

Privacy Enhancing Technologies

Mohamed Layouni, Phd student
Crypto & Quantum Info Lab
McGill University

CQIL

- Crypto & Quantum Info Lab of McGill Univ.:
 - 3 professors
 - 12 students (7PhD + 5 MSc)
- Wide research spectrum with special focus areas:
 - Quantum cryptography
 - Privacy-enhancing technologies
- My thesis supervisors:
 - Dr. Stefan Brands (principal supervisor)
 - Dr. Claude Crepeau

Outline

- Motivations
- Different kinds of anonymity
- Anonymity at the connection-level
- Anonymity at the application-level
- Sample applications
- Research directions

Motivations

- Users are concerned about their privacy (e.g., remaining anonymous, hiding queries to a database...)
 - Service providers are concerned about their own privacy (competitive data) towards central parties
 - Governments and organizations are concerned about possible abuses resulting from anonymity.
- ⇒ Research in privacy-enhancing technologies aims to satisfy all these seemingly **conflicting interests**.

Forms of Anonymity

- Anonymity is a primary requirement in many applications
 - E-voting
 - E-health
 - ...
- *Unconditional* is better than *computational* anonymity:
 - Cannot be broken via backdoors or retroactively
- In some applications, anonymity might be abused; here, we need mechanisms for *controlled* anonymity
 - Detect fraudulent behavior
 - Trace fraudsters
 - Blacklist fraudsters
 - Contain fraud

Anonymity at different levels

Anonymity can be provided at:

- The *data transport* level
 - E.g., preventing receiver from learning IP address
- The *protocol* level
 - Ensure that no tracing is possible by analyzing the data sent around as part of the protocol flow

Our research in McGill on privacy-enhancing technologies focuses on privacy at the protocol level, notably e-voting and e-health.

Data-transport anonymity

A connection is anonymous if:

- It hides the users' network identity information (e.g., IP address) from the receiving computer,
- It prevents all other entities on the network from telling that a particular initiator is communicating with a particular responder

Data-transport anonymity

(contd.)

Threat Model:

- The adversary is modelled as a number of collaborating local observers of the network.
- Practical solutions for anonymous communication rely on the assumption that there exists a number of entities that can be trusted not to collaborate.

Recipient vs. Sender Anonymity

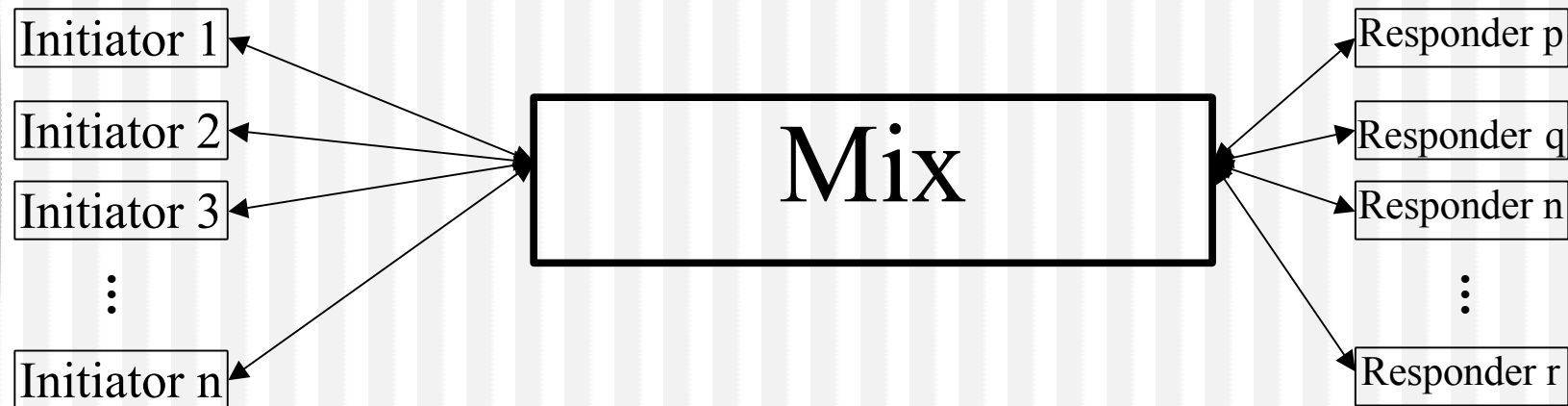
A variety of techniques to achieve anonymity:

- Recipient Anonymity
 - Broadcast networks + public key encryption
 - Bulletin boards...
- Sender Anonymity
 - Mix-based networks
 - Padding, inserting dummy traffic, latency...
 - reordering, multiplexing,
 - encryption, substitution, compression...

Data-transport anonymity: sample solutions

Mix Networks:

Invented by Chaum (1982)

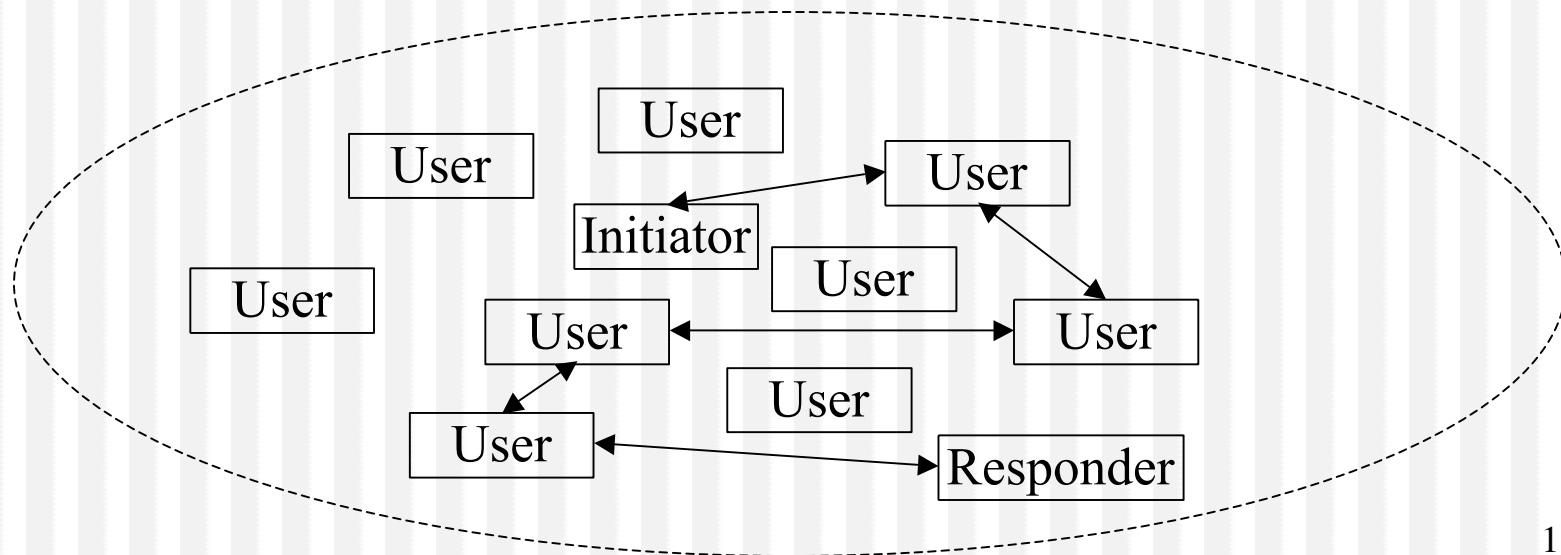


- ✓ Order is hidden by padding, reordering, and delaying traffic...
- ✓ Mix networks may be inappropriate for real-time applications

Data-transport anonymity: sample solutions (contd.)

Crowds-like anonymous connections :

- Random path from the initiator to the responder.
- Sender anonymity (users ignore their positions on the path)



Data-transport anonymity: sample solutions (contd.)

Onion Routing:

$\text{http://A/} \left\{ K_A, \text{http://B/} \left\{ K_B, \dots \text{http://Z/} \left\{ K_Z, \text{http://www.anonymity.org} \right\}_{K_Z} \right\}_{K_B} \right\}_{K_A}$

- The path is established in a nested way – The initiator prepares a layered request (called onion) that contains information for each router.
- The information contained in an onion consists of a cryptographic key, the identity of the next hop, and an encrypted onion for the next hop.

Protocol-level anonymity

- Anonymity at the protocol level consists of avoiding any information flow that could be used to trace (i.e., data anonymization)

Generic Anonymity Goals:

- Untraceability: transactions contain no identifiable information
- Unlinkability: no relation between anonymous actions can be determined

Protocol-level anonymity (contd.)

Threat Model:

- The adversary who is trying to trace someone has infinite computing power
- There should be no need to trust any third party, nor a collusion of them, to be ensured of your own anonymity

Areas of Application:

- E-voting, e-health, e-payment, privacy-preserving data mining, private information retrieval, etc.

Example I:

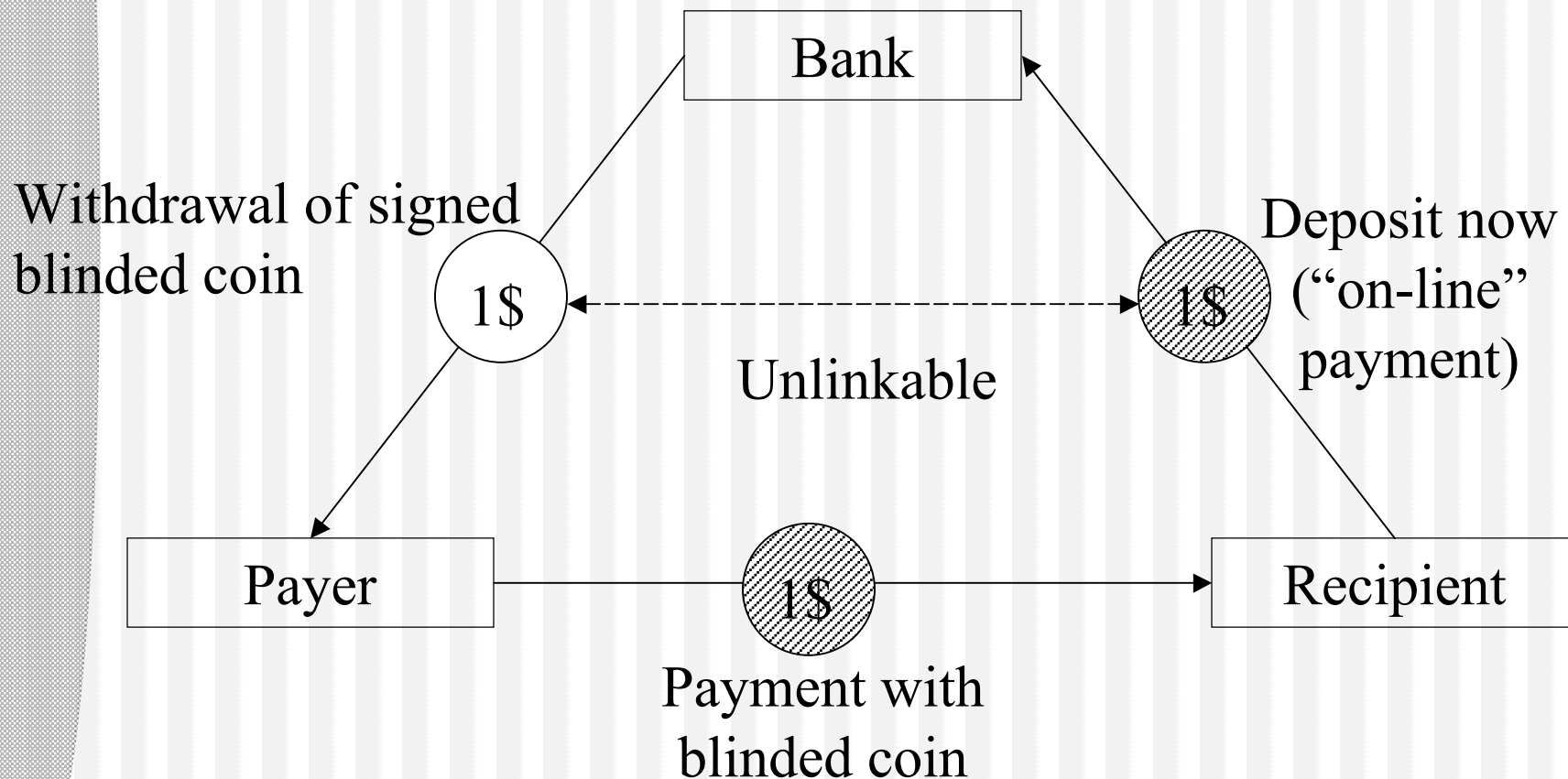
Anonymous payment systems

- Who is anonymous? In general the customer
- In what actions? Generally payments, but not withdrawals.
- To what degree? Transactions should be untraceable and unlinkable.
- Security needs?
 - Double-spending & double-depositing should be protected against

Anonymous payment systems

(contd.)

Simplistic scheme with Blind Signatures: on-line



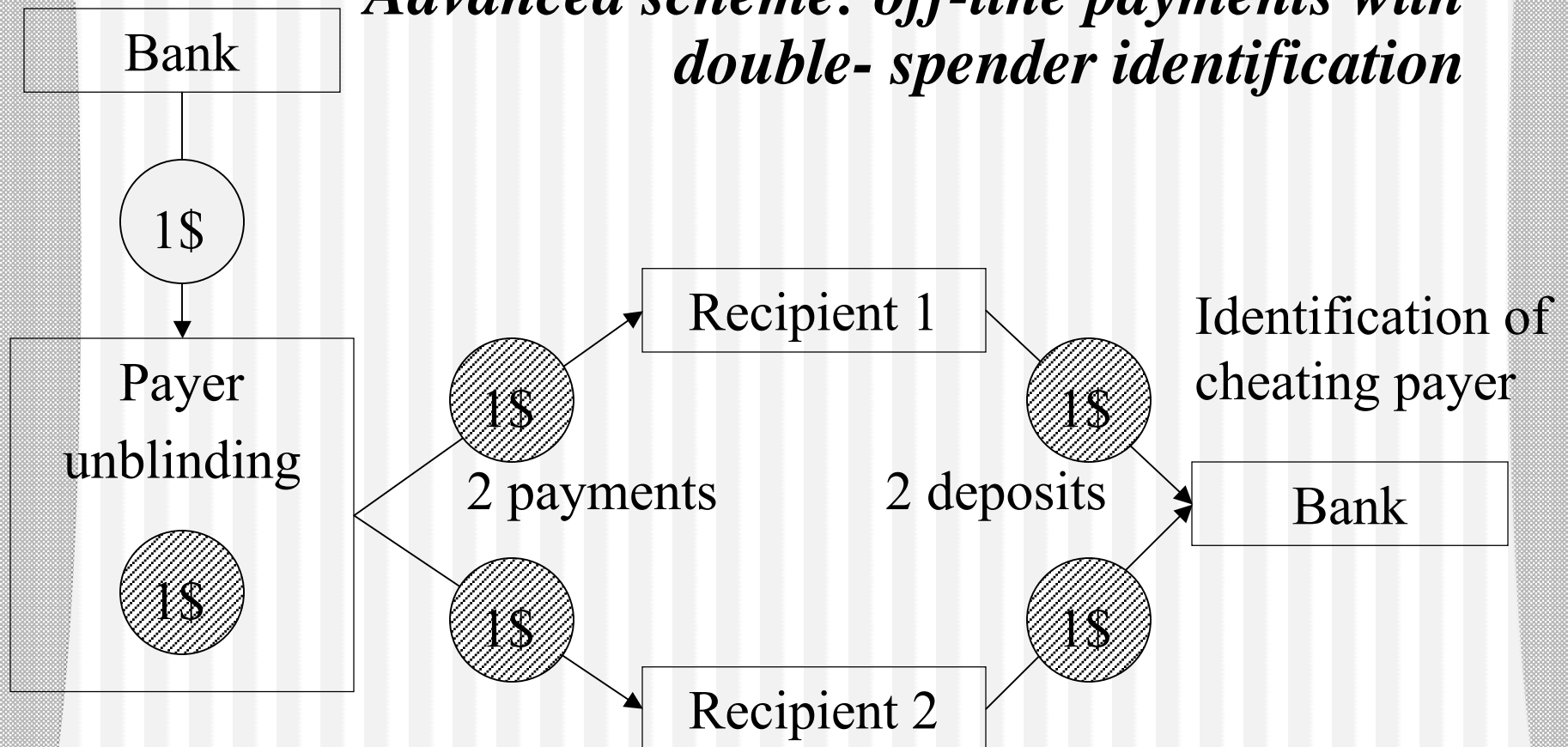
Invented by Chaum (1983)

Source: B. Pfitzmann 2000' course notes ¹⁶

Anonymous payment systems

(contd.)

Advanced scheme: off-line payments with double-spender identification



1st scheme: Chaum/Fiat /Naor (1988) Source: B. Pfitzmann 2000' course notes
State-of-the-art scheme: Brands (1993)

Example II: Electronic Elections

Specific requirements:

- Eligibility and proper access control
- Universal verifiability (vs. individual)
 - Integrity/non-alteration of votes
 - Dispute resolving, cheaters dismissal...
- Privacy (information-theoretic/computational)
- Robustness (against a reasonably sized coalition)
- No vote duplication (copying ...)
- Receipt-freeness (w.r.t. coercibility and vote-buying)

Example II: Electronic Elections (contd.)

Three main approaches [Hirt&Sako'2000]:

- Schemes using Mix-nets
- Schemes using homomorphic encryption
- Schemes using blind signatures

A Number of *legally-binding* trials:

- CyberVote project in Germany, France and Sweden.
- Similar trials in Japan, New-Zealand, UK...
- SERVE system proposed for use in US general elections this year (plans suspended ...)

Example II: Electronic Elections (contd.)

Many concerns...

- Strong opposition to the use of current technology because it does not achieve the required level of security and privacy.
- Concerns about conspiracy between technology providers and politicians
- Problems of software authenticity, privacy, integrity, verifiability...

Example III: Anonymous Databases

- Databases must preserve the privacy of data subjects.
- Several building blocks may be deployed:
 - *Anonymous Registration*: confidential, identity-hiding, verifiable registration.
 - *Data Anonymization*: adding noise while keeping the data exploitable (e.g., by preserving the mean and variance...)
 - *Private Information Retrieval (PIR)*: allowing users to query a database while hiding the identity of the data-items being searched (e.g., patent databases, stock quotes)

Example III: Anonymous Databases (contd.)

Information-Theoretic PIR:

- Trivial solution:
 - Copy whole database, pick-up required info
 - Inefficient communication complexity $O(n)$
- Non-trivial solutions:
 - K-database scheme (queries are sent to DBs)
 - DB managers get no info about user's query
 - User reconstructs desired record from answers
 - Communication complexity in $O(n^{1/\Omega(k)})$

Example III:

Anonymous Databases (contd.)

Computational PIR:

- Servers are assumed computationally bounded
- Better communication complexity
 - Trivial solution with comm. complexity of $O(n^\epsilon)$, for any $\epsilon > 0$.
 - Multiple-server solution with poly-logarithmic communication complexity...
- Other issues concern:
 - Time complexity, number of rounds needed
 - Fault-resiliency, e.g., against Byzantine-faults...

McGill Research Directions

- Modeling of privacy-oriented applications and their environments (E-voting, E-health, ...)
- Identification of common security threats, and possible attacks.
- Formalization of the intended security goals.
- Combination of state-of-the-art technologies, and development of new primitives to achieve the required security properties.