Plan of the course

Quantitative Information Flow
- Motivation, application examples
- Secrets and vulnerability
- Channels and leakage
- Multiplicative Bayes-Capacity
- Comparing systems, the lattice of information
- Applications and exercises
Hyper distributions

\[ \pi \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad C \begin{bmatrix} \frac{2}{4} & \frac{2}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \quad \text{joint} \begin{bmatrix} \frac{2}{8} & \frac{2}{8} \\ \frac{1}{8} & \frac{3}{8} \end{bmatrix} \]

Postersiors: \( \sigma^{y_1} = (2/3, 1/3), \sigma^{y_2} = (2/5, 3/5) \)

Output distribution \( \delta \in \mathcal{PY} \):

\[ \delta_y = p(y) = \sum_x p(x, y) = \sum_x \pi_x c_{xy} \quad \text{hence} \quad \delta = \pi c = (\frac{3}{8}, \frac{5}{8}) \]

Hyper distribution \( \delta \in \mathcal{P}^2 \mathcal{X} \):

\[ x_1 \begin{bmatrix} \frac{3}{8} & 5/8 \\ \frac{2}{3} & \frac{2}{5} \\ \frac{1}{3} & \frac{3}{5} \end{bmatrix} \]
Hyper distributions

\([\pi, C]\): hyper distribution obtained from \(\pi, C\)

\(\delta \in \mathcal{PY}\): outer distribution (output)

\(\sigma^y \in \mathcal{PX}\): inner distributions (posteriors)

\(\pi\) can be obtained by averaging the inners

\[\pi = \sum_y \delta_y \sigma^y\]

Abstract channel \(C\): mapping \(\mathcal{PX} \rightarrow \mathcal{P}^2\mathcal{X}\):

\[C(\pi) = [\pi, C]\]
After running $C$ on $\pi$ we get $[\pi, C]$

How vulnerable is $[\pi, C]$?

We have vulnerability measures on $\mathcal{P}\chi$, we need to extend them to $\mathcal{PP}\chi$

Natural choice: averaging:

$$V[\pi, C] = \sum_{y} \delta_y V(\sigma^y)$$

Natural geometric view

Low probability observations are not considered important!
Posterior vulnerability \( V[\pi, C] = \sum_y \delta_y V(\sigma^y) \)

Notation: \([\pi] = \) hyper assigning prob 1 to \( \pi \)

Then \( V[\pi] = V(\pi) \)

For Bayes-vulnerability:
\[
V_b[\pi, C] = \sum_y \max_x \pi_x C_{xy}
\]

For \( g \)-vulnerability:
\[
V_g[\pi, C] = \sum_y \max_w \sum_x \pi_x C_{xy} g(w, x)
\]
Leakage

$V[\pi, C]$ might be high because of $\pi$, not because of $C$!

eg. easy to guess password

What we care about is how much more vulnerable our secret becomes because of $C$

So we compare $V[\pi]$ and $V[\pi, C]$

Additive case:

$$\mathcal{L}^+(\pi, C) = V[\pi, C] - V[\pi]$$

Multiplicative case:

$$\mathcal{L}^\times(\pi, C) = \frac{V[\pi, C]}{V[\pi]}$$
Example

\[ \pi \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \quad C \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} y_1 & y_2 \\ 2/4 & 2/4 \\ 1/4 & 3/4 \end{bmatrix} \quad \text{hyper} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} 3/8 & 5/8 \\ 2/3 & 2/5 \\ 1/3 & 3/5 \end{bmatrix} \]

Prior vulnerability: \( V[\pi] = 1/2 \)

Individual posteriors: \( V[\sigma^1] = 2/3, V[\sigma^2] = 3/5 \)

Posterior vuln.: \( V[\pi, C] = 3/8 \cdot 2/3 + 5/8 \cdot 3/5 = 5/8 \)

Leakage: \( \mathcal{L}^+(\pi, C) = 1/8, \mathcal{L}^x = 5/4 \)
Zero Leakage

**Non-interfering channel**: all rows are the same, i.e. $C_{xy} = C_{x'y}$ for all $x, x', y$

Fix some arbitrary **full-support** prior $\pi$

NI iff input and output are **independent**

If NI then $\mathcal{L}^+(\pi, C) = 0$

Is the converse true?
Imperfect Cancer Test

\[
\begin{array}{c|cc}
\pi & Y & N \\
\hline
Y & 0.008 & 0.992 \\
N & 0.07 & 0.93 \\
\end{array}
\]

\[
\begin{array}{c|cc}
C & Y & N \\
\hline
Y & 0.90 & 0.10 \\
N & 0.07 & 0.93 \\
\end{array}
\]

\[
\begin{array}{c|cc}
\text{joint} & Y & N \\
\hline
Y & 0.00720 & 0.00080 \\
N & 0.06944 & 0.92256 \\
\end{array}
\]

Assuming a “generally healthy” population, the probability of cancer is low.

The probability of getting \( Y \) as a false positive is higher than that of a true positive!

Hence the best guess is always to guess no cancer! (the test by itself is useless)

Geometric view: posteriors fall on the same hyperplane!
Often we want to abstract from a specific prior

Capacity: maximize leakage over $\pi$

$$\mathcal{ML}^+(C) = \max_\pi \mathcal{L}^+(\pi, C)$$

$$\mathcal{ML}^x(C) = \max_\pi \mathcal{L}^x(\pi, C)$$
Multiplicative Bayes-Capacity

\[ \mathcal{M}\mathcal{L}_b^x(C) = \max_\pi \frac{V_b[\pi, C]}{V_b[\pi]} \]

Given for a **uniform prior**

**Theorem:** for any channel \( C \):

\[ \mathcal{M}\mathcal{L}_b^x(C) = \mathcal{L}_b^x(\pi_u, C) = \sum_y \max_x C_{xy} \]
Multiplicative Bayes-Capacity

Mult. Bayes-Capacity is 1 iff $C$ is non-interfering

\[ \begin{bmatrix} Y & N \\ Y & 0.5 & 0.5 \\ N & 0.5 \end{bmatrix} \begin{bmatrix} 0.90 & 0.10 \\ 0.07 & 0.93 \end{bmatrix} \begin{bmatrix} Y & N \\ Y & 0.45 & 0.05 \\ N & 0.035 & 0.465 \end{bmatrix} \]

Geometric view: the posterior must be the same as the prior
A universal upper bound for leakage!

**Theorem**
For any channel $C$, prior $\pi$ and gain function $g$:

$$ \mathcal{L}_g^\times (\pi, C) \leq \mathcal{M} \mathcal{L}_b^\times (C) = \sum_y \max_x C_{xy} $$
Let us construct the channel matrix

Input: \( K_1 K_2 K_3 \in \{000, 001, \ldots, 111\} \)

Output: \( out \in \{\text{OK, FAIL}\} \)

Assume the user string is \( x_1 x_2 x_3 = 110 \)

Different values of \( x_1 x_2 x_3 \) give different channel matrices, but they all have this kind of shape (seven inputs map to Fail, one maps to OK)
Password-checker 2

Assume the user string is \( x_1 x_2 x_3 = 110 \)

Assume the adversary can measure the execution time

Let us construct the channel matrix

Input: \( K_1 K_2 K_3 \in \{000, 001, \ldots, 111\} \)

Output: \( out \in \{\text{OK}, (\text{FAIL}, 1), (\text{FAIL}, 2), (\text{FAIL}, 3)\} \)

\[
\begin{array}{cccc|c}
000 & 1 & 0 & 0 & 0 \\
001 & 1 & 0 & 0 & 0 \\
010 & 1 & 0 & 0 & 0 \\
011 & 1 & 0 & 0 & 0 \\
100 & 0 & 1 & 0 & 0 \\
101 & 0 & 1 & 0 & 0 \\
110 & 0 & 0 & 0 & 1 \\
111 & 0 & 0 & 1 & 0 \\
\end{array}
\]
Let $C$ be a channel from $\mathcal{X}$ to $\mathcal{Y}$.

5.1 Show that for any prior $\pi$ and gain function $g$:

$$L_g^\times(\pi, C) \leq |\mathcal{Y}| \quad \text{and} \quad L_g^\times(\pi, C) \leq |\mathcal{X}|$$

5.2 Let $\pi_u$ be the uniform prior. Show that

$$(\forall g : L_g^\times(\pi_u, C) = 1)$$

if and only if $C$ is non-interfering.