Privacy-Enhancing Technologies: Anonymous Credentials and Pseudonym Systems

Anja Lehmann
IBM Research – Zurich
ROADMAP

- Anonymous Credentials
  - privacy-preserving (user) authentication

- Pseudonym Systems
  - privacy-preserving & auditable data exchange
Strong User Authentication

- Strong (user) authentication via certificates / attribute-based credentials
  - Many European countries have or will introduce eID cards
  - Desirable for security, but detrimental for privacy
  - Existing schemes require full information disclosure & user is linkable in all transactions

→ This is a privacy and security problem!
  - Linkability enables tracking & profiling of users
  - Acquired personal data requires protection
Strong & Privacy-Preserving User Authentication

- Envisioned by Chaum in 1981, first full scheme by Camenisch & Lysyanskaya in 2001
  - User can **selectively disclose** each attribute
  - User can prove **predicates over the attributes**, e.g., “I'm over 18”
  - **Unlinkable** authentication as default, linkability as an option
Envisioned by Chaum in 1981, first full scheme by Camenisch & Lysyanskaya in 2001

- User can **selectively disclose** each attribute
- User can prove **predicates over the attributes**, e.g., “I'm over 18”
- **Unlinkable** authentication as default, linkability as an option
Privacy-Enhancing Credentials | Existing Solutions

- Most prominent core-credential/signature schemes:

  **Identity Mixer (IBM)**
  - Multi-use credentials
  - Zero-Knowledge Proofs
  - Strong RSA, pairings (LRSW, qSDH)

  **U-Prove (Microsoft)**
  - One-time use credentials (multi-use via batch-issuance)
  - Blind Signatures
    - RSA, DL
Privacy-Enhancing Credentials | Extended Features

- Many more extensions & properties:
  - Revocation, multi-credential proofs, issuance with carry-over attributes, conditional disclosure, „symmetric“ credentials

- Various cryptographic realizations

---

Hard to deploy: complex, multitude of protocols, features, inconsistent naming, ...
Privacy-Enhancing Credentials | Generic Framework

- Technology-independent & „easy-to-use“ framework
  - Comprehensive & standardized language framework
  - Technology-agnostic credential & policy handling on top of crypto engine
  - Generic, automated crypto engine

www.zurich.ibm.com/idemix
Privacy-Enhancing Credentials | New Applications

- V2X communication (vehicles (V2V) and infrastructure (V2I))
  - Security needs: authentication & privacy
  - Current approach: pseudonym CA
  - Privacy-credentials fit perfectly! (almost)

- Hardware-based device/user attestation (DAA)
  - Draft for FIDO standard
  - FIDO ("Fast IDentity Online") Alliance
    = industry consortium developing standardized strong user/device authentication

- Blockchain: “eternal” and public transaction ledger
  - Privacy credentials needed to avoid privacy nightmare
  - Identity Mixer being integrated into Hyperledger Fabric
  - IBM joined the Sovrin Foundation – decentralized digital identity network
ROADMAP

- Anonymous Credentials
  - privacy-preserving (user) authentication

- Pseudonym Systems
  - privacy-preserving & auditable data exchange

Pseudonym System | Motivation

How to exchange and correlate (pseudonymous) data?
– E.g., eHealth records, social security system
– User-centric conversion inconvenient & unreliable
**Pseudonym System | Globally Unique Pseudonyms**

- Data gets associated with globally unique identifiers / pseudonyms
  - E.g., social security number in US, Belgium, Sweden, ...

- Unique identifiers are **security & privacy risk**
  - no control about data exchange & usage
  - if associated data is lost, all pieces can be linked together
  - linkability of data allows re-identification of “anonymized” data (e.g. Netflix challenge)
Pseudonym System | Local Pseudonyms & Trusted Converter

- User data is associated with random looking local identifiers – the *pseudonyms*
- Only central entity – the *converter* – can link & convert pseudonyms

+ control about data exchange
+ if records are lost, pieces cannot be linked together

<table>
<thead>
<tr>
<th>Main ID</th>
<th>Doctor A</th>
<th>Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice.1210</td>
<td>Hba02</td>
<td>7twnG</td>
</tr>
<tr>
<td>Bob.0411</td>
<td>P89dy</td>
<td>ML3m5</td>
</tr>
<tr>
<td>Carol.2503</td>
<td>912uj</td>
<td>sD7Ab</td>
</tr>
</tbody>
</table>

Doctor A

- ID: Hba02
- ID: P89dy
- ID: 912uj

Hospital

- ID: ML3m5
- ID: sD7Ab
- ID: y2B4m

new Japan eID / social security number system (?)
User data is associated with random looking local identifiers – the *pseudonyms*.

Only central entity – the *converter* – can link & convert pseudonyms.

+ control about data exchange
+ if records are lost, pieces cannot be linked together
+ converter can provide audit logs to users (*GDPR*-requirement)
  - converter learns all request & knows all correlations
Pseudonym System | Local Pseudonyms & Oblivious Converter

- User data is associated with random looking local identifiers – the *pseudonyms*
- Only central entity – the *converter* – can link & convert pseudonyms

User Portal for Bob.0411

Doctor A → Hospital. 02/26/2017...

Converter

Doctor A

<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hba02</td>
<td></td>
</tr>
<tr>
<td>P89dy</td>
<td></td>
</tr>
<tr>
<td>912uj</td>
<td></td>
</tr>
</tbody>
</table>

Hospital

<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML3m5</td>
<td></td>
</tr>
<tr>
<td>sD7Ab</td>
<td></td>
</tr>
<tr>
<td>y2B4m</td>
<td></td>
</tr>
</tbody>
</table>

+ control about data exchange
+ if records are lost, pieces cannot be linked together
+ converter can provide audit logs to users *(GDPR-requirement)*
  - converter learns all requests & knows all correlations
(Un)linkable Pseudonyms | Pseudonym Generation

- User, converter & server jointly derive pseudonyms from unique identifiers

- [CL15] generation triggered by converter, knows unique IDs
- [CL17] oblivious pseudonym generation triggered by user
(Un)linkable Pseudonyms | Pseudonym Conversion

- Only converter can link & convert pseudonyms, but does so in a blind way
(Un)linkable Pseudonyms | Consistency

- pseudonym generation is deterministic & consistent with blind conversion
(Un)linkable Pseudonyms | Consistency

- pseudonym conversions are transitive, unlinkable data can be aggregated
(Un)linkable Pseudonyms | User Audits

- [CL17] every pseudonym conversion triggers blind generation of audit log entry

<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hba02</td>
<td></td>
</tr>
<tr>
<td>P89dy</td>
<td></td>
</tr>
<tr>
<td>912uj</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML3m5</td>
<td></td>
</tr>
<tr>
<td>sD7Ab</td>
<td></td>
</tr>
<tr>
<td>y2B4m</td>
<td></td>
</tr>
</tbody>
</table>

Audit Bulletin Board

Converter

Doctor A

Hospital

Unique ID

Bob.0411

Doctor A → Hospital. 02/26/2017
(Un)linkable Pseudonyms | Security Model

- Universal composability (UC) model convenient & **simple** for privacy-preserving systems

\[\begin{align*}
\text{NymGen, } S_A & \quad \text{NymGen, } S_B \\
U_i & \quad \text{NymGen, } S_B \\
\text{NymGen, } S_A, S_B & \quad \text{NymGen, } S_A \\
\text{NymGen, } nym_{i,A} & \quad \text{NymGen, } nym_{i,B} \\
\text{NymGen, } nym_{i,A} & \quad \text{NymGen, } nym_{i,B} \\
\text{NymGen, } nym_{i,A}, S_A, qid & \quad \text{NymGen, } nym_{i,B}, S_A, qid \\
\text{NymGen, } nym_{i,A}, S_A, qid & \quad \text{NymGen, } nym_{i,B}, S_A, qid \\
\end{align*}\]
Our Protocol

- high-level idea of convertible pseudonyms
- adding (efficient) auditability
- security against active adversaries
High-level Idea | Pseudonym Generation

Core Idea
Generation: $X$ blindly computes $nym_{i,A} \leftarrow PRF(k,uid_i)^{x_A}$

1. $X$ and $U_i$ jointly compute $z_i \leftarrow PRF(k,uid_i)$
2. $U_i$ encrypts $z_i$ for $S_A$
   
   $C_{nym} \leftarrow Enc(pk_A,z_i)$

3. $X$ blindly computes $nym_{i,A}$
   
   $C'_{nym} \leftarrow C_{nym}^{x_A}$

4. $S_A$ decrypts pseudonym
   
   $nym_{i,A} \leftarrow Dec(sk_A,C'_{nym})$
   
   $nym_{i,A} = PRF(k,uid_i)^{x_A}$
High-level Idea | Pseudonym Conversion

Core Idea
Generation: $X$ blindly computes $nym_{i,A} \leftarrow \text{PRF}(k, uid_i)^{x_A}$
Conversion: $X$ blindly computes $nym_{i,B} \leftarrow nym_{i,A}^{x_B / x_A}$

[1] $S_A$ encrypts $nym_{i,A}$ under $S_B$'s key
$C \leftarrow \text{Enc}(pk_B, nym_{i,A})$

[2] $X$ blindly transforms encrypted pseudonym
$C' \leftarrow C^\Delta$ with $\Delta = x_B / x_A$
$C' = \text{Enc}(pk_B, nym_{i,A})^{x_B / x_A}$
$= \text{Enc}(pk_B, \text{PRF}(k, uid_i)^{x_A})^{x_B / x_A}$
$= \text{Enc}(pk_B, \text{PRF}(k, uid_i)^{x_B})$
$= \text{Enc}(pk_B, nym_{i,B})$

[3] $S_B$ decrypts converted pseudonym
$nym_{i,B} \leftarrow \text{Dec}(sk_B, C')$
$nym_{i,B} = \text{PRF}(k, uid_i)^{x_B}$
High-level Idea | Overview

Generation

Conversion

Converter \( X \)  

NymRequest

NymResponse → Server A

nym_{i,A}

ConvRequest → Server A

nym_{i,A}

ConvResponse → Server B

nym_{i,B}
High-level Idea | Adding Auditability

Decrypt all audit ciphertexts: \( info \leftarrow \text{Dec}(usk, C^*) \) ?

\[
\text{NymRequest, } upk' \quad \rightarrow \quad \text{Converter } \chi
\]

\[
\text{NymResponse, } upk' \quad \rightarrow \quad \text{Server A}
\]

\[
nym_{i,A, upk}' \quad \rightarrow \quad \text{Server A}
\]

\[

\begin{align*}
\text{C}^* & \leftarrow \text{Enc}(upk', info) \\
\text{ConvRequest, } upk'' & \rightarrow \quad \text{Server A} \\
\text{ConvResponse, } upk''' & \rightarrow \quad \text{Server B}
\end{align*}
\]

\[
nym_{i,B, upk}''' \quad \rightarrow \quad \text{Server B}
\]

upk is randomizable encryption key
\( upk' \leftarrow \text{RAND}(upk) \)
High-level Idea | Adding Efficient Auditability (via Audit Tags)

decrypt ciphertext for $T_A$: 
$\text{info} \leftarrow \text{Dec(usk,}C^*)$

$\text{NymRequest, upk', }C_T$

$\text{Converter }X$

$\text{NymResponse, upk', }C_T$

$\text{Server A}$

$nym_{i,A}, \text{upk', }T_A$

$\text{T}_A \leftarrow \text{Dec(sk}_A, C_T)$

$\text{Generation}$

$\text{Conversion}$

$\text{Audit Bulletin Board}$

$T_A, C^*$

$\ldots$

$\text{C'} \leftarrow \text{Enc(upk'', info)}$

$\text{ConvRequest, upk'', }T_A$

$\text{Server A}$

$nym_{i,A}, \text{upk', }T_A$

$\text{ConvResponse, upk'''}$

$\text{Server B}$

$nym_{i,B}, \text{upk'''}$
High-level Idea | Adding *Efficient* Auditability (via Audit Tags)

Decrypt ciphertext for $T_A$:
\[
info \leftarrow \text{Dec}(\text{usk}, C^*)
\]

Get new audit tags for $T_A$:
\[
T_B \leftarrow \text{Dec}(\text{usk}, C^*_{TB})
\]

**Generation**

**Conversion**

$C^* \leftarrow \text{Enc}(upk'', info)$

$C^*_{TB} \leftarrow \text{Enc}(upk'', T_B)$ ...

for random $T_B$
High-level Idea | Adding Efficient Auditability (via Audit Tags)

decrypt ciphertext for $T_A$:

$$\text{info} \leftarrow \text{Dec}(\text{usk}, C^*)$$

get new audit tags for $T_A$:

$$T_B \leftarrow \text{Dec}(\text{usk}, C^*_{TB})$$

$$T'_A \leftarrow \text{Dec}(\text{usk}, C^*_{TA})$$

NymRequest, upk', $C_T$

Server A

NymResponse, upk', $C_T$

Server A

ConvRequest, upk'', $T_A$, $C^*_{TA}$

Server A

ConvResponse, upk'''

Server B

$T_B \leftarrow \text{Dec}(\text{usk}, C^*_{TB})$

$C^*_{TB} \leftarrow \text{Enc}(\text{upk''', } T_B) \ldots$ for random $T_B$

$T_A \leftarrow \text{Dec}(\text{sk}_A, C_T)$

Tag Chain:

$$C^* \leftarrow \text{Enc}(\text{upk''', } \text{info})$$

Converter $\mathcal{X}$

Audit Bulletin Board

$T_A$, $C^*$

Tag Chain:

$T_A$, $C^*_{TB}$

$T_A$, $C^*_{TA}$
High-level Idea | Security against Active Adversaries

decrypt ciphertext for \( T_A \):
\[ \text{info} \leftarrow \text{Dec}(\text{usk}, C^*) \]

get new audit tags for \( T_A \):
\[ T_B \leftarrow \text{Dec}(\text{usk}, C^*_T B) \]
\[ T'_A \leftarrow \text{Dec}(\text{usk}, C^*_T A) \]

Signature scheme for homomorphic encodings
(Un)linkable & Auditable Pseudonyms | Security & Efficiency

- Provably secure construction in the Universal Composability (UC) framework based on
  - homomorphic encryption scheme (ElGamal encryption)
  - homomorphic encryption scheme with re-randomizable public keys (ElGamal-based)
  - oblivious pseudorandom function with committed outputs (based on Dodis-Yampolskiy-PRF)
  - signature scheme for homomorphic encoding functions (based on Groth signature scheme)
  - zero-knowledge proofs (Fiat-Shamir NIZKs)
  - commitment scheme (ElGamal based)
  - DDH

- Secure against actively corrupt users & servers, and honest-but-curious converter
  - (w/o audits even fully corrupt converter [CL15])

- Concrete instantiation ~50ms computational time per party for conversion
Summary

- Mature privacy-enhancing technologies exist – privacy and functionality are not exclusive
- Linkability crucial for utility, but also weakens privacy
  - Paradigm shift: unlinkability per default, linkability only when necessary
  - Controlled, selective linkability & enforced transparency
- GDPR creates a great practical demand for privacy-preserving mechanisms – data minimisation, consent enforcement, auditability, ...
- „Crypto Magic“ needs education and dissemination!

Thanks! Questions?
anj@zurich.ibm.com