r)) \leq 32\rho\varepsilon \cdot 2^{k-m} \quad (3)$$

If we consider truncation of the mechanism, i.e. $M' \triangleq \text{trunc}(M)$, then the truncated mechanism is still ε-DP and ρ-accurate since $0 \leq q(X_i) \leq 4\rho$ on $X_0, X_1, \cdots, X_{8\rho\varepsilon}$.

By ρ-accuracy, using the definition,

$$\alpha = \mathbb{E}_R[f^*(U) - g^*(U)] \geq (4\rho - \rho) - \rho = 2\rho.$$

Also, using (3),

$$\alpha \leq \Pr(f^*(U) \neq g^*(U)) \cdot [\max(g^*) - \min(f^*)] \leq (32\rho\varepsilon 2^{k-m}) \cdot 4\rho.$$

$$\Rightarrow 2\rho \leq (64\rho\varepsilon) \cdot (2\rho) \cdot 2^{k-m} \Rightarrow k \geq m - log(64\rho\varepsilon) = m - log(\rho\varepsilon) - \Omega(1)$$

□

In particular, non-trivial mechanism for $\mathcal{Q}_1$ w.r.t. $\text{Weak}_k(m)$ requires $k \geq m - \Omega(1)$.

Project 1 *Extend the impossibility (or show possibility) for (k, m) block source when $k < m - \log(\rho\varepsilon) - \Omega(1)$.*

What about block sources where either $k > m - \log(\rho\varepsilon) - \Omega(1)$ or $m \leq \log(\rho\varepsilon)$, e.g. $\gamma - \mathsf{SV}$? Let's try to extend the negative result using Theorem 6 of Lecture 3, $\forall \mathcal{R} \in \mathsf{SV}(\gamma, N)$,

$$\mathsf{SD}(f(R), g(R)) \leq \varepsilon \Rightarrow \Pr_{r \xleftarrow{r} \mathcal{U}_N}(f(r) \neq g(r)) \leq \frac{2\varepsilon}{\gamma}. \tag{4}$$

For two neighboring databases, $\Delta(X, X') = 1$, $\mathsf{RD}(f(R), g(R)) \leq \varepsilon \Rightarrow \mathsf{SD}(f(R), g(R)) \leq \varepsilon$, where $f(R) = M(X, q; R)$ and $g(R) = M(X', q; R)$. By (4),

$$\Pr(f(r) \neq g(r)) \leq \frac{2\varepsilon}{\gamma} = \Omega(\varepsilon).$$

Hence, for a non-trivial difference between functional jumps should be less than $\Omega(\frac{1}{\varepsilon})$. But there is no accuracy guarantee since $\rho > \frac{1}{\varepsilon}$. Therefore we cannot extend impossibility.

In the next lecture we will define a stronger notion of consistency and ε-consistent sampling to investigate the differential privacy for SV sources.

References

[1] Yevgeniy Dodis, Adriana Lpez-Alt, Ilya Mironov and Salil P. Vadhan. Differential Privacy with Imperfect Randomness. In *CRYPTO 2012: 497-516*