Towards Regulating Large-Scale Multi-Enterprise Environments with Confidentiality Guarantees

Mohammad Javad Amiri
University of Pennsylvania

Amr El Abbadi
UC Santa Barbara

Divy Agrawal
UC Santa Barbara

Boon Thau Loo
U Pennsylvania

Tristan Allard
Univ Rennes
A set of known **mutually distrustful** entities
Multi-Enterprise Environments

Manufacturer

Supplier

Logistics Provider

Transportation

Hospitals

Supply Chain Management

Multi-Platform Crowdworking Environment

- Require collaboration among a set of mutually distrustful entities
- Internal and global regulations need to be enforced
- The confidentiality of data is paramount
• Supports Local and global transactions
  • Global transactions are visible to all enterprises
  • Local transactions of each enterprise are confidential

What if a subset of enterprises are involved in a confidential collaboration?
Qanaat: Confidential Collaborations across Enterprises [VLDB’22]

• A hierarchical data model consisting of a set of data collections
• Operational primitives
  • **Write**: transactions of $d_X$ write only on the records of $d_X$
  • **Read**: transactions of $d_X$ can read the records of $d_Y$ if $X \subseteq Y$ (order-dependency)

What if the infrastructure includes malicious nodes?
Confidential Data Leakage Prevention

• Malicious nodes can violate data confidentiality
  • leaking requests, replies, or data stored and processed
• Privacy firewall mechanism
  • Separates ordering node from execution nodes
    • $3f + 1$ ordering nodes and $2g + 1$ execution nodes
      • Assuming $f$ faulty ordering and $g$ faulty execution nodes
  • Adds a privacy firewall in between
    • Consists of a set of $h + 1$ rows of $h + 1$ filters ($h$ faulty filters)
  • Network configuration physically restricts communication paths between ordering nodes, filters, and execution nodes
  • A malicious node can either access confidential data or communicate freely with clients but not both

Regulating Large-Scale Multi-Enterprise Environments
Data Verifiability

• Qanaat supports **Read** and **Write** operations
  • Write: the same data collection
  • Read: superset data collections (order-dependency)

What if we need to verify private data?
Crowdworking Environment

- Envisioned as key technological components of the future of work
“Whereas universal and lasting peace can be established only if it is based upon social justice; ... for example, by the regulation of the hours of work ...”

preamble of the constitution of the International Labor Organization
[Commission on International Labor Legislation, 1919]
FLSA: Total work hours of a worker per week may not exceed 40 hours

In California, Assembly Bill 5 (AB5) entitles workers to greater labor protections, such as minimum wage laws, sick leave, and unemployment and workers' compensation benefits.

CA Proposition 22 imposes its set of regulations, e.g., requires a worker to work at least 25 hours per week to qualify for healthcare subsidies.
There is more than one platform …

- Workers often work on several platforms
- Requesters submit tasks on multiple platforms
Privacy of Participants

• No participant obtains or infers any information beyond what is needed
  • A driver who works for both Uber and Lyft, does not want either of them know that she works for the other.

• How to enforce regulations?
  • Reconcile transparency with privacy
Problems

- Guarantee the compliance of crowdworking platforms with regulations
- Local (per platform) regulations exist: maximum driving time per day
- Transparent and Privacy-preserving regulation enforcement
- Collaboration among mutually distrustful platforms
- Enforcement of global regulations
- Complex tasks that may need multiple contributions
Our Vision for Future Regulated Multi-Enterprise Systems [WWW’21]

- **Goal:** Enforce regulations on multi-platform crowdworking environments while preserving privacy

- Three main design dimensions
  - **D1:** Type of supported regulations
    - Express as SQL constraints over a universal table
    - e.g., aggregate or not/ has join or not
    - Verifiable vs. enforceable
  - **D2:** Privacy guarantees given to participants
    - Pluggable disclosures (received/involved)
  - **D3:** Architecture of the system
    - Centralized registration authority
    - Decentralized state management
A Simple Token-Based System

- Inspired by e-cash systems, regulations are implemented by managing budgets per participant
- Lightweight, single-use, and anonymous tokens

The registration authority refreshes participants tokens periodically

- **GENERATE**: initializing the budgets and refilling them
  - Enforceable and Verifiable tokens
- **SPEND**: spending portions of the budgets
- **PROVE**: providing proof for verifiable regulations to a third party
- **CHECK**: checking whether a given spending is allowed or not
- **ALERT**: reporting dubious spending
Execution Sequence

Tasks:
Internal
Cross-Platform

Transactions:
Submission
Claim
Verification
Reaching Consensus [SIGMOD’21]

Local Consensus: pluggable and depends on the failure model of nodes

Cross-Platform Consensus: Among the involved platforms

Global Consensus: Requires the participation of all platforms

<table>
<thead>
<tr>
<th>Transaction/Task</th>
<th>Internal</th>
<th>Cross-Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submission</td>
<td>Local</td>
<td>Cross-Platform</td>
</tr>
<tr>
<td>Claim</td>
<td>Local</td>
<td>Cross-Platform</td>
</tr>
<tr>
<td>Verification</td>
<td>Global</td>
<td>Global</td>
</tr>
</tbody>
</table>
Enforcing regulations on a set of mutually distrustful enterprises

Preserving the privacy of participants
- Hierarchical data model

Confidential data leakage prevention
- Privacy firewall mechanism

Collaboration among enterprises
- Distributed consensus protocols

Expressing and modeling regulations
- SQL constraints over a universal table

Private data verification
- Token-based systems or zero-knowledge proofs
Women. Life. Freedom
#MahsaAmini
Questions?

I’m on (academic) job market this year!