CYBER SECURITY
COOPERATIVE
RESEARCH
CENTRE

UNSW
SYDNEY

DIMY: Enabling Privacy-preserving Contact Tracing

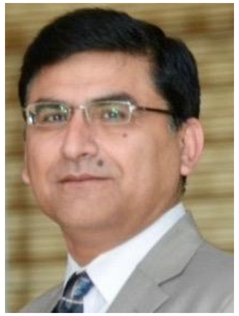
Dr Nadeem Ahmed

Senior Research Fellow

CYBER SECURITY
COOPERATIVE
RESEARCH
CENTRE

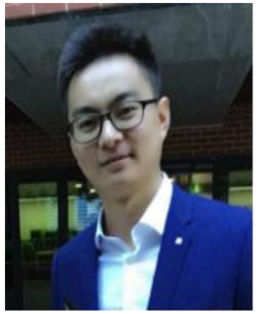
Authors and Contributors

Nadeem Ahmed



Regio Michelin

Wanli Xue



Guntur Dharma Putra

Wei Song

Sushmita Ruj

Prof. Salil S. Kanhere

Prof. Sanjay Jha

Theme1.2

cybersecuritycrc.org.au

Agenda

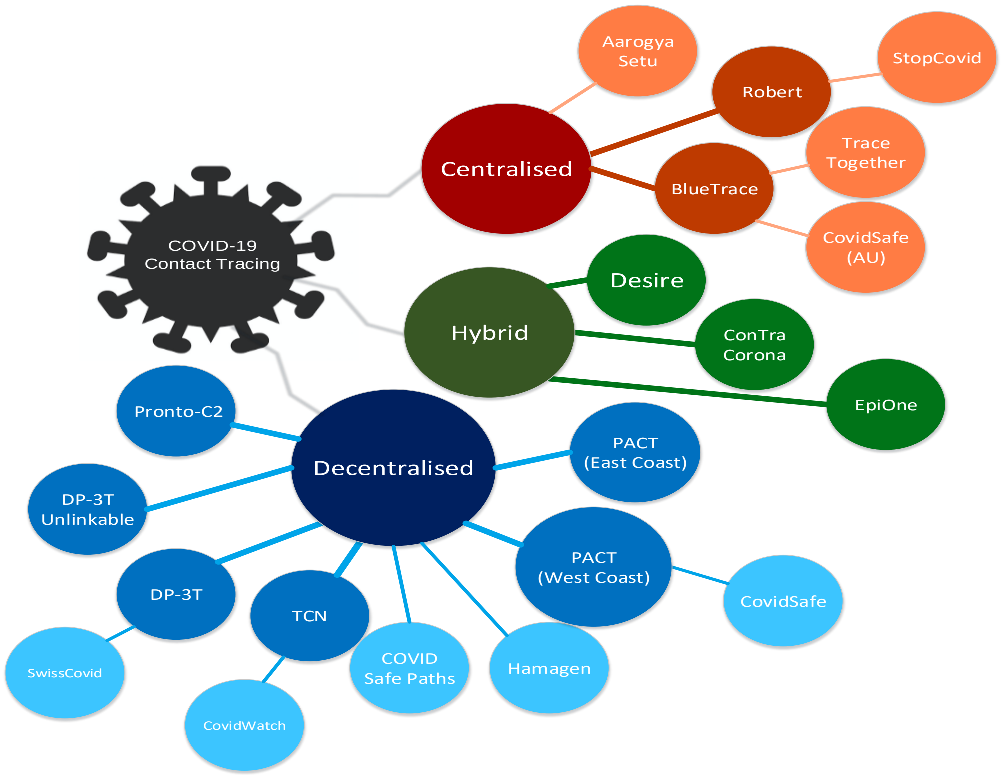
- Contact tracing
- Motivation of this work
- Building blocks of DIMY
- Demo
- Performance evaluation
- Q/A

Contact Tracing in Pandemics

- Case investigation technique
  - Establish the close contacts of an infected person to break the chain of infection
  - Experience with previous pandemics
- Manual contact tracing has some limitations
  - Requires a large, trained workforce to cope with the caseload
  - Hard to remember everyone met while infected in the last 2-3 weeks
  - A person may have met people that are strangers
  - Reactive approach
- Proactive digital contact tracing

Digital Contact Tracing

- Use of modern technologies such as smart phone apps, wearables and QR codes etc.
- More than 47 smart phones based digital contact tracing apps [1]
  - Majority employing BLE message exchanges between smart phones to capture the digital handshake



[1] P. H. O'Neill et. al, "A flood of coronavirus apps are tracking us. now it's time to keep track of them", https://www.technologyreview.com/2020/05/07/1000961/ launching-mittr-covid-tracing-tracker/.

Digital Contact Tracing

- Three commonly used architectures

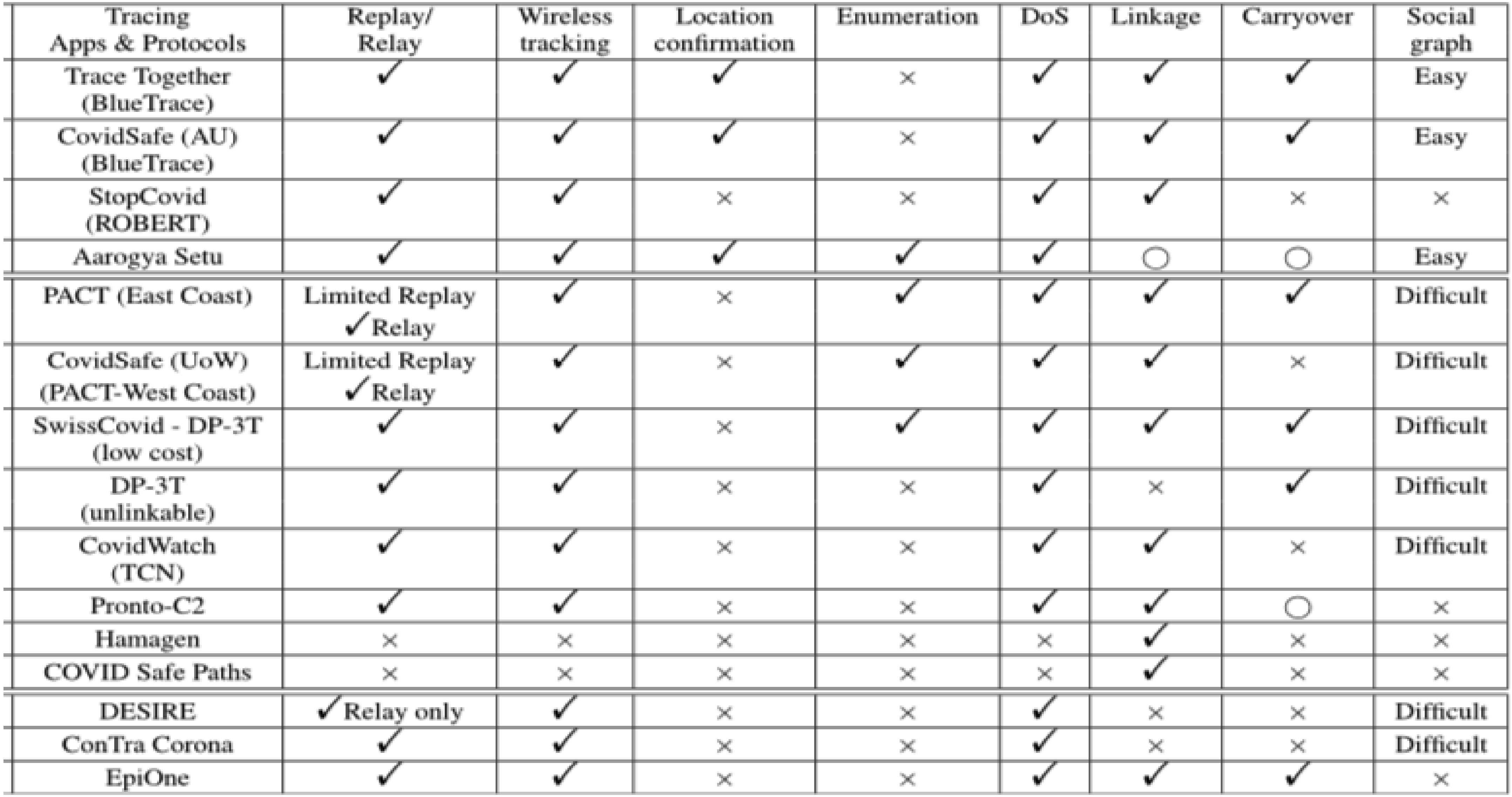
Functionality	Centralised	Decentralised	Hybrid
Ephemeral ID generation	Backend	Client devices	Client devices
Contact risk analysis and notification	Backend	Client devices	Backend
Data stored on client devices	IDs received from the backend and Encounter messages from close contacts	Seeds of positive cases received from the backend + own generated seeds	Encounter tokens and IDs generated
Data stored on the backend	List of all positive cases + their close contacts	Seeds from all positive cases	Encounter and query tokens

Motivation: Privacy and Security Concerns

- Security and privacy analysis of contact tracing apps revealed several risks and issues [2][3]
  - Different trust models for different architectures
  - Apps based on centralised architecture are vulnerable to server-side breaches and malicious *function creep* at the backend
  - Several apps are vulnerable to linkage attacks where real identities of positive cases can be easily established
  - High communication, processing and storage costs

[2] S. Vaudenay, "Centralized or decentralized? The contact tracing dilemma",IACR Cryptol. ePrint Arch. 2020 (2020) 531.
[3] N. Ahmed, R. A. Michelin, W. Xue, S. Ruj, R. Malaney, S. S. Kanhere, A. Seneviratne, W. Hu, H. Janicke, S. K. Jha, "A Survey of COVID-19 Contact Tracing Apps," IEEE Access 8 (2020) 134577–134601.

Privacy and Security Concerns



	Tracing Apps & Protocols	Replay/ Relay	Wireless tracking	Location confirmation	Enumeration	DoS	Linkage	Carryover	Social graph
Centralised	Trace Together (BlueTrace)	✓	✓	✓	×	✓	✓	✓	Easy
	CovidSafe (AU) (BlueTrace)	✓	✓	✓	×	✓	✓	✓	Easy
	StopCovid (ROBERT)	✓	✓	×	×	✓	✓	×	×
	Aarogya Setu	✓	✓	✓	✓	✓	○	○	Easy
Decentralised	PACT (East Coast)	Limited Replay ✓Relay	✓	×	✓	✓	✓	✓	Difficult
	CovidSafe (UoW) (PACT-West Coast)	Limited Replay ✓Relay	✓	×	✓	✓	✓	×	Difficult
	SwissCovid - DP-3T (low cost)	✓	✓	×	✓	✓	✓	✓	Difficult
	DP-3T (unlinkable)	✓	✓	×	×	✓	×	✓	Difficult
	CovidWatch (TCN)	✓	✓	×	×	✓	✓	×	Difficult
	Pronto-C2	✓	✓	×	×	✓	✓	○	×
	Hamagen	×	×	×	×	×	✓	×	×
	COVID Safe Paths	×	×	×	×	×	✓	×	×
Hybrid	DESIRE	✓Relay only	✓	×	×	✓	×	×	Difficult
	ConTra Corona	✓	✓	×	×	✓	×	×	Difficult
	EpiOne	✓	✓	×	×	✓	✓	✓	×

Did I Meet You (DIMY)
Privacy-Preserving Digital Contact Tracing

- Addressing the privacy, security and performance issues associated with existing digital contact tracing apps

- DIMY [4] provides:
  - Full life cycle data privacy protection
  - Resilience against many well-known attacks while introducing negligible overheads
  - Lower footprint as compared with existing state-of the art apps

- Integration of key technologies
  - Diffie-Hellman key exchange
  - Shamir secret sharing mechanism
  - Bloom Filters
  - Blockchain

[4] Ahmed, N et.al, “DIMY: Enabling Privacy-preserving Contact Tracing”
https://arxiv.org/abs/2103.05873

Building Blocks for DIMY

Diffie Hellman Key Distribution

- Share a common key over an insecure channel
- An eavesdropper cannot reconstruct the shared secret in a computationally feasible context even if they have heard all the messages exchanged
- For our work, the shared secret key is treated as the encounter ID

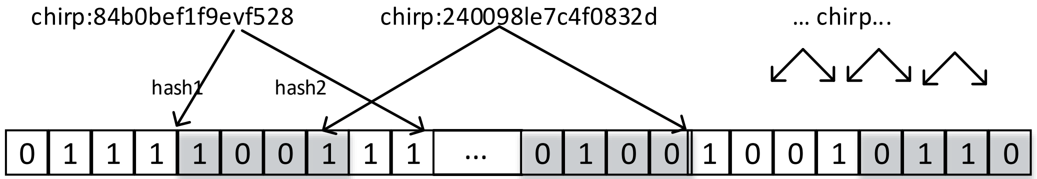
Shamir Secret Sharing

- Make *n* shares of the secret such that the secret can be reconstructed given any *k* shares (*k*<=*n*)
- No information can be known about the secret given any number of shares less than *k*
- Diffie-Hellman messages are exchanged using *k*-out-of-*n* secret sharing

Building Blocks for DIMY

Bloom Filters

- A probabilistic set membership representation that supports efficient membership queries
- False positives are possible but false negatives are not.

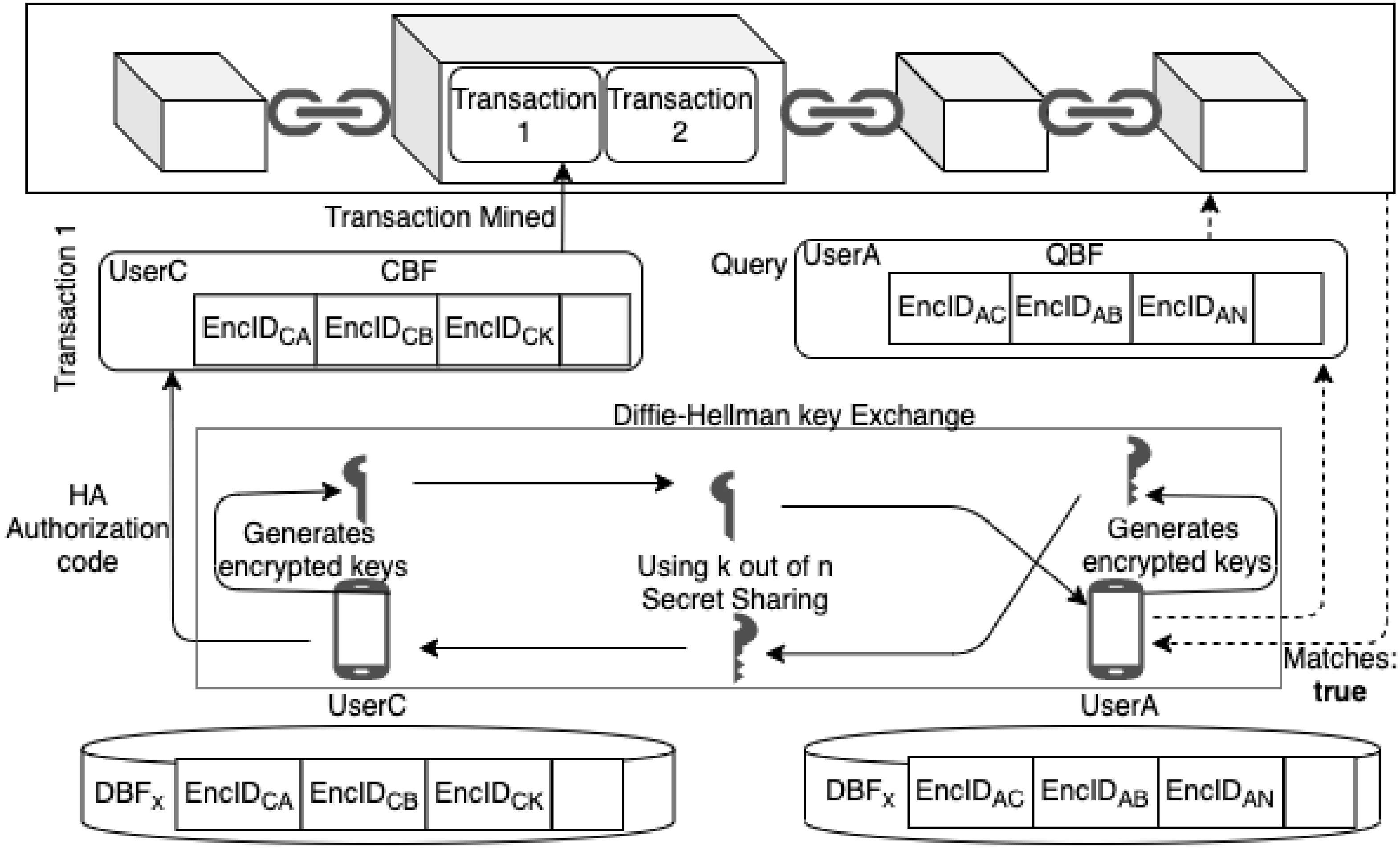


0	1	1	1	1	0	0	1	1	1	...	0	1	0	0	1	0	0	1	0	1	1	0

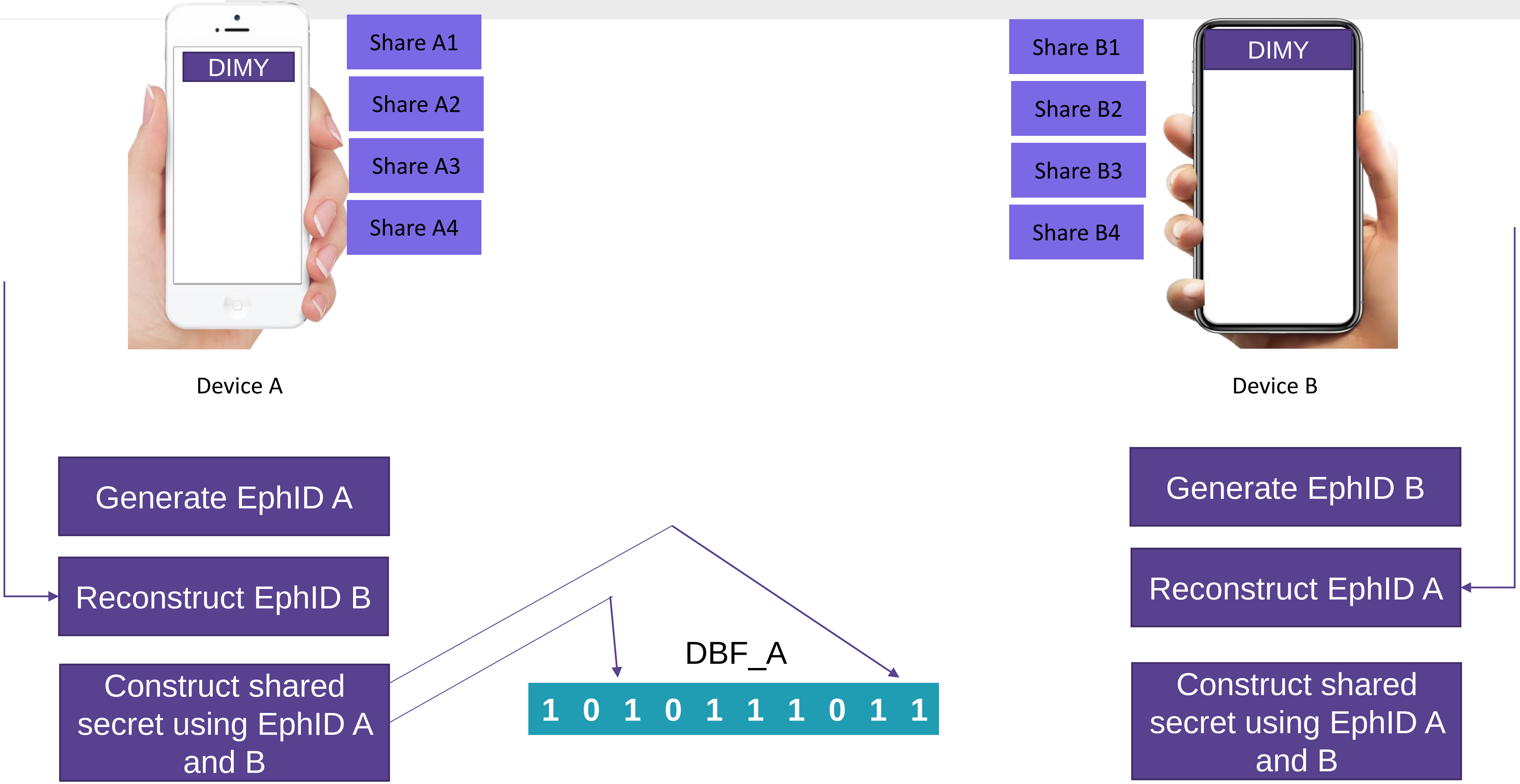
Blockchain

- Chronologically sequential immutable blocks linked together by hashing of previous blocks
- Provides data integrity, transparency of operations and the decentralized storage

Did I Meet You (DIMY) Architecture



1. DIMY Contact Representation



Share A1

Share A2

Share A3

Share A4

Device A

Share B1

Share B2

Share B3

Share B4

Device B

Generate EphID A

Reconstruct EphID B

Construct shared secret using EphID A and B

DBF_A

1 0 1 0 1 1 1 0 1 1

Generate EphID B

Reconstruct EphID A

Construct shared secret using EphID A and B

CYBER SECURITY
COOPERATIVE
RESEARCH
CENTRE

2. Daily BF, Contact BF and Query BF

Device A

DBF1	1	0	1	0	0	1	0	0	1	1
DBF2	0	0	1	0	1	1	0	0	1	0
DBF3	1	0	1	0	1	1	1	0	1	1
DBF4	1	0	1	0	1	0	1	0	0	0

Bit-wise OR

QBF/CBF	1	0	1	0	1	1	1	0	1	1

Confirmation /Result of matching

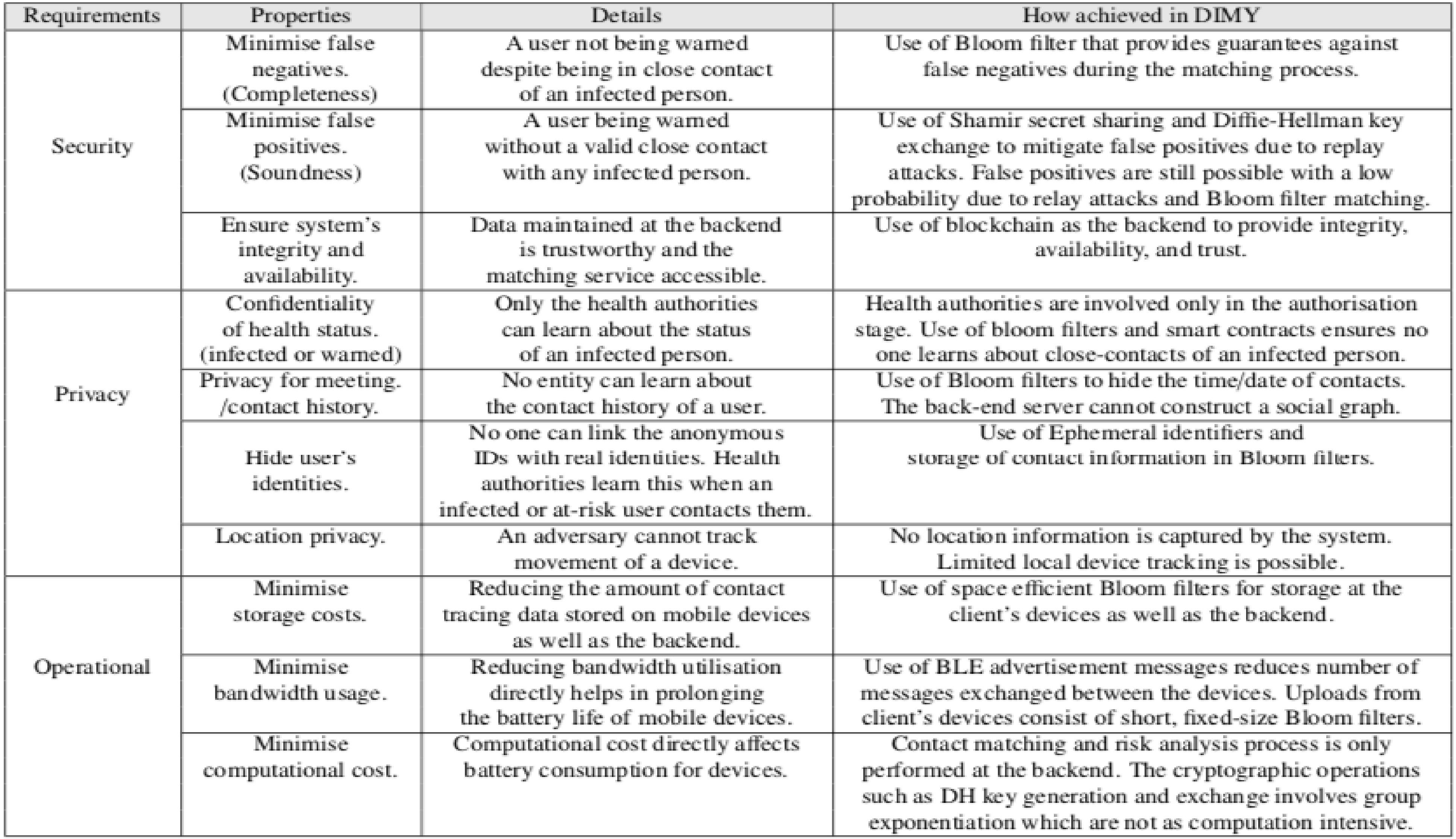
Resilience against attacks

Actors considered in the threat model:

- App users
- External actors
- Backend administrators
- Government
- Health Officials

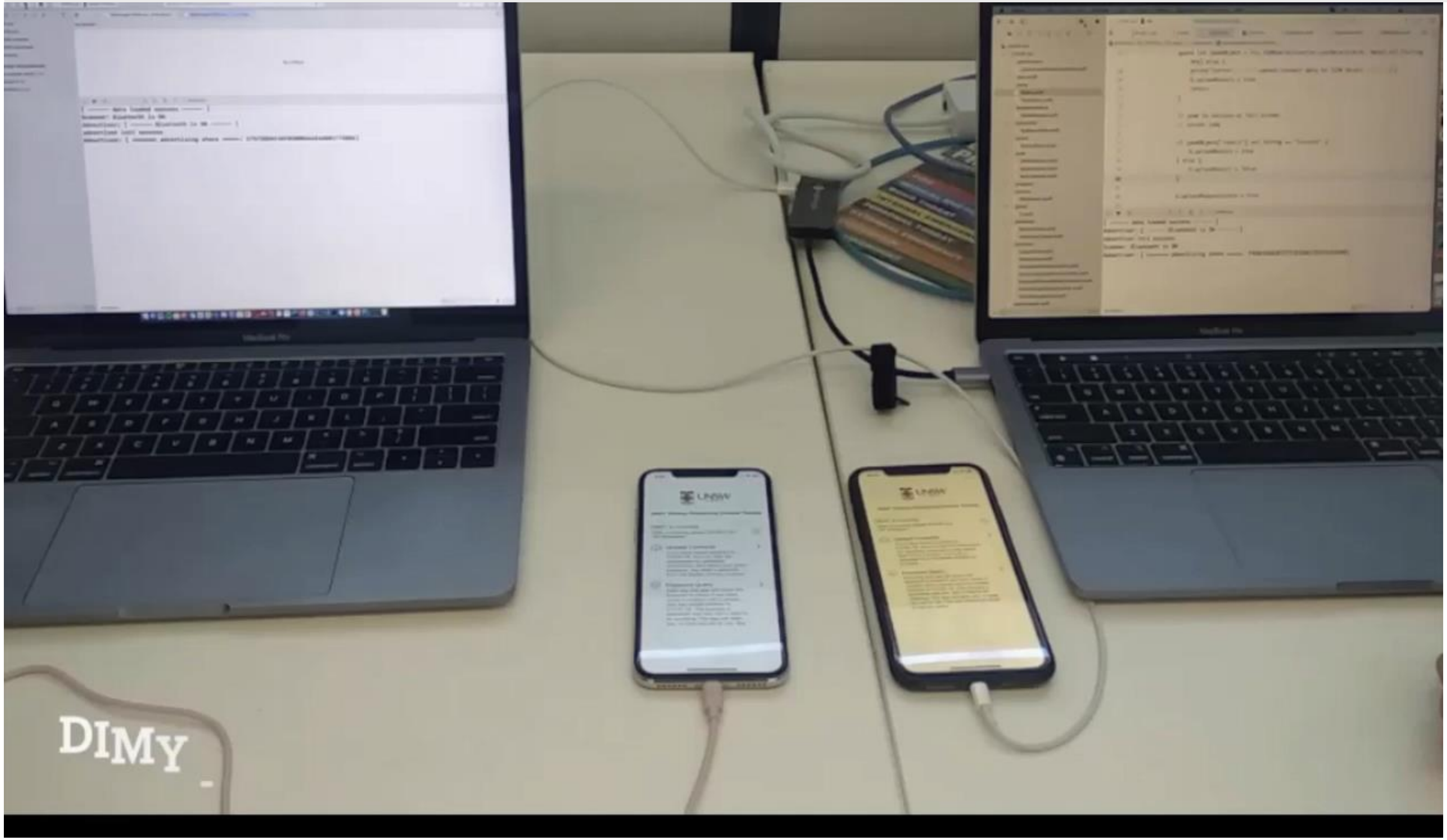
Attacks	DIMY
Replay	X
Relay	✓
Device Tracking	✓
Carryover	✓
Location confirmation	X
Enumeration	X
Denial of service	✓
Linkage	X
Social graph	X

Security, Privacy and Operational Requirements



Requirements	Properties	Details	How achieved in DIMY
Security	Minimise false negatives. (Completeness)	A user not being warned despite being in close contact of an infected person.	Use of Bloom filter that provides guarantees against false negatives during the matching process.
	Minimise false positives. (Soundness)	A user being warned without a valid close contact with any infected person.	Use of Shamir secret sharing and Diffie-Hellman key exchange to mitigate false positives due to replay attacks. False positives are still possible with a low probability due to relay attacks and Bloom filter matching.
	Ensure system's integrity and availability.	Data maintained at the backend is trustworthy and the matching service accessible.	Use of blockchain as the backend to provide integrity, availability, and trust.
Privacy	Confidentiality of health status. (infected or warned)	Only the health authorities can learn about the status of an infected person.	Health authorities are involved only in the authorisation stage. Use of bloom filters and smart contracts ensures no one learns about close-contacts of an infected person.
	Privacy for meeting. /contact history.	No entity can learn about the contact history of a user.	Use of Bloom filters to hide the time/date of contacts. The back-end server cannot construct a social graph.
	Hide user's identities.	No one can link the anonymous IDs with real identities. Health authorities learn this when an infected or at-risk user contacts them.	Use of Ephemeral identifiers and storage of contact information in Bloom filters.
	Location privacy.	An adversary cannot track movement of a device.	No location information is captured by the system. Limited local device tracking is possible.
Operational	Minimise storage costs.	Reducing the amount of contact tracing data stored on mobile devices as well as the backend.	Use of space efficient Bloom filters for storage at the client's devices as well as the backend.
	Minimise bandwidth usage.	Reducing bandwidth utilisation directly helps in prolonging the battery life of mobile devices.	Use of BLE advertisement messages reduces number of messages exchanged between the devices. Uploads from client's devices consist of short, fixed-size Bloom filters.
	Minimise computational cost.	Computational cost directly affects battery consumption for devices.	Contact matching and risk analysis process is only performed at the backend. The cryptographic operations such as DH key generation and exchange involves group exponentiation which are not as computation intensive.

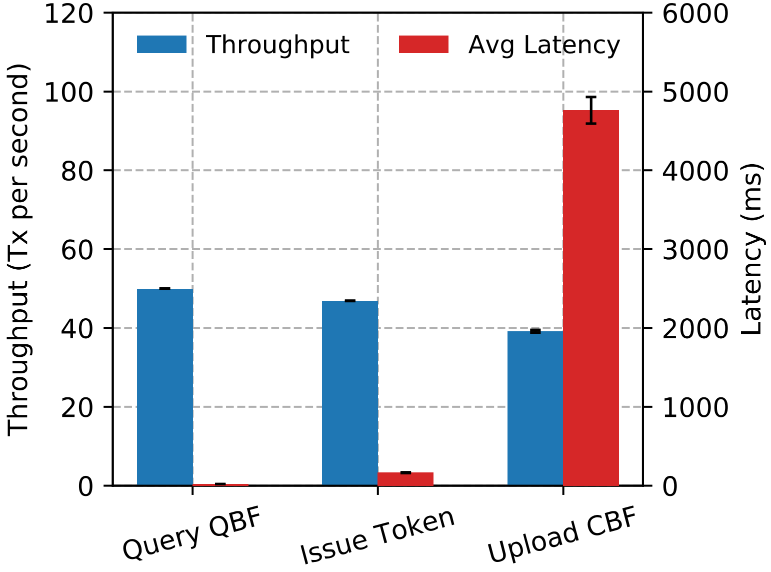
Did I Meet You (DIMY)
Demo

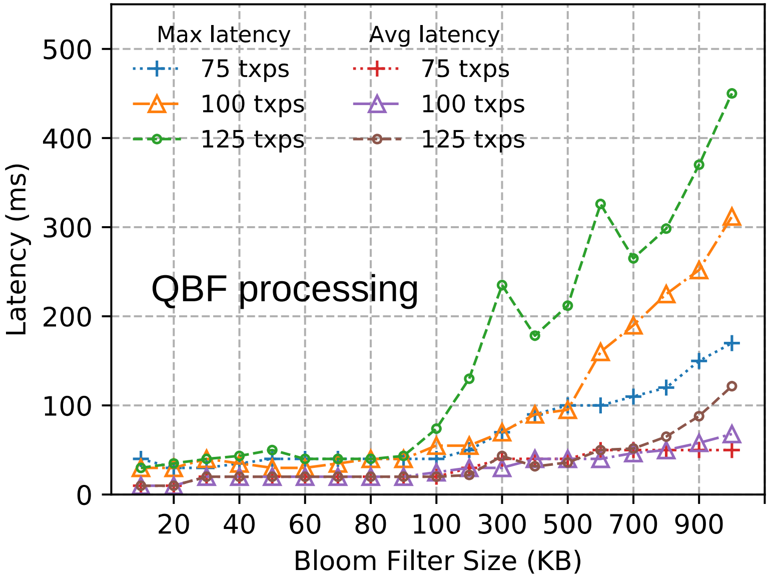


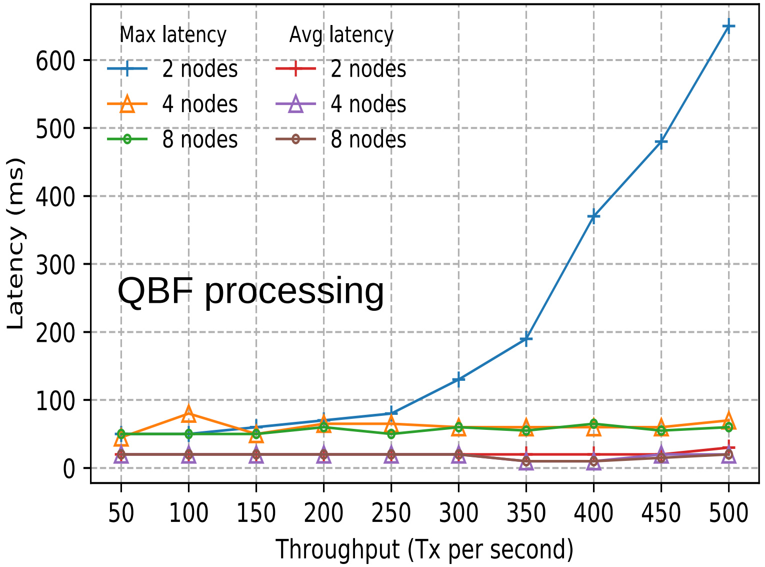
CYBER SECURITY
COOPERATIVE
RESEARCH
CENTRE

Performance evaluation

HyperLedger implemented on a local GPU server (12 cores and 64GB of RAM)



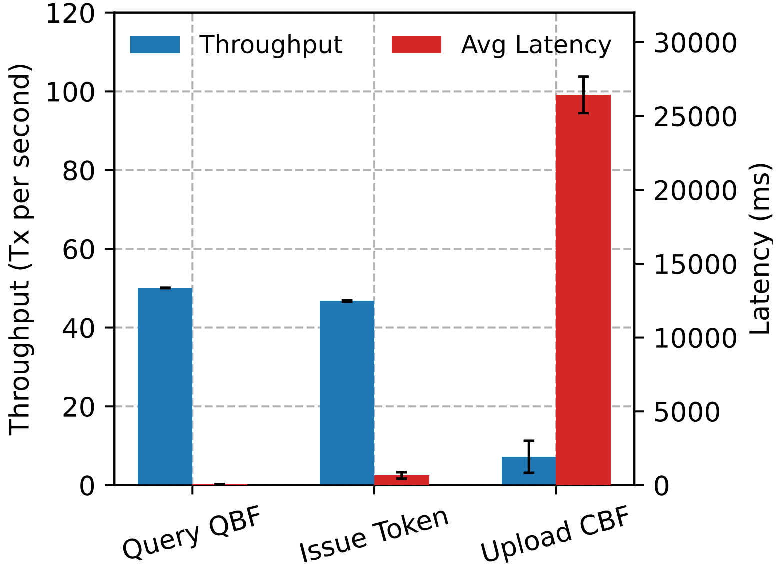


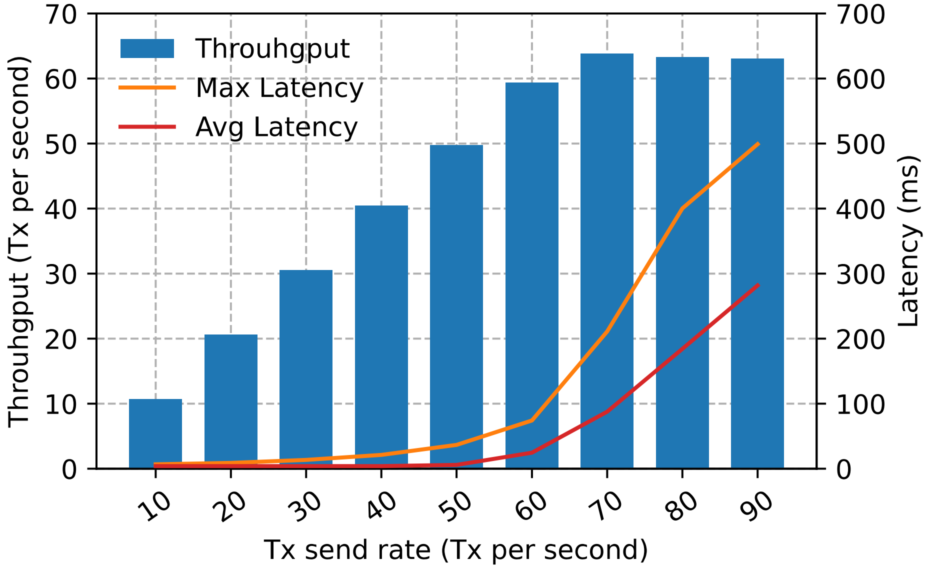


cybersecuritycrc.org.au

Performance evaluation

Backend on AWS:
A single t2.small node with 2.4Ghz CPU and 2GB of RAM
Two HyperLedger nodes and one orderer node as Docker containers





Thank You

nadeem.ahmed@unsw.edu.au