Decentralized Inverse Transparency With Blockchain

TUM Blockchain Salon

Valentin Ziegmeier

11th May 2023
Use case: Secure usage logging

If data are made available to *data consumers*…

…make all accesses visible (transparent) to *data owners*.
Decentralized inverse transparency

If data are made available to data consumers directly…

…generate non-repudiable evidence of the exchange…

…to record the usage for data owners.

Legend

D Datum
E Evidence
L Access log
Motivation: Decentralized inverse transparency

- **Problem:** Having to trust any third party means manipulation is always a possibility

- **Blockchain as supporting technology:**
  - **Advantages:** immutable and decentralized ⇒ forward security, no trusted third party
  - **Drawback:** Not correctible, no arbiter

- **Solution:** KOVACS data exchange and usage logging system
  - Non-repudiable data exchange ⇒ accountability
  - Decentralized and private usage logs in blockchain ⇒ proof of ownership and unlinkability

- **Impact:** KOVACS enables fully decentralized inverse transparency
  - GDPR-compliant solution
  - Independent of utilized blockchain software
Requirements: summary

- Forward security: Ensured by blockchain ✓
- Identity verification △
- Non-repudiable data exchange △
- GDPR compliance △
Identity verification

- Needed to attribute logs to people
- Utilizes existing IdP
- Self-sovereign identities are requested once and reused for all future communications

**Implications:**
- IdP knows of the existence of nodes
- IdP does not know who communicates with whom

See:
New-usage protocol

- **Start:** o(wner) holds datum (d), c(onsumer) holds nothing

- **Perform new-usage protocol**
  - core: protocol by Markowitch & Roggemann
  - adapted for blockchain context:
    - c and o generate individual pseudonym
    - o creates usage log and sends blockchain update

- **Result:**
  - both hold non-repudiation evidence (of origin / receipt)
  - usage is logged

See:
Time-asymmetric encryption

Pre-computed

Random password

Random salt

bcrypt

~3s

Encryption key

At request time

Encryption key

AES

>1s

Encrypted data

Plaintext data
Time-asymmetric decryption

See:
- Dworkin M. 2007. “Recommendation for block cipher modes of operation: Galois/counter mode (GCM) and GMAC”. NIST Special Publication 800-38D
Reduced confidentiality ↔ requirement to protect personal data

Immutability ↔ right to erasure

Problem

- GDPR only applies to personally identifiable information
- Pseudonymized data are...
  - personally identifiable if a link pseudonym ↔ real-world identity exists
  - anonymous otherwise

Theory

⇒ Users self-provision pseudonyms guaranteeing unlinkability and proof of ownership

Solution
P³ pseudonym provisioning

Generate RSA public/private key pair

Hash public key with BLAKE2s

Hash = one-time pseudonym

Resulting guarantees:

• **Unlinkability** (from BLAKE2s)
• **Proof of ownership** (via underlying key pair)

See:

Block structure

Block

- Last Hash
- Nonce
- Payload

Payload

- $p(c_i)$
- $p(o_j)$
- $\text{enc}_c(u_{ij}(c_i \rightarrow o_j))$
- $\text{enc}_o(u_{ij}(c_i \rightarrow o_j))$

- Pseudonym of consumer
- Pseudonym of owner
- Usage log encrypted by consumer
- Usage log encrypted by owner
KOVACS: Deployment

- Fully decentralized deployment
- Each node has own copy of blockchain
- Peer-to-peer data exchange
  - Blockchain updates
  - Data exchange
Summary: KOVACS system model
Analysis: Adversarial model

Data owners

\[ O = \{ o_1, o_2, \ldots, o_n \} \]

Logged usage \( u_{ij1}(c_i \rightarrow o_j) \)

Logged usage \( u_{ij2}(c_i \rightarrow o_j) \)

Fake usage \( u_{ij3}(c_i \rightarrow o_j) \)

Repudiate \( u_{ij1} \)

Derive identity of \( c_i \) or \( o_j \)

Derive association between \( u_{ij1} \) and \( u_{ij2} \)

Leak identity of \( c_i \) with \( \alpha = o_j \)

Fabricate \( u_{ij3} \)

Adversary \( \alpha \)

Data consumers

\[ C = \{ c_1, c_2, \ldots, c_n \} \]
Analysis

Robustness against attacks

1. Repudiate usage
   ⇒ M&R hardness
   ⇒ technically infeasible
2. Derive identity
   ⇒ BLAKE2 hardness
   ⇒ technically infeasible
3. Associating usages
   ⇒ BLAKE2 + RSA hardness
   ⇒ technically infeasible

Protocol confidentiality

P2P, encrypted, no TTP
⇒ confidential

Optional hardening:
• Fake chatter (next slide)
• Random block publication

GDPR compliance

Encrypted payload
enables confidentiality ✔

Unlinkability & proof of ownership
enable right to erasure ✔
Fake chatter

- Hide relationship of c to o
- Additional fake exchanges
- **Effect:** Communication hidden
Benchmarks

Performance: exchange duration

Network with 50 nodes | 2 second timeout | Ethereum (PoW)

(a) retrieve single log

(b) retrieve all logs
Summary

KOVACS enables...

• secure non-repudiable data exchanges
• fully decentralized deployment
• independence of the underlying blockchain solution

Practical implication: GDPR-compliant and scalable usage log storage on any blockchain

Academic impact: Paper published in *ACM Distributed Ledger Technologies* journal
Thank you for your attention.

Decentralized Inverse Transparency With Blockchain
Valentin Ziegelmeyer

Garching, 11th May 2023

Read the paper: https://mediatum.ub.tum.de/node?id=1706624