IV.1 Pointwise Maximal Leakage

1) Some subjects that could be studied for a wider understanding

- Differential Privacy (DP)
- Local Differential Privacy (LDP)
- Information Privacy
- Differential Indistinguishability
- Quantitative Information Flow
- g-leakage Framework (studied above)
- ML Framework (studied above)
- Privacy Loss Random Variable

2) Understanding the limitation of ML that has been pointed out

"The maximal leakage is defined for the average outcome Y"

From what I understand, is that ML is a characteristic of the channel, but does not tell how much a specific outcome for Y leaks.

Example: We consider $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$, and the following channels:

$$P_{Y|X} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \qquad Q_{Y|X} = \begin{bmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

→ Same ML equal to $\log(2)$

But if Y=3 then this is an insane leakage for $P_{Y|X}$ but not for $Q_{Y|X}$.

3) Threat Model Setup : the guessing adversary

Suppose X is a random variable over $\mathcal{X}$. The output of the channel is given by Y of the finite alphabet $\mathcal{Y}$. This channel is called privacy mechanism.

An adversary is interested in guessing $U \in \mathcal{U}$ that is a possibly randomized function of X characterized by $P_{U|X}$. The adversary observes the outcome $y \in \text{Supp}(Y)$, and deduces $\hat{U} \in \mathcal{U}$, by some $P_{\hat{U}|Y}$. We define:

$$\ell_u(X \to y) = \log \frac{\sup_{P_{\hat{U}|Y}} \mathbb{P}(\hat{U}=U \mid Y=y)}{\max_{u \in \mathcal{U}} P_U(u)}$$

Guessing observing y

Guessing without any observation

A natural inference is to take the supremum over $P_{U|X}$:

$$\ell(X \to y) = \sup_{P_{U|X}} \ell_u(X \to y) = \log \sup_{P_{U|X}} \frac{\sup_{P_{\hat{U}|Y}} \mathbb{P}(\hat{U}=U \mid Y=y)}{\max_{u \in \mathcal{U}} P_U(u)}$$

We just defined the Pointwise Maximal Leakage.

Theorem 1:

$$\ell(X \to y) = \log \max_{x \in \text{Supp}(X)} \frac{P_{X|Y}(x|y)}{P_X(x)}$$

Proof: In the proof, we use the following lemma. Let define as Issa did:

$$\mathcal{L}(X \to Y)[U] = \sup_{P_{\hat{U}|Y}} \log \frac{\sum_y \mathbb{P}(\hat{U}=U|Y=y) P_Y(y)}{\max_{u \in \mathcal{U}} P_U(u)} = \log \frac{\sum_y \max_{u \in \mathcal{U}} P_{UY}(u,y)}{\max_{u \in \mathcal{U}} P_U(u)}$$

Lemma 1: $e^{\mathcal{L}(X \to Y)[u]} = \sum_y \mathbb{P}_Y(y)\, e^{l_u(X \to y)}$

Proof: The only thing that is not trivial, is to show that

$$\sum_y \mathbb{P}_Y(y) \sup_{\mathbb{P}_{\hat{U}|Y}} \mathbb{P}(\hat{U}=U \mid Y=y) = \sup_{\mathbb{P}_{\hat{U}|Y}} \sum_y \mathbb{P}(\hat{U}=U|Y=y)\mathbb{P}_Y(y)$$

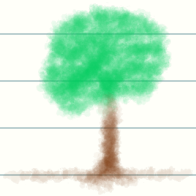
But the key is to notice that the first sup should be $\sup_{\mathbb{P}_{\hat{U}|Y=y}}$.

Taking the supremum over U, we get $e^{\mathcal{L}(X \to Y)} = \sum_y \mathbb{P}_Y(y)\, e^{l(X \to y)}$

Therefore, $\sum_y \underbrace{\mathbb{P}_Y(y)\, e^{l(X \to y)}}_{(1)} = \sum_y \underbrace{\max_{x \in \mathrm{Supp}(X)} \mathbb{P}_{Y|X}(y|x)}_{(2)}$.

But for all $y \in \mathcal{Y}$, $(1) \leq (2)$. Indeed, we know that the supremum is obtained when $U = X$ and $\hat{U}$ is the best likelihood estimator, so (1) becomes:

$\frac{\max_{x \in \mathcal{X}} \mathbb{P}_{XY}(x,y)}{\max_{x \in \mathcal{X}} \mathbb{P}_X(x)}$ and (2) is $\max_{x \in \mathcal{X}} \mathbb{P}_{Y|X}(y|x)$.



Moreover, for all $\bar{x} \in \mathcal{X}$, $\mathbb{P}_X(\bar{x}) \leq \max_{x \in \mathcal{X}} \mathbb{P}_X(x)$

$$\text{then, } \mathbb{P}_X(\bar{x})\, \mathbb{P}_{Y|X}(y|\bar{x}) \leq \max_{x \in \mathcal{X}} \mathbb{P}_X(x)\, \mathbb{P}_{Y|X}(y|\bar{x})$$

$$\text{and maximizing over } \bar{x}, \quad \max_{x \in \mathcal{X}} \mathbb{P}_{XY}(x,y) \leq \max_{x \in \mathcal{X}} \mathbb{P}_X(x) \max_{x \in \mathcal{X}} \mathbb{P}_{Y|X}(y|x)$$

We instantly get $(1) \leq (2)$ for all $y \in \mathcal{Y}$. Since $\sum(1) = \sum(2)$, we get $(1) = (2)$ and finally,

$$l(X \to y) = \log \frac{\max_{x \in \mathrm{Supp}(X)} \mathbb{P}_{Y|X}(y|x)}{\mathbb{P}_Y(y)} = \log \max_{x \in \mathrm{Supp}(X)} \frac{\mathbb{P}_{X|Y}(x|y)}{\mathbb{P}_X(x)}.$$

$$= \log\left(\frac{1}{\mathbb{P}_Y(y)}\right) - \log\left(\frac{1}{\max_{x \in \mathrm{Supp}(X)} \mathbb{P}_{Y|X}(y|x)}\right)$$

Interpretation: i) $PML(y, \pi)$ decreases with $\mathbb{P}_Y(y)$

The bigger $\mathbb{P}_Y(y)$ is, the lower the leakage about y. (sounds logic: knowing that y happens does not give us a lot of information)

ii) $PML(y, \pi)$ increases with $\max_{x \in Supp(X)} \mathbb{P}_{Y|X}(y|x)$

Here, $\mathbb{P}_{Y|X}(y|x)$ is the "maximum possible leakage" one can get about y.

Remark: We know that ML is divergence-based, where

$$\mathcal{L}(X \to Y) = D_\infty(\mathbb{P}_X \mathbb{P}_Y^* \mid \mathbb{P}_{X,Y})$$

Then PML also is divergence-based:

$$l(X \to y) = D_\infty(\mathbb{P}_{X|Y=y} \mid \mathbb{P}_X)$$

→ This gives us the perfect intuition about PML.

4) Threat Model Setup: the g-leakage framework

As seen before, the posterior vulnerability in the g-leakage framework is given by $V_g(X|Y) = \mathbb{E}_{\hat{Y}}[V_g(X|Y=\hat{Y})]$.

Naturally, a good choice for the pointwise posterior vulnerability is $V_g(X|Y=y)$, so that we can define the dynamic g-leakage: $l_g(X \to y) = \log \frac{V_g(X|Y=y)}{V_g(X)}$.

In the paper we have $\max_w \mathbb{E}[g(w, X) \mid Y=y]$ which is exactly the same.

This second expression makes the link with the randomized function POV expression, namely $l_u(X \to y)$. The dynamic min-entropy leakage is given by $l_{id}(X \to y)$.

5) Equivalence theorem

Theorem 2: The g-leakage and the U-leakage are equivalent:

$$\forall P_{U|X},\ \exists g: \mathcal{X}^2 \to [0,1],\quad \ell_U \equiv \ell_g$$

$$\forall g: \mathcal{X}^2 \to [0,1],\ \exists P_{U|X},\quad \ell_g \equiv \ell_U$$

Corollary 1: New definition of PML:

$$\ell(X \to y) = \sup_{g:\mathcal{X}^2 \to [0,1]} \ell_g(X \to y)$$

6) Properties

We can make $\ell(X \to y)$ conditional over a random variable Z. Each outcome $z \in \mathcal{Z}$ defines a conditional channel $P_{Y|X,Z=y}$. Then, the conditional leakage is easily defined as

$$\ell(X \to y \mid z) = \log \max_x \frac{P_{X|YZ}(x|y,z)}{P_{X|Z}(x|z)}.$$

For the denominator to be non zero, the max is taken over $\mathrm{Supp}(X|Z=z)$.

Remark: $\ell(X \to y \mid z) = D_\infty(P_{X|Y=y,Z=z} \,\|\, P_{X|Z=z}) = \max_x i_{XY|Z=z}(x,y)$

The mutual information density of a point $(x,y) \in \mathcal{X} \times \mathcal{Y}$ is given by the formula

$$i_{XY}(x;y) = \begin{cases} 0 & \text{if } P_X(x)P_Y(y) = 0 \\ \log \dfrac{P_{XY}(x,y)}{P_X(x)P_Y(y)} & \text{otherwise} \end{cases} \quad \in \mathbb{R} \cup \{-\infty\}.$$

Results:

$$\ell(X \to y) = \max_x i_{XY}(x;y) \qquad I(X,Y) = \mathbb{E}(i(X;Y))$$

Properties: i) $0 \leq \ell(X \to y) \leq \max_x \log\left(\frac{1}{P_X(x)}\right)$

ii) If $X \to Y$ is totally random (i.e. independency) then $\forall y \in \mathcal{Y}$, $\ell(X \to y) = 0$.

iii) If $X \to Y$ is deterministic, then $\forall y \in \mathcal{Y}$, $\ell(X \to y) = \log\left(\frac{1}{P_Y(y)}\right)$.

iv) For two successive channels $X \to Y \to Z$, $\ell(X \to z) \leq \ell(Y \to z)$.

v) For two successive channels $X \to Y \to Z$, $\ell(X \to z) \leq \max_y \ell(X \to y)$.

vi) If the Markov graph $\begin{matrix} & & Z \\ & \nearrow & \uparrow \\ X & & \to Y \end{matrix}$ holds, then

$$\ell(X \to y \mid z) = \ell(X \to y) - i(y; z).$$

vii) For the double channel $X \to Y, Z$, $\ell(X \to y, z) \leq \ell(X \to z) + \ell(X \to y \mid z)$.

Remark: iii) Imagine a deterministic channel that can be seen as a surjective function $\mathcal{X} \to \mathcal{Y}$. Then the leakage equals $\log\left(\frac{1}{P_Y(y)}\right)$ which is coherent since if $P_Y(y)$ is lower then the leakage should be higher.

Summary: We defined the dynamic g-leakage $\ell_g(X \to y) = \log \frac{V_g(X \mid Y=y)}{V_g(X)}$, from which we derive:

- The dynamic min-entropy leakage : $\ell_{id}(X \to y)$
- The pointwise maximal leakage : $\sup_g \ell_g(X \to y)$

7) Privacy guarantees

Instead of a fixed number $\mathcal{L}(X \to Y) = \log\left(\mathbb{E}_Y\left(e^{\ell(X \to Y)}\right)\right)$, we consider the random variable $\ell(X \to Y)$ which has much more information

about the leakage. $\ell(X \to Y)$ is like a very rich measure of the leakage of a channel $X \to Y$. We then can construct lots of different guarantees:

Maximal Leakage guaranty: $\mathbb{E}_Y(e^{\ell(X\to Y)}) \leq e^{\varepsilon}$ (also called average case)

Almost-sure guaranty: $\ell(X \to y) \leq \varepsilon$ almost everywhere

or, equivalently, $\operatorname{ess\,sup} \ell(X \to y) \leq \varepsilon$

↳ We say that $X \to Y$ satisfies ε-PML guaranty.

Tail bound guaranty: $\mathbb{P}_Y(\ell(X \to Y) > \varepsilon) \leq \delta$

↳ We say that $X \to Y$ satisfies (ε, δ)-PML guaranty.

8) Some Data-Processing

Let A be a property of a channel. We say that A is:

- transitive if for any successive channels $X \to Y \to Z$, we have

$$A(X \to Y) \text{ and } A(Y \to Z) \Rightarrow A(X \to Z)$$

- closed under postprocessing if for any successive channels $X \to Y \to Z$:

$$A(X \to Y) \Rightarrow A(X \to Z)$$

- closed under preprocessing if for any successive channels $X \to Y \to Z$:

$$A(Y \to Z) \Rightarrow A(X \to Z)$$

- closed if for any successive channels $X \to Y \to Z$:

$$A(X \to Y) \text{ or } A(Y \to Z) \Rightarrow A(X \to Z)$$

Examples: i) Almost-sure guaranty is closed

ii) Tail bound guaranty is closed under post-processing.

Let $\mathcal{E}$ be a measurable subset of $\mathcal{Y}$. Then the event maximal leakage (EML) is defined by $\ell(X \to \mathcal{E}) = \max_x \log \frac{\mathbb{P}_{Y|X=x}(\mathcal{E})}{\mathbb{P}_Y(\mathcal{E})} = \max_x \log \mathbb{E}_Y[i(x;Y)] - \log \mathbb{P}_Y(\mathcal{E})$.

Another way to define it is by defining the following deterministic channel $Y \to Z$: $\mathcal{Z} = \{0,1\}$ and $Z = \mathbb{1}_{\mathcal{E}}(Y)$. This way, $\ell(X \to \mathcal{E}) = \ell(X \to 1)$.

Proof: $e^{\ell(X \to 1)} = \max_x \frac{\mathbb{P}_{Z|X}(1|x)}{\mathbb{P}_Z(1)}$ but $\mathbb{P}_Z(1) = \mathbb{P}_Y(\mathcal{E})$

and $\mathbb{P}_{Z|X=x}(1) = \mathbb{P}_{Y|X=x}(\mathcal{E})$ so we got the result.