ABS Perturbation Methodology
Through the Lens of Differential Privacy

James Bailie and Joseph (Chien-Hung) Chien*

ABS Methodology Division

29 October 2019

Acknowledgement

we would like to thank contributions from the ABS colleagues: Kelly Chiu, Daniel Elazar, Edwin Lu, Bernadette Giuffrida, Jenny Pocknee, Aparjit Singh and Ric Clarke
Recent developments

- US Census Bureau reconstruction attack (Garfinkel et al. 2018).
- A preliminary ABS assessment on the reconstruction attacks.
  - Different legislation requirements.
  - Different available information.
  - The probability of a successful attack is low.
- Office for National Statistics and Statistics New Zealand are currently exploring DP.
- DP focuses on producing safe outputs.
How do we assess reconstruction attacks on the perturbed outputs?

<table>
<thead>
<tr>
<th></th>
<th>True counts</th>
<th>Possible with ±1 noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4</td>
<td>3 3 4 4 4 4 5 5 5 ... 4 5 5 5 3 ... 5</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>5 6 7 5 6 7 5 6 7 ... 7 5 6 7 5 ... 7</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>9 9 9 9 9 9 9 9 9 ... 10 10 10 10 11 ... 11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Published</th>
<th>Possible with ±1 noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5</td>
<td>4 4 4 5 5 5 6 6 6 ... 5 6 6 6 4 ... 6</td>
</tr>
<tr>
<td>Female</td>
<td>7</td>
<td>6 7 8 6 7 8 6 7 8 ... 8 6 7 8 6 ... 8</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>9 9 9 9 9 9 10 ... 10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Published</th>
<th>Possible with ±1 noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3</td>
<td>2 2 3 3 4 4 4 4 ... 3 4 4 4 2 ... 4</td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>4 5 6 4 5 6 4 5 6 ... 6 4 5 6 4 ... 6</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>9 9 9 9 9 9 10 ... 10</td>
</tr>
</tbody>
</table>
Selected results for single cell success probabilities on perturbed outputs

Table 1: Probability of the attacker determining a single cell with a true count of one, where the number of variables $m$ ranges from two to six, and the number of categories $c$ in each variable ranges from two to ten.
ABS TableBuilder

- **5 safes** governance framework.
- dynamic tables.
- same cell same noise to maintain user trust of statistical outputs.

- **entropy maximisation** maximises entropy subject to desired variance and minimum bias (Marley and Leaver 2011).
- **protections** include sparsity settings, field exclusion rules and auditing/monitoring.
Dwork and Roth (2014) define a publishing method $M$ as $(\epsilon, \delta)$-differential privacy if for two neighbouring datasets $D$ and $D'$

$$Pr(M(D) \in S) \leq \exp(\epsilon) \times Pr(M(D') \in S) + \delta,$$

where $D$ and $D'$ differing at most one cell, all perturbed outputs $S \subseteq \text{Range}(M)$.

The publishing method $M$ adds noise derived from some distribution (Laplace is a popular choice).
Calculation of $\delta$ for a single query in TableBuilder
Calculation of $\epsilon$ for a single query in TableBuilder

$$\epsilon \geq \max \left[ \log \left( \frac{\Pr(M(n) = S)}{\Pr(M(n-1) = S)} \right), \log \left( \frac{\Pr(M(n) = S)}{\Pr(M(n+1) = S)} \right) \right] \text{ in region } > \delta$$
### TableBuilder in the lens of DP

<table>
<thead>
<tr>
<th>cell counts &lt; 7</th>
<th>cell counts &gt; 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon$</td>
<td>1.253</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>0.693</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
</tr>
</tbody>
</table>
Lessons learnt and future directions

- Perturbation adds complexity to perform reconstruction attacks.
- We can derive $\epsilon$ and $\delta$ in the case of a SINGLE query.
- Collaborative opportunities to address the following questions:
  - How can DP best combine with the 5-safes governance framework?
  - How do we optimally calculate $(\epsilon, \delta)$ in a dynamic query environment?
  - How can we incorporate DP to generate perturbation distributions?


Extra slides
DP adaptation of entropy maximisation

- let $M(n) = n + e$ where $n$ is the original count and $e$ is the perturbation value.

- Fraser and Wooton (2005) and Marley and Leaver (2011) discuss perturbation values are generated by maximising entropy:

$$\text{MAX}_e = - \sum_e p_{e|n} \log p_{e|n}$$

subject to the constraints:

- the perturbation values cannot add bias i.e. $E_{e|n}(e) = 0$.
- the variance of the perturbation values cannot exceed a specified threshold $\text{var}_{e|n} \leq \nu_n$.
- the confidentialised cell values cannot be negative i.e. $M(n) \geq 0$.
- probability distribution adds to 1 e.g. $\sum_e p_{e|n} = 1$ and $p_{e|n} \geq 0$. 
Adding DP constraints to optimiser

- Constraining the region of $\delta$ by specifying $p_{e|n}$ and $p_{e|n-1}$ as

$$\sum_{e=-\infty}^{l_{n-1}-2} p_{e|n} + \sum_{e=u_{n-1}}^{\infty} p_{e|n} \leq \delta, \forall n \in \mathbb{Z}^{\geq 0},$$

where

$$l_{n} = \min \{ e | p_{e|n} > 0 \} \text{ and } u_{n} = \max \{ e | p_{e|n} > 0 \}.$$

Specifying $p_{e|n}$ and $p_{e|n+1}$ as

$$\sum_{e=-\infty}^{l_{n+1}} p_{e|n} + \sum_{e=u_{n+1}+2}^{\infty} p_{e|n} \leq \delta, \forall n \in \mathbb{Z}^{\geq 0},$$

- Constraining the height of $\epsilon$ by specifying

$$\log \left( \frac{p_{e|n}}{p_{e+1|n-1}} \right) \leq \epsilon, \forall n \in \mathbb{Z}^{\geq 0} \text{ with } p_{e+1|n-1} \neq 0 \text{ and } p_{e|n} > \delta.$$

$$\log \left( \frac{p_{e|n}}{p_{e-1|n+1}} \right) \leq \epsilon, \forall n \in \mathbb{Z}^{\geq 0} \text{ with } p_{e-1|n+1} \neq 0 \text{ and } p_{e|n} > \delta.$$