CS 599: Formal Methods in Security and Privacy
Quantitative Information Flow

Marco Gaboardi
gaboardi@bu.edu

Alley Stoughton
stough@bu.edu
Wigderson Named Turing Awardee for Decisive Work on Randomness

By Neil Savage

Posted Apr 10 2024
We want to quantify the confidential information that leaks in what is considered nonconfidential.
Quantitative Information Flow Control

Quantitative information flow has been used for:

• Analyzing distributed protocols and scheme,
• Analyzing side-channel vulnerabilities and preventions.
• Analyzing crypto protocols,
• Analyze election protocols
• Analyze differential privacy mechanisms
• …
• The adversary has some prior $\pi_R$ on $R$ and it updates it after seeing $U$. 
Shannon Entropy

\[ H(X) = \sum_{x \in \mathcal{X}} \Pr[X = x] \log \left( \frac{1}{\Pr[X = x]} \right) = \mathbb{E} \left[ \log \left( \frac{1}{\Pr[X = x]} \right) \right] \]

- uncertainty about X
- expected amount of information gain by observing the value of the random variable,
- average number of bits required to transmit X optimally
Conditional Entropy

\[ H(X|Y) = \sum_y \Pr[Y = y] \cdot \sum_{x \in \mathcal{X}} \Pr[X = x|Y = y] \log\left(\frac{1}{\Pr[X = x|Y = y]}\right) \]

- If C is constant \( H(R|U) = H(R) \).
- If C is non constant and deterministic \( H(R|U) = 0 \), so:

\[ \text{Leakage}(U) = H(R) \]
Information leakage

Information leaked = initial uncertainty - remaining uncertainty

- Which could be

\[ \text{Leakage}(U) = H(R) - H(R | U) \]

- This is the mutual information between R and U
Shannon Entropy

We could think that:

“If a secret $X$ has distribution $\pi$, then an adversary’s probability of guessing the value of $X$ correctly in one try is at most $2^{-H(\pi)}$”

- This is false. E.g. for this distribution $H(\pi) \sim 2.44$, and $2^{-H(\pi)} \sim 0.18$
Same issue on conditional entropy

- Assume that $R$ is a uniformly distributed $8k$-bit integer with range $0 \leq R < 2^{8k}$, where $k \geq 2$. Hence $H(R) = 8k$.

- Consider these two programs:

  ```plaintext
  if $R \mod 8 = 0$ then $U := R$ else $U := 1$
  ```

  And

  ```plaintext
  $U := R \&\& 0^{7k-1}1^{k+1}$
  ```

- In both cases $H(R|U) \sim 7k-1$ suggesting that the number of guesses needed to guess $R$ is $2^{-(7k-1)}$
Bayes Vulnerability

\[ V(X) = \max_{x \in X} \Pr[X = x] \]

- In our case it is the max probability assigned by the prior \( \pi_R \).
- **Best choice** for a rational adversary to guess the secret in one try.
Bayes Vulnerability examples

\[ V(X) = \max_{x \in \mathcal{X}} \Pr[X = x] \]

- Consider \( \pi_R \) to be a uniform distribution over \( n \) outcomes. Then, \( V(\pi_R) = \frac{1}{n} \)

- Consider \( \pi_R \) to be the following distribution again, we have \( V(\pi_R) = 0.5 \)
We can use Bayes vulnerability to define a notion of entropy.

\[ H_{\text{min}}(X) = \log \frac{1}{V(X)} \]

This is actually known as min entropy, and it can be seen as the greatest lower bound of the information content in bits of observations of X.
Conditional Min Entropy

- We can have a conditional version of the previous notions

\[ H_{\text{min}}(X \mid Y) = \log \frac{1}{V(X \mid Y)} \]

- Where

\[ V(X \mid Y) = \sum_{y \in Y} \Pr[Y = y] \max_{x \in \mathcal{X}} \Pr[X = x \mid Y = y] \]
Information leakage v2

Information leaked = initial uncertainty - remaining uncertainty

• Which could be

\[ \text{Leakage}(U) = H_{\text{min}}(R) - H_{\text{min}}(R | U) \]
Bayes vulnerability and min entropy

We have:

\[ V(R \mid U) = 2^{H_{\text{min}}(R \mid U)} \]

- The expected probability that the adversary could guess R given U decreases exponentially with \( H_{\text{min}}(R \mid U) \).
Conditional Min Entropy

• Assume that \( R \) is a uniformly distributed 8k-bit integer with range \( 0 \leq R < 2^{8k} \), where \( k \geq 2 \). Hence \( H(R) = 8k \).

• Consider these two programs:

\[
\text{if } R \mod 8 = 0 \text{ then } U := R \text{ else } U := 1
\]

And

\[
U := R \&\& 0^{7k-1}1^{k+1}
\]

• For the first we have \( H_{\text{min}}(R|U) \sim 3 \) while for the second is still \( H_{\text{min}}(R|U) \sim 7k-1 \).
Conditional Min Entropy

- Assume that $R$ is a uniformly distributed $8k$-bit integer with range $0 \leq R < 2^{8k}$, where $k \geq 2$. Hence $H(R) = 8k$.

- Consider these two programs:
  
  if $R \mod 8 = 0$ then $U := R$ else $U := 1$

  And
  
  $U := R \ || \ 0^{8k-3}1^3$

- For both of them we have $H_{\min}(R|U) \approx 3$.

Is this reasonable?
Can we have a more general approach?
Gain function

• Suppose we have a set of secrets $X$ and a set of actions $W$, then a gain function $g$ is a function of type:

$$g : X \times W \rightarrow \mathbb{R}$$

• We can think about $g$ as a scoring function for actions on a secret
Gain function

• Suppose we have a set of secrets $X$ and a set of actions $W$, then a gain function $g$ is a function of type:

$$g : X \times W \rightarrow \mathbb{R}$$

• We can think about $g$ as a scoring function for actions on a secret

We could have a similar definition based on losses.
The best action for a rational adversary is the one that maximizes the expected gain.

\[ V_g(X) = \max_{w \in W} \sum_{x \in X} \Pr[X = x] \cdot g(w, x) \]
Example 3.3 With $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{W} = \{w_1, w_2, w_3, w_4, w_5\}$, let gain function $g$ have the (rather arbitrarily chosen) values shown in the following matrix:

$$
\begin{array}{c|cc}
\text{G} & x_1 & x_2 \\
\hline
w_1 & -1.0 & 1.0 \\
w_2 & 0.0 & 0.5 \\
w_3 & 0.4 & 0.1 \\
w_4 & 0.8 & -0.9 \\
w_5 & 0.1 & 0.2 \\
\end{array}
$$

To compute the value of $V_g$ on (say) $\pi = (0.3, 0.7)$, we must compute the expected gain for each possible action $w$ in $\mathcal{W}$, given by the expression $\sum_{x \in \mathcal{X}} \pi_x g(w, x)$ for each one, to see which of them is best. The results are as follows.

\[
\begin{align*}
\pi x_1 g(w_1, x_1) + \pi x_2 g(w_1, x_2) &= 0.3 \cdot (-1.0) + 0.7 \cdot 1.0 &= 0.40 \\
\pi x_1 g(w_2, x_1) + \pi x_2 g(w_2, x_2) &= 0.3 \cdot 0.0 + 0.7 \cdot 0.5 &= 0.35 \\
\pi x_1 g(w_3, x_1) + \pi x_2 g(w_3, x_2) &= 0.3 \cdot 0.4 + 0.7 \cdot 0.1 &= 0.19 \\
\pi x_1 g(w_4, x_1) + \pi x_2 g(w_4, x_2) &= 0.3 \cdot 0.8 + 0.7 \cdot (-0.9) &= -0.39 \\
\pi x_1 g(w_5, x_1) + \pi x_2 g(w_5, x_2) &= 0.3 \cdot 0.1 + 0.7 \cdot 0.2 &= 0.17 \\
\end{align*}
\]

Thus we find that $w_1$ is the best action and $V_g(\pi) = 0.4$ .
g-Vulnerability example
Interesting gain functions

• Identity gain function: \( g(w, x) = 1 \) if \( x = w \) and 0 otherwise.

• Gain functions induced by a metric \( d \): \( g(w, x) = d(w, x) \)

• Binary gain functions \( g(w, x) = 1 \) if \( x \in w \) and 0 otherwise.

• Penalty gain functions \( g(w, x) = 1 \) if \( x = w \), 0 if \( w = \bot \), -1 otherwise.

• Loss functions \( l(w, x) = -\log(w(x)) \) where \( w \) is a distribution
Gain function properties

• We can show that for every gain function $g$, the $g$-vulnerability $V_g$ is a convex function.

• Algebraic structure on gain functions translate to algebraic structure on the associated $g$-vulnerability.

\[
V_{g \times k}(X) = k \times V_g(X) \quad \text{for } k \geq 0
\]

\[
V_{g+r}(X) = V_g(X) + r
\]
Information leakage v2

Information leaked = initial uncertainty - remaining uncertainty
• We can abstract programs over finite data types \( c \) to stochastic matrices.

<table>
<thead>
<tr>
<th>C</th>
<th>( y_1 )</th>
<th>( y_2 )</th>
<th>( y_3 )</th>
<th>( y_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>0</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{4} )</td>
</tr>
<tr>
<td>( x_3 )</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{6} )</td>
<td>0</td>
</tr>
</tbody>
</table>

• where \( C_{xy} = \Pr[c(X) = y | X = x] \)
Bayes Theorem

\[ \Pr[x \mid y] = \frac{\Pr(y \mid x) \Pr(x)}{\Pr(y)} \]

- We can use Bayes’ theorem and a channel to compute the posterior given a prior.
Given $\pi = \left[ \frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3} \right]$ And

We can compute the joint channel:

<table>
<thead>
<tr>
<th>C</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
<th>$y_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{6}$</td>
<td>$\frac{1}{3}$</td>
<td>$0$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$0$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{2}{3}$</td>
<td>$0$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>$0$</td>
<td>$\frac{1}{2}$</td>
<td>$0$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$x_4$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{2}$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

And with this, renormalizing:

|   | $p_{X|y_1}$ | $p_{X|y_2}$ | $p_{X|y_3}$ |
|---|-------------|-------------|-------------|
| $x_1$ | $\frac{2}{3}$ | $\frac{2}{9}$ | $\frac{2}{9}$ |
| $x_2$ | $0$ | $\frac{4}{9}$ | $\frac{4}{9}$ |
| $x_3$ | $0$ | $0$ | $0$ |
| $x_4$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |
Hyper-distribution

Consider this set of posteriors

|       | $p_X|_{y_1}$ | $p_X|_{y_2}$ | $p_X|_{y_3}$ |
|-------|-------------|-------------|-------------|
| $x_1$ | 2/3         | 2/9         | 2/9         |
| $x_2$ | 0           | 4/9         | 4/9         |
| $x_3$ | 0           | 0           | 0           |
| $x_4$ | 1/3         | 1/3         | 1/3         |

We could think about it as a distribution over posteriors

<table>
<thead>
<tr>
<th>[π &gt; C]</th>
<th>1/4</th>
<th>3/4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2/3</td>
<td>2/9</td>
</tr>
<tr>
<td>$x_2$</td>
<td>0</td>
<td>4/9</td>
</tr>
<tr>
<td>$x_3$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$x_4$</td>
<td>1/3</td>
<td>1/3</td>
</tr>
</tbody>
</table>

This is what we call a hyper-distribution, read as π through C.
Hyper-distribution

We can write a hyper-distribution as:

\[
[\pi \triangleright C] = \sum_i a_i[\delta^i]
\]
Abstract channels

We can think about channels as essentially mapping priors to hyper-distributions.

The abstract channel $C$ of a channel $C$ is the mapping:

$$\pi \rightarrow [\pi \triangleright C]$$

We can think about this as the semantics of $C$

$$[[C]] = \lambda \pi . [\pi \triangleright C]$$

We can write a hyper-distribution as:

$$[\pi \triangleright C] = \sum a_i[\delta^i]$$
Properties

- C satisfies **non-interference** if its abstract channel is a lifting:

  \[
  [[C]] = \lambda \pi. \text{unit } \pi
  \]

- We can identify **canonical forms** for abstract channels and characterize abstract channels properties through properties about their functions.

- We can also take convex combinations of abstract channel and **compose** them in other abstract channels.
Posterior g-Vulnerability

\[
V_g[\pi \triangleright C] = \sum_i a_i V_g(\delta^i)
\]

Assuming

\[
[\pi \triangleright C] = \sum_i a_i \delta^i
\]

• Expected value of g-vulnerabilities.
Example 3.3 With $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{W} = \{w_1, w_2, w_3, w_4, w_5\}$, let gain function $g$ have the (rather arbitrarily chosen) values shown in the following matrix:

<table>
<thead>
<tr>
<th>$G$</th>
<th>$x_1$</th>
<th>$x_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_1$</td>
<td>$-1.0$</td>
<td>$1.0$</td>
</tr>
<tr>
<td>$w_2$</td>
<td>$0.0$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>$w_3$</td>
<td>$0.4$</td>
<td>$0.1$</td>
</tr>
<tr>
<td>$w_4$</td>
<td>$0.8$</td>
<td>$-0.9$</td>
</tr>
<tr>
<td>$w_5$</td>
<td>$0.1$</td>
<td>$0.2$</td>
</tr>
</tbody>
</table>

To compute the value of $V_g$ on (say) $\pi = (0.3, 0.7)$, we must compute the expected gain for each possible action $w$ in $\mathcal{W}$, given by the expression $\sum_{x \in \mathcal{X}} \pi_x g(w, x)$ for each one, to see which of them is best. The results are as follows.

\[
\begin{align*}
\pi_{x_1} g(w_1, x_1) + \pi_{x_2} g(w_1, x_2) &= 0.3 \cdot (-1.0) + 0.7 \cdot 1.0 = 0.40 \\
\pi_{x_1} g(w_2, x_1) + \pi_{x_2} g(w_2, x_2) &= 0.3 \cdot 0.0 + 0.7 \cdot 0.5 = 0.35 \\
\pi_{x_1} g(w_3, x_1) + \pi_{x_2} g(w_3, x_2) &= 0.3 \cdot 0.4 + 0.7 \cdot 0.1 = 0.19 \\
\pi_{x_1} g(w_4, x_1) + \pi_{x_2} g(w_4, x_2) &= 0.3 \cdot 0.8 + 0.7 \cdot (-0.9) = -0.39 \\
\pi_{x_1} g(w_5, x_1) + \pi_{x_2} g(w_5, x_2) &= 0.3 \cdot 0.1 + 0.7 \cdot 0.2 = 0.17
\end{align*}
\]

Thus we find that $w_1$ is the best action and $V_g(\pi) = 0.4$.
Posterior g-Vulnerability example

Let’s consider this channel

<table>
<thead>
<tr>
<th>C</th>
<th>y1</th>
<th>y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x₁</td>
<td>0.75</td>
<td>0.25</td>
</tr>
<tr>
<td>x₂</td>
<td>0.25</td>
<td>0.75</td>
</tr>
</tbody>
</table>

With prior (0.3,0.7) we get:

\[ a₁ = 0.4 \]
\[ a₂ = 0.6 \]
\[ δ₁ = (0.5625, 0.4375) \]
\[ δ₂ = (0.5625, 0.4375) \]

\[ V₉[π ∴ C] = 0.5575 \]
Example 5.5
To get a fuller understanding of posterior g-vulnerability for the gain function $g$ and channel matrix $C$ considered in Example 5.4, let us now graph $V_g[πC]$ as a function of a general prior $π=(x, 1-x)$, where $0 \leq x \leq 1$. Here we get the hyper-distribution $[πC]_{1+2x^4-2x^4x^1+2x^3-2x^3-2x}$.

Figure 5.1 compares the graph of $V_g(π)$ (which was seen already in Fig. 3.1) and $V_g[πC]$ with $[πC]$. As can be seen, $V_g[πC]$ is often (but not always) greater than $V_g(π)$. This is to be expected intuitively, since channel $C$ increases the adversary's knowledge about $X$, enabling a better choice of actions. But the particular line segments in the graph of $V_g[πC]$ might need some clarification.

Recall that in the prior situation, the adversary's choice of which action to take is guided only by $π$ itself — she does not consider the channel. But in the posterior situation, her choice can be guided by both $π$ and the output of channel $C$, so that she can choose one action if the output is $y_1$ and another if the output is $y_2$. If we let $w_iw_j$ denote the strategy of choosing action $w_i$ on output $y_1$ and $w_j$ on output $y_2$, then we see that she has $|W|^2=25$ possible strategies. Note that those of the form $w_iw_i$ are "degenerate" strategies, since they choose action $w_i$ regardless of the output: we therefore write them simply as $w_i$.

Figure 5.2 plots the expected gain for seven of these strategies: they are $w_1$, $w_2w_1$, $w_3w_1$, $w_4w_1$, $w_4w_3$, $w_4w_2$, and $w_4$. (Note that the graphs of the "degenerate" strategies $w_1$ and $w_4$ were seen already in Fig. 3.2.) The expected gains for the remaining $25-7=18$ strategies are not shown, because each of them is dominated by the other seven, in the sense that on any $π$ at least one of the seven gives an expected gain that is at least as large. (Recall that in Fig. 3.2 the strategy $w_5$ is dominated in the same way.)
Many other topics

- How to apply it in practical analyses
- How to use program logics to reason about this framework
- Geometric properties
- Stochastic properties
- …