Faster Private Set Intersection based on OT Extension

Michael Zohner (TU Darmstadt)

Joint work with
Benny Pinkas (Bar Ilan University)
Thomas Schneider (TU Darmstadt)
Private Set Intersection (PSI)
Applications

- Secure database join
- Common contacts
- Botnet detection
- Cheater detection in online games
- Testing human genomes
- Relationship path discovery
A naïve PSI protocol

**Input:** $x_1, ..., x_n$

$H(x_1), ..., H(x_n)$

$H(x_i) \neq H(y_j)$, for $0 < i, j < n$

**Input:** $y_1, ..., y_n$

$H(y_1), ..., H(y_n)$

- **Pro:** fast, little communication

- **Con:** can leak privacy of Bob's inputs
Our Contributions

- Survey major results on semi-honest PSI
- Optimize existing PSI protocols
- Present a new PSI scheme
- Compare performance of all schemes
Existing PSI Protocols

- Public-key Cryptography
  - DH-based Protocol [M86], $O(n)$ pk-crypto
  - Blind RSA Protocol [CT10], $O(n)$ pk-crypto

- Generic Secure Computation
  - Based on Yao's garbled circuits, GMW
  - Circuit in [HEK12], $O(n \log n)$ sym-crypto

- Oblivious Transfer
  - Only sym-crypto via OT extension [IKNP03]
  - Bloom-filter [DCW13], $O(n)$ sym-crypto
Oblivious Transfer (OT)

- **Input**: Bob holds two strings \((s_0, s_1)\), Alice holds a choice bit \(c\)

- **Output**: Alice receives \(s_c\) but learns nothing about \(s_{1-c}\), Bob learns nothing about \(c\)
- Improve circuit-based PSI of [HEK12] using GMW
  - Multiplexer complexity independent of bit-length
  - Reduce computation / communication by factor 2
  - Also applicable to other functionalities

- Randomize Garbled Bloom filter of [DCW13]
  - Reduce computation by factor 3
  - Reduce communication by factor 4
  - Whole protocol can be parallelized
PSI based on OT

Private Equality Test:

Private Set Inclusion:

Private Set Intersection:
- **Input:** Alice has $x$, Bob has $y$. **Output:** $x \overset{?}{=} y$

- **Example:** $x = 001, y = 011$

- Bob sends $\lambda$-bit mask $\bigoplus \lambda \bigoplus \lambda \bigoplus \lambda$ to Alice

- Alice computes $\bigoplus \lambda \bigoplus \lambda \bigoplus \lambda$ and compares
PSI based on OT (Set Inclusion)

- **Input:** Alice has $x$, Bob has $Y = \{y_1, \ldots, y_n\}$. **Output:** $x \in Y$

- Run $n$ Private Equality Tests in parallel
  - Alice's OT choices for all $y_1, \ldots, y_n$ are the same
  - Send $n\lambda$ bits from Bob to Alice
PSI based on OT (Set Intersection)

- Input: Alice has $X = \{x_1, \ldots, x_n\}$, Bob has $Y = \{y_1, \ldots, y_n\}$.
- Output: $X \cap Y$.

- Run $n$ Private Set Inclusions in parallel
  - Requires $n^2$ comparisons, hence not an option
Hashing

- Hash elements to bins to reduce comparisons

- **Example:** Alice holds $X=\{x_1, \ldots, x_4\}$, Bob holds $Y=\{y_1, \ldots, y_4\}$

  \[ H(x_4), H(x_1), H(x_3) \]

  - Reduces comparisons from $n^2$ to $O(n \log n)$
Comparison Results

- PSI on $n = 2^{18}$ elements of 32-bit length for 128-bit security on Gbit LAN

![Comparison Diagram]
Comparison Results

- PSI on $n = 2^{18}$ elements of 32-bit length for 128-bit security on Gbit LAN

PK-Based:
- high run-time for large security parameters
+ best communication
Comparison Results

- PSI on $n = 2^{18}$ elements of 32-bit length for 128-bit security on Gbit LAN

Circuit-Based:
- high run-time & communication
+ easily extensible to arbitrary functions

22.08.14 | Faster PSI based on OT extension | Michael Zohner | Slide 16
- PSI on $n = 2^{18}$ elements of 32-bit length for 128-bit security on Gbit LAN
Conclusion

Rule of Thumb:

- OT-based protocols in general case
- DH-based ECC if communication is bottleneck
- Circuit-based protocols for easy extension

Goal: PSI on million element sets in less than 1 second
Faster Private Set Intersection based on OT Extension

Thank you for your attention
References


- [IKNP03] Y. Ishai, J. Kilian, K. Nissim, E. Petrank: Extending Oblivious Transfers Efficiently. In CRYPTO'03.
Protocol Overview

Special Purpose Protocols

- Homomorphic Encryption
- Public Key Crypto

Generic Protocols

- Arithmetic Circuit
- Boolean Circuit

- Yao
- GMW

Symmetric Crypto >> One-Time Pad

OT

22.08.14 | Faster PSI based on OT extension | Michael Zohner | Slide 21
Faster Private Set Intersection based on OT Extension

Private Set Intersection

Applications
- Secure Database Join
- Common Contacts
- Relationship Path Discovery

Techniques
- Public-key Cryptography
- Generic Secure Computation
- Oblivious Transfer

Summary
- Optimize & Implement
- New Protocol
- Comparison

1 Mio. elements in 20 seconds on PCs
## Results

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Communication</th>
<th>Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Hashing</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>DH FFC</td>
<td>192</td>
<td>1224</td>
</tr>
<tr>
<td>DH ECC</td>
<td>26</td>
<td>416</td>
</tr>
<tr>
<td>Blind RSA</td>
<td>132</td>
<td>1982</td>
</tr>
<tr>
<td>Circuit + GMW</td>
<td>23400</td>
<td>1304</td>
</tr>
<tr>
<td>Circuit + Opt GMW</td>
<td>14040</td>
<td>762</td>
</tr>
<tr>
<td>Yao</td>
<td>20736</td>
<td>609</td>
</tr>
<tr>
<td>Garbled Bloom</td>
<td>1393</td>
<td>154</td>
</tr>
<tr>
<td>Opt. GBF</td>
<td>740</td>
<td>68</td>
</tr>
<tr>
<td>OT + hashing</td>
<td>78</td>
<td>14</td>
</tr>
</tbody>
</table>