XRD: A Scalable Messaging System with Cryptographic Privacy

David Lu
Mentor: Albert Kwon

PRIMES Computer Science Conference
October 13, 2018
Acknowledgements

Thank you to Albert Kwon for mentoring me

Thank you to Prof. Devadas for PRIMES-CS

Thank you to Dr. Gerovitch for the PRIMES program

Thank you to my parents for their support
Motivation and Background
Motivation

Alice’s hides message content through encryption.

However, Alice still leaks metadata:
- Identities
- Timing
- Size

Alice

Bob
## Prior work

<table>
<thead>
<tr>
<th>System</th>
<th>Strong privacy guarantee</th>
<th>Scalable to millions of users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tor</td>
<td>✗ (traffic analysis)</td>
<td>✓</td>
</tr>
<tr>
<td>Mix-nets &amp; DC-nets</td>
<td>✓</td>
<td>✗ (messages go through one server or all users)</td>
</tr>
<tr>
<td>Stadium and Karaoke</td>
<td>▲ (differential privacy)</td>
<td>✓</td>
</tr>
<tr>
<td>Our goal</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Privacy guarantee

- Provide metadata private messaging against powerful adversaries
Deployment and threat model

- Global network adversary
- Fraction of the servers are malicious
- Large number of malicious users
XRD Base Design
Each message is either “loopback” or conversation message

If A and B are not communicating
Each message is either “loopback” or conversation message.

A = message sent to A
B = message sent to B
C = message sent to C

If A and B are communicating...
Security argument of base design

- Every mailbox gets exactly one message
  - Mailboxes are identical
- The origin of the message is hidden by mix-nets (because there is at least one honest server)
  - Hides swap-or-not
Active attacks

Each message is either “loopback” or conversation message

A = message sent to A
b = message sent to B
c = message sent to C

If A and B are communicating
Stopping active attacks: zero-knowledge proofs

- Each server generates a zero-knowledge proof
  - Proofs prove valid decryption and shuffle
- Thwarts attacks because tampered or dropped messages are caught
Scaling the Base Design
XRD: scaling the simple design

1. Send messages to $l$ chains
2. Mix and decrypt messages
3. Forward messages to mailboxes

$l = 2$
XRD: scaling the simple design

1. Send messages to $l$ chains
2. Mix and decrypt messages
3. Forward messages to mailboxes

If 1 and 4 are talking to each other with $l = 2$
XRD: scaling the simple design

1. Send messages to $l$ chains
2. Mix and decrypt messages
3. Forward messages to mailboxes

If 1 and 4 are not talking to each other with $l = 2$
Security argument

- Every mailbox gets exactly $\ell$ messages
  - Mailboxes are identical
- Every pair of users intersects
  - Hides which users are talking with each other
- The origin of the message is hidden by mix-nets (because there is at least one honest server per mix-net)
  - Hides swap-or-not
XRD: scaling the simple design

1. Send messages to $\ell$ chains
2. Mix and decrypt messages
3. Forward messages to mailboxes

If 1 and 5 are talking to each other $\ell = 4$
Scalability properties

For $m$ users and $n$ chains,

- We can make sure all users intersect with $\ell = \sqrt{2n}$
- Each chain handles $m*\ell/n = (\sqrt{2})m/(\sqrt{n})$ messages
  - If you increase $n$, the load per chain goes down (scalable)
XRD Results
Experimental set-up

- Benchmark time for decryption, shuffle, proof, and verification
- Using the numbers from our benchmark, we simulated what the numbers would be for a different number of users and servers
Latency vs. number of users

- 800 servers
- 3 servers per chain
Latency vs. number of servers

- 2M users
- 3 servers per chain
Summary

- XRD is a scalable messaging system with cryptographic privacy
- Latency decreases with the square root of the number of the servers
- 78 second latency for 2M users and 800 servers
Backup
Future Work

- Increasing XRD speed
- Protecting against active attacks using a different method than zero-knowledge proofs
- Realistic evaluation