Fuzzi: A Three Level Logic for Differential Privacy

Hengchu Zhang, Edo Roth, Andreas Haeberlen, Benjamin C. Pierce, Aaron Roth

University of Pennsylvania
Differential Privacy is Useful

Census Bureau Adopts Cutting Edge Privacy Protections for 2020 Census

From Feb 15 2019
Written by: Dr. Ron Jarmin, Deputy Director and Coo

Research » Publications

RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response

Learning with Privacy at Scale

Vol. 1, Issue 8 • December 2017
by Differential Privacy Team
Differential Privacy is Useful

\[ \forall \sim \quad f(\) is \((\epsilon, \delta)\) close to \(f(\) \]

Census Bureau Adopts Cutting Edge Privacy Protections for 2020 Census

Fri Feb 15 2019
WRITTEN BY: DR. RON JARMIN, DEPUTY DIRECTOR AND COO
Privacy Parameters

Parameter $\varepsilon$ bounds the multiplicative difference in probability
Differential Privacy in an imperative programming language?
Fuzzi and its Three Levels

Type System for differentially private, imperative programs

Advanced Probabilistic Couplings for Differential Privacy
Barthe et al. 2016.

Language Semantics

Base Logic

apRHL

Type System

automation

abstraction

manual proofs
An Example Fuzzi Program

partition

100.0, 205.0, 1000.0, 2500.0,
99999.0, 10000.0, ...

100.0, 205.0, ...

1000.0, 2500.0, ...

99999.0, 10000.0, ...

sum

1000000.0

10000000.0

9000000000.0

(8999900000.0)

sum

125759.1

50075392.6

90025315.9 (900042943.8)

laplace noise

100.0, 205.0, ...

1000.0, 2500.0, ...

99999.0, 10000.0, ...

sum

1000000.0

10000000.0

9000000000.0

(8999900000.0)

sum

125759.1

50075392.6

90025315.9 (900042943.8)
// income : 1 {float}
// epsilon=0.0, delta=0.0
income_groups = partition(income, ...);

// income_groups : 1 [{float}]
// epsilon=0.0, delta=0.0
t_part = income_groups[0];
income_sum = bsum(t_part, 1000.0);

// low_income_sum : 1000.0 {float}
// epsilon=1.0, delta=0.0
income_sum = laplace(income_sum, 1000.0);

An Example Fuzzi Program
Fuzzi Type System

\[
\begin{align*}
\{\Gamma_1\} & \ c \ \{\Gamma_2, (\epsilon, \delta)\} \\
\{\Gamma_2\} & \ c' \ \{\Gamma_3, (\epsilon', \delta')\} \\
\hline
\{\Gamma_1\} & \ c ; c' \ \{\Gamma_3, (\epsilon + \epsilon', \delta + \delta')\}
\end{align*}
\]
Type System as an Interface to apRHL

Hoare Logic

\[ P(M) \]

\[ \text{prog} \]

\[ Q(N) \]

Relational

\[ P(M, M') \]

\[ \text{prog, prog'} \]

\[ Q(N, N') \]

Approximate Relational Hoare Logic

\[ P(M, M') \]

\[ \text{prob-prog, prob-prog'} \]

\[ Q(N, N') \]

\[ \epsilon, \delta \]

\[ P, Q := x\langle 1 \rangle = x\langle 2 \rangle \land y\langle 1 \rangle = 5 \]
Type System as an Interface to apRHL

\[
\{ \Gamma_1 \} \ c \ \{ \Gamma_2, (\epsilon, \delta) \}
\]

\[\left[ x : s \ \text{int} \right] = \left| x\langle 1 \rangle - x\langle 2 \rangle \right| \leq s\]
Packaging Manual Proofs for Mechanisms

\[ P(M, M') \]

\[ \delta, \epsilon \]

\[ Q(N, N') \]
## Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Differentially Private</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.84 (11.02, 10e-6)</td>
<td>MNIST</td>
</tr>
<tr>
<td>Ensemble of Logistic Regression</td>
<td>0.82 (20.0, 0.0)</td>
<td>MNIST (partitioned)</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>0.69 (7.70, 0.0)</td>
<td>Spambase</td>
</tr>
<tr>
<td>K-Means</td>
<td>0.55 - 0.9, median 0.69 (21.0, 0.0)</td>
<td>Iris</td>
</tr>
</tbody>
</table>


A very modal model of a modern, major, general type system. Appel et al. 2007.


A very modal model of a modern, major, general type system. Appel et al. 2007.

Conclusion

1. We propose a high-level sensitivity type system for tracking differential privacy
   a. We establish soundness through straightforward embedding into apRHL;
   b. The type system is expressive enough for verification conditions of manual differential privacy proofs in apRHL.

2. We show how to push manual proof results of DP back into sensitivity type system
   a. We develop manual proofs of bag-map, bag-sum, partition, advanced composition.

3. We evaluate Fuzzi by implementing 4 textbook machine learning algorithms
   a. We build a prototype of Fuzzi in Haskell
   b. We translate Fuzzi program into Python3 for execution
Fuzzi: A Three Level Logic for Differential Privacy

Hengchu Zhang, Edo Roth, Andreas Haeberlen, Benjamin C. Pierce, Aaron Roth

University of Pennsylvania
A Privacy Type System for Simple While Programs

\[ \Gamma ::= \emptyset \mid \Gamma, x :_{s} \tau \]

Plus

\[ \Gamma \vdash e_{l} :_{s} \text{int} \quad \Gamma \vdash e_{r} :_{t} \text{int} \]

\[ \Gamma \vdash e_{l} + e_{r} :_{s+t} \text{int} \]
A Privacy Type System for Simple While Programs

\[
\begin{align*}
\text{Plus} & \\
\Gamma \vdash e_l : s \text{ int} & \quad \Gamma \vdash e_r : t \text{ int} \\
\hline
\Gamma \vdash e_l + e_r : s + t \text{ int}
\end{align*}
\]

Laplace

\[
\begin{align*}
\Gamma \vdash e : s \text{ float} \\
\{\Gamma\} x = \text{laplace}(e, w) \{\Gamma[x \rightarrow 0], (s/w, 0)\}
\end{align*}
\]
Properties of Differential Privacy

1. Compositional
   ✓ Given $f_1$ $(\epsilon_1, \delta_1)$-DP, and $f_2$ $(\epsilon_2, \delta_2)$-DP
   ✓ Running $f_1$ followed by $f_2$ is $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-DP

2. Robust to post-processing
   ✓ Further analysis on the results of $f$ does not weaken its DP guarantees
Differential Privacy is Subtle

Understanding the Sparse Vector Technique for Differential Privacy
Min et al. 2016.

On the Privacy Properties of Variants on the Sparse Vector Technique
Chen and Machanavajjhala. 2015.