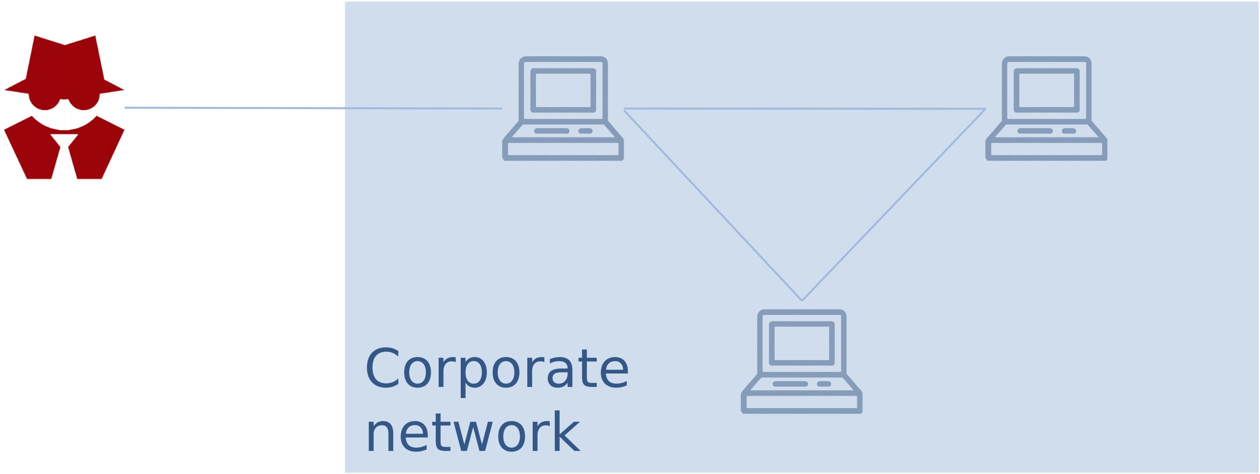
Privacy-preserving storage of enterprise logdata using pseudonymisation and threshold cryptosystems

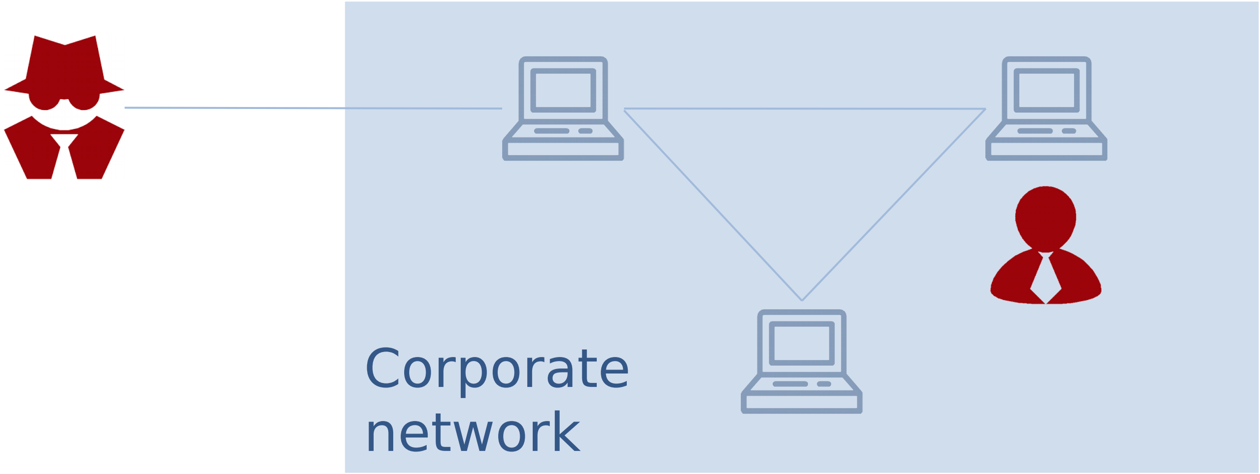
Tom Petersen

Research group on Security in Distributed Systems (SVS)
Department of Computer Science
University of Hamburg

Insider attacks (I)



Insider attacks (I)



Bitkom Spezialstudie Wirtschaftsschutz (2016):

- about 60% of incidents *in the area of data theft, industrial spying or sabotage* have a connection to current or former employees

Insider attacks (II)

Firewall, IDS, …

Icons from: https://www.shareicon.net (Flaticon basic license)

Insider attacks (II)

Firewall, IDS, ...

Icons from: https://www.shareicon.net (Flaticon basic license)

Insider attacks (II)

Firewall, IDS, …

Anomaly-based detection

Icons from: https://www.shareicon.net (Flaticon basic license)

Insider attacks (II)

Firewall, IDS, …

Anomaly-based detection

§32 BDSG

Icons from: https://www.shareicon.net (Flaticon basic license)

Approach (I)

Pseudonymisation

+

Threshold cryptosystems

Approach (I)

Pseudonymisation

- Using an identifier different from the real identifier of a subject

Alice : 0x2003

- Assignment e.g. by using functions like hash algorithms or **mapping tables**

+

Threshold cryptosystems

Approach (I)

Pseudonymisation

- Using an identifier different from the real identifier of a subject

Alice : 0x2003

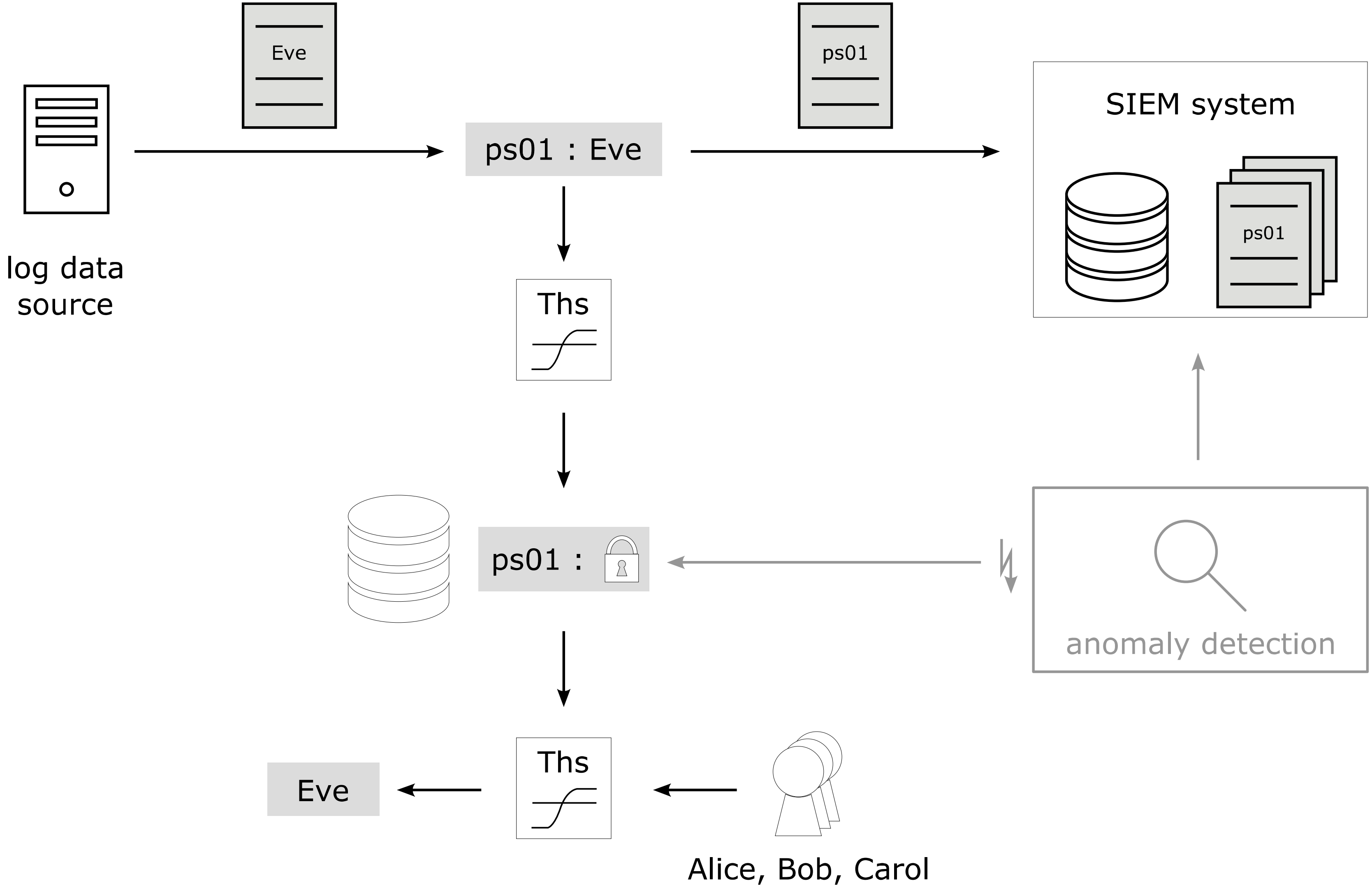
- Assignment e.g. by using functions like hash algorithms or **mapping tables**

+

Threshold cryptosystems

- (t,n) **threshold decryption**
- Key generation
  1. (pk, shares)
  2. Message encryption (pk)
  3. Computing partial decryptions
  4. Decrypting message
- Secret key is NOT restored
- Distributed trust

Approach (II)



log data source

Eve

ps01 : Eve

ps01

SIEM system

ps01

Ths

ps01 :

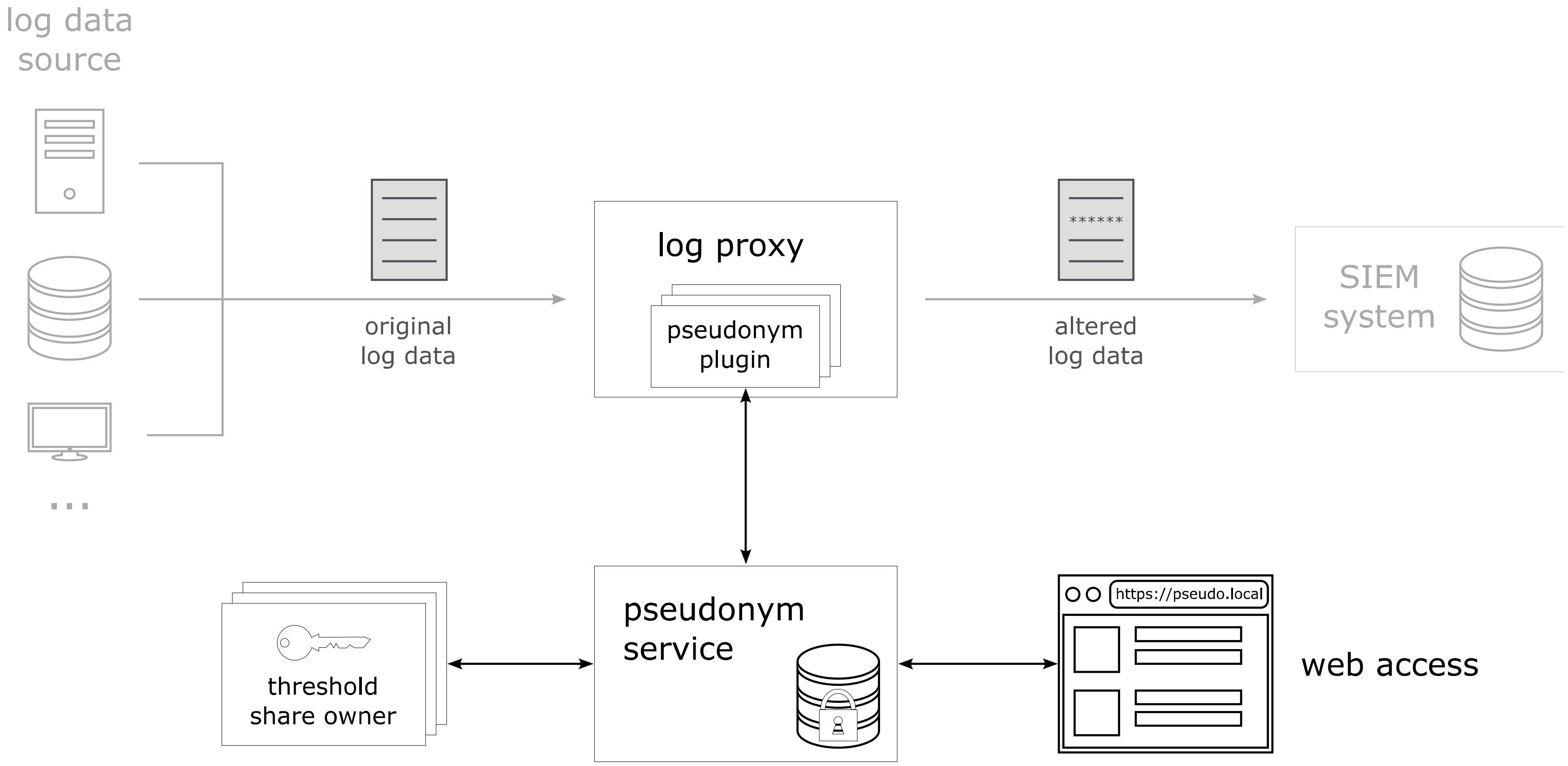
anomaly detection

Ths

Eve

Alice, Bob, Carol

Prototypical implementation



Key points

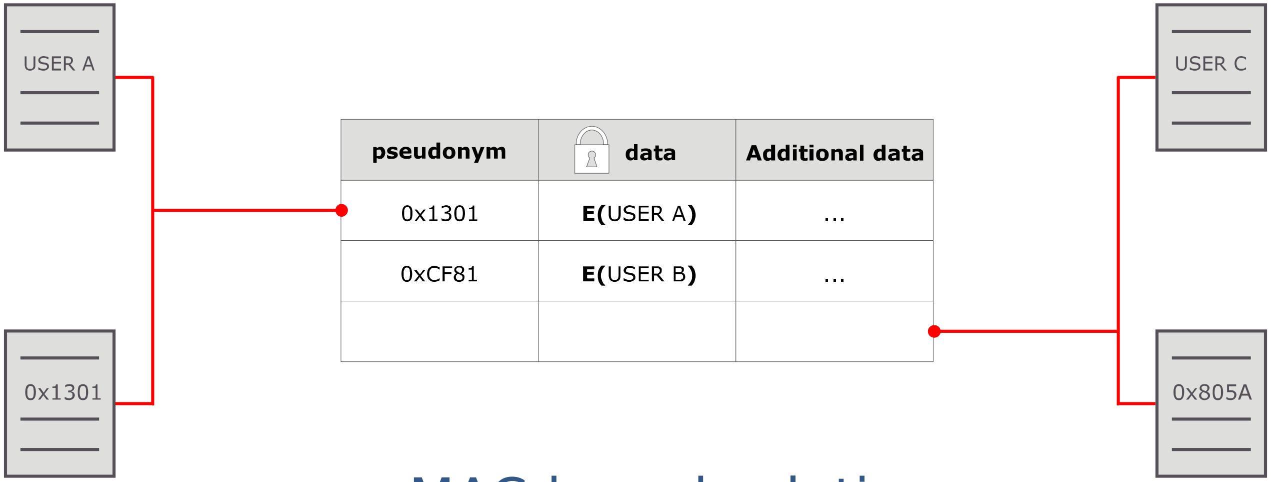
- Parameter-dependent pseudonymization: time and usage based → influence on linkability

Key points

- Parameter-dependent pseudonymization: time and usage based → influence on linkability
- Threshold public key decryption combining ElGamal and Shamir’s secret sharing → own implementation

Key points

- Parameter-dependent pseudonymization: time and usage based → influence on linkability
- Threshold public key decryption combining ElGamal and Shamir’s secret sharing → own implementation
- Identifying created pseudonyms:



pseudonym	🔒 data	Additional data
0x1301	**E(**USER A**)**	…
0xCF81	**E(**USER B**)**	…

→ MAC-based solution

Conclusion and future work

- Well-suited approach for the stated problem
  - Parameter-dependent linkability for pseudonymous log data
    → privacy-preserving anomaly detection
  - Disclosure of pseudonym owner secured by four-eye principle

Conclusion and future work

- Well-suited approach for the stated problem
  - Parameter-dependent linkability for pseudonymous log data
    → privacy-preserving anomaly detection
  - Disclosure of pseudonym owner secured by four-eye principle

- Open questions
  - Effects of different parameter choices
    - on anomaly detection
    - on (unintentionally) disclosing pseudonym owners
  - Further (context-dependent) parameters

Thank you!