Decentralizing Privacy:
Using Blockchain to Protect Personal Data

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7163223

Kevin Kredit and Vineet James

Agenda

- Blockchain basics
- Paper walkthrough
  - Introduction
  - The Privacy Problem
  - Proposed Solution
  - The Network Protocol
  - Discussion of Future Extensions
  - Conclusion
- Updates
- Problems

Blockchain basics

Blockchain basics

- Important properties
  - Open -- public, distributed transaction history
  - Secure -- incorruptible history
  - Safety in numbers -- inviolable rules, infeasible to hack
- Infographics
  - https://blockchainhub.net/blog/infographics/what-is-a-blockchain/
  - https://followmyvote.com/infographics/blockchain-technology-breakdown-infographic/
  - (Bitcoin) https://visual.ly/community/infographic/technology/bitcoin-infographic

Paper walkthrough

Part 1: Introduction

Why do Companies and organizations collect our data ?

- personalize services
- optimize the corporate decision making process
- predict future trends
- and more...

Data is a valuable asset and sharing data has it's benefits however, there is a growing public concern about user privacy.

In the current model, Users have no or little control over the data that's collected by centralized Third-parties and how that data is used.

Part 1: Introduction

Related Work

- **OpenPDS** (Open personal data store) - Returns only answers instead of raw data
- **OAuth** - Organizations acting as a centralized trusted authority themselves and providing access control.
- **Data Anonymization methods**
  - **k-anonymity -** Ensures each record is indistinguishable from other records in the set.
  - **l-diversity -** Ensures sensitive data is represented by a diverse enough set of possible values
  - **t-closeness -** looks at the distribution of sensitive data
- **Differential Policy -** adds noise to the computational process prior to sharing the data
- **FHE (fully homomorphic encryption) -** allow running computations and queries over encrypted data

Part 1: Introduction

Proposal

- A decentralized personal data management system that ensures users own and control their data.
- A protocol which turns blockchain into an automated access-control manager that does not require trust in a third party.

Contribution

- Combining blockchain and off-blockchain storage to construct a personal data management platform focused on privacy.
- Illustrate future improvements to the technology, how blockchains could become a vital resource in trusted-computing.

Part 2: The Privacy Problem

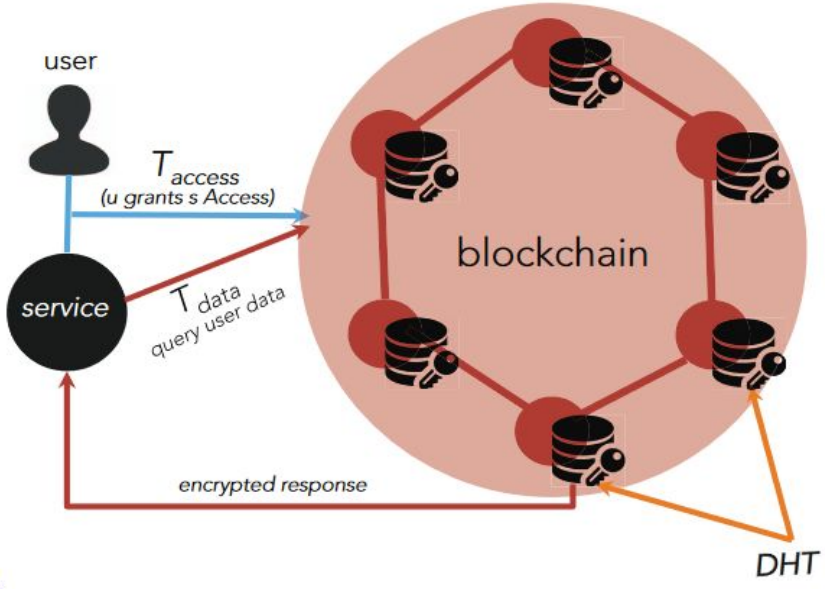
- **Project goals**:

  Focus specifically on mobile platforms and address Privacy issues like.

  - **Data ownership** - User owns and controls their data
  - **Data transparency and auditability** - User has control over what data is being collected
  - **Fine-grained access control** - User has control over how it's being used

Part 3: Proposed Solution

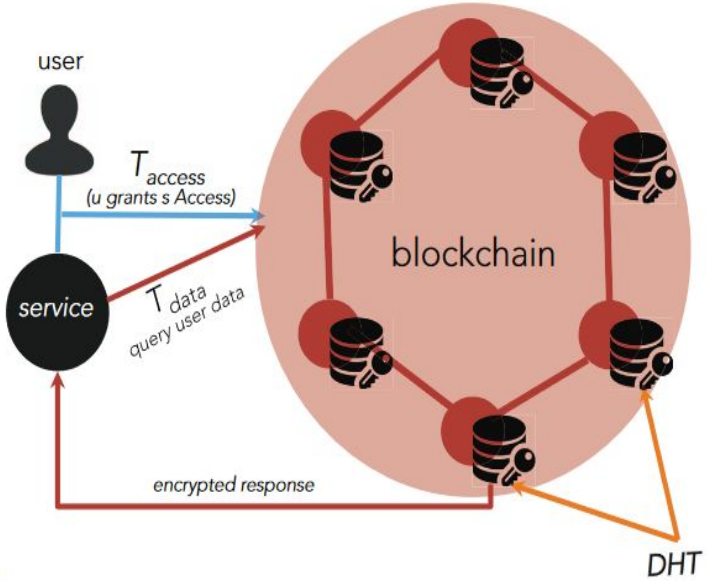
- **Idea** | Fleshing Out | Full Implementation Details | Prototype | Product
- **Main Entities**
  - Users
  - Services
  - Nodes (Blockchain & DHT)
- **Blockchain Transactions**
  - T_{access}
  - T_{data}



Part 3: Proposed Solution

How this platform would work

- the user installs the application and signs up for the first time
- a new shared (user, service) identity is created
- Associated permissions are sent to the blockchain in a T_{access}
- Data collected is encrypted using a shared encryption key
- Encrypted data is sent to the blockchain in a T_{data}
- Data is stored off-chain in a DHT (distributed hash table)
- Blockchain only has the pointer key(256 SHA)on the ledger
- Both the service and user can now query data using T_{data}
- Blockchain verifies the digital signature (User/Service).
- Blockchain also verifies access permissions for Service.
- User can change permissions granted to a service any time

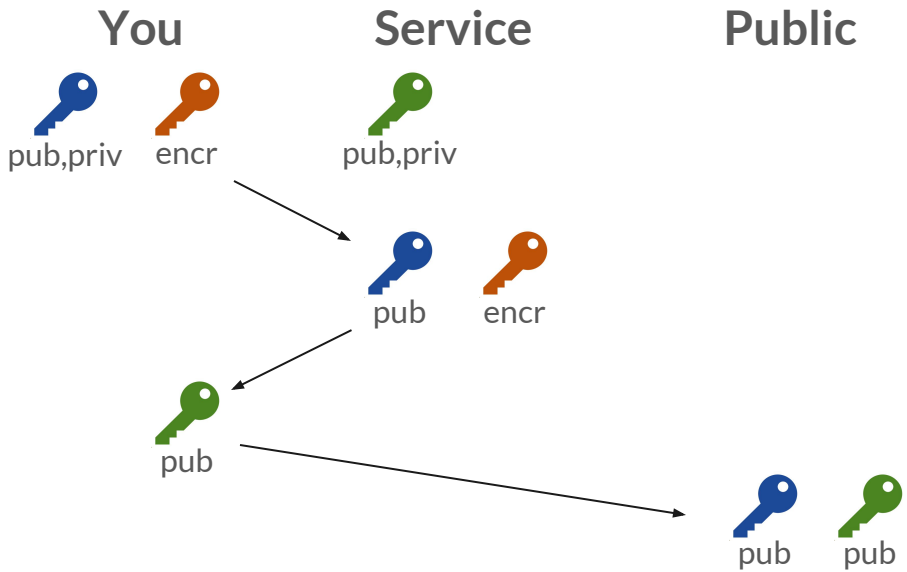


Part 4: The Network Protocol

- Idea | **Fleshing Out** | Full Implementation Details | Prototype | Product
- Building blocks
  - Identities
  - Policies
  - Protocols
- Analysis

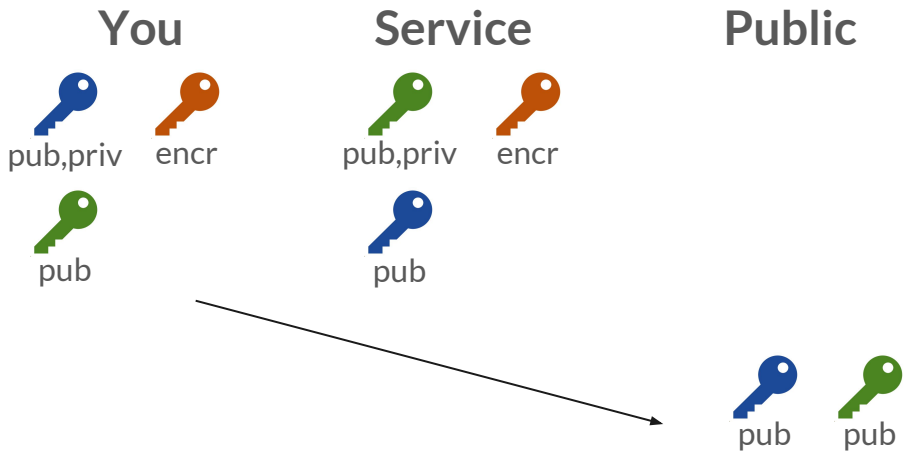
Part 4: The Network Protocol

- Building blocks
  - **Identities**
  - Policies
  - Protocols



Part 4: The Network Protocol

- Building blocks
  - Identities
  - **Policies**
  - Protocols



New policy: service Green is allowed to access user Blue's location and contacts.

Part 4: The Network Protocol

- Building blocks
  - Identities
  - Policies
  - **Protocols**

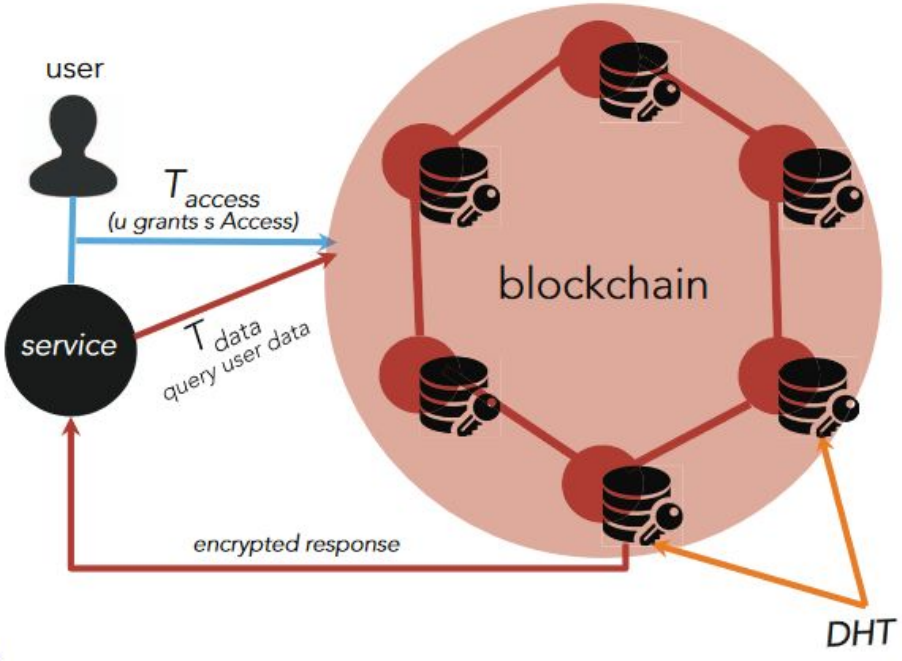
Protocol 3 Access Control Protocol

1: **procedure** HANDLEACCESSTX(pk^k_{sig}, m)
2: $s \leftarrow 0$
3: $pk^{u,s}_{sig}, pk^{s,u}_{sig}, POLICY_{u,s} = Parse(m)$
4: **if** $pk^k_{sig} = pk^{u,s}_{sig}$ **then**
5: $L[\mathcal{H}(pk^k_{sig})] = m$
6: $s \leftarrow 1$
7: **end if**
8: **return** s
9: **end procedure**

"If message is really from the user, then save the new policy on the blockchain; else, fail."

Part 4: The Network Protocol

- Building blocks
  - Identities
  - Policies
  - **Protocols**



Part 4: The Network Protocol

- Building blocks
  - Identities
  - Policies
  - **Protocols**

Protocol 4 Storing or Loading Data

1: **procedure** HANDLEDATATX(pk^k_{sig}, m)
2: $c, x_p, rw = Parse(m)$
3: **if** $CheckPolicy(pk^k_{sig}, x_p) = \texttt{True}$ **then**
4: $pk^{u,s}_{sig}, pk^{s,u}_{sig}, POLICY_{u,s} \leftarrow Parse(L[\mathcal{H}(pk^{u,s}_{sig})])$
5: $a_{x_p} = \mathcal{H}(pk^{u,s}_{sig} \parallel x_p)$
6: **if** $rw = 0$ **then** ▷ rw=0 for write, 1 for read
7: $h_c = \mathcal{H}(c)$
8: $L[a_{x_p}] \leftarrow L[a_{x_p}] \cup h_c$
9: (DHT) $ds[h_c] \leftarrow c$
10: **return** h_c
11: **else if** $c \in L[a_{x_p}]$ **then**
12: (DHT) **return** $ds[h_c]$
13: **end if**
14: **end if**
15: **return** $\emptyset$
16: **end procedure**

"If transaction is allowed by the policy: if writing, calculate hash of new data, store hash on blockchain, add data to DHT; if reading, fetch data from DHT."

Part 4: The Network Protocol

- Analysis
  - Assumptions:
    - Tamper free blockchain, sufficiently large network of nodes
    - Private keys are securely managed
  - Original goals:
    - Data ownership
      - ✔ impersonation is impossible, and policies inviolable
      - ✔ data is encrypted on the DHT
    - Data transparency and auditability
      - ✔ policies define precisely what is collected
    - Fine-grained access control
      - ✔ policies can be arbitrarily fine and changed at any time

Part 5: Discussion of Future Extensions

- From storage to processing: solving the data *usage* problem
  - Data transparency and auditability
    - ✕ you can't control how your data will be controlled after it's accessed
  - Solution: do processing on blockchain
    - PDS: provide answers to questions, not raw data
    - FHE: perform operations directly on encrypted data
- Trust in blockchains: equitably assigning trust and reducing energy consumption
  - Proof-of-**Work** vs Proof-of-**Behavior** or Proof-of-**Stake**
  - https://blockgeeks.com/guides/proof-of-work-vs-proof-of-stake/

Part 6: Conclusion

- Claim success
- Briefly mention coding laws directly into the blockchain

Updates and problems

Updates

- Several wallet services have been hacked; private key management is still an issue
  - https://khannasecurity.com/blog/can-my-blockchain-wallet-be-hacked/
- Bitcoin's value has crashed; cannot assume and should not rely on a stable cryptocurrency
  - https://www.fool.com/investing/2018/02/06/the-cryptocurrency-crash-is-here-and-this-is-whats.aspx
- There have been tons of "blockchain" startups maybe one is doing something like this
  - https://www.nhbr.com/2018/12/21/when-blockchain-meets-data-privacy-and-security/
  - https://medium.com/inside-r3/blockchain-approaches-to-data-privacy-in-healthcare-e6e7f114094c
  - https://www.cmswire.com/information-management/why-enterprises-are-looking-to-blockchain-for-better-data-privacy/

Problems

- Power consumption (they address it some, but not enough)
  - Blockchains are extremely power hungry
  - We need orders of magnitude more efficiency
  - https://arstechnica.com/tech-policy/2017/12/bitcoins-insane-energy-consumption-explained/
- Incentivization
  - Without a relatively stable underlying cryptocurrency, what's the miners' motivation?
  - Does a blockchain require constant processing to remain secure? I.e, if you take over a network, can you change *past* events as well as future events? -> Yes, but requires more resources.
- Control over your data
  - You still can't say *how* your data is used; this only gives you control over who can see your data

Questions?

- TIP: For those studying sensor privacy, the references in this paper could be of use to you
- RESOURCES:
  Blockchain visual demo - https://www.youtube.com/watch?v=_160oMzblY8
  (https://anders.com/blockchain/coinbase.html)