Robust Traceability of Trace Amounts

Cynthia Dwork
Adam Smith
Thomas Steinke
Salil Vadhan
Jonathan Ullman

Microsoft Penn State Harvard Northeastern

Thanks to Chen Sun for help with data
Sometimes summaries reveal a lot

Relax — it can only see metadata.
Sometimes summaries reveal a lot

- [Homer et al. (2008)] showed exact high-dimensional summaries allow an attacker with knowledge of population to test membership in a data set
  - Can also find out whether participant was case or control, or...
  - Not specific to genetic data

- This paper: strengthened membership tests
  - Approximate statistics
  - Less side information
This talk

- Background
- An abstract setting
- Results
Abstract setting

• Data $X : x_1, x_2, ..., x_n \in \{0,1\}^d$
  ➢ $d$ binary attributes for each person
  ➢ Think: $d$ big and $n$ moderate

• Summary statistics
  ➢ Column averages $\bar{x}(j) = \sum_i x_i(j)$, for $j = 1, ..., d$.

• Actual output
  ➢ Estimates $q(j) \in \bar{x}(j) \pm \alpha$

• Goal:
  
  given $q$ and a “target person” $z \in \{0,1\}^d$, determine if $z \in X$

• Assumptions:
  ➢ $x_1, ..., x_n$ i.i.d. from distribution $P$
  ➢ Attributes are independent
    • $P = P_{\mu}$ is a described by vector $\mu_1, ..., \mu_d$
    
    $E_{X \sim P_{\mu}}(X) = \mu$
  ➢ $Z$ either uniform in sample $X$ or fresh from $P$

Two applications
• Deanonymization
• Forensics

Impossible without some assumptions
Relation to Previous work

- **Membership tests** [Sankararaman et al., Nature Genomics 2009] assume
  - Exact statistics are published ($\alpha = 0$)
  - Nearly-exact knowledge of distribution

- **Fingerprinting codes** [Tardos 2003, Bun, Ullman, Vadhan 2014, Steinke, Ullman 2015] assume
  - Robust to perturbed statistics ($\alpha < 1/2$)
  - Artificial distribution, exactly known

- **This work**
  - Robust to perturbation: analysis for arbitrary $\alpha < 1/2$
    - Same test works for all perturbation mechanisms
    - Mathematically, very different from “normal” hypothesis testing
  - Limited side information
    - Reference sample of size $m \geq 1$ from the population

Graphical Model

Out

In
Graphical Model: This Work

“Out”

“In”
## Comparison

<table>
<thead>
<tr>
<th></th>
<th>Previous work</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tracer knowledge about</strong> $P_p$</td>
<td>Exact parameters or large sample from $P$ (2$n$ points)</td>
<td>$m \geq 1$ fresh samples from $P$</td>
</tr>
<tr>
<td><strong>Mechanism</strong></td>
<td>$q(X) = \bar{X}$</td>
<td>$q(j) \in \bar{X}(j) \pm \alpha$ (for $\alpha$ constant) and $\mu_i \sim p_i$ where the $p_i$ are “smooth” (e.g. uniform, Lipschitz differentiable density)</td>
</tr>
<tr>
<td><strong>Dimension of released data</strong></td>
<td>$d &gt; n$</td>
<td>$d &gt; n + \alpha^2 n^2 + n^2/m$</td>
</tr>
<tr>
<td><strong>Success probability (max of FP and FN rates)</strong></td>
<td>$1 - \exp\left(\frac{C d}{n}\right)$</td>
<td>$1 - \exp\left(-C \frac{d}{n + \alpha^2 n^2 + n^2/m}\right)$ if we assume $q$ depends only on $\bar{x}$ $\Omega(\alpha^2)$ in general</td>
</tr>
</tbody>
</table>

- Simple test; same test works in many settings
- Matches asymptotic accuracy of differentially private release: $\alpha \approx \sqrt{d}/(\epsilon n)$ so $d \approx \alpha^2 (\epsilon n)^2$
Tracing algorithm

• Given $q \in [0,1]^d$ and $z, y_1, \ldots, y_m \in \{0,1\}^d$ and $\delta > 0$
  ➢ Compute
  
  $$T = \langle z - y_1, q - y_{-1} \rangle$$
  
  ➢ If $T > 3\alpha \sqrt{d \log(1/\delta)}$, return “In”
  Else return “Out”

• Theorems [see paper]: Under various conditions,
  $\Pr(Tracer \ says \ “In” \mid OUT) < \delta$, and
  $\Pr(Tracer \ says \ “In” \mid IN) > 1 - \exp(...).$

Previous work: Likelihood ratio test

$$T \approx \left( z, \log \left( \frac{q_j}{1-q_j} \right) - \log \left( \frac{p_j}{1-p_j} \right) \right)$$
Proof Idea

$T_{out}$

$T_{in}$
This talk

• Background

• An abstract setting

• Results
Increasing the dimension

- Simulated data
  - Independent columns ("linkage equilibrium")
- Means drawn from actual distribution on allele frequencies (Hapmap CEU)
  - Following set up from Sankararaman et al.
- \( n = 100 \)
- \( m = 200 \)
- Published statistics rounded down to multiple of 0.1

Conclusion: Results fit roughly to theory
Robustness to perturbation

- $n = 100$
- $m = 200$
- $d = 5,000$
- **Two tests**
  - LR [Sankararam et al]
  - IP [this work]
- **Two publication mechanisms**
  - Rounded to nearest multiple of 0.1 (red / green)
  - Exact statistics (yellow / blue)

**Conclusion:** IP test is robust. Calibrating LR test seems difficult.
**Shrinking the reference pool**

- Rounding to 0.1
- \( n = 100 \) and \( d = 5,000 \)
- Get reliable signal for \( m \) above about 25
What happens when $m = 1$?

- Here $n = 100$ and $m = 1$
- Mechanism rounds down to multiples of 0.1
- Still get a reliable signal for individual’s presence
  - As predicted, much larger dimension is necessary
Future Work

• Real data
• Optimal test
  ➢ Application: calibrating competitions
• Other types of statistics
  ➢ Preliminary results on pairwise frequencies

Bigger questions
• How common are these problems “in the wild”?
• How should policies adjust?