Blockchain Privacy and Its Applications in Healthcare

Guang Gong

Department of Electrical and Computer Engineering
University of Waterloo
CANADA
https://uwaterloo.ca/scholar/ggong

University of Waterloo Conference: Cybersecurity, Privacy, and Artificial Intelligence in Health Data
Delta City Centre Hotel, Ottawa, May 5, 2023

Outline

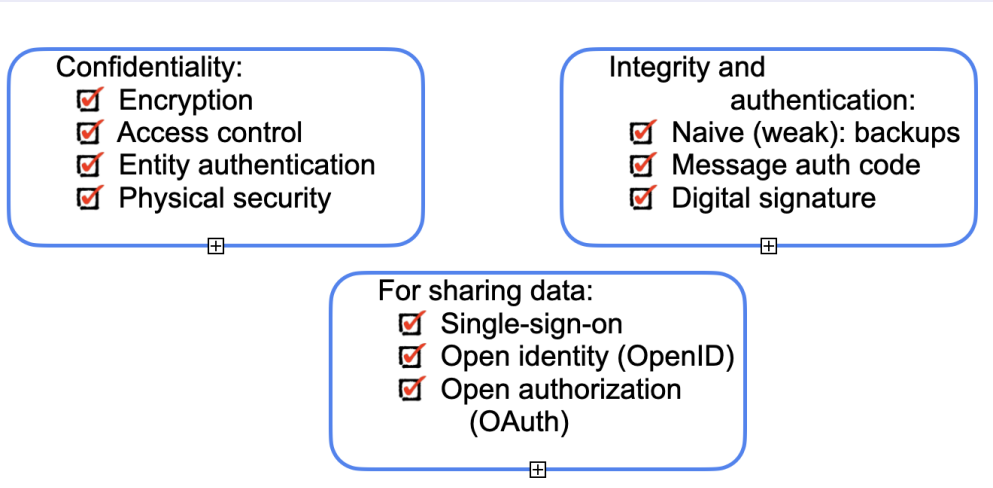
- Problems and challenges in implementing security and privacy in health data
- Proposed solutions
  - Introducing new technology: blockchain and Bitcoin network
  - Blockchain privacy and zero-knowledge (ZK) proofs in cryptography
  - New findings
- Concluding remarks on the adoption of blockchain technology in health data

Problems and Challenges

How can we secure and preserving privacy of massive (big) health data records?

Current Tech

First, we need implementing existing security mechanisms and technologies.



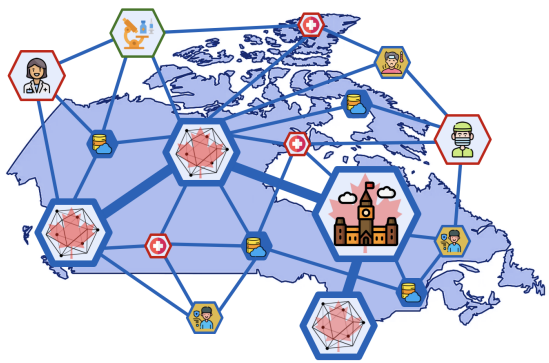
Are they enough? Absolutely **NOT** due to the facts that the Internet of Medical Things, smart devices, and information systems become the backbones of healthcare systems, AI for providing data analytics and cloud service for sustaining storage.

New security and privacy protection mechanisms needed

- Health data has a vast degree of diversity: from IoMT, smart devices, and AI.
- How to realize security and privacy in **Person-Centred Data by Design** enabled healthcare systems?
- How to **retrieve and share** them cross multi-stakeholders (for individual, clinical, and analytical access and use) and **cross provinces**?
  - Health data themselves must be **encrypted** during storage and transfer, since encryption is necessary for confidentiality of the data, so that the contents of the data are secure from corruption or damage or unauthorized access and from malware and cybersecurity threats.
- How to preserve patients' **privacy** (such as location, load, traffic of interaction with doctors/care providers, etc.) or doctors' privacy (e.g., proprietary treatments)

Challenges

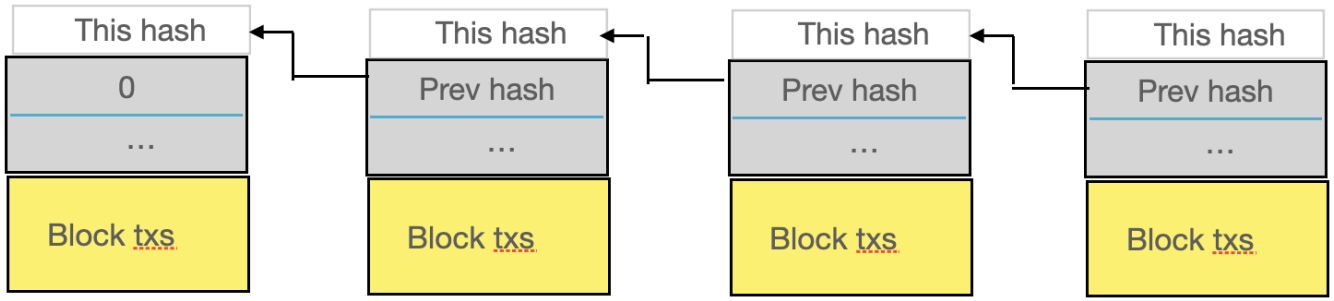
Those constitute extremely **challenging** problems! $\longrightarrow$ new technologies which can implement those functionalities are in request.



Blockchain structure

- A **blockchain** can be described as a **database shared** among multiple distributed nodes, and the data stored on it is highly **unforgeable, verifiable**, **access traceable, and transparent**, without the need for a public trusted authority.
- Blockchains can provide **trusted** consensus, computation, and immutable data between untrusted entities.
- There are three basic components in the blockchain structure:
  - **Transaction**: refers to all operations on the ledger.
  - **Block**: It is used to record all the transactions and states within a specified time interval, and can be considered as a consensus result for the current state of the ledger.
  - **Chain**: A set of blocks which are listed in chronological order, and can be considered as a historical log of all the time nodes of the whole ledger.

A blockchain is a singly linked list of blocks by a hash function succeedingly applied to each block.



- **How does each node agree on the block?** – it runs a consensus protocol (PBFT, PoW, PoS, ...)
- There are four types of blockchain networks. The core difference between them is the read and write access and the degree of decentralization.

Comparison of Different Types of Blockchain

Type	Ledger-keeping **rights**	**Degree** of decentralization	Performance
Public	all	high	low
Permissioned	partial	high	low
Consortium	partial	medium	medium
Private	private	high	low

Bitcoin Blockchain

- How to perform **digital cash transactions** ? Traditionally, this is conducted by the bank.
- **Replacing banks.** Basic functions that a bank provides are
  - Identity management
  - Transactions
  - Prevents double spending
- How can we enforce these properties cryptographically?

Transactions in Bitcoin Blockchain

- When user A with her public-key pk_A (as their valet account, i.e., identity), wants to send amt=\$3 in Bitcoin currency to B with his public-key sk_B, she forms a transaction $Tx = (pk_A, pk_B, 3)$ and digitally signs that, i.e., $Sig_A(pk_A, pk_B, 3)$.
- The signed data is **broadcast** to the network and, if valid, it ends up in a block in the blockchain, with the consensus among all the miners!
- A bitcoin is a chain of transactions tracing its flow from mining up until the current owner address, i.e., **coin is a chain of digital signatures** $\longrightarrow$ no one can forge it, but everyone can verify that!
- Validating a transaction uses **cryptography** (i.e., signature, authorization and spending once)

Blockchain Privacy

Transparency

Bitcoin blockchain as well as most of cryptocurrencies, including permissioned and consortium blockchains **publish all transaction data** to all miners or validators to enable consensus validation.

Privacy in demand

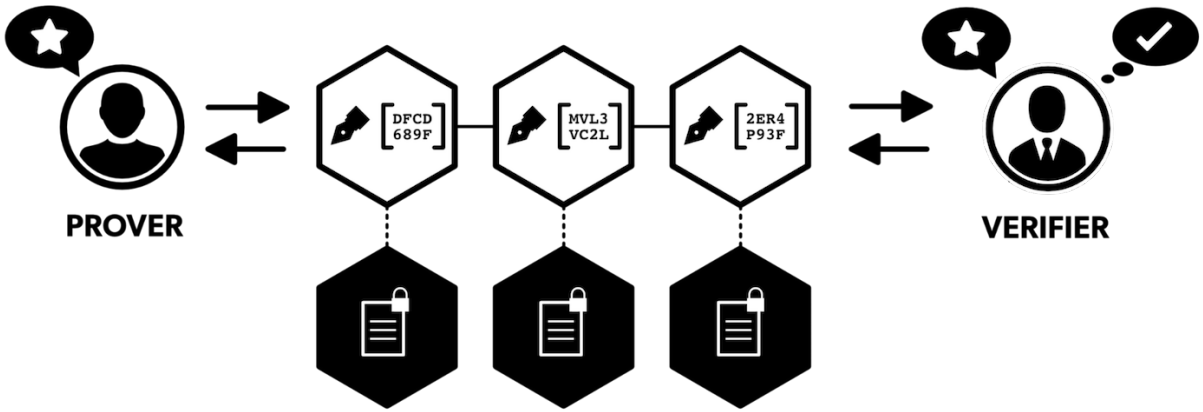
How to providing **privacy** of senders, transaction data, and receivers for the applications with sensitive data?

Tool

Zero-knowledge proof systems!

Zero-Knowledge Proofs

Loosely speaking, zero-knowledge proofs are **proofs** that yields **nothing** beyond the validity of the assertion.



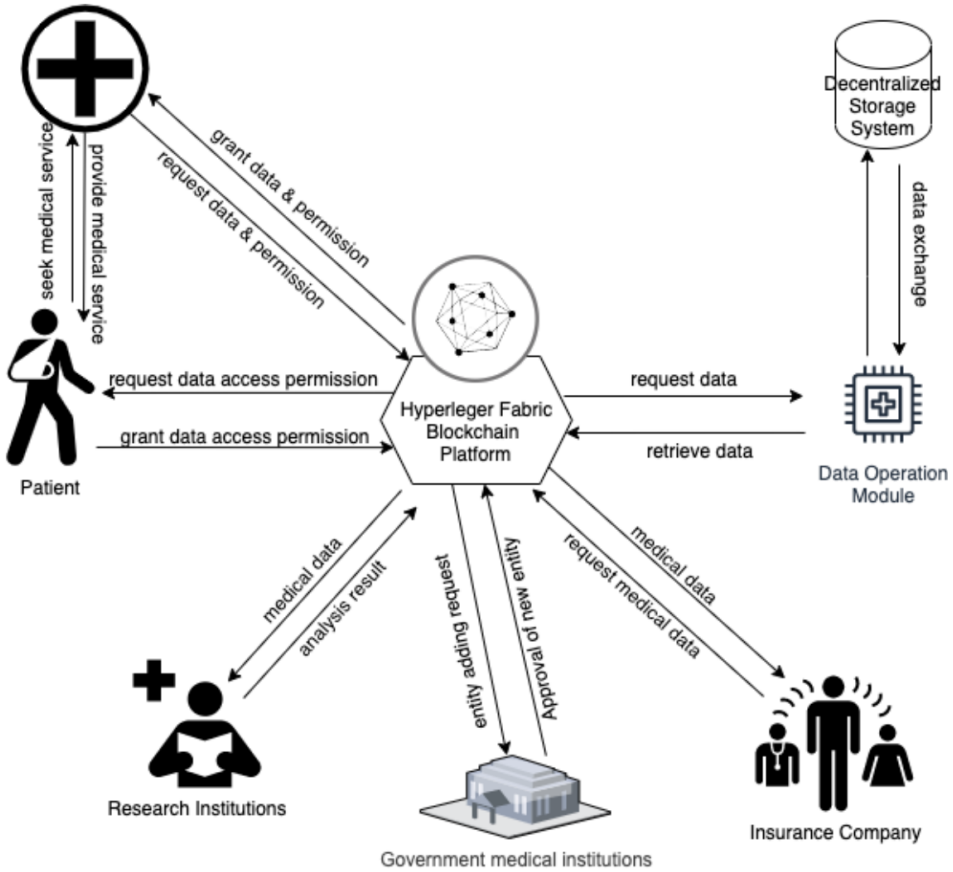
zkSNARK

zero-**k**nowledge **S**uccinct **N**on-interactive **AR**gument of **K**nowledge is designed for arithmetic circuit proof systems.

Our solutions

- We have proposed **Polaris (HG22)**, a new zkSNARK.
  - It is currently the most efficient system.
  - It is post-quantum secure, compared with those based on hard problems: GGPR13, Groth16, Bullet proof (BCCGP2016), Marlin (CHMMVW20), ...
  - **Polaris** has improved verifier's efficiency compared with the most promising ones: Stark (BBHR18), Ligero/Ligero++ (AHIV17, 2021), Aurora (BCRSVW19), Spartan$_{OR}$ (Setty20), ...
- We have investigated Zupply: a new anonymously auditing and validating the health data record (MG2023). This scheme proposed:
  - **Blockchain Platform (BP) (e.g., Hyperleger, Ethereum)** to mint an authentication and authorization token (AAT) for the first entity (could be patients, doctors, etc.) of a data log and save the Merkle root of all minted AATs on Blockchain. Authorizing token's usage or transferring is done via smart contracts. Authenticity of the token is established by ZK proofs.
- We have proposed a **new efficient algorithm** (LG2023) to speed-up the computation in Groth16 zkSNARK, implemented in Zcash and adopted by Ethereum. Our new algorithm can gain from 3%-20% improvement over the current systems.

Proposed Blockchain healthcare model



- In blockchain enabled healthcare, **hospitals and medical institutions** can be entities (i.e., miners/validators) in the system, while these institutions have to collaborate with insurance companies and scientific institutions, in addition it may be regulated by government departments.

Concluding Remarks

Our work

- Polaris is a new tool to achieve practical blockchain privacy.
- Zupply is a new blockchain enabled health data management scheme with anonymous authentication and authorization.
- A new algorithm to speed-up zero-knowledge proofs in Zcash and Ethereum.

Privacy is in need in all types of blockchains

- None of the existing work on applying blockchain networks to health data, either in theory or in practice, has considered the **privacy problem**.
- This **resembles** the situation of 40 years ago for the **password login system**. The system stored all the passwords in plaintext, which caused a majority of break-ins. Due to the tremendous efforts have been put, currently, none of the secure web authentication stores passwords in plaintext.

Suitability

- Evidently, transferring and sharing **authentication and authorization credentials** of health data can be done via smart contracts due to its automation feature.
- For **Person-Centred Data by Design**, blockchain enabled health record systems are more efficient than any centralized system because of the uniqueness of blockchains which can let mutually-distrusting entities agree on history and can provide **trusted** consensus and immutable data between untrusted entities.
- It is suitable for sharing data in the health data systems among multiple stakeholders and in heterogeneous environments. We believe that it is well suited to adopt the joint design of **Consortium chain** and public chain.

Barrie to implement blockchain privacy in health data

- ZK proofs were proposed only few years later than public-key cryptography. However, they have been only considered as theoretical interests in cryptography for more than 40 years due to their heaviness in crypto operations.
- The barrie to the adoption to heath data is the **efficiency** of ZK proofs, i.e., how to efficiently implement and optimize those proofs at **hardware, software, and protocol layers** (much more complicated than a simple encryption)! In the recent three years, the implementation of zkSNARK for blockchain privacy are rapidly developed because of the need in pratice.

References

- Mohammadtaghi Badakhshan and Guang Gong, Anonymously Maintained Decentralized DAG Data Record Over Public Blockchains, to be submitted to PETs 2023, May 2023.
- Guiwen Luo, Shihu Fu, and Guang Gong, Speeding Up Multi-Scalar Multiplication over Fixed Points Towards Efficient zkSNARKs, IACR Transactions on Cryptographic Hardware and Embedded Systems (CHES), Volume 2023, Issue 2, 2023.
- Shihui Fu and Guang Gong, Polaris: Transparent Succinct Zero-Knowledge Arguments for R1CS with Efficient Verifier, the Proceedings on Privacy Enhancing Technologies, 2022 (1), pp. 544 - 564.
- Aiden Feng, A Design of Electronic Medical Record System based on Permissioned Blockchain, Master Thesis, UW, 2022.

Thanks! Questions?