First article: Pointwise Maximal Leakage (PML)

I) Motivations

→ Side-channel: Any unconventional way of getting a slight information about a secret.

→ Mutual information as a leakage measure? Not very adapted here.

II) Maximal Leakage

1) Threat Model and Definition

Setup of a guessing adversary:

- X is a secret
- U is what actually interests the adversary
- Y is the observation of X from which the adversary wants to guess U.

The SSH example

- X are the real keystroke timings
- Y are the observed keystroke timings
- U is the actual password.

We assume we have the Markov chain $U - X - Y$.

Reminder: We say that a Markov graph $G=(V,E)$ holds, if for any separating set S separating V_1 from V_2 (where $V_1 \sqcup S \sqcup V_2 = V$) then

$$V_1 \mid S \perp\!\!\!\perp V_2 \mid S$$

Let $\hat{U}(Y)$ (or even simpler $\hat{U}$) the best estimator of U given Y, i.e. that minimizes $\mathbb{P}(\hat{U}=U)$.

Before observing Y, the probability of guessing U with no clue is $\max_{Z \perp\!\!\!\perp U} \mathbb{P}(Z=U)$ that can be reduced to $\max_{u \in \mathcal{U}} \mathbb{P}_U(u)$ where $\mathcal{U}$ is the set of all the possible values taken by U.

The maximal leakage for this specific example is the ratio $\frac{\mathbb{P}(\hat{U}=U)}{\max_{u \in \mathcal{U}} P_U(u)}$.

It means:

"By how much can I multiply my prior probability by observing Y"

It is always preferable to take the log, and $\log_2$ means:

"With a leakage of n bits, I can increase my probability by 2^n"

For a general definition, we would need to fix what the attacker wants. But measuring the leakage of Y over X, this makes no sense that we need to set what is U. In our case, U could be:

- The entire password
- The entire password + more information about the computer
- The first letter of the password
- Whether it is a vowel
- etc...

The SDPI (Strong Processing Data Inequality) allows to maximize over all the possible setups:

Let X and Y be two random variables on respectively $\mathcal{X}$ and $\mathcal{Y}$.

The Maximal Leakage from X to Y is defined by:

$$\mathcal{L}(X \to Y) = \sup_{\substack{U - X - Y - \hat{U} \\ U, \hat{U} \in \mathcal{U}}} \log \frac{\mathbb{P}(\hat{U}=U)}{\max_{u \in \mathcal{U}} P_U(u)}$$

Where $\mathcal{U}$ is any set.

2) Main result

The form provided above is more for the intuition and less for the computation. Indeed, we will show that

$$\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X} \\ P_X(x) > 0}} P_{Y|X}(y|x)$$

This is, by definition, the Sibson mutual information of order infinity, $I_\infty(X;Y)$.
The definition is valid only is $\mathcal{X}$ and $\mathcal{Y}$ are finites alphabets.
If $\mathcal{X} = \{0, 1\}$ then

$$\mathcal{L}(X \to Y) = \log_2\left(1 + \frac{1}{2}\left\|P_{Y|X}(\cdot|0) - P_{Y|X}(\cdot|0)\right\|_1\right)$$

where $\|\cdot\|_1$ is the L_1 distance (Manhattan).

Some properties:

i) If $X - Y - Z$ holds, then $\mathcal{L}(X \to Z) \leq \min\left(\mathcal{L}(X \to Y), \mathcal{L}(Y \to Z)\right)$
↳ Data Processing Inequality

ii) $\mathcal{L}(X \to X) = H_0(X) = \log |\mathrm{Supp}(X)|$

iii) $\mathcal{L}(X \to Y) \leq \min\left(\log|\mathcal{X}|, \log|\mathcal{Y}|\right)$

iv) $\mathcal{L}(X \to Y) \geqslant I(X;Y) \geqslant 0$

v) $\mathcal{L}(X \to Y) = 0 \iff X \perp\!\!\!\perp Y$

vi) If $(X_i, Y_i)_{1 \leqslant i \leqslant l}$ are mutually independent, then

$$\mathcal{L}\left((X_i) \to (Y_i)\right) = \sum_{i=1}^{l} \mathcal{L}(X_i \to Y_i)$$

vii) $\mathcal{L}(X \to Y)$ only depends on $\mathrm{Supp}(X) \subseteq \mathcal{X}$ and $P_{Y|X} \in [0,1]^{\mathcal{Y} \times \mathrm{Supp}(X)}$.

viii) If $\mathrm{Supp}(X)$ is fixed, then $e^{\mathcal{L}(X \to Y)}$ is convex in $P_{Y|X}$.

For a leakage measure, we often consider i), v) and vi) as axiomatic.
Mutual information does check this.

3) Basic notions

It's never a bad idea to get back to the basics.

For a finitely valued random variable X, the entropy of X (or more Shannon's entropy) is defined by:

$$H(X) = \sum_{x \in Supp(X)} p(x) \log\left(\frac{1}{p(x)}\right) = \mathbb{E}\left(\log\left(\frac{1}{p(X)}\right)\right)$$

where p is the distribution of X.

This is the only continuous function $H : \mathcal{X}_{<\infty} \to \mathbb{R}$, that satisfies:

i) $0 \leq H(X) \leq H\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ where $n = |Supp(X)|$

ii) $H(p_1, \dots, p_n, 0) = H(p_1, \dots, p_n)$

iii) $H(X,Y) = H(X) + H(Y|X)$ this is the Chain Rule

iv) $H\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = \log(n)$ (we can choose any base $b > 1$ we want)

or, equivalently:

i) $\forall \sigma \in S_n$, $H(p_1, \dots, p_n) = H(p_{\sigma(1)}, \dots, p_{\sigma(n)})$

ii) For any $p_n = q_1 + q_2$, $H(p_1, \dots, q_1, q_2) = H(p_1, \dots, p_n) + p_n H\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}\right)$

As a measure of uncertainty, we can also mention Rényi entropy for $\alpha \in \mathbb{R}^+ \setminus \{1\}$:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_{x \in Supp(X)} p(x)^\alpha \right)$$

However, we only have a weak chain rule: "If $X \perp\!\!\!\perp Y$, then $H_\alpha(X,Y) = H_\alpha(X) + H_\alpha(Y)$"

The mutual information measures how much information do two variables share:

$$I(X \to Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X)$$

However the fact that it is symmetric is not very meaningful for channels.

We always have that:

i) $0 \leq I(X,Y) \leq H(X), H(Y)$ ii) $I(X,Y)=0 \Leftrightarrow X \perp\!\!\!\perp Y$

This is more a measure of what X and Y have in common than a leakage of the channel $X \to Y$.

Remark: We can try to expand the mutual information:

$$I(X,Y) = H(X) + H(Y) - H(X,Y) = \sum_x \mathbb{P}_X(x) \log\left(\frac{1}{\mathbb{P}_X(x)}\right) + \sum_y \mathbb{P}_Y(y) \log\left(\frac{1}{\mathbb{P}_Y(y)}\right)$$

$$- \sum_{x,y} \mathbb{P}_{X,Y}(x,y) \log\left(\frac{1}{\mathbb{P}_{X,Y}(x,y)}\right) = \sum_{x,y} \mathbb{P}_{X,Y}(x,y) \log\left(\frac{1}{\mathbb{P}_X(x)\mathbb{P}_Y(y)}\right) \sum_{x,y} \mathbb{P}_{X,Y}(x,y) \log\left(\frac{1}{\mathbb{P}_{X,Y}(x,y)}\right)$$

$$\boxed{I(X,Y) = \sum_{x,y} \mathbb{P}_{X,Y}(x,y) \log\left(\frac{\mathbb{P}_{X,Y}(x,y)}{\mathbb{P}_X(x)\mathbb{P}_Y(y)}\right)}$$

A Discrete Memory Channel (DMC) has an input of $\mathcal{X}$ and an output of $\mathcal{Y}$, and induces a transition function $W: \mathcal{Y}|\mathcal{X} \to \mathbb{R}^+$.

In this setup, mutual information can be written as:

$$I(X,Y) = \sum_{x,y} \mathbb{P}_X(x) W(y|x) \log\left(\frac{W(y|x)}{\sum_x \mathbb{P}_X(x) W(y|x)}\right)$$

The capacity is the value of the mutual information, when the input distribution maximizes it: $C = \max_{\mathbb{P}_X} I(X,Y)$.

As we have seen before, the maximal leakage only depends on $\text{Supp}(\mathbb{P}_X)$:

$$\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X} \\ \mathbb{P}_X(x)>0}} W(y|x)$$

What is the relationship between capacity and maximal leakage?

The capacity is a useful tool for the normal framework where a message U is encoded $X^n = f_n(U)$, then passes trough the channel, and is finally decoded : $\hat{U} = g_n(Y^n)$.

Let M_n be an increasing sequence of numbers. We say that $(M_n)_{n\in\mathbb{N}}$ is accepted by the channel if :

$$\sup_{U \in [M_n]} \mathbb{P}(\hat{U} \neq U \mid X^n = f_n(U)) \xrightarrow{n\to+\infty} 0$$

The rate of an accepted $(M_n)_{n\in\mathbb{N}}$, is the value $R = \lim_{n\to+\infty} \frac{1}{n} \log_2(M_n)$ in bits per channel use. The capacity is therefore the supremum over all accepted rates :

$$C = \sup_{(M_n) \text{ accepted}} \lim_{n\to+\infty} \frac{1}{n} \log_2(M_n) \ .$$

If we want to be more formal, we should replace (M_n) by (f_n, g_n, M_n). This is called a "code".

We have the incredible fact that : $C = \max_{P_X \in \Delta(\mathcal{X})} I(X,Y)$ which is constant for a fixed channel. → Discrete...

If we define $C_\alpha = \max_{P_X \in \Delta(\mathcal{X})} I_\alpha(X,Y)$, then we have $\mathcal{L}(X \to Y) = C_\infty$.

Since $I(X,Y) \leq I_\infty(X,Y)$ we therefore have that $\mathcal{L}(X \to Y) \geq C$, always assuming that $\mathrm{Supp}(X) = \mathcal{X}$.

→ For security aspects, Maximal Leakage is more sensitive than Shannon's capacity.

When we define a distance between laws, a good way of measuring the leakage is by measuring the distance between the real joint law and the product of $\mathbb{P}_X$ by another law in $\mathcal{U}$. This way, if $D(\cdot\|\cdot)$ measures the distance between two laws, then:

$$I(X,Y) = \inf_{\mathbb{P}_Y^* \in \mathcal{U}} D(\mathbb{P}_{X,Y} \| \mathbb{P}_X \mathbb{P}_Y^*)$$

If $D(\cdot\|\cdot)$ is the Kullback-Leibler divergence, then $I(X,Y)$ is the mutual information and the symmetry gives us $\mathbb{P}_Y^* = \mathbb{P}_Y = \sum_x \mathbb{P}_{X,Y}(x,\cdot)$.

However, the symmetry does not hold in general. For instance, if we take Rényi divergences, then a good way to define a leakage measure is:

$$I_\alpha(X,Y) = \inf_{\mathbb{P}_Y^* \in \mathcal{U}} D(\mathbb{P}_{X,Y} \| \mathbb{P}_X \mathbb{P}_Y^*)$$

This is the real case. The distance tells us how much information Y gives about X.

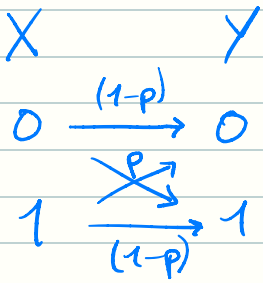
This is the set of all possibilities where Y gives no information about X.

This is not symmetric in general anymore. The maximal leakage is also given by $\mathcal{L}(X \to Y) = I_\infty(X,Y)$ that derives from a distance for laws. Wonderful isn't it?

4) Some examples

Ex 1: $X \sim \mathcal{B}(q)$ $0 < q < 1$ $p \leq \frac{1}{2}$

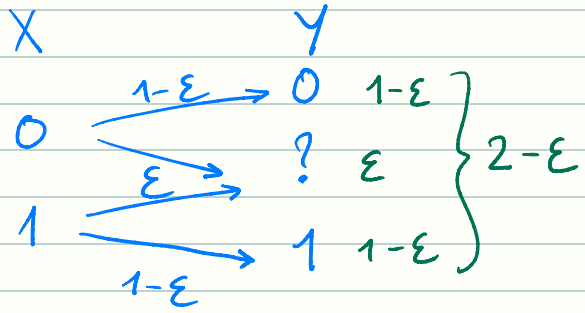
$$\mathcal{L}(X \to Y) = \log(2(1-p)) = 1 + \log(1-p)$$



→ If $p = 0$, I can increase the guessing probability by $2^1 = 2$.

→ If $p = \frac{1}{2}$, I cannot increase my guessing probability.

Ex 2: $X \sim B(q)$ $0<q<1$ $0 \leq \varepsilon < 1$



$$\mathcal{L}(X \to Y) = \log(2-\varepsilon)$$

Remark: i) For any deterministic channel, $\mathcal{L}(X \to Y) = \log(|Supp(Y)|)$. We also have that $|Supp(Y)| \leq |Supp(X)|$.

ii) In the definition of maximal leakage, a more intuitive version could be:

$$\mathcal{L}(X \to Y) = \sup_{\substack{u=f(x) \\ \hat{u}=f(y)}} \log \frac{\mathbb{P}(u=\hat{u})}{\max_{u \in \mathcal{U}} P_U(u)}$$

5) Proof of the main result

We want to prove the following:

$$\sup_{\substack{u-x-y-\hat{u} \\ u,\hat{u} \in \mathcal{U}}} \log \frac{\mathbb{P}(\hat{u}=u)}{\max_{u \in \mathcal{U}} P_U(u)} = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X} \\ P_X(x)>0}} P_{Y|X}(y|x)$$

Proof: Let $\mathcal{L}(X \to Y)[U] = \log \frac{\sum_y \max_{u \in \mathcal{U}} P_{UY}(u,y)}{\max_{u \in \mathcal{U}} P_U(u)}$. This is clearly the best estimator we can get, and therefore $\mathcal{L}(X \to Y) = \sup_{U-X-Y} \mathcal{L}(X \to Y)[U]$.

The probability that the best estimator actually gives the best result can be developed by the Markov chain $U-X-Y$:

$$\sum_y \max_{u \in \mathcal{U}} P_{UY}(u,y) = \sum_y \max_{u \in \mathcal{U}} \sum_x P_X(x) P_{U|X}(u,x) P_{Y|X}(y,x)$$

$$\leq \sum_y \max_{u \in \mathcal{U}} \sum_x P_X(x) P_{U|X}(u,x) \left[\max_{x' \in \mathcal{X}} P_{Y|X}(y,x')\right]$$

$$= \sum_y \left[\max_{x' \in \mathcal{X}} P_{Y|X}(y,x')\right] \max_{u \in \mathcal{U}} \sum_x P_X(x) P_{U|X}(u,x)$$

$$= \max_{u \in \mathcal{U}} P_U(u) \sum_y \left[\max_{x \in \mathcal{X}} P_{Y|X}(y,x)\right]$$

So LHS $\leq$ RHS. For the reversed inequality, we construct a $P_{U|X}$ that reaches the RHS.

Let $p^* = \min_{x \in \mathcal{X}} P_X(x) > 0$ (we can assume $\text{Supp}(X) = \mathcal{X}$ without any loss of generality). For each $x \in \mathcal{X}$ let $k(x) = \frac{P_X(x)}{p^*}$.

We first define $\mathcal{U} = \mathcal{X} \times [\![1, \lceil k(x) \rceil]\!]$, or more formally, $\mathcal{U}$ is defined by $\bigcup_{x \in \mathcal{X}} \{(x,1), \dots, (x, \lceil k(x) \rceil)\}$.

For $x \in \mathcal{X}$, U returns an element of $\{(x,1), \dots, (x, \lfloor k(x) \rfloor)\}$ with the same probability equal to $\frac{p^*}{P_X(x)}$ and returns $(x, \lceil k(x) \rceil)$ with the resting probability $1 - \lfloor k(x) \rfloor \frac{p^*}{P_X(x)}$.

We have $\max_{u \in \mathcal{U}} P_U(u) = p^*$ obtained for the value $x \in \mathcal{X}$ such that $P_X(x) = p^*$.
Moreover, $\sum_y \max_{u \in \mathcal{U}} P_{UY}(u,y) = \sum_y \max_{u \in \mathcal{U}} \sum_x P_X(x) P_{U|X}(u|x) P_{Y|X}(y|x)$

$$= \sum_y \max_{x \in \mathcal{X}} P_X(x) P_{U|X}((x,1)|x) P_{Y|X}(y|x)$$

$$= p^* \sum_y \max_{x \in \mathcal{X}} P_{Y|X}(y|x)$$

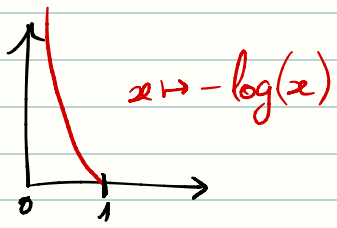
We just proved that LHS $\geq$ RHS.

III/ The g-leakage

1) Min-entropy leakage

Given a random variable X, the min-entropy is defined as:

$$H_\infty(X) = -\log\left(\max_x \mathbb{P}(X=x)\right)$$



The bigger it is, the harder the guessing is.

Let consider X and Y, two random variables, the the conditional min-entropy is given by:

$$H_\infty(X|Y) = -\log\left(\max_x \mathbb{P}(X=x|Y)\right)$$

↓ Take the average over all possible values of y.

Knowing Y, it should be easier for the attacker to guess the correct value of X. Therefore $H_\infty(X) \geq H_\infty(X|Y)$.

The min-entropy leakage is finally defined by:

$$L_\infty(X,Y) = H_\infty(X) - H_\infty(X|Y)$$

Remark: H_∞ has some entropy properties: the flatter the distribution is, the bigger H_∞ is. It measures the randomness by considering the best 1-try possible attacker.

Min-entropy leakage can be written as:

$$L_\infty(X,Y) = \log\left(\frac{\max_x \mathbb{P}(X=x|Y)}{\max_x \mathbb{P}(X=x)}\right)$$

2) The main problem with min-leakage

Example: Let X contain 1000 different passwords 10-bits long.
Let Y be one of these 1000 passwords.

$L_\infty(X,Y) = 10$ bits

We now consider a user having one of these 1000 passwords. We want to measure how this user is in trouble. The probability of guessing it is 2^{-10} without the leak, and $0{,}001 + 0{,}999 \cdot 2^{-10} = 0{,}00198\ldots$

Then $L_\infty(X_{user}, Y) = \log\left(\frac{0{,}00198}{2^{-10}}\right) \approx 1{,}016$ bits

And the attacker might be very happy!

But a whole password has been completely leaked! Min-entropy is not sufficient for this type of threats. The g-leakage is a generalization of the min-entropy leakage.

3) The g-function

Let $(\mathcal{X}, \mathcal{Y}, \mathbb{P}_{Y|X})$ be a channel. We call $\pi := \mathbb{P}_X$ the prior, and the posterior is given by $\pi(x)\mathbb{P}_{Y|X}(y,x)$. The prior vulnerability is given by $V(X) = \max_x \pi(x)$, and the posterior vulnerability is written $V(X|Y) = \sum_{y \in \mathcal{Y}} \max_x \pi(x)\mathbb{P}_{Y|X}(y,x)$.

As a reminder, $H_\infty(X) = -\log(V(X))$, $H_\infty(X|Y) = -\log(V(X|Y))$ and $L_\infty(X,Y) = \log\frac{V(X|Y)}{V(X)}$.
The min-capacity is then defined as $ML_\infty(\mathbb{P}_{Y|X}) = \sup_{\mathbb{P}_X} L_\infty(X,Y)$, realized for the uniform prior.

Let $x \in \mathcal{X}$ be the secret. Given a guess $w \in \mathcal{X}$, we define between 0 and 1 how much is the attacker happy with w by the gain function $g : \mathcal{X}^2 \to [0,1]$.

Example: In the min-entropy case, g is defined by $g : (w,x) \mapsto \begin{cases} 1 \text{ if } x = w \\ 0 \text{ otherwise} \end{cases}$
This is called the identity g function, g_{id}.

This is better for clarity if we use $\mathcal{W}$ for the space of guesses instead of $\mathcal{X}$ but this doesn't change the core idea. Formally, $g : \mathcal{W} \times \mathcal{X} \to [0,1]$.

- The prior g-vulnerability is defined by $V_g(X) = \max_w \mathbb{E}[g(w,X)]$.
- The posterior g-vulnerability is defined by

$$V_g(X|Y) = \sum_y \max_w \mathbb{E}[C(X,y)\, g(w,X)]$$

$$= \mathbb{E}_Y\left[\max_w \mathbb{E}_X\left[\frac{\mathbb{P}_{X|Y}(X,Y)}{\mathbb{P}_X(X)}\, g(w,X)\right]\right]$$

Other forms are shown on the paper but this one makes the link with PML.

$$\begin{aligned} V_g(\pi, C) &= \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} \pi[x]C[x,y]g(w,x) \\ &= \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x,y)g(w,x) \\ &= \sum_{y\in\mathcal{Y}} p(y) V_g(p_{X|y}) \end{aligned}$$

- The g-entropy is defined by $H_g(X) = -\log(V_g(X))$
- The g-leakage is defined by $L_g(X,Y) = H_g(X) - H_g(X|Y)$.
- The g-capacity is defined by $ML_g(\mathbb{P}_{Y|X}) = \sup_{\mathbb{P}_X} L_g(X,Y)$.

Theorem:

$\forall w \in \{\text{vulnerability, entropy, leakage, capacity}\}$, gid-w coincides with min-w.

4) Special g-functions

Let $d : \mathcal{X} \to \mathbb{R}_+$ be a distance. Since $\mathcal{X}$ is finite, let $d^* = \max_{c\in\mathcal{X}^2} d(c)$.

A gain function induced from d could be $1 - \frac{d(x,y)}{d^*}$, or $e^{-d(x,y)}$.

Let $\mathcal{W} \subseteq \mathcal{P}(\mathcal{X})$ be the set of possible guesses. Then it is usual that $\mathcal{W}$ is a partition of $\mathcal{X}$ and that we want to know on which set the secret x is.

Let $k \in \mathbb{N}^*$. We then define $\mathcal{W} = \{W \in \mathcal{P}(\mathcal{X}) \mid |W| = k\}$. Let us compute the prior of $g_{\mathcal{W}}$-vulnerability: $\max_{W} \mathbb{E}[g_{\mathcal{W}}(w, X)] = \max_{x_0 \neq x_1 \neq x_2} \pi(x_0) + \pi(x_1) + \pi(x_2)$.

This is exactly what we expect of a prior vulnerability within k tries!

We then easily define:

- the k-tries prior vulnerability
- the k-tries posterior vulnerability
- the k-tries min-entropy
- the k-tries min-leakage
- the k-tries min-capacity

Going further:

IV. MATHEMATICAL PROPERTIES OF g-VULNERABILITY AND g-LEAKAGE

V. RESULTS ON CHANNEL CAPACITY

VI. COMPARING CHANNELS

VII. RELATED WORK

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

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

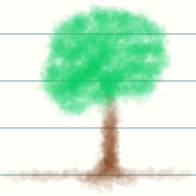
But the key is to notice that the first sup should be $\sup_{\mathbb{P}_{\hat{U}|Y=y}}$.

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) $PML(y, \varepsilon)$ 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) $PML(y, \varepsilon)$ 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:

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

→ 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

$$V_g(X|Y) = \mathbb{E}_Y\left[\max_w \mathbb{E}_X\left[\frac{\mathbb{P}_{X|Y}(X,Y)}{\mathbb{P}_X(X)} g(w,X)\right]\right].$$

Naturally, a good choice for the pointwise posterior vulnerability could be something like

$$\ell_g(X \to y) = \max_w \mathbb{E}_X\left[\frac{\mathbb{P}_{X|Y}(X,y)}{\mathbb{P}_X(X)} g(w,X)\right].$$

In the paper we have $\sup_{\mathbb{P}_{W|Y=y}} \mathbb{E}_{X,W}[g(W,X) \mid Y=y]$ which is exactly the same.

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

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 $\begin{matrix} & Z \\ & / \\ X & \to Y \end{matrix}$ holds, then

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

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

7) Going further...

III. PRIVACY GUARANTEES

IV. RELATIONSHIP TO OTHER PRIVACY/STATISTICAL NOTIONS

V. CONCLUSIONS

V) Rethinking Disclosure Prevention with PML

1) Differential Privacy

How to keep anonymity when computing statistics about medical records, personal incomes, survey responses, ... ?

Let $\mathcal{X}$ be the set of values that individuals can have, and let $\mathcal{D} = \mathcal{X}^N$ where N is the number of individuals.
d and d' of $\mathcal{D}$ are called neighbors if they differ in only one individual.

Let $\mathcal{Y}$ be the set of possibly outputted values and let Y be the outputted value characterized by $\mathbb{P}_{Y|D}$. Since $\mathcal{Y}$ can be infinite, we call $\mathcal{T}$ the set of measurable sets by any $\mathbb{P}_{Y|D=d}$. We call $\mathcal{M}(d)$ the random variable $Y|D=d$. We say that $\mathcal{M}$ is ε-differentially private iff $\forall S \in \mathcal{T}, \log\left(\frac{\mathbb{P}(\mathcal{M}(d) \in S)}{\mathbb{P}(\mathcal{M}(d') \in S)}\right) \leq \varepsilon$, for any pair of neighbors (d, d').

k-anonymity characterizes a dataset $d \in \mathcal{D}$ and it is just the fact that $\forall x \in \mathcal{X}, |\{i \in [N] \mid d_i = x\}|$ is either bigger than k or equal to zero.

2) Preliminaries

Again we recall the setup:

useful — secret — output — guess

$$U - X - Y - \hat{U}$$

$\mathbb{P}_{U|X}$ $\mathbb{P}_X$ $\mathbb{P}_{Y|X}$ $\mathbb{P}_{\hat{U}|Y}$ → Imagine those as deterministic.

The min-entropy is defined as $H_\infty(X) = -\log(\max_x \mathbb{P}_X(x))$ or, accordingly to its name, $H_\infty(X) = \log(\min_x \frac{1}{\mathbb{P}_X(x)})$.

The min-leakage is given by $L_\infty(X,Y) = H_\infty(Y) - H_\infty(Y|X)$.

The min-capacity is given by $C_\infty(\mathbb{P}_{Y|X}) = \sup_{\mathbb{P}_X} L_\infty(X,Y)$.

Let $(\Omega, \mathcal{T}, \mu)$ be a measure space and let P and Q random variables $\ll \mu$ of Ω with respective pmf p, q. The Renyi Divergence of order ∞ of P from Q is defined by $D_\infty(P\|Q) = \operatorname{ess\,sup}\left[\log(p) - \log(q)\right]$ assuming that $(-\infty)-(-\infty) = 0$.

We recall that Maximal Leakage is given by:

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

↳ Surprisingly, for a fixed support of X, this only depends on $\mathbb{P}_{Y|X}$

Well Pointwise Maximal Leakage is given by:

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

Let $\mathcal{P}$ be a subset of all possible distributions for X. A channel $X\to Y$ is said to satisfy $(\varepsilon, \mathcal{P})$-PML if $\forall \mathbb{P}_X \in \mathcal{P}$, $\mathbb{P}_Y(\{y\in\mathcal{Y} \mid \ell(X\to y) \le \varepsilon\}) = 1$

or, equivalently, $\sup_{\mathbb{P}_X\in\mathcal{P}} D_\infty(\mathbb{P}_{XY}\|\mathbb{P}_X\mathbb{P}_Y) \le \varepsilon$.

Proof of the equivalence: $D_\infty(\mathbb{P}_{XY}\|\mathbb{P}_X\mathbb{P}_Y) = \operatorname{esssup} \log i(x;y)$ and $\ell(X\to y) = \max_x i(x,y)$, so we just follow the definitions.

If the density of Y is continuous for a given $X=x$, then the two conditions above also are equivalent to $\forall \mathbb{P}_X\in\mathcal{P}$, $\forall y\in\mathcal{Y}$, $\ell(X\to y)\le\varepsilon$.

A channel $X\to Y$ is said to satisfy ε-PML if $\max_{x\in\mathcal{X}} i(x,y) \le \varepsilon$ almost everywhere on $\mathcal{Y}$.

The leakage capacity can be defined as

$$C_\mathcal{L}(X \to Y) = \max_{x,x'} \sup_y \log \frac{\mathbb{P}_{Y|X=x}(y)}{\mathbb{P}_{Y|X=x'}(y)}$$

Why sup instead of essup ?

Can be infinite

Theorem: $\sup_{\mathbb{P}_X} D_\infty(\mathbb{P}_{XY} \| \mathbb{P}_X \mathbb{P}_Y) = C_\mathcal{L}(X \to Y)$

Proof: On the LHS, we have $\log \frac{\mathbb{P}_{Y|X=x}(y)}{\mathbb{P}_Y(y)}$ (1). On the RHS, we have $\log \frac{\mathbb{P}_{Y|X=x}(y)}{\mathbb{P}_Y(y)} - \log \frac{\mathbb{P}_{Y|X=x'}(y)}{\mathbb{P}_Y(y)}$ (2). LHS ≤ RHS is clear since we can easily find an x' such that $\mathbb{P}_{Y|X}(y|x') \leq \mathbb{P}_Y(y)$. We now want LHS ≥ RHS.

Let $y^* \in \mathcal{Y}$ as close to the supremum as we need, and let x_{min} and x_{max} the values maximizing the expression. Then we chose the following $\mathbb{P}_X$: $\begin{cases} \mathbb{P}_X(x_{min}) = 1-\alpha \\ \mathbb{P}_X(x_{max}) = \alpha \end{cases}$

Then LHS ≥ $\max_x \log \frac{\mathbb{P}_{Y|X}(y^*|x)}{\mathbb{P}_Y(y^*)} = \log \frac{\mathbb{P}_{Y|X}(y^*|x_{max})}{\alpha \mathbb{P}_{Y|X}(y^*|x_{max}) + (1-\alpha)\mathbb{P}_{Y|X}(y^*|x_{min})}$ → Definition of x_{max}

The limit when $\alpha \to 0$ is RHS □

(1) Depends on $\mathbb{P}_X$ (2) Does not depend on $\mathbb{P}_X$

3) Impossibility of Absolute Disclosure Prevention

Global = Small entropy i.e. less useful information that is commonly known.