CERIAS
The Center for Education and Research in Information Assurance and Security

PURDUE UNIVERSITY

Differential Identifiability§

Jaewoo Lee and Chris Clifton
Department of Computer Science, Purdue University

Legal Privacy Requirement: Individual Identifiability

HIPAA Safe Harbor Rule

- removal of 18 **identifiers** (e.g., name, ssn, etc.)
- dates → year (merge all ages ≥ **90**)
- geographic units of at least **20,000 people**

HIPAA safe harbor as probabilistic bound

- bound on the probability of identifying individuals
- 1.7% of US population is male & age≥85 (2010 US Census)
- knowing the age, gender and geographic unit can limit to 68 people

$$\Pr[I(i) \in DB \mid Release] \leq \frac{1}{68} \cong 1.5\%$$

Privacy goal : $\Pr[$  $\in DB \mid R] \leq \rho$

Differential Privacy: The Hot New Definition

- For all datasets D_1 and D_2 differing at most one element,

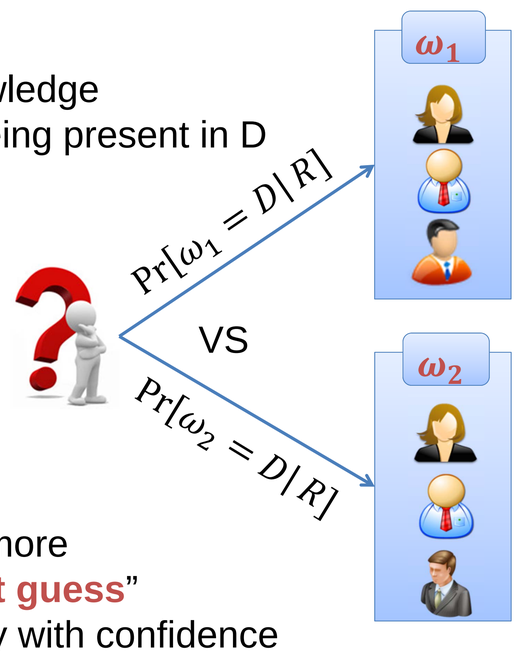
$$\frac{Pr[\mathcal{M}_f(D_1) = R]}{Pr[\mathcal{M}_f(D_2) = R]} \leq e^{\epsilon}$$

- hides individuals in the dataset by making two distributions **indistinguishable** within a factor of e^{ϵ}
- *unclear **how to set** ϵ to meet the privacy goal*

Probabilistic Attack¶

- given $\mathbf{R} = \mathcal{M}_{mean}(DB)$
- adversary's background knowledge
  - people already known as being present in D
    D'={  ,  }
- two possibilities

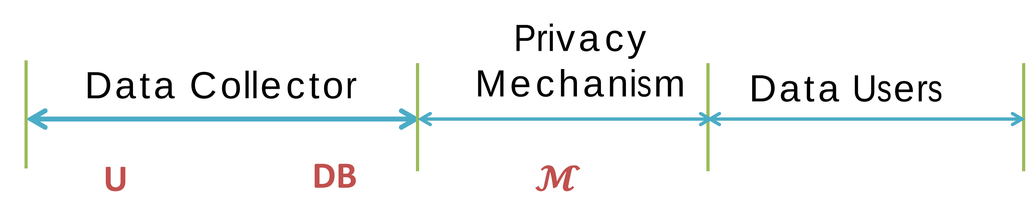
Possible worlds	
ω_1	ω_2
$f(\omega_1) = 2$	$f(\omega_2) = 5$

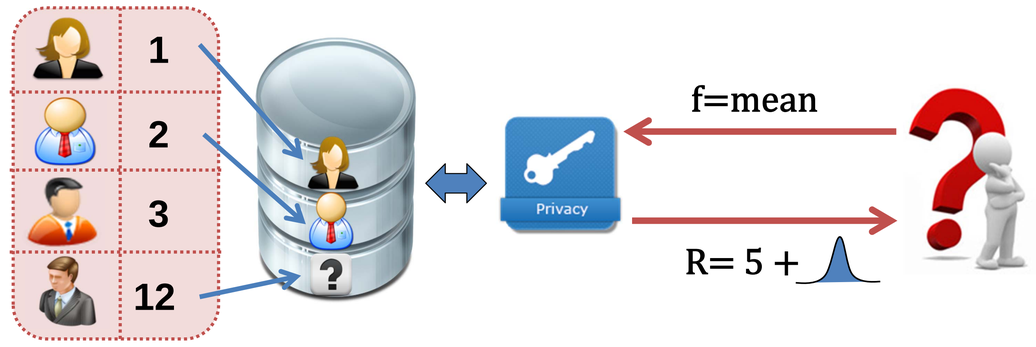


- choose one from which R is more likely to be produced as "**best guess**"
- if adversary guesses correctly with confidence higher than ρ, privacy is breached
- For differential privacy, adversary's confidence for two different scenarios is different, even with same ϵ

¶ : appears in ISC 2011
§ : submitted to KDD '12 and under review

Interactive Privacy Mechanism





Differential Identifiability§

- limits the probability of identifying an individual in the database to $\boldsymbol{\rho}$
- formally defined as

$$\forall D' = D - \{i\}, \quad \forall i \in U - D'$$
$$\Pr[I(i) \in I_D \mid \mathcal{M}_f(D) = R, D'] \leq \rho$$

Adversary Model

- What is known to the adversary
  - **U** : people in the universe & their data
  - ***D'*** : |D|-1 rows
  - $\mathcal{M}$: (randomized) privacy mechanism (+ noise distribution)
- What the adversary does
  - generates a set of **possible worlds** Ψ
    - *Assumed to be equally likely*
  - $\boldsymbol{\Psi} = \{\boldsymbol{D'} \cup \{\boldsymbol{i}\} | \boldsymbol{i} \in \boldsymbol{U}\}$
  - For each $\boldsymbol{\omega} \in \boldsymbol{\Psi}$, computes $\boldsymbol{Pr}[\boldsymbol{\omega} = \boldsymbol{D} \mid \boldsymbol{R} = \boldsymbol{\mathcal{M}_f}(\boldsymbol{D})]$

Sensitive Range

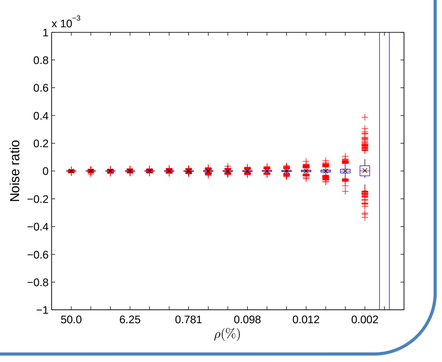
- the largest contribution an individual can make to the output of function

$$S(f) = \max_{\omega, \omega' \in \Psi} |f(\omega) - f(\omega')|$$

Calibrating noise

- $\Gamma(i) = \Pr[I(i) \in I_D | \ \mathcal{M}_f(D) = R, D']$
- To enforce $\max_i \Gamma(i) \leq \rho$

$$\lambda \geq \frac{S(f)}{ln\frac{(m-1)\rho}{1-\rho}}$$



* Partial support for this work was provided by MURI award FA9550-08-1-0265 from the Air Force Office of Scientific Research

CERIAS

Discovery Park e-Enterprise Center