Differential Privacy with Bounded Priors:

Reconciling Utility and Privacy in Genome-Wide Association Studies

Florian Tramèr, Zhicong Huang, Erman Ayday, Jean-Pierre Hubaux

ACM CCS 2015
Denver, Colorado, USA
October 15, 2015

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

İhsan Doğramacı Bilkent Üniversitesi

Outline

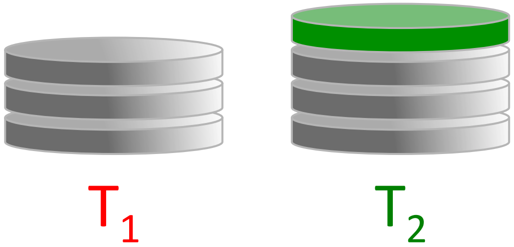
- Data Privacy and Membership Disclosure
  - Differential Privacy
  - Positive Membership Privacy
  - Prior-Belief Families and Equivalence between DP and PMP
- Bounded Priors
  - Modeling Adversaries with Limited Background Knowledge
  - Example: Inference Attacks for Genome-Wide Association Studies
- Evaluation
  - Perturbation Mechanisms for GWAS
  - Trading Privacy, Medical Utility and Cost

Differential Privacy[1,2]

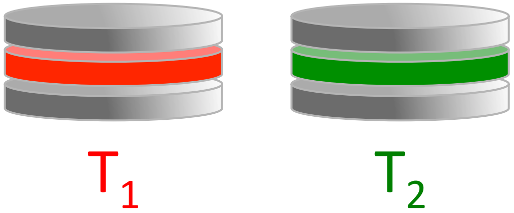
- Belonging to a dataset ≈ Not belonging to it
- A mechanism $\mathcal{A}$ provides ε-DP iff for any datasets T_1 and T_2 *differing in a single element*, and any $S \subseteq$ range($\mathcal{A}$), we have:

$$\Pr[\mathcal{A}(T_1) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{A}(T_2) \in S]$$

Unbounded DP



Bounded DP

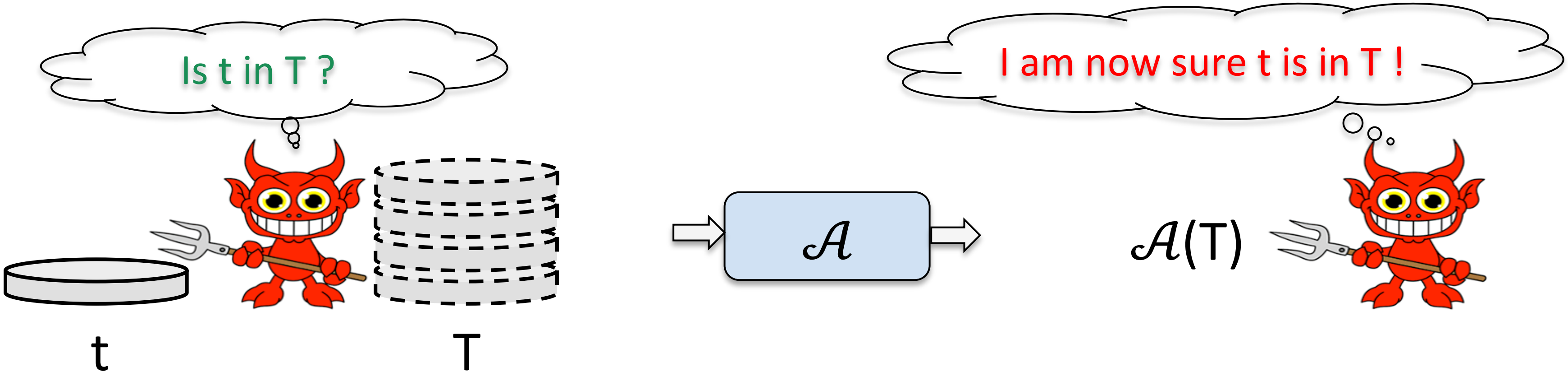


[1] Dwork. "Differential privacy". Automata, languages and programming. 2006

[2] Dwork et al. "Calibrating Noise to Sensitivity in Private Data Analysis". TCC'06. 2006

Positive Membership Privacy[1]

- Data Privacy: protection against **membership disclosure**
  - Adversary should not learn whether an entity from a universe $\mathcal{U} = \{t_1, t_2, \ldots\}$ belongs to the dataset T



- Privacy: posterior belief ≈ prior belief for all entities
- **Impossible in general! (no free lunch)**

[1] Li et al. “Membership privacy: a unifying framework for privacy definitions”. CCS ’13. 2013

Prior Belief Families[1]

- Adversary's prior belief: Distribution $\mathcal{D}$ over $2^{\mathcal{U}}$
- Range of adversaries: Distribution family $\mathbb{D}$
- A mechanism $\mathcal{A}$ satisfies (ε, $\mathbb{D}$)-PMP iff for any S ⊆ range($\mathcal{A}$), any prior distribution $\mathcal{D} \in \mathbb{D}$, and any entity t ∈ $\mathcal{U}$, we have

$$\Pr[t \in T \mid \mathcal{A}(T) \in S] \leq e^{\epsilon} \cdot \Pr[t \in T]$$

$$\Pr[t \notin T \mid \mathcal{A}(T) \in S] \geq e^{-\epsilon} \cdot \Pr[t \notin T]$$

[1] Li et al. "Membership privacy: a unifying framework for privacy definitions". CCS '13. 2013

PMP ⇔ DP[1]

- Mutually Independent Distributions:
  - $\mathcal{D} \in \mathbb{D}_I$: each entity t is in T, **independently** with probability p_t
  - $\mathcal{D} \in \mathbb{D}_B$: Same as above, conditioned on $|T|=k$, for some k

    ⇒ Adversary also **knows the size** of the dataset T

- Theorem:

$$\epsilon\text{ - unbounded - DP} \Leftrightarrow (\epsilon, \mathbb{D}_I)\text{ - PMP}$$

$$\epsilon\text{ - bounded - DP} \Leftrightarrow (\epsilon, \mathbb{D}_B)\text{ - PMP}$$

- We focus on bounded DP (results hold for unbounded case)

[1] Li et al. "Membership privacy: a unifying framework for privacy definitions". CCS '13. 2013

Outline

- Data Privacy and Membership Disclosure
  - Differential Privacy
  - Positive Membership Privacy
  - Prior Belief Families and Equivalence between DP and PMP
- Bounded Priors
  - Modeling Adversaries with Limited Background Knowledge
  - Example: Inference Attacks for Genome-Wide Association Studies
- Evaluation
  - Perturbation Mechanisms for GWAS
  - Trading Privacy, Medical Utility and Cost

Bounded Priors

- Observation: $\mathbb{D}_B$ includes adversarial priors with **arbitrarily high certainty** about all entities:

$$\Pr[t \in T] \in \{0, 1\}, \ \forall t \neq t' \in \mathcal{U}$$
$$\Pr[t' \in T] \in (0, 1)$$

- Do we care about such strong adversaries?
  - All entities except t’ have **no privacy a priori** (w.r.t membership in T)
  - The membership status of t’ can also be **known with high certainty**
    - Membership is **extremely rare / extremely likely**
    - Or adversary has **strong background knowledge**
  - How do we model an adversary with **limited a priori knowledge**?

Bounded Priors

- We consider adversaries with the following priors:
  - Entities are independent (size of dataset possibly known)
  - $\Pr[t \in T] \in \{0,1\}$ for some entities
    - Adversary might **know** membership status of some entities
  - $a \leq \Pr[t \in T] \leq b$ for other entities, where $a>0$ and $b<1$
    - For an "unknown" entity, **membership status is uncertain a priori**
  - Denoted $\mathbb{D}_B^{[a,b]}$ (or $\mathbb{D}_B^{a}$ if $a=b$)
- Questions:
  - Is the model **relevant in practice** ?
  - What **utility** can we gain by considering a **relaxed adversarial setting** ?

Bounded Priors In Practice: Example

- Genome-Wide Association Studies:
  - **Case-Control study** (typically $N_{case} = N_{ctrl}$)
  - Membership in case group ⟺ patient has some disease
  - Find out which genetic variations (SNPs) are **associated with disease**
    - Ex: χ^2 test for each SNP (low p-value ⟺ conclude SNP is probably associated)
- Re-identification attacks[1,2]:
  - Collect **published aggregate statistics** for the case/control groups
  - Use a **victim's DNA sample** & **statistical testing** to distinguish between:
    - $\mathbf{H_0}$: victim is not in case group
    - $\mathbf{H_1}$: victim is in case group (victim has the disease)
  - **Assumptions** (some implicit):
    - N_{case} & N_{ctrl} are known (usually published)
    - Entities are **independent**
    - Prior: $\Pr[t \in T] = N_{case} / (N_{case} + N_{ctrl}) \Longrightarrow$ **typically ½ in attack evaluations**
  - **Attacks taken seriously!** (some statistics removed from open databases)[3]

[1] Homer et al. "Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays". PLoS genetics. 2008
[2] Wang et al. "Learning Your Identity and Disease from Research Papers: Information Leaks in Genome Wide Association Study". CCS '09. 2009
[3] Zerhouni and Nabel. "Protecting aggregate genomic data". Science. 2008

Achieving PMP for Bounded Priors

- Recall:

$$\epsilon\text{-DP} \Leftrightarrow (\epsilon, \mathbb{D}_B)\text{-PMP}$$

- (ε, $\mathbb{D}_B$)-PMP:

$$\Pr[t \in T \mid \mathcal{A}(T) \in S] \leq e^{\epsilon} \cdot \Pr[t \in T]$$
$$\Pr[t \notin T \mid \mathcal{A}(T) \in S] \geq e^{-\epsilon} \cdot \Pr[t \notin T]$$

- These inequalities are **tight** iff Pr[t ∈ T] ∈ {0,1}
  - For bounded priors (Pr[t ∈ T] ∈ [a,b]) we have:

$$\epsilon\text{-DP} \Rightarrow (\epsilon', {\mathbb{D}_B}^{[a,b]})\text{-PMP}, \text{ where } \epsilon' < \epsilon$$

  - Perturbation required to achieve ε-PMP depends on [a,b]
  - Minimal perturbation required when a = b = ½

Privacy – Utility Tradeoff

- If we consider **bounded adversaries** with prior in $\mathbb{D}_B^{[a,b]}$ instead of **adversaries** with prior in $\mathbb{D}_B$:
  - **Are we still protecting against relevant threats?** ✓

    ⇒ Attacks proposed on GWAS
  - **Can we gain in utility?** ✓

    ⇒ Less data perturbation required

    ⇒ Actual gain to be evaluated

Outline

- Data Privacy and Membership Disclosure
  - Differential Privacy
  - Positive Membership Privacy
  - Prior Belief Families and Equivalence between DP and PMP
- Bounded Priors
  - Modeling Adversaries with Limited Background Knowledge
  - Example: Inference Attacks for Genome-Wide Association Studies
- Evaluation
  - Perturbation Mechanisms for GWAS
  - Trading Privacy, Medical Utility and Cost

Evaluation

- Statistical Privacy for GWAS:
  - Laplace / Exponential mechanisms based on χ^2 – scores[1,2]
  - Exponential mechanism with specialized distance metric[3]
- Tradeoffs:
  1. **Privacy** Mitigate inference attacks
  2. **Output Utility** Associated SNPs should be output
  3. **Dataset Size** Privacy and **Cost** depend on number of patients
- What we want to achieve:
  1. **ε-PMP** for:
     - The adversarial setting of *Homer et al.*, *Wang et al.*
     - Compared to an unbounded adversary
  2. **High probability** of outputting the correct SNPs
  3. **Also for small studies** (N ≃ 2000)[4]

[1] Uhler, Slavkovic, and Fienberg. “Privacy-Preserving Data Sharing for Genome-Wide Association Studies”. Journal of Privacy and Confidentiality. 2013

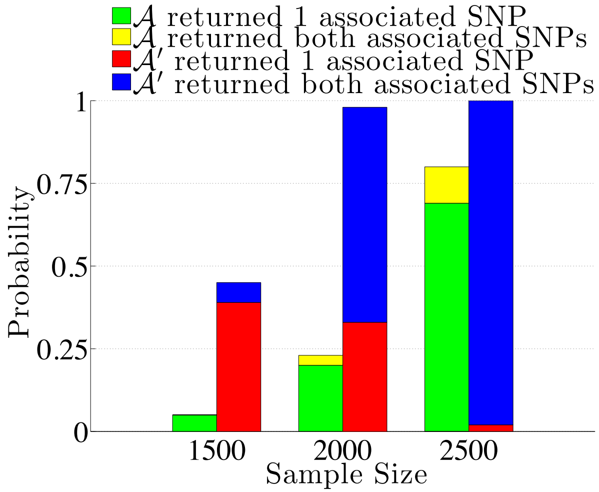
[2] Yu et al. “Scalable privacy-preserving data sharing methodology for genome-wide association studies”. Journal of biomedical informatics. 2014

[3] Johnson and Shmatikov. “Privacy-preserving Data Exploration in Genome-wide Association Studies”. KDD ’13. 2013

[4] Spencer et al. “Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip”. PLoS genetics. 2009

Evaluation

- GWAS simulation with 8532 SNPs, 2 associated SNPs
  - Variable sample size N ($N_{case} = N_{ctrl}$)
  - Satisfy PMP for $\varepsilon = \ln(1.5)$
  - Mechanism $\mathcal{A}$ protects against adversary with unbounded prior $\mathbb{D}_B$
    - $\mathcal{A}$ must satisfy ε-DP
  - Mechanism $\mathcal{A}'$ protects against adversary with bounded prior $\mathbb{D}_B^{½}$
    - It is sufficient for $\mathcal{A}'$ to satisfy ε'-DP for $\varepsilon' = \ln(2)$
  - Exponential mechanism from[1] :



[1] Johnson and Shmatikov. "Privacy-preserving Data Exploration in Genome-wide Association Studies". KDD '13. 2013

Conclusion

- Membership privacy is **easier to guarantee** for adversaries with **bounded priors**
  - **Less perturbation ⟹ Higher utility**
  - For GWAS: **Better tradeoff** between **dataset size** and **utility of output**
- We can **tailor privacy mechanisms** to **specific attacks/threats**
  - Can we make reasonable assumptions on the adversary's prior beliefs?
  - For GWAS: known attacks implicitly rely on such assumptions
  - Compute **appropriate level of noise** to guarantee **bounds on adversary's posterior beliefs**
- Future Work:
  - Can we build **stronger** inference attacks on GWAS?
    - ⟹ Infer "rare" membership (disease status is typically rare in a population)
    - ⟹ Known attacks are less successful when prior Pr[$t \in T$] is very small[1]
  - Direct comparison: **attack success rate** vs. **data perturbation (utility)**[2]
    - ⟹ Promote a "practice-oriented" study of statistical privacy

[1] Sankararaman et al. "Genomic privacy and limits of individual detection in a pool." Nature genetics. 2009

[2] Fredrikson et al. "Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing." Proceedings of USENIX Security. 2014