Poirot: Private Contact Summary Aggregation

Chenghong Wang\textsuperscript{1}, David Pujol\textsuperscript{1}, Yanping Zhang\textsuperscript{1}, Johes Bater\textsuperscript{1}, Matthew Lentz\textsuperscript{1,4}, Ashwin Machanavajjhala\textsuperscript{1}, Kartik Nayak\textsuperscript{1}, Lavanya Vasudevan\textsuperscript{1,2,3} and Jun Yang\textsuperscript{1}

\textsuperscript{1}Computer Science Department, Duke University
\textsuperscript{2}Department of Family Medicine and Community Health, Duke University
\textsuperscript{3}Duke Global Health Institute
\textsuperscript{4}VMware Research
Poirot: In a Nutshell

Physical distancing between individuals is key to preventing the spread of a disease such as COVID-19

We want:

• Functionality: Measure physical interactions through “contact events”
• Privacy: Ensure that the resulting data cannot be linked back to an individual
How will Poirot be used?

Provide actionable information to individual users and decision makers in a privacy-preserving manner.
**Threat Model**

- **Users**
  - semi-honest
  - Learns their own #contacts with locations and times plus differentially-private aggregated statistics

- **Admins**
  - untrusted
  - Untrusted administrators: learn differentially-private aggregate statistics

- **Auth Server**
  - semi-honest
  - Learns the set of participating users

- **Poirot Server N**
  - semi-honest, assume non-collusion
  - Learns the set of participating users + some metadata.
Poirot Design

Collection

Permission
-1 Month
Auth Server >> Poirrot Server
ID, {Attr}
Code
(1) Code
(2) Poirrot Server
(3) Code
(4) Poirrot Server
(5)

Discover Contacts
Cont.
Advertisements
< [Time, Loc, Dist>
< [Time, Loc, Dist>

Upload Summary
Daily
# of contacts per location bin
Summary
Period (1hr): 9/22 10:00
Data: [2, 0, 5, …]

Processing
What is the average number of contacts for <location, time> pair?

Poirrot Server 1...
Poirrot Server N

Users
Admins

MPC
Noise
f

Individuals
Administrators

Average Contacts Per Day
By Location
Dining Hall
High
Med
Low

Collection Processing Usage (University Example)
Poirot Design -> Data Collection-> Private Permissioning

~1 Month

Auth Server

ID, {Attr}

Code

Poirot Server

Code

{Token T}
Poirot Design -> Data Collection -> Discover Contacts
Poirot Design -> Data Collection -> Upload Summary

Servers only learn metadata about contact summaries
Multiparty Computation (MPC) allows computing on secret-shared data, Differential Privacy ensures released statistics do not reveal individual’s data.
Poirot Design-> Data Processing

Raw statistics

Distorted statistics
Poirot-> Evaluation-Accuracy

• Dataset: Copenhagen Network Study dataset
Poirot-> Evaluation-Performance

<table>
<thead>
<tr>
<th>Case</th>
<th># of Locations</th>
<th>Time</th>
<th>User Population</th>
<th>App execution time (ms)</th>
<th>Server execution time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duke</td>
<td>256</td>
<td>24</td>
<td>20K</td>
<td>366.1</td>
<td>94.4</td>
</tr>
<tr>
<td>NC</td>
<td>100</td>
<td>1</td>
<td>10M</td>
<td>6.0</td>
<td>776.1</td>
</tr>
<tr>
<td>Copenhagen</td>
<td>1</td>
<td>24</td>
<td>705</td>
<td>1.68</td>
<td>0.015</td>
</tr>
</tbody>
</table>
Conclusion

• Provide accurate information about physical interactions.
• Guarantees individual’s contact privacy
• Our system scales to large, realistic deployment scenarios.

https://poirot.cs.duke.edu/