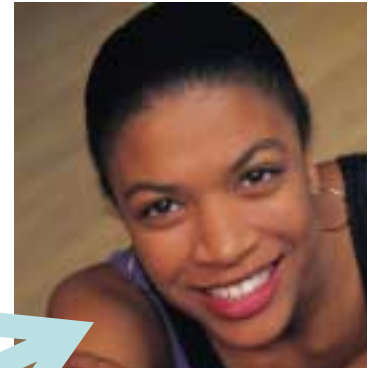
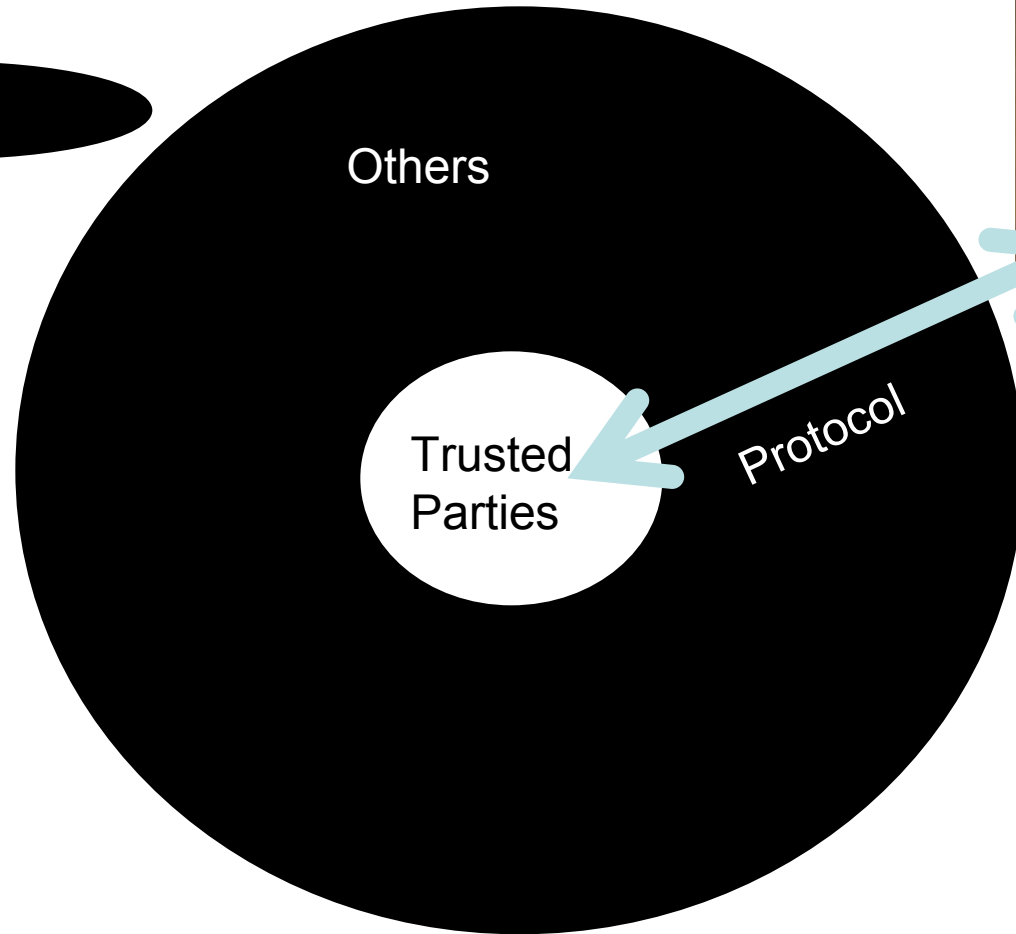
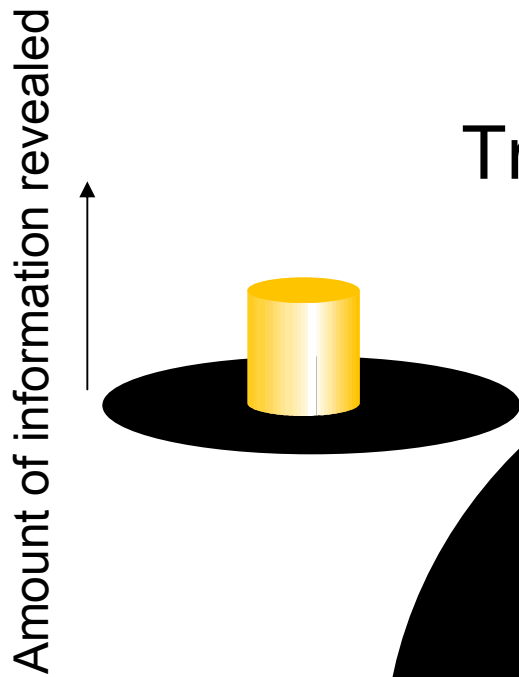
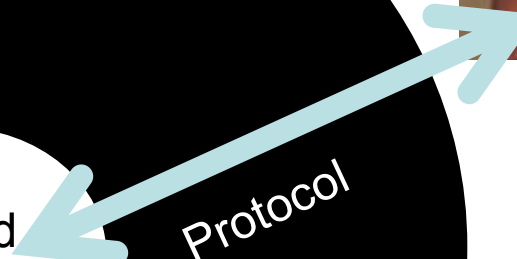


Traditional security model



Alice



Traditional theory of security

Desirable protocols do not leak any information to non-trusted parties

Information-theoretically perfect secrecy:

a priori and *a posteriori* pdfs identical

- no information leakage to any adversary

Computationally perfect secrecy:

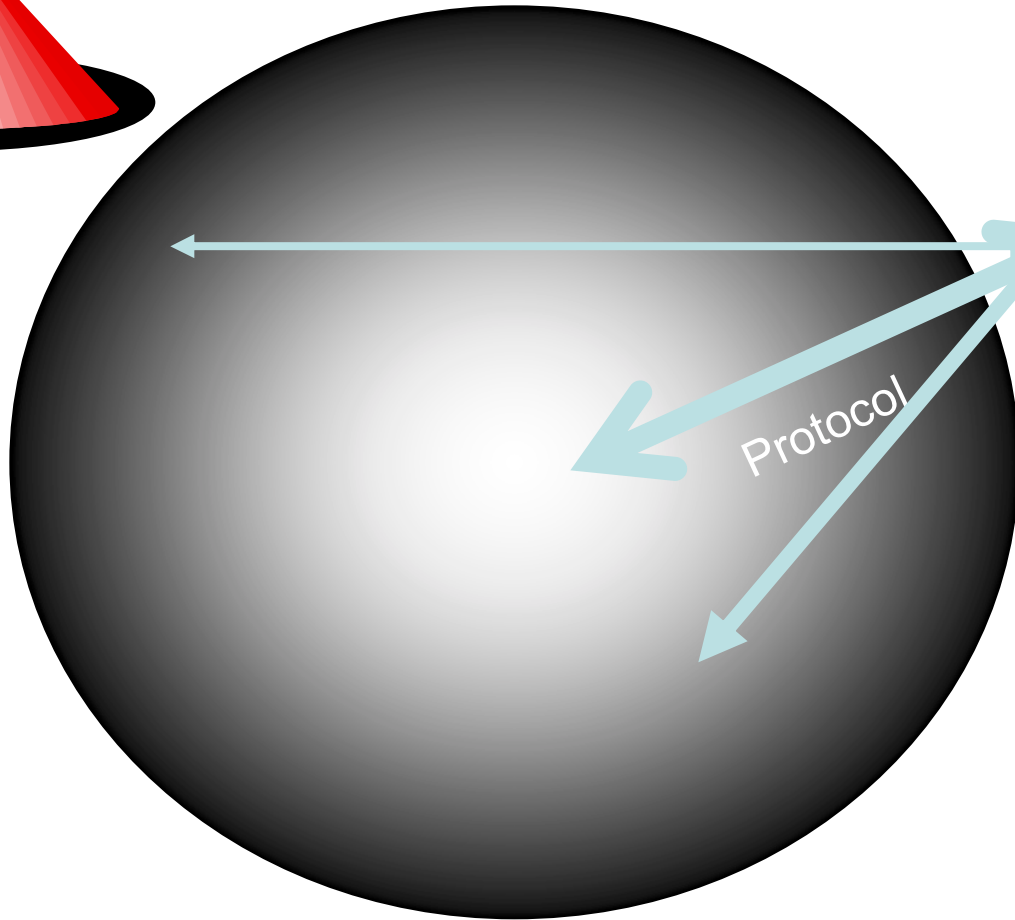
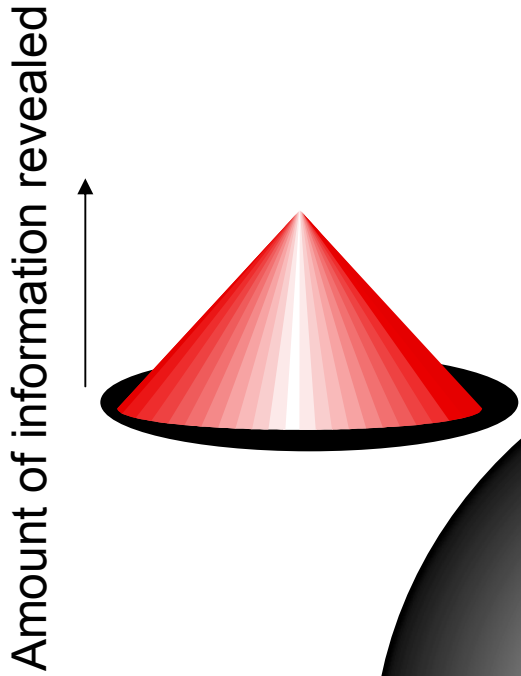
a probabilistic polynomial-time algorithm cannot distinguish between prior and posterior

- no information leakage to realistic adversary

Problems not addressed by perfectly secret protocols

- Need to leak statistics in:
 - Markets
 - Statistical databases
 - Collaborative filtering
- Need another model for communities
- There is an existing market for personal information
 - Safeway cards for 10% discount
 - Extra for unlisted phone numbers
 - Need an understanding of “amount of privacy” to study the value of privacy in this market

The privacy world

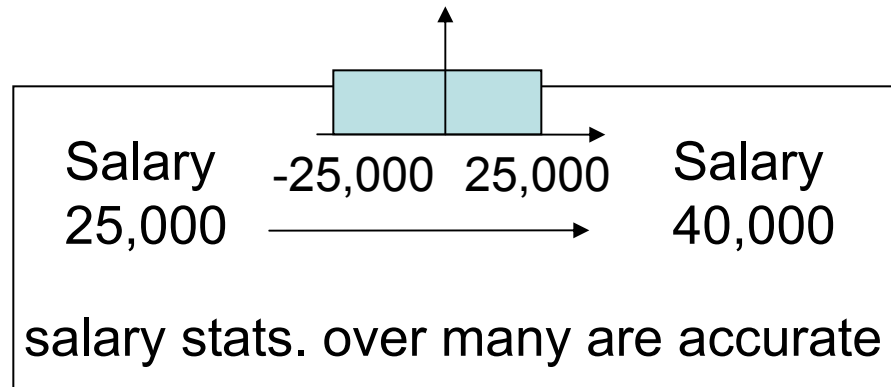


Alice

An intentionally non-perfect protocol

- **Randomization** (probabilistic perturbation of data)
 - provides statistics to data collector, privacy to individual
- **Current Uses:**
 - Public health surveys (20+ years)
 - Statistical database security (20+ years)
 - IBM application for personal privacy protection on data collection websites (6 months)
- **Potential Use, with Alice's participation**
 - Interaction with parties neither trusted nor untrusted (e.g. virtual communities)
 - Collaborative filtering with privacy
 - Negotiations

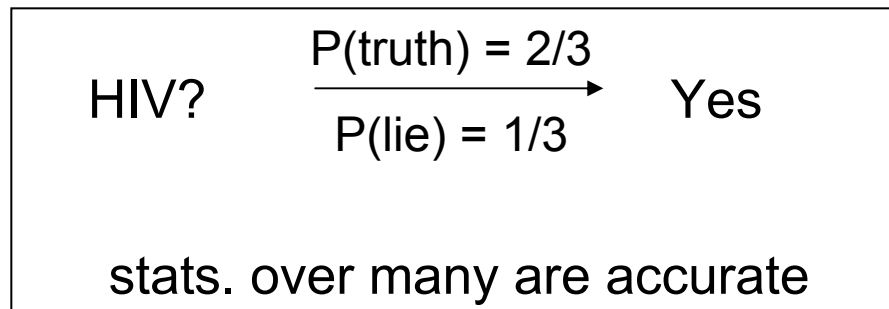
Randomization: continuous-valued



The output now decreases possible salary range:

15-65K

Randomization: binary-valued



After the protocol, the possibilities are skewed
the answer is most likely to be correct

The statistical database security problem

- Data collector asks for:

$$f_i(x_1, x_2, x_3, \dots) = A_i$$

- Can simultaneously solve above
- (perfect zk protocols do not leak additional information about x_i , but A_i are revealed; thus not a traditional cryptographic problem)
- If x_i perturbed each time, the equations are inconsistent
$$f_i(x_1 + \Delta_{1i}, x_2 + \Delta_{2i}, x_3 + \Delta_{3i}, \dots) = A_i + \Delta_i$$
- Security and attack characterization open problem for 20+ years; though many attempts (Denning, Adams, Duncan, ... Landers).

Variable Privacy

Definition 1: “variable privacy” is the use of non-perfect protocols with Alice’s participation in choice of protocol parameters

Natural consequence of the definition of privacy in a world that includes non-perfect protocols

Need a framework for “variable privacy”

- What is a measure of the privacy provided by randomization?
- Can it be related to the “security” of randomization?

Our privacy model

1. Alice and Bob determine a level of information leakage, $P(Y|X)$
2. Bob requests a data point X from Alice, she reveals Y according to $P(Y|X)$
3. Bob provides something to Alice in return
 - Dishonest Bob can use the information leakage to find out more than Alice intended
 - The cost to Dishonest Bob is a measure of protocol privacy

Would provide a framework for “variable privacy”, and an understanding of the security of randomization, an open problem for 20 years in statistical databases

Protocol as channel

Protocol Input: The truth value of “X has HIV”

Output: Perturbed value of the bit.

Probabilities: of truth: $2/3$, of lie: $1/3$

Communication channel with probability of error $1/3$

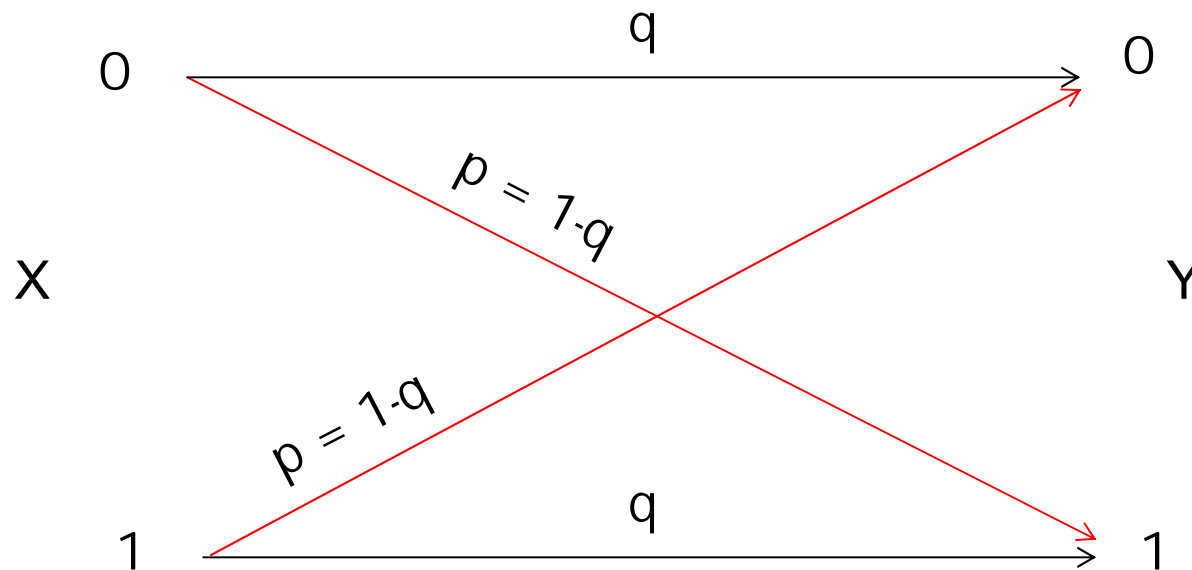
Formally: Protocols as communication channels

$$\varphi: X \rightarrow Y$$

$$\varphi(X) = Y$$

- X is the set of all possible values of user personal information, plaintext
- Y is the set of all possible values of observable information from a single instance of the protocol or the attack, ciphertext
- Unlike channels in communication theory, the purpose of φ is to limit communication of X .
- $\Phi = (X, P(Y|X), Y)$

Binary Symmetric Randomization Protocol



$$\Phi = (\{0, 1\}, \{0, 1\}, P(Y|X))$$

$$P(Y | X) = q, Y = X; = 1 - q, Y \neq X$$

Typical query sequence for attack

message bit 1: female?

message bit 2: over 40?

plaintext bit 1: Losing Calcium?

plaintext bit 2: Graying?

plaintext bit 3: Balding?

plaintext bit 4: Gaining weight?

Rate defined as

$\log(\text{no of possible messages})/\text{plaintext length}$

Rate (efficiency) of above attack = $\log(4)/4 = 0.5$

PRP

Definition 2:

A *plaintext* is a string of bits each a function of bits in the database: $p = (f_1(a)_{a \in A_1 \subset D}, f_2(a)_{a \in A_2 \subset D}, \dots, f_n(a)_{a \in A_n \subset D})$

Definition 3:

A (M, n) *probabilistically-related plaintext* is a plaintext of length n having non-zero mutual information with M possible equal-length messages. Its rate is $\log_2 M/n$

$p = (p_1, p_2, \dots, p_n)$ a (M, n) PRP

$\Leftrightarrow \exists m = (m_1, m_2, \dots, m_k)$ such that $H(m|p) \neq H(m)$

(uncertainty in m decreases on knowing p)

Attacks on randomization - repeated plaintext

- An attack: asking the same question many times
- Can be thwarted by
 - never answering the same question twice, or
 - always answering it the same.
- Plaintext repetition:
 - corresponds to an error-correcting code word

a a a a a a a

Error

Known that:

- tracker can reduce estimation error indefinitely
- by increasing the number of repeated plaintext bits indefinitely;

$$n \rightarrow \infty \Rightarrow \varepsilon^n \rightarrow 0$$

- and that this is the best he can do with repeated plaintext

$$\varepsilon^n \rightarrow 0 \Rightarrow \text{cost per message bit} = n/1 \rightarrow \infty$$

Is the following an attack?

- plaintext bit 1: “location = North”;
- plaintext bit 2: “virus X test = positive”;
- plaintext bit 3: “gender = male” AND “condition A = present”

If

(location = North) \oplus (virus X test = positive)

\Leftrightarrow (gender = male) AND (condition A = present)

Then: $A_3 = A_1 \oplus A_2$; **check-sum bit**

Not easily recognized as attack

Attacks

Definition 5: An (M, n) *attack* for binary protocol Φ is:

- an (M, n) plaintext
- and an estimation map $\Psi: \Sigma^n \rightarrow \Sigma^k$ for
 - estimating the message $m(\gamma)$ from
 - the ciphertext (randomized bits) $\Phi(p(\gamma))$.

Its rate is $\log_2 M/n$.

Code

Definition 6: An (M, n) code for set of messages M and binary channel Φ is:

- A coding function f from M to code words of size n ,
$$f: M \rightarrow \Sigma^n$$
- and a decoding function $g: \Sigma^n \rightarrow M$ for
 - estimating the message m from
 - the randomized bits $\Phi(f(m))$.

Its rate is $\log_2 M/n$.

Efficiency of attacks: repeated plaintext attack

- Definition 7: A *small error attack* is one in which $\epsilon^n \rightarrow 0$ as $n \rightarrow \infty$
- Plaintext repetition:
 - corresponds to an error-correcting code word
 $a a a a a a a$
 - probability of error is monotonic decreasing with n for n -symbol code words
 - rate of code = $1/n$
 - sacrifice rate for accuracy; rate of small error attack $\rightarrow 0$
- Are DRPs more efficient?

Reliable Attacks

Definition 8: A *reliable attack* of rate R is a small error attack of fixed rate R

Definition 9: A small error attack of asymptotic rate R_∞ is a small error attack with rate $\rightarrow R_\infty$

Do small error attacks of non-zero asymptotic rates exist?

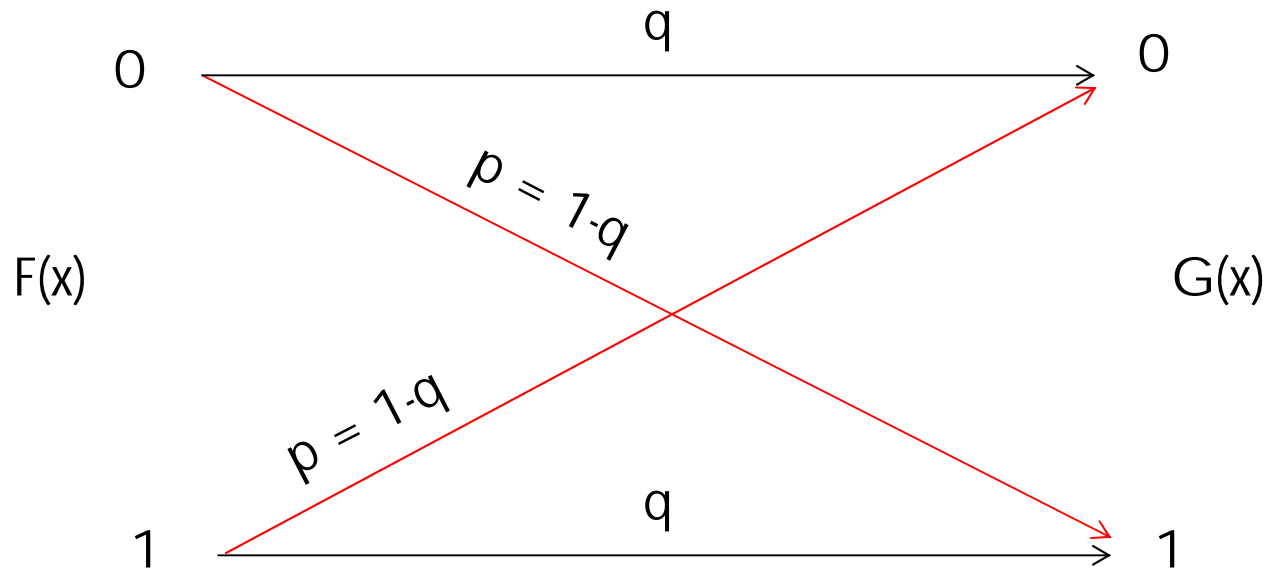
Do reliable attacks exist?

Proposed measure of privacy of randomization

The privacy of randomization is the **tight lower bound on the asymptotic length of plaintext per message, per bit of message entropy**, for a stationary message sequence and a small error attack

Corollary: **Privacy (Φ) = $1/C(\Phi)$**

Application: Binary Symmetric Protocol



$$C = 1 + p \log_2 p + (1-p) \log_2 (1-p)$$

$$= 0 \text{ if } p = 0.5;$$

$$\approx 4\beta^2 / \ln 2 \text{ if } p = 0.5 \pm \beta; \beta \ll 0$$

Application to binary randomization

Binary symmetric protocols for small bias β have channel capacity $O(\beta^2)$.

Corollary: Plaintext length required, per bit of message entropy, for a small error attack in the binary randomization protocol with small bias β is $O(1/\beta^2)$ and *independent of the error*

The privacy of binary randomization with small bias β is $O(1/\beta^2)$

We have shown that

Dishonest Bob can do better by increasing the number of points combined in a single query

i.e. there exist attacks for which

$$\begin{aligned} n \rightarrow \infty &\Rightarrow \varepsilon_n \rightarrow 0 \\ \varepsilon_n \rightarrow 0 &\not\Rightarrow n/k \rightarrow \infty \end{aligned}$$

There is a tight lower bound on the limit of n/k such that

$$n \rightarrow \infty \Rightarrow \varepsilon_n \rightarrow 0$$

i.e., $(n \rightarrow \infty \Rightarrow \varepsilon_n \rightarrow 0) \Rightarrow \lim n/k > 1/C$

The variable privacy big picture

- Alice can use randomization as a privacy protocol
 - designing the channel capacity
 - based on knowledge that error correcting codes are attacks
- Dishonest Bob cannot approach rates higher than channel capacity
- Randomization is a *game* between Alice and Bob
- In this world, *maximum privacy exists when Alice gets maximum benefit for a piece of revealed information*