Differentially Private Release and Learning of Threshold Functions

October 19, 2015

Mark Bun
Kobbi Nissim
Uri Stemmer
Salil Vadhan

Harvard University
Harvard U. and Ben-Gurion U.
Ben-Gurion University
Harvard University
Median Estimation

Data domain $[T] = \{1, \ldots, T\}$
Median Estimation

Data domain $[T] = \{1, \ldots, T\}$
Median Estimation

Data domain $[T] = \{1, \ldots, T\}$

$n = \text{constant \# of samples (independent of } T\text{)}$

...but naïve computation may violate privacy!

$\text{Vitamin } C$ level

$1 2 3 \ldots$ m $m'$ ... $\ldots x_n$
Privacy-Preserving Data Analysis

Want curators that are: • Private  • Accurate  • Efficient
Privacy-Preserving Data Analysis

Want curators that are:

- Differentially Private
- Accurate for "Threshold Tasks"
- Sample Efficient
This Talk

• **Sample complexity** of threshold tasks with approx. differential privacy
  - At most $8^{\log^* T}$ [Beimel-Nissim-Stemmer13]

• **MAIN RESULT**: These tasks have **equivalent** sample complexity, with
  - $n \geq \log^* T$

• Based on reductions to the simpler “interior point problem”
Differential Privacy

DN03+Dwork, DN04, BDMN05,
Dwork-McSherry-Nissim-Smith06, Dwork-Kenthapadi-McSherry-Mironov-Naor06

\[ D \]

\[
\begin{array}{c}
\vdots \\
x_n \\
\end{array}
\]

\[ D' \]

\[
\begin{array}{c}
\vdots \\
x_n' \\
\end{array}
\]

\[ M \]

\[ D \text{ and } D' \text{ are neighbors if they differ on one row} \]

small const., e.g. \( \varepsilon = 0.1 \)

“cryptographically small” require \( \delta << 1/n \), often \( \delta = \text{negl}(n) \)

\[ M \text{ is } (\varepsilon,\delta)-\text{differentially private} \text{ if for all neighbors } D, D' \text{ and } S \subseteq \text{Range}(M): \]

\[
\Pr[M(D') \in S] \leq (1+\varepsilon)\Pr[M(D) \in S] + \delta
\]
Accuracy for Approx. Medians

\[ \mathcal{D} = \text{unknown distribution over } [T] \text{ with median } m \]

M is accurate if
\[ \Pr_{x \sim \mathcal{D}} [x \in [m, m')] < 0.05 \]
(w.p. 99% over sample, coins(M))
Private Approx. Medians

[McSherry-Talwar07]

• $(\varepsilon, 0)$-differential privacy: Changing one person’s data alters PMF by factor of $(1+\varepsilon)$

• Accuracy: $n = O(\log T)$ samples suffice to produce approx. median

<table>
<thead>
<tr>
<th>Privacy</th>
<th>Accuracy</th>
<th>Sample Complexity</th>
</tr>
</thead>
</table>

Sample $m'$ with exponentially decaying PMF:
Accuracy for Threshold Tasks

\( f_t : [T] \rightarrow \{0,1\} \) with \( f_t(x) = 1 \) if \( x \leq t \), \( = 0 \) if \( x > t \)

Threshold Estimation

For each \( t \in [T] \): “What fraction of dist. \( \mathcal{D} \) satisfies the threshold property \( f_t \)?”

\( M \) is accurate if

\[ |a_t - f_t(\mathcal{D})| < 0.05 \] for every threshold \( t \)

(Properly) PAC Learning Thresholds [Val84]

“What threshold function generalizes labeled examples from \( \mathcal{D} \)?”

\( M \) is accurate if

\( f_s(x) = f_t(x) \) w.p. > 0.95. over \( x \sim \mathcal{D} \)

Privacy  Accuracy  Sample Complexity
Accuracy for Threshold Tasks

\( f_t : [T] \rightarrow \{0,1\} \) with \( f_t(x) = 1 \) if \( x \leq t \), \( = 0 \) if \( x > t \)

Threshold Estimation

\( = \) CDF Learning

Threshold Estimation

Non-private sample complexity = \( O(1) \) [DKW56]

M is accurate if

\[ |\alpha_t - f_t(D)| < 0.05 \]

for every threshold \( t \)

Quantile Estimation

\( \hat{D} \rightarrow \)

\( x_1, x_2, \ldots, x_n \)

(Properly) PAC Learning

Thresholds [Val84]

Non-private sample complexity = \( O(1) \) [BEHW86]

M is accurate if

\[ f_s(x) = f_t(x) \text{ w.p. } > 0.95 \]

empirical risk

\( = \) Empirical Risk Minimization

Privacy

Accuracy

Sample Complexity
Sample Complexity for Diff. Privacy

How big does \( n \) have to be to guarantee accuracy and privacy for threshold tasks?

Question: Is there an additional price of diff. privacy over statistical accuracy alone?

\[
(0.1, 1/n^2)\text{-diff. priv.}
\]
# Sample Complexity for Diff. Privacy

<table>
<thead>
<tr>
<th>No privacy</th>
<th>Approximate Medians / Releasing Thresholds</th>
<th>(Proper) PAC Learning of Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = \Theta(1)$</td>
<td>$n = \Theta(1)$</td>
</tr>
<tr>
<td></td>
<td>[DKW56]</td>
<td>[BEHW86]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(0.1, $1/n^2$)-diff. privacy</th>
<th>Upper bounds:</th>
<th>Lower bound:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mathcal{O}(T^{1/2})$</td>
<td>$\Omega(\log^* T)$</td>
</tr>
<tr>
<td></td>
<td>[DN03, DN04, BDMN05, DMNS06]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mathcal{O}(\log T)$</td>
<td>$\Omega(\log^* T)$</td>
</tr>
<tr>
<td></td>
<td>[BLR08, DNPR10, CSS10, DNRR15]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$8^{\log^* T(1+o(1))}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[BNS13]</td>
<td></td>
</tr>
</tbody>
</table>

**First separation** (generalizes to higher VC-dim)

**OUR WORK:** Plus somewhat improved upper bounds

- Privacy
- Accuracy
- Sample Complexity
Lower Bounds in Diff. Privacy

- Tight lower bounds known for $(\varepsilon, 0)$-diff. privacy [HT10, Hardt11], but break even for $\delta = \text{negl}(n)$ [De11, BNS13]

- Lower bounds for $(\varepsilon, \delta)$-diff. privacy based on reconstruction attacks [DN03, Roth10] and fingerprinting codes [BUV14 et seq.] exploit high dimensionality of queries/data

- This work: Lower bounds for $(\varepsilon, \delta)$-diff. privacy even for simple queries (i.e. VC-dimension = 1)
Our Techniques

• Reductions between threshold tasks and the “Interior Point Problem”
  – These tasks have equivalent sample complexity

• New upper and lower bounds for solving IPP with approx. differential privacy
  \[ \log^* T \leq n \leq 2^{\log^* T} \]
Interior Point Problem

- **Input:** Database $D = (x_1, \ldots, x_n) \in [T]^n$
- **Output:** Any $p \in [T]$ with $\min_i x_i \leq p \leq \max_i x_i$

Want $(\varepsilon, \delta)$-diff. privacy + success w.p. $2/3$
General Reductions

Threshold Estimation
“What fraction of $D$ satisfies the threshold property $f_t$?”

= Approximate Medians / Quantile Estimation

= CDF Learning

Interior Point Problem

(Properly) PAC Learning Thresholds
“What threshold function generalizes labeled examples from $D$?”

= Empirical Risk Minimization
Results for Interior Point

• **Lower bound:** Sample complexity of IPP is
  \[ n \geq \Omega(\log^* T) \]

• **Upper bound:** Sample complexity of IPP is
  \[ n \leq 2^{\log^* T(1+o(1))} \]

- Simpler algorithm inspired by lower bound construction
- Better dependence on error
Interior Point Lower Bound

- Recursively construct hard distributions $\mathcal{D}_n$ on domain size $T(n) \approx \text{tower}(n)$ \Rightarrow \ n \geq \log^* T$

- **Base case:** For $n = 1$, set $T(1) = 2$

\[
\begin{array}{c}
1 \quad 2 \\
\text{Output 1 w.p.} \geq \frac{2}{3}
\end{array}
\]  

- **Inductive case:**
  Suppose M solves IPP on $\mathcal{D}_{n+1}$ over domain $[T(n+1)]$
  \Rightarrow construct $M'$ for IPP on $\mathcal{D}_n$ over $[T(n)]$

\[
\begin{array}{c}
1 \quad 2 \\
\text{Output 1 w.p.} \geq \frac{(2/3) - \delta}{1 + \epsilon} > \frac{1}{3}
\end{array}
\]
Interior Point Lower Bound

To sample $D_{n+1}$ from $\mathcal{D}_{n+1}$:
1. Sample $D_n = (x_1, \ldots, x_n)$ from $\mathcal{D}_n$
2. Sample $y_0 \in [b^{T(n)}]$ at random
3. For $i = 1, \ldots, n$, sample $y_i$ that agrees with $y_0$ up to base b-“digit” $x_i$

\[
D_n = \begin{array}{c}
  x_1 \\
  x_2 \\
  x_3 \\
  \vdots \\
  x_n \\
\end{array} \quad \Rightarrow \quad D_{n+1} = \begin{array}{c}
  y_0 \\
  y_1 \\
  y_2 \\
  \vdots \\
  y_n \\
\end{array}
\]
To sample $D_{n+1}$ from $\mathcal{D}_{n+1}$:

1. Sample $D_n = (x_1, \ldots, x_n)$ from $\mathcal{D}_n$
2. Sample $y_0 \in [b^{T(n)}]$ at random
3. For $i = 1, \ldots, n$, sample $y_i$ that agrees with $y_0$ up to base $b$-“digit” $x_i$

E.g.

$D_n = \begin{array}{c}
\hline
x_1 = 3 \\
x_2 = 5 \\
x_3 = 3 \\
\vdots \\
x_n = 4 \\
\hline
\end{array}$

$D_{n+1} = \begin{array}{c}
\hline
y_0 = 8675309 \\
y_1 = 8674812 \\
y_2 = 8675365 \\
y_3 = 8671863 \\
\vdots \\
y_n = 8675052 \\
\hline
\end{array}$
Interior Point Lower Bound

$D_{n+1}$ is a hard distribution by contradiction:

M solves IPP on $D_{n+1}$

$\Rightarrow M'$ solves IPP on $D_n$

\[
\begin{align*}
D_n &= \begin{array}{c}
x_1 = 3 \\
x_2 = 5 \\
\vdots \\
x_n = 4
\end{array} \\
D_{n+1} &= M
\end{align*}
\]

\[
\begin{align*}
D_n &= \begin{array}{c}
y_0 = 8675309 \\
y_1 = 8674812 \\
y_2 = 8675365 \\
\vdots \\
y_n = 8675052
\end{array} \\
M' &= \text{x = max agreement between } y \text{ and } y_0
\end{align*}
\]

$y = 8675113$

$x = 4$
Interior Point Lower Bound

If $M$ succeeds,
- $x \geq \min_i x_i$ by construction
- $x \leq \max_i x_i$ by privacy (whp)

$\Rightarrow M'$ succeeds

$x_1 = 3$
$x_2 = 5$
$\vdots$
$x_n = 4$

$D_n =$

$y_0 = 8675309$
$y_1 = 8674812$
$y_2 = 8675365$
$\vdots$
$y_n = 8675052$

$D_{n+1} =$

$y = 8675113$

$x = 4$

$x = \text{max agreement between } y \text{ and } y_0$
If $M$ succeeds,
• $x \geq \min_i x_i$ by construction
• $x \leq \max_i x_i$ by privacy (whp)

$\Rightarrow M' \text{ succeeds}$

$D_n = \begin{pmatrix}
  x_1 = 3 \\
  x_2 = 5 \\
  \vdots \\
  x_n = 4 
\end{pmatrix}$

$D_{n+1} = \begin{pmatrix}
  y_0 = 8675309 \\
  y_1 = 8674812 \\
  y_2 = 8675365 \\
  \vdots \\
  y_n = 8675052 
\end{pmatrix}$

$x = \text{max agreement between } y \text{ and } y_0$

$x^* = 6$

$y^* = 8675305$
Interior Point Lower Bound

If $M$ succeeds,

- $x \geq \min_i x_i$ by construction
- $x \leq \max_i x_i$ by privacy (whp)

$=> M'$ succeeds

$D_n = \begin{array}{c}
  x_1 = 3 \\
  x_2 = 5 \\
  \vdots \\
  x_n = 4 
\end{array}$

$D_{n+1} = \begin{array}{c}
  y_0 = 8675309 \\
  y_1 = 8674812 \\
  y_2 = 8675365 \\
  \vdots \\
  y_n = 8675052 
\end{array}$

Since $y_0[6]$ random,
$\Pr[y^*[6] = y_0[6]] \leq (1+\epsilon)/b + \delta$

Max agreement between $y$ and $y_0$

$x^* = 6$

$y^* = 8675305$
Interior Point Lower Bound

- Recursively construct hard distributions $\mathcal{D}_n$ on domain size $T(n) \approx \text{tower}(n)$ $\Rightarrow$ $n \geq \log^* T$

- **Base case:** For $n = 1$, set $T(1) = 2$

- **Inductive case:**
  Suppose $M$ solves IPP on $\mathcal{D}_{n+1}$ over domain $[T(n+1)]$ $\Rightarrow$ construct $M'$ for IPP on $\mathcal{D}_n$ over $[T(n)]$

Output 1 w.p. $\geq \frac{2}{3}$

Output 1 w.p. $\geq \frac{(2/3) - \delta}{1 + \epsilon} > \frac{1}{3}$
Conclusions

• Diff. privacy-preserving reductions between threshold tasks

• Price of $(\varepsilon, \delta)$-diff. privacy for simple statistics

• Open questions:
  – Combinatorial characterization of sample complexity?
    [e.g. HT10, Har11, NTZ13, BNS13]
  – Sample complexity of improper PAC learning?
    [e.g. BKN10, FX14]

Thank you!