Riposte: An Anonymous Messaging System Handling Millions of Users

Henry Corrigan-Gibbs,
Dan Boneh, and David Mazières
Stanford University

IEEE Security and Privacy
18 May 2015
With encryption, we can hide the data... 

...but does that hide enough?

(pk, sk)

VUIC9zZW5zaXRpdmU

pk
<table>
<thead>
<tr>
<th>Time</th>
<th>From</th>
<th>To</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:12</td>
<td>Alice</td>
<td>Bob</td>
<td>2543 B</td>
</tr>
<tr>
<td>10:27</td>
<td>Carol</td>
<td>Alice</td>
<td>567 B</td>
</tr>
<tr>
<td>10:32</td>
<td>Alice</td>
<td>Bob</td>
<td>450 B</td>
</tr>
<tr>
<td>10:35</td>
<td>Bob</td>
<td>Alice</td>
<td>9382 B</td>
</tr>
</tbody>
</table>

[cf. Ed Felten’s testimony before the House Judiciary Committee, 2 Oct 2013]
<table>
<thead>
<tr>
<th>Time</th>
<th>From</th>
<th>To</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:12</td>
<td>Alice</td>
<td><a href="mailto:taxfraud@stanford.edu">taxfraud@stanford.edu</a></td>
<td>2543 B</td>
</tr>
<tr>
<td>10:27</td>
<td>Carol</td>
<td>Alice</td>
<td>567 B</td>
</tr>
<tr>
<td>10:32</td>
<td>Alice</td>
<td>Bob</td>
<td>450 B</td>
</tr>
<tr>
<td>10:35</td>
<td>Bob</td>
<td>Alice</td>
<td>9382 B</td>
</tr>
</tbody>
</table>

Hiding the data is necessary, but not sufficient

[cf. Ed Felten’s testimony before the House Judiciary Committee, 2 Oct 2013]
Goal

The “Anonymity Set”
Goal
Goal
To: taxfraud@stanford.edu

Protest will be held tomo…

See my cat photos at w…

DBs do not learn who wrote which message
Building block for systems related to “hiding the metadata”
→ Anonymous Twitter
→ Anonymous surveys
→ Private messaging, etc.
Low-latency anonymity systems (e.g., Tor) … do not protect against a global adversary

Mix-nets … require expensive ZKPs to protect against active attacks

Riposte is an anonymous messaging system that:
• protects against a near-global active adversary
• handles millions of users in an “anonymous Twitter” system
Outline

• Motivation
• A “Straw man” scheme
• Technical challenges
• Evaluation
"Straw man" Scheme
[Chaum ‘88]
"Straw man" Scheme
Write msg $m_A$ into DB row 3

$m_A \in \mathbb{F}$

“Straw man” Scheme
“Straw man” Scheme
### "Straw man" Scheme

<table>
<thead>
<tr>
<th>$S_X$</th>
<th>$S_Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m_A$</th>
<th>$r_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$r_2$</td>
</tr>
<tr>
<td>0</td>
<td>$r_3$</td>
</tr>
<tr>
<td>0</td>
<td>$r_4$</td>
</tr>
<tr>
<td>0</td>
<td>$r_5$</td>
</tr>
</tbody>
</table>
**Scheme**

<table>
<thead>
<tr>
<th>$S_X$</th>
<th>$S_Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

```
\[ \begin{align*}
S_X & = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix} \\
S_Y & = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix}
\end{align*} \]
```

- $m_A$

raw man” Scheme

```
\[ \begin{align*}
0 & - r_1 \\
0 & - r_2 \\
m_A & - r_3 \\
0 & - r_4 \\
0 & - r_5 \\
\end{align*} \]
```

```
"Straw man" Scheme

$S_X$

\[
\begin{array}{c}
0 \\
0 \\
0 \\
0 \\
0
\end{array}
\]

$S_Y$

\[
\begin{array}{c}
0 \\
0 \\
0 \\
0 \\
0
\end{array}
\]

raw man" Scheme
"Straw man" Scheme
\[ S_X \]

\[
\begin{array}{c}
  r_1 \\
r_2 \\
r_3 \\
r_4 \\
r_5 \\
\end{array}
\]

\[ S_Y \]

\[
\begin{array}{c}
  -r_1 \\
  -r_2 \\
  -r_3 + m_A \\
  -r_4 \\
  -r_5 \\
\end{array}
\]

“Straw man” Scheme
The image depicts two sets, $S_X$ and $S_Y$, with elements $r_1, r_2, r_3, r_4, r_5$ in $S_X$ and $-r_1, -r_2, -r_3 + m_A, -r_4, -r_5$ in $S_Y$. Below $S_X$, there is a column labeled $m_B$. The text on the page reads ""Straw man" Scheme."
The document contains a table labeled $S_X$ and another labeled $S_Y$. The table entries are as follows:

**$S_X$**

<table>
<thead>
<tr>
<th>$r_1$</th>
<th>$r_2$</th>
<th>$r_3$</th>
<th>$r_4$</th>
<th>$r_5$</th>
</tr>
</thead>
</table>

**$S_Y$**

| $-r_1$ | $-r_2$ | $-r_3 + m_A$ | $-r_4$ | $-r_5$ |

Below the tables, there is a diagram showing a person labeled "raw man" performing a subtraction operation. The result of the subtraction is equal to another set of entries:

<table>
<thead>
<tr>
<th>$m_B$</th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
<th>$s_4$</th>
<th>$s_5$</th>
</tr>
</thead>
</table>

The result of the subtraction is $m_B - s_5$. The document is discussing a "raw man" Scheme.
Straw man

Scheme

$S_X$

$S_Y$

$-r_1$

$-r_2$

$-r_3 + m_A$

$-r_4$

$-r_5$

$s_1$

$s_2$

$s_3$

$s_4$

$s_5$

$m_B - s_5$

raw man”

Scheme
\[ S_X \]

\[
\begin{array}{c|c}
 r_1 & s_1 \\
r_2 & s_2 \\
r_3 & s_3 \\
r_4 & s_4 \\
r_5 & s_5 \\
\end{array}
\]

\[ S_Y \]

\[
\begin{array}{c|c}
 -r_1 & -s_1 \\
-r_2 & -s_2 \\
-r_3 + m_A & -s_3 \\
-r_4 & -s_4 \\
-r_5 & m_B - s_5 \\
\end{array}
\]

“Straw man” Scheme
### S\(_X\)

<table>
<thead>
<tr>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_1 + s_1 )</td>
</tr>
<tr>
<td>( r_2 + s_2 )</td>
</tr>
<tr>
<td>( r_3 + s_3 )</td>
</tr>
<tr>
<td>( r_4 + s_4 )</td>
</tr>
<tr>
<td>( r_5 + s_5 )</td>
</tr>
</tbody>
</table>

### S\(_Y\)

<table>
<thead>
<tr>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-r_1 - s_1)</td>
</tr>
<tr>
<td>(-r_2 - s_2)</td>
</tr>
<tr>
<td>(-r_3 - s_3 + m_A)</td>
</tr>
<tr>
<td>(-r_4 - s_4)</td>
</tr>
<tr>
<td>(-r_5 - s_5 - m_B)</td>
</tr>
</tbody>
</table>

**“Straw man” Scheme**
"Straw man" Scheme
\[
S_X
\begin{array}{|c|}
\hline
r_1 + s_1 \\
\hline
r_2 + s_2 \\
\hline
r_3 + s_3 \\
\hline
r_4 + s_4 \\
\hline
r_5 + s_5 \\
\hline
\end{array}
\]

\[
S_Y
\begin{array}{|c|}
\hline
-r_1 - s_1 \\
\hline
-r_2 - s_2 \\
\hline
-r_3 - s_3 + m_A \\
\hline
-r_4 - s_4 \\
\hline
-r_5 - s_5 - m_B \\
\hline
\end{array}
\]

"Straw man" Scheme
<table>
<thead>
<tr>
<th>$S_X$</th>
<th>$S_Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_1 + s_1$</td>
<td>$-r_1 - s_1$</td>
</tr>
<tr>
<td>$r_2 + s_2$</td>
<td>$-r_2 - s_2$</td>
</tr>
<tr>
<td>$r_3 + s_3$</td>
<td>$-r_3 - s_3 + m_A$</td>
</tr>
<tr>
<td>$r_4 + s_4$</td>
<td>$-r_4 - s_4$</td>
</tr>
<tr>
<td>$r_5 + s_5$</td>
<td>$-r_5 - s_5 - m_B$</td>
</tr>
</tbody>
</table>

“Straw man” Scheme
At the end of the day, servers combine DBs to reveal plaintext.

"Straw man" Scheme
First-Attempt Scheme: Properties

“Perfect” anonymity as long as servers don’t collude

- Can use $k$ servers to protect against $k-1$ collusions

Practical efficiency: almost no “heavy” computation involved

Unlike a mix-net, storage cost is constant in the anonymity set size
Outline

• Motivation
• A “Straw man” scheme
• Technical challenges
• Evaluation
Outline

• Motivation
• A “Straw man” scheme
• Technical challenges
  – Collisions
  – Malicious clients
  – $O(L)$ communication cost
• Evaluation
Outline

• Motivation
• A “Straw man” scheme
• Technical challenges
  – Collisions
  – Malicious clients
  – $O(L)$ communication cost
• Evaluation

in the paper
Challenge: Bandwidth Efficiency

In “straw man” design, client sends DB-sized vector to each server

**Idea**: use a **cryptographic trick** to compress the vectors

→ Based on PIR protocols

[Ostrovsky and Shoup 1997]
Distributed Point Function

KeyGen ($m, \ell$) → Eval ($k_1$, $x_1$) → Eval ($k_2$, $x_2$) → … → Eval ($k_n$, $x_n$) → $m = 0 0 0 0$

[Gilboa and Ishai 2014]
Distributed Point Function

\[(m, \ell)\]

KeyGen

\[k_1 \rightarrow \text{Eval} \rightarrow x_1\]

\[k_2 \rightarrow \text{Eval} \rightarrow x_2\]

\[\ldots \rightarrow \text{Eval} \rightarrow \ldots\]

\[k_n \rightarrow \text{Eval} \rightarrow x_n\]

\[x_1 + x_2 + \ldots + x_n = m\]

Privacy: A subset of keys leaks nothing about message or \(\ell\)
DPFs Reduce Bandwidth Cost
DPFs Reduce Bandwidth Cost
Alice sends $L^{1/2}$ bits (instead of $L$)

- Two-server version just uses AES (no public-key crypto)
- With fancier crypto, privacy holds even if all but one server is malicious

[Chor and Gilboa 1997]
[Gilboa and Ishai 2014]
Outline

• Motivation
• Definitions and a “Straw man” scheme
• Technical challenges
• Evaluation
Bottom-Line Result

• Implemented the protocol in Go
• For a DB with 65,000 Tweet-length rows, can process **30 writes/second**
• Can process **1,000,000 writes** in 8 hours on a single server

➔ **Completely parallelizable workload**
At large table sizes, AES cost dominates.
<table>
<thead>
<tr>
<th>Time</th>
<th>From</th>
<th>To</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:12</td>
<td>Alice</td>
<td><a href="mailto:taxfraud@stanford.edu">taxfraud@stanford.edu</a></td>
<td>2543 B</td>
</tr>
<tr>
<td>10:15</td>
<td>Bob</td>
<td>Alice</td>
<td>567   B</td>
</tr>
<tr>
<td>10:17</td>
<td>Carol</td>
<td>Bob</td>
<td>450   B</td>
</tr>
<tr>
<td>10:22</td>
<td>Dave</td>
<td>Alice</td>
<td>9382  B</td>
</tr>
<tr>
<td>Time</td>
<td>From</td>
<td>To</td>
<td>Size</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>10:12</td>
<td>Alice</td>
<td>Riposte Server</td>
<td>207 KB</td>
</tr>
<tr>
<td>10:15</td>
<td>Bob</td>
<td>Riposte Server</td>
<td>207 KB</td>
</tr>
<tr>
<td>10:17</td>
<td>Carol</td>
<td>Riposte Server</td>
<td>207 KB</td>
</tr>
<tr>
<td>10:22</td>
<td>Dave</td>
<td>Riposte Server</td>
<td>207 KB</td>
</tr>
</tbody>
</table>
Conclusion

In many contexts, “hiding the metadata” is as important as hiding the data

Combination of crypto tools with systems design $\Rightarrow$ 1,000,000-user anonymity sets

Next step: Better performance at scale