| | | | |
|---|---|---|---|
| Distributing Trust and Blockchains | | Date: | *November $11^{th}$, 2019* |
| Instructor: *Sujit Prakash Gujar* | | Scribes: | Tarun Mohandas, Kishan Singhal |

# Lecture 19

## 1 Computing Privacy Loss

A randomized algorithm $M$ is said to be $\epsilon$ differentially private if $\forall E \subseteq Range(M)$ and $\forall x, y$ such that $\|x - y\| \leq 1$ then

$$P(M(x) \in E) \leq e^{\epsilon} P(M(y) \in E) \, \forall E \subseteq S$$

This can be modified for computing privacy loss in the following way. First we consider a certain output $O \in E$ and $P(M(x) = O)$. We have

$$P(M(x) = O) \leq e^{\epsilon} P(M(y) = O)$$

$$\frac{P(M(x) = O)}{P(M(y) = O)} \leq e^{\epsilon}$$

$$ln \frac{P(M(x) = O)}{P(M(y) = O)} \leq ln \, e^{\epsilon}$$

$$ln \frac{P(M(x) = O)}{P(M(y) = O)} \leq \epsilon$$

Therefore, a mechanism $M(x)$ is said to be $\epsilon$ differentially private if

$$ln \frac{P(M(x) = O)}{P(M(y) = O)} \leq \epsilon$$

## 2 Calculating Privacy Loss for Mechanisms

### 2.1 Coin Toss Mechanism

Now let us take a mechanism and try to compute the privacy loss from the formula we derived above. Let us consider a database that has $n$ types of elements: $x = (x_1, x_2, x_3 ... x_n)$ and $y$ differs from $x$ only on one element: $y = (x_1, x_2, x_3, ... \bar{x}_i .... x_n)$, that is $\|x - y\|_1$. The mechanism $M$ can be described as follows:

- Toss a Coin: If heads, respond $x_i$ the true value.
- If tails, toss the coin again. If heads, return 1 and if tails return 0.

Let $\tilde{x_i}$ be the output $O$ returned by the mechanism. Privacy loss can be considered for this mechanism as follows:

$$ln \frac{P(\tilde{x_i} = 1 | x_i = 1)}{P(\tilde{x_i} = 1 | x_i = 0)}$$

$P(\tilde{x_i} = 1 | x_i = 1)$ is probability of first coin toss being heads OR probability of first coin toss being tails and second coin toss being heads. Therefore $P(\tilde{x_i} = 1 | x_i = 0) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$. There is only one possibility out of 4 for $P(\tilde{x_i} = 1 | x_i = 0)$ to be 1. Therefore $P(\tilde{x_i} = 1 | x_i = 0) = \frac{1}{4}$

$$ln \frac{P(\tilde{x_i} = 1 | x_i = 1)}{P(\tilde{x_i} = 1 | x_i = 0)} = ln \frac{\frac{3}{4}}{\frac{1}{4}} = ln 3$$

Therefore this mechanism is $ln3$ differentially private mechanism.

### 2.2 $\epsilon$-differentially private mechanism

Let us consider another mechanism $M$ as follows: If $\tilde{x_i} = O$ is the output of the mechanism. The following describes getting $\tilde{x_i}$ with corresponding probability.

$$\tilde{x_i} = \begin{cases} x_i & w.p.\ \frac{e^\epsilon}{e^\epsilon+1} \\ \bar{x_i} & w.p.\ \frac{1}{e^\epsilon+1} \end{cases}$$

For this mechanism, we can clearly see that:

$$P(\tilde{x_i} = 1 | x_i = 1) = \frac{e^\epsilon}{e^\epsilon + 1}$$

$$P(\tilde{x_i} = 1 | x_i = 0) = \frac{1}{e^\epsilon + 1}$$

$$ln \frac{P(\tilde{x_i} = 1 | x_i = 1)}{P(\tilde{x_i} = 1 | x_i = 0)} = ln \frac{\frac{e^\epsilon}{e^\epsilon+1}}{\frac{1}{e^\epsilon+1}} = ln\, e^\epsilon = \epsilon$$

Therefore this is a $\epsilon$-differentially private mechanism.

# 3 Composition Theorem

**Theorem 1.** *If $f : S \to S'$ be any deterministic mapping on output of $\epsilon$-differentially private mechanism, then $f \circ M$ is also $\epsilon$-differentially private.*

**Given**: $P(M(x) \in E_2) \leq e^\epsilon P(M(y) \in E_2)\ \forall E_2 \subseteq S$
**To Prove**: $P(f \circ M(x) \in E_1) \leq e^\epsilon P(f \circ M(y) \in E_1)\ \forall E_1 \subseteq S'$

*Proof.*

$$E = \{x \in S | f(x) \in E_1\}$$

This is a direct result of taking pre-image of $E_1$ i.e. $f^{-1}(E_1)$

$$P(f \circ M(x) \in E_1) = P(M(x) \in E_2) \leq e^\epsilon P(M(y) \in E_2) = e^\epsilon P(f \circ M(y) \in E_1)$$

□

Any $\epsilon$-differentially private mechanism is $k\epsilon$-differentially private for a group of size $k$. That is: $\|x - y\|_1 \leq k \Rightarrow$ its $k\epsilon$ differentially private.

# 4 Advantages of $\epsilon$-differential privacy

- Neutralize linkages attacks - protects against reidentification
- Quantification of privacy loss
- Post processing does not increase privacy loss
- Group privacy
- Composition - if two mechanisms $\epsilon_1$ and $\epsilon_2$ are composed, then resulting mechanism is $\epsilon_1$ and $\epsilon_2$