Cloud Intelligence
A VLDB 2012 Workshop

An Anonymization-Based Approach For Privacy-Preserving BPaaS

Mehdi Bentounsi, S. Benbernou, M.J. Atallah and C. Deme
Université Paris Descartes, France
mehdi.bentounsi@parisdescartes.fr

August 31, 2012. Istanbul, Turkey
Outline

✓ Introduction and motivation
✓ Overview of BPaaS by example
✓ Formal model
✓ Privacy-preserving BPaaS
Cloud computing

- Outsourcing of IT infrastructure to providers.

- Cloud services: IaaS, PaaS, and SaaS.

- Clients: SME... to take advantage of the economic model.
Cloud Federation [Papazoglou, SSAE’12]

Comprises an assembly of **IT resources** spanning various cloud providers and **diverse clouds** where each part can be independently selected by the developer and then **put together**, on demand, by the cloud provisioning system over the Internet.
Business Process as a Service

• Why?
  – Ease of deployment and fast provisioning.
  – Pay-per-use (reduce costs).
  – Add more capacity at peak demand (elasticity).
  – Many providers: Amazon, Google, Salesforce.com, etc.

• The next step forward in the evolution of cloud computing. [Bittman, Gartner ‘11]
Motivation

• Large **multi-tenant** repositories of business process models in the Cloud.
• Duplicate fragments (clones) appear in repositories. [Uba,BPM’11]
• Reusing process fragments as **Web services** to develop new **process-based service compositions**. [Schumm,ICWE’12]

✓ Reduce development complexity!
Requirement

• Business processes are direct generators of revenue and key enablers of strategy.

• Innovations in business process design represent a significant proportion of costs.

✔ Reusing process fragments in the Cloud must preserve the business activities of their owners!
Outline

- Introduction and motivation
- Overview of BPaaS by example
- Formal model
- Privacy-preserving Business Process Outsourcing
Decomposition of processes  [Khalaf,ICWS’06]

1. A business process is decomposed into a set of process fragments suitable for re-use.

Hospital Business Process
Identification of fragments [Caetano, BIS’10] and [Huang, WI’10]

1. A business process is decomposed into a set of process fragments suitable for re-use.

2. Each process fragment is identified and made available to be reused as a Web Service.
Development of processes [Schumm, ICWE’12]

3. A process-based service compositions is developed by reusing process fragments deployed in the Cloud.
Problem

- An adversary (a curious) can discover the provenance of the reused process fragments.

- Can infer connections between end-users and organizations that outsource fragments to the Cloud.

✓ No related work!
Outline

✓ Introduction and motivation
✓ Overview of BPaaS by example
✓ Formal model
✓ Privacy-preserving BPaaS
Formal model

- **Business Process**: business graph. [Beeri, VLDB’06]
- **Process Fragment**: business subgraph.

- **BPaaS**: a finite set of business processes.
- **Reusing Function.**
Let $P$ a business process to be developed in the BPaaS

For each Process Fragment $F$ in $P$

Verify if $F$ exists in the BPaaS

$P'$ a process-based service compositions developed in the BPaaS

Reuse $F$ to develop $P$
Outline

✓ Introduction and motivation
✓ Overview of BPaaS by example
✓ Formal model
✓ Privacy-preserving BPaaS
Privacy-Preserving BPaaS

- **First attempt**: Hide the process fragment URL and owner ID.

- **Example**: Let us consider a BPaaS with:
  - \( L \) organizations using it,
  - \( L \) process fragments implementing an activity \( A \), and
  - An adversary may desire to reuse a process fragment implementing \( A \).

- ✓ A curious can make link with external informations (list of organizations using the BPaaS) and discover the provenance of the process fragment.
Not preserve the privacy!

The Probability to discover the owner of a process fragment in the worst case is 1.
Anonymous Views

- **View on BPaaS:**
  A set of process fragments having the same Activity (called clones).

- **Anonymous View on BPaaS:**
  View on BPaaS having at most K clones. [Samarati, TKDE’01]

✓ Make it hard for an adversary to know the provenance of a reused process fragment!
Anonyfrag

• **Quasi-identifier fragments**: A set \( A = \{A_1, A_2, ..., A_i\} \) of Abstract Fragments (activities) in BPaaS \( S \) is called Quasi-identifier fragments or QIF, if these AFs can be used to learn the tenants business activities.

• **Kl−anonyfrag requirement**: Kl−anonyfrag requirement is each view \( V_a \) on BPaaS w.r.t. \( a \in QIF \) must contain at most \( K \) clones.

• **Kl−anonyfrag**: Given a BPaaS \( S \) used by \( l \) tenants; and an AF \( a \) with a set of at most \( K \) business subgraphs or clones in \( S \). An adversary knows that it exists at most \( K \) clones implementing \( a \) are deployed in \( S \); and doesn’t know:
  - Exactly the number of tenants that own the \( K \) business subgraphs among \( l \) tenants.
  - Which tenants exactly have deployed the PF in \( S \).
Example: $2^4$-anonyfrag

The Probability to discover the owner of a process fragment in the worst case is $2/4 = 0.5$!
Summary

- Business process outsourcing to Business Process as a Service to take advantage of the Cloud computing model.

- Reusing process fragments to easily develop process-based service compositions and privacy risks.

- Formal model of BPaaS.

- Anonymous View on BPaaS for privacy-preserving.
Ongoing and future work

- Implementation of **Anonyfrag-platform** with real data sets. [Al-Masri,ICCCN’07] and [Al-Masri,WWW’08]

- Use **Diverse Views** to treat process fragments availability issues.

- **Monitoring** of Security and Privacy (detect attackers based on tenants’ feedbacks).
Ongoing and future work

- Implementation of **Anonyfrag-platform** with real data sets. [Al-Masri,ICCCN’07] and [Al-Masri,WWW’08]

- Use **Diverse Views** to treat process fragments availability issues.

- **Monitoring** of Security and Privacy (detect attackers based on tenants’ feedbacks).
Acknowledgment

• This work is supported by SOMONE and French Association Nationale de la Recherche et de la Technologie under CIFRE contract 1169/2010.

• Portions of this work were supported by Université Paris Descartes collaborative Project Brainshare.

• Portions of this work were supported by National Science Foundation Grants CNS-0915436, CNS-0913875, Science and Technology Center CCF-0939370; by an NPRP grant from the Qatar National Research Fund; by Grant FA9550-09-1-0223 from the Air Force Office of Scientific Research; and by sponsors of the Center for Education and Research in Information Assurance and Security.
References

Thank you!