Does additional information \textit{always} reduce anonymity?

\textbf{Claudia Diaz, Carmela Troncoso and George Danezis}

K.U.Leuven ESAT/COSIC
Anonymity

- “State of being not identifiable within a set of subjects, the anonymity set” [PH00]
  - “Set of all possible subjects who might cause an action”
  - Communication systems: sender / recipient / relationship anonymity
- The anonymity adversary typically obtains a probability distribution linking subjects and objects/actions
- The entropy of the probability distribution obtained by the attacker gives a measure of his uncertainty on the identity of the subject (i.e., that subject’s anonymity [SD02,DCSP02])

\[
H(X) = -\sum_{i=1}^{N} p_i \log_2 p_i; \quad p_i = \Pr(X = x_i)
\]

- The uncertainty (entropy) increases with the size of the anonymity set (N) and with the uniformity of the probability distribution
- Can be applied to mix traffic traces, or to user profiles (examples later)
Anonymity when only Alice’s profile is known

- $Y$: Random variable describing Alice’s sending profile
- Adversary: only observes that Alice sends a message
- Uncertainty of attacker on recipient is given by the entropy of $Y$
- $H(Y) = 1.78$ bits
Anonymity when only mix communication trace is observed

- Goal of attacker: find the recipient of Alice’s message (no profile is known)
- Pool mix with threshold $T=2$ and pool $P=1$
- $X$: Random variable describing communication trace observations: $P_X(R_i)=2^{-i}$
- $H(X) = 2$ bits
Combining several sources of information - The claim

- “Let \( X \) and \( Y \) be probability distributions of the application layer and the network layer. One can measure anonymity \( H(X) \) and \( H(Y) \). [T]he attacker could build a combined model by introducing the circumstances of communication as attributes in the application layer model. Due to the fact that new information can only reduce the cardinality of the set of suspects the resulting probability distribution gets more unequal, i.e., entropy decreases.” [CS06]

- Intuitively consistent with Shannon’s result on conditional entropy: \( H(Y|X) \leq H(Y) \)
Assumptions
- Only Alice’s profile is known
- Alice sends only one message
- Alice’s friends appear (only once each) as the first five recipients after she sent the message

Five possibilities for Alice’s message:

1. Message was for Bob \((P_Y(B)=0.04)\) and it was immediately sent by the mix \((P_X(1)=2^{-1})\)
2. Message was for Charlie \((P_Y(C)=0.06)\) and it spent one round in the mix before being sent \((P_X(2)=2^{-2})\)
3. Message was for Dave \((P_Y(D)=0.1)\) and it spent two rounds in the mix before being sent \((P_X(3)=2^{-3})\)
4. Message was for Els \((P_Y(E)=0.3)\) and it spent three rounds in the mix before being sent \((P_X(4)=2^{-4})\)
5. Message was for Fred \((P_Y(F)=0.5)\) and it spent four rounds in the mix before being sent \((P_X(5)=2^{-5})\)
Combining several sources of information- The (counter)example

- We define a random variable $Z$ that combines both Alice’s profile ($Y$) and the observed communication trace ($X$)
  - $Z$ takes values $\{z_i\} = \{y_i\} = \{B,C,D,E,F\}$ with $P_Z(z_i)$

$$P_Z(z_i) = \frac{P_Y(y_i)P_X(i)}{\sum_{j=1}^{5} P_Y(y_j)P_X(j)}$$

$P_Z(B)=0.25, P_Z(C)=0.18, P_Z(D)=0.15, P_Z(E)=0.23, P_Z(F)=0.19$

$H(Z) = 2.3$ bits $> H(X)= 2$ bits $; > H(Y) = 1.78$ bits !!
Relationship between attacker uncertainty and conditional entropy

Given a traffic trace \( x_j \), the uncertainty of the attacker on Alice’s recipient choice \( y_i \) is given by the entropy \( H_j(Z) \):

\[
H_j(Z) = - \sum_i \text{Pr}(y_i \mid x_j) \log_2 \text{Pr}(y_i \mid x_j)
\]

The conditional entropy \( H(Y \mid X) \) is defined as:

\[
H(Y \mid X) = - \sum_{i,j} \text{Pr}(y_i, x_j) \log_2 \text{Pr}(y_i \mid x_j)
\]

\[
H(Y \mid X) = - \sum_j \text{Pr}(x_j) \sum_i \text{Pr}(y_i \mid x_j) \log_2 \text{Pr}(y_i \mid x_j)
\]

Therefore:

\[
H(Y \mid X) = \sum_j \text{Pr}(x_j) H_j(Z)
\]

May not be possible to compute!
Conclusions

- Computing anonymity when several sources of information are available is not yet well understood
  - We have shown how to do it in a toy example
  - It may be complex to generalize
- We have shown that the attacker uncertainty **might** increase if the information from different sources is “contradictory”
- The uncertainty of the attacker given a traffic observation is **not** given by the conditional entropy
- More research is needed to understand the relationship between attacker uncertainty (anonymity) and the entropy of the random variables in the system (e.g., profiles or mix mapping)
Thank you!