Cloud-Enabled Big-Data Analytics: The Sky's the Limit

Managing Trust in the Cloud

Kinan Dak Albab       Rawane Issa

PhD Students, Boston University
Hariri Institute for Computing
Software Application & Innovation Lab
https://multiparty.org/
Trust and Computer Science

1. Public to trust correctness, security, and privacy of our solutions.
2. Data contributors to trust our algorithms and tools.
3. Policy makers to trust correctness and accountability of our algorithms and models.
4. Computer Scientists and researches to trust the community and each other.
5. Business leaders, service providers, and resource providers to trust each other.
Can we compute the salary gap among genders and ethnicities without violating companies privacy concerns and revealing employee's records?
During a global “cyber-attack” like the WannaCry ransomware attack, can we tell how far (in a network) we are from an infected PC without revealing if we are infected ourselves?
Can we create a dating app which checks if two people are interested in each other but does not reveal to any party possible unreturned romantic interest?
Can we conduct statistical analysis on medical data without revealing patient’s privacy?
YES we can.
Trust Management in the Cloud is Critical!

1. Sensitive / Valuable ("Big") Data
2. Multiple computing parties / service providers
3. Multiple maintainers / managers of machines
4. Security of Data and Service
Who to trust?

**Data Contributor:**

1. Algorithm / Code
2. Service providers
3. Cloud providers
4. Code delivery mechanisms and CDNs
5. Computing stack

**Service Provider:**

1. Developers / Software
2. Data Contributor
3. Cloud Providers
4. Code delivery
5. Computing stack
Everybody has to trust everybody else!
Cryptography is as much a Social Science as it is a Mathematical/Computer Science.
Traditional “Solutions”

1. Trust blindly
2. Contract, regulation and policy
3. Anonymization
Secure Multi-Party Computation (MPC)

Sharing knowledge without sharing data*

\[ K = f (\text{Confidential}, \text{Confidential}, \text{Confidential}, \ldots) \]

* under certain security assumptions
Sum all numbers without revealing them
Compute secret shares

 Shares of “50”
Send shares to corresponding parties

One share from each secret
Sum received shares

-572
1173
-441
Sum resulting three shares

\[\begin{align*}
-572 \\
1173 \\
-441
\end{align*}\]

\[\text{Sum} = 160\]
Pay Equity

1. Use MPC to aggregate salary and employee information from many companies in the Boston area.
2. Analyse the aggregate to study pay equity between genders and ethnicities.
Contributor A

Service Provider (e.g., BU)
(web server/database)

Analyst (e.g., BWWC)
(client running web browser)

Contributor B

---

true data A + random mask A = masked data A

masked data A + masked data B = masked aggregate data

Analyst does not access this data

random mask A + random mask B = aggregate mask

true aggregate data

---

Public-key Encrypted Storage
only Analyst has key;
no one else (including the S.P.)
can read the content of this data
## Current Applications

<table>
<thead>
<tr>
<th>Partner(s)</th>
<th>Application(s)</th>
<th>Stage</th>
</tr>
</thead>
</table>
| **CITY of BOSTON**                            | - Secure aggregation of tabular and multiple choice response data from multiple companies  
- In upcoming deployment: correlations between multiple choice responses | • Deployed twice (5/2015, 6/2016)  
• Actual result data published in report by BWWC (1/2017)  
• Upcoming deployment (9/2017)            |
| Greater Boston Chamber of Commerce          | • Secure aggregation and analysis of tabular data from multiple companies | • Implementation ready  
• Deployment planned (early 2018)                                                        |
| Eastern Bank                                 | • Secure aggregation of tabular data and multiple choice response data from multiple banking organizations | • Implementation ready  
• Awaiting deployment timeline                                                            |
| MBAs Massachusetts Bankers Association       | • Machine learning over sensitive consumer data subject to consumer-specified policies to enhance web services (e.g., route recommendation) | • Prototypes under development                                                            |
| Honda Research Institute                     | • Enhancement of mobile health intervention apps and data sharing tools with secure aggregate to add value for users and clinicians/researchers | • Application features under development  
• Deployment planned (2018)                                                               |
| Hey.Charlie                                  | • Secure aggregation of tabular data from multiple organizations | • In discussions with stakeholders                                                        |
| Pardee School Initiative on Forced Migration and Human Trafficking |                                                                                                                                                                                                          |                                                                                           |
## Current & Past collaborators

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinan Dak Albab</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Ran Canetti</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Azer Bestavros</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Eric Dunton</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Ben Getchell</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Kyle Holzinger</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Rawane Issa</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Frederick Jansen</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Rose Kelly</td>
<td>UMass</td>
<td></td>
</tr>
<tr>
<td>Andrei Lapets</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Shannon Roberts</td>
<td>UMass</td>
<td></td>
</tr>
<tr>
<td>Malte Schwarzkopf</td>
<td>MIT</td>
<td></td>
</tr>
<tr>
<td>Mayank Varia</td>
<td>BU</td>
<td></td>
</tr>
<tr>
<td>Nikolaj Volgushev</td>
<td>Alexandra Institute</td>
<td></td>
</tr>
<tr>
<td>Jacqueline You</td>
<td>BU</td>
<td></td>
</tr>
</tbody>
</table>
Thank you!

babman@bu.edu  ra1issa@bu.edu