

The Laboratory for Education and Research on Security Assured Information Systems (LERSAIS)



### INFSCI 2150 / TEL 2810 Information Security and Privacy Blockchain Technology in Health Informatics Domain (Lecture 13)

#### Runhua Xu

PhD Student, School of Computing and Information University of Pittsburgh runhua.xu@pitt.edu

### Outline

- Blockchain Technology Overview
  - Cryptographic Concepts
  - Chaining and Block Chaining
  - Networks, trust, consensus
- Smart Contracts
- Applications in Health Informatics Domain
- Labs Overview

## **Blockchain Overview**

John Schmidt-Surlehn gr-6- 93 Conha Subber 15 By Sunday for 2/3 die on London & Cheer Li 12 0-Sun 14 By Sollinghisto for after Harmon sto lender and by from Belliman game 31 By alatiments 21512 154-21512 February 12 Sefends for 119 dagars 147 190 6 185 -164 56. 04 256 nu Sept 7 Joshalf for 20 Bunter thety Coum 10 to & por 30 Bunter & 190-6-006 6033 160 6 3169 13510 4 269 115 Chul 29 To Goods for 58 clouns 100 6 13 Aug 31 By Cash & affer. Harman & Linder -(under 30 By a Be alaterments for bolows dugt - 111. 11 1 18 May 31 By Junding & offic lender bigg. 18 V 640 duct 5 12. 2 -May 15 Sistaffs & 20 Runder Kates We pary 12 to a: . 15 & Schope of Summers 206 64 272 150-2612. 39548 1167 241 2010 112.3 ash July & Togerale for be dans 30548 5 Califor Ti By Sundrus & alf . B. A Gate chimal + & & & & Bur Son 5 115 1123 -M Much SI By Sundrus & War Chilleregne London + & & & day 113 10 131- 213 10 Speed 30 By Cash & alf in Harman & A & 224.14.0 272 11.9 584 Septem 13 A Shifts for 22/Bundles 1994.19.0 299 10015 31 To Stati for 110 daws 2643 117 abul 25 Pogerde for 192 dans allober 92 to a Hadden ston alos adead by here to pay theme 187 18158 NT Och In Productor & 2 Billion lander. Y & Hoda 215.17 \_ 3034 283 37 283 38 Decen 31 Toleros Decen 31 Toleros au Guar 21 To Diradon for after Sed so landen. Conna Count 31 By Shaffy the 39 41 della traft this year no Ocume 31 By Junierow 12.11 10 0 4 1/2 82 Sith Hores 3 --173 151 12075 100-1125-465 05 . 100 Sefler 29 % Cash 17513 (Dearn 31 to Abatements for diret time 413 15 Quan 30 By Shiffs for 58 4 Sillo 17513 16412 6 256 223 10711 285 166 9 nis huter 31 / Cash parthun 936 256 Course 31 To Chernal for The Ob aluit 272 7145 nis ruly 20 a back pain them 1214dille St IN V august 31 To 1-33 Older 31 Je F 3010 Luna 31 To M 5 LEDGERS 271 00 100 nis Gugt 31 to Cash part them 40 10boratory for Education and Research on Security Assured Information Systems (LERSAIS) To abalanced to Bast united 171/11 264

### Transaction Costs and Digital Ledgers

- In 1937, Ronald Coase published an article titled "The Nature of the Firm"
  - Transaction costs in an open market include search, negotiation, execution, and policing or enforcement
  - The basic premise of the paper is that as the costs associated with transactions decreased, firms would shrink in size.
- Digital systems reduce the costs of many types of transactions by orders of magnitude.
- Distributed digital ledgers can be used to further reduce the costs of the ledger that provides for secure trusted transactions.

### What is a Distributed Ledger

- At the most basic level
  - It is simply a replicated data structure e.g. one or more database tables.
- The concept of a distributed public permissionless ledger
  - as first introduced by Bitcoin, was given the name "blockchain".
- The things that make a blockchain special are:
  - It is distributed there are multiple copies that are synchronized
  - It is "public" and "permisionless" anyone can work on it
  - It is tamper resistant additions to the data structure are made in a way that provides reasonable assurance they are not fraudulent
  - There are mechanisms that provide incentives for organizations to contribute to the maintenance of the data structure
- There are other approaches
  - hash trees or Merkle trees or hashgraphs

### The BitCoin Proposal (2008)

- There has been a long search for digital cash
  - a way to exchange funds the same way we do with cash
    - Non-centralized scenario, double pay?
- In 2008, Satoshi Nakamoto published a paper entitled "Bitcoin: A Peer-to-Peer Electronic Cash System"
  - To read the original paper on bitcoin, see <u>https://bitcoin.org/bitcoin.pdf</u>
- The bitcoin proposal provided a mechanism for:
  - "Private" transactions of the kind we normally use cash for.
  - "Trusted" transactions of the kind normally managed by a bank.

### Challenges Ahead – Hype vs Value

#### Gartner Hype Cycle for Emerging Technologies, 2017



#### gartner.com/SmarterWithGartner

Source: Gartner (July 2017) © 2017 Gartner, Inc. and/or its affiliates. All rights reserved.



### **Blockchain Categorization**

- Blockchain networks can be categorized based on their permission model
  - who can maintain them (e.g., publish blocks).
    - If anyone can publish a new block, it is permissionless.
    - If only particular users can publish blocks, it is permissioned
  - a permissioned blockchain network is like a corporate intranet that is controlled, while a permissionless blockchain network is like the public internet, where anyone can participate.

### Blockchain Categorization - Permissionless

- Decentralized ledger platforms open to anyone publishing blocks, without needing permission from any authority
  - Open source software, freely available
  - Any blockchain network user within a permissionless blockchain network can read and write to the ledger
  - Malicious users may attempt to publish blocks in a way that subverts the system
    - To prevent this, permissionless blockchain networks often utilize a multiparty agreement or 'consensus' system
      - Proof of work mthod
      - Proof of stake method
    - Usually, promote non-malicious behavior through rewarding the publishers of protocolconforming blocks with a native cryptocurrency.

### **Blockchain Categorization - Permissioned**

- Permissioned blockchain networks are ones where users publishing blocks must be authorized by some authority (be it centralized or decentralized)
  - may thus allow anyone to read the blockchain or they may restrict read access to authorize
  - may allow anyone to submit transactions to be included in the blockchain or, again, they may restrict this access only to authorized individuals
  - can have the same traceability of digital assets as they pass through the blockchain
  - may also be used by organizations that need to more tightly control and protect their blockchain

## Blockchain – Components

### Cryptography - Cryptology



### Cryptography – Secret Key Encryption



### Secret Key Encryption

- Sender, receiver share common key
  - Keys may be the same, or trivial to derive from one another
  - Sometimes called symmetric cryptography
- Two basic types
  - Transposition ciphers
  - Substitution ciphers
- Product ciphers
  - Combinations of the two basic types

### Secret Key Encryption

- $y = E_k(x)$  : Ciphertext  $\rightarrow$  Encryption
- $x = D_k(y)$  : Plaintext  $\rightarrow$  Decryption
- k = encryption and decryption key
- The functions  $E_k()$  and  $D_k()$  must be inverses of one another
  - $E_k(D_k(y)) = ?$
  - $D_k(E_k(x)) = ?$
  - $E_k(D_k(x)) = ?$

### Why and Why not Public Key Cryptography

- Key establishment
- Non-repudiation
  - Signatures
- Computational effort



### Public Key Cryptography

- Two keys
  - Private key known only to individual
  - Public key available to anyone
- Idea
  - Confidentiality:
    - encipher using public key,
    - decipher using private key
  - Integrity/authentication:
    - encipher using private key,
    - decipher using public one

### **Public Key Encryption**









### Key Size for Security

- 80 bit keys are "safe" to use today, but 128 bit keys are recommended
- Common to see 2048-bit RSA
- Elliptic Curve Cryptography is becoming popular for mobile devices

### **"Secure" Hash Functions**

- One-way (random oracle?)
- Takes any arbitrary input and produces a fixed size output
  - 256 bits for security today
  - Show SHA-256: <u>http://passwordsgenerator.net/sha256-hash-generator/</u>
  - Takes ~2<sup>128</sup> trials to find a collision with probability 0.5
- Cryptographic Nonce
  - cryptographic nonce is an arbitrary number that is only used once
  - hash (data + nonce) = digest

### Transactions

- A transaction represents an interaction between parties.
  - e.g., a transaction could be a way of recording activities occurring on digital or physical assets.
  - e.g., a notional example of a cryptocurrency transaction



### Addresses and Address Derivation

Address

- which is a short, alphanumeric string of characters derived from the blockchain network user's public key
  - using a cryptographic hash function
  - along with some additional data (e.g., version number, checksums)
- Addresses are shorter than the public keys and are not secret.
- How to generate
  - public key  $\rightarrow$  cryptographic hash function  $\rightarrow$  address
  - Each blockchain implementation may implement a different method to derive an address.
- Private Key Storage
  - users must manage and securely store their own private keys.
  - Software -- wallet

### Blocks

- Blockchain network users submit candidate transactions to the blockchain network via software.
- The software sends these transactions to a node or nodes within the blockchain network.
- The chosen nodes may be
  - non-publishing full nodes
  - as well as publishing nodes.
- The submitted transactions are then propagated to the other nodes in the network, but this by itself does not place the transaction in the blockchain.
  - It usually then wait in a queue until it is added to the blockchain by a publishing node.

### Blocks

- A block contains a block header and block data
  - The block header contains metadata for this block.
  - The block data contains a list of validated and authentic transactions which have been submitted to the blockchain network.
    - Validity and authenticity is ensured by checking that the transaction is correctly formatted and that the providers of digital assets in each transaction
    - The transaction have each cryptographically signed the transaction.

### **Chaining Blocks**

- Blocks are chained together through each block containing the hash digest of the previous block's header, thus forming the blockchain.
  - If a previously published block were changed, it would have a different hash.
  - This in turn would cause all subsequent blocks to also have different hashes since they include the hash of the previous block.
  - This makes it possible to easily detect and reject altered blocks



## Blockchain – Consensus Models

### **Consensus Models**

- A key aspect of blockchain technology is determining which node publishes the next block.
  - why would a node propagate a block that another node is attempting to publish?
  - who resolves conflicts when multiple nodes publish a block at approximately the same time?
  - blockchain technologies use consensus models to enable a group of mutually distrusting nodes to work together.

### **Some Properties**

- The initial state of the system is agreed upon
  - e.g., the genesis block
- Users agree to the consensus model by which blocks are added to the system.
- Every block is linked to the previous block by including the previous block header's hash digest.
- Users can verify every block independently.
- A key feature
  - there is no need to have a trusted third party provide the state of the system
  - every user within the system can verify the system's integrity

### Proof of Work Consensus Model

- A user publishes the next block by being the first to solve a computationally intensive puzzle.
  - The solution to this puzzle is the "proof" they have performed work
  - The puzzle is designed such that solving the puzzle is difficult but checking that a solution is valid is easy.
    - enables all other full nodes to easily validate any proposed next blocks
    - any proposed block that did not satisfy the puzzle would be rejected.

### Proof of Work Consensus Model

- A common puzzle method
  - the hash digest of a block header be less than a target value.
    - All minor nodes make many small changes to their block header
    - trying to find a hash digest that meets the requirement
  - The target value may be modified over time to adjust the difficulty (up or down) to influence how often blocks are being published.

### Proof of Work Consensus Model

SHA256("blockchain" + Nonce) = Hash Digest starting with "000000"

```
SHA256("blockchain0") =
0xbd4824d8ee63fc82392a6441444166d22ed84eaa6dab11d4923075975acab938
(not solved)
```

```
SHA256("blockchain1") =
0xdb0b9c1cb5e9c680dfff7482f1a8efad0e786f41b6b89a758fb26d9e223e0a10
(not solved)
```

```
SHA256("blockchain10730895") =
0x00000ca1415e0bec568f6f605fcc83d18cac7a4e6c219a957c10c6879d67587
(solved)
```

To solve this puzzle, it took 10,730,896 guesses

...

### Proof of Stake Consensus Model

- Based on the idea that the more stake a user has invested into the system
  - the more likely they will want the system to succeed, and the less likely they will want to subvert it
  - Stake is often an amount of cryptocurrency that the blockchain network user has invested into the system
  - Proof of stake blockchain networks use the amount of stake a user has as a determining factor for publishing new blocks
    - the likelihood of a user publishing a new block ← → the ratio of their stake to the overall blockchain network amount of staked cryptocurrency

### **Other Consensus Models**

- Round Robin Consensus Model
- Proof of Authority/Proof of Identity Consensus Model
- Proof of Elapsed Time Consensus Model

### Ledger Conflicts and Resolutions

- node\_A creates block\_n (A) with transactions #1, 2 and 3. node\_A distributes it to some nodes.
- node\_B creates block\_n (B) with transactions #1, 2 and 4. node\_B distributes it to some nodes.
- There is a conflict.
  - o block\_n will not be the same across the network.
    - block\_n(A) contains transaction #3, but not transaction #4.
    - block\_n (B) contains transaction #4, but not transaction #3.

### Ledger Conflicts and Resolutions



Most blockchain networks will wait until the next block is published and use that chain as the "official" blockchain, thus adopting the "longer blockchain".

### **Blockchain Limitations and Misconceptions**

- Immutability
- Users Involved in Blockchain Governance
- Blockchain Death never be fully shut down
- Cybersecurity
- Cyber and Network-based Attacks
- Malicious Users
- Resource Usage

### **Smart Contracts**

### Overview

- Smart contracts extend and leverage blockchain technology.
- A smart contract is a collection of
  - code and data (sometimes referred to as functions and state)
    - that is deployed using cryptographically signed transactions on the blockchain network
      - (e.g., Ethereum's smart contracts, Hyperledger Fabric's chaincode).
    - Smart contracts must be deterministic
  - The smart contract is executed by nodes within the blockchain network
    - all nodes that execute the smart contract must derive the same results from the execution
    - the results of execution are recorded on the blockchain.

### Definition

# A smart contract is a **computer program** executed in a **secure environment** that **directly controls digital assets**

### **Computer Program**

A computer program is a collection of instructions that performs a specific task when executed by a computer.

A computer requires programs to function, and typically executes the program's instructions in a central processing unit. if HAS\_EVENT\_X\_HAPPENED() is true:
 send(party\_A, 1000)
else:
 send(party\_B, 1000)

-- Wikipedia

### Properties of Secure Environments

- Correctness of execution
  - The execution is done correctly, is not tampered
- Integrity of code and data
- Optional properties
  - Confidentiality of code and data
  - Verifiability of execution
  - Availability for the programs running inside

# Examples of Secure Environments

- Servers run by trusted parties
- Decentralized computer network (ie. blockchains)
- Quasi-decentralized computer network (ie. consortium blockchains)
- Servers secured by trusted hardware (e.g. SGX)



### **Directly Control - Example**

- Legal contract
  - "I promise to send you \$100 if my lecture is rated 1\*"
- Smart contract:
  - "I send \$100 into a computer program executed in a secure environment which sends \$100 to you if the rating of my lecture is 1\*, otherwise it eventually sends \$100 back to me"

### What are digital assets?

- A broad category
  - Domain name
  - Website
  - Money
  - Anything tokenisable (e.g. gold, silver, stock share etc)
  - Game items
  - Network bandwidth, computation cycles

### **Digital Assets**

Rank \$	Project ÷	Category ÷	Platform ÷	Campaign end date	Campaign target +	Amount raised +
1	Star Citizen	Video game	Kickstarter, independent	Ongoing	\$500,000	\$90,009,649
2	Elio Motors	Automotive - Low-cost, high mileage vehicle	Independent	Ongoing	-	\$21,161,869
3	Pebble Time	Smartwatch	Kickstarter	Mar 27, 2015	\$500,000	\$20,338,986
4 🤇	Ethereum	Cryptocurrency	Bitcoin, Independent	Sep 2, 2014	-	\$18,439,086
5	Coolest Cooler	Product Design	Kickstarter	Aug 29, 2014	\$50,000	\$13,285,226

Ethereum Foundation sold 60,102,206 digital tokens which will be useful in a decentralized network



#### Ethereum: the first blockchain-based smart contract platform

- Blockchain with expressive programming language
  - Programming language makes it ideal for smart contracts
- Why?
  - Most public blockchains are cryptocurrencies
    - Can only transfer coins between users
  - Smart contracts enable much more applications

### How Ethereum Works

- Two types of account:
  - Normal account like in Bitcoin
    - has balance and address
  - Smart Contract account
    - like an object: containing (i) code, and (ii) private storage (key-value storage)
    - Code can
      - Send ETH to other accounts
      - Read/write storage
      - Call (ie. start execution in) other contracts

### A Simple Smart Contract

```
contract SimpleStorage {
    uint storedData;
    function set(uint x) public {
        storedData = x;
    }
    function get() public view returns (uint) {
        return storedData;
    }
}
```

### **Ethereum Languages**



### Example

What other see on the blockchain



### Smart Contract – Privacy Issues

- Ethereum blockchain guarantees correctness and availability, not privacy for smart contracts
  - Everything on the Ethereum blockchain is public
    - Cannot execute on private data (e.g. death will remains secret until the owner dies)
- Transactions are traceable



### Smart Contract – Privacy

- Hawk (Kosba et al. IEEE S&P'16)
  - Privacy-Preserving Smart Contracts
  - Execute confidential, fair, multiparty protocols
- ZeroCash over Ethereum
  - Mixing coins with others

### Smart Contract – Scalability

- Resources on blockchain are expensive
  - Full nodes perform the same on-chain computations
  - Full nodes store the same data
- Gas-limit is relatively small
  - Can't run an OS on blockchain
  - Can't increase gas-limit: DoS vector

### **Blockchain in Healthcare**

### Blockchain: Opportunities for Health Care

A New Model for Health Information Exchanges

	HIE Pain Points	Blockchain Opportunities
00	<b>Establishing a Trust Network</b> depends on the HIE as an intermediary to establish point-to-point sharing and "book-keeping" of what data was exchanged.	<b>Disintermediation of Trust</b> likely would not require an HIE operator because all participants would have access to the distributed ledger to maintain a secure exchange without complex brokered trust.
\$	<b>Cost Per Transaction</b> , given low transaction volumes, reduces the business case for central systems or new edge networks for participating groups.	<b>Reduced Transaction Costs</b> due to disintermediation, as well as near-real time processing, would make the system more efficient.
٩	Master Patient Index (MPI) challenges arise from the need to synchronize multiple patient identifiers between systems while securing patient privacy.	<b>Distributed framework for patient digital identities,</b> which uses private and public identifiers secured through cryptography, creates a singular, more secure method of protecting patient identity.
	Varying Data Standards reduce interoperability because records are not compatible between systems.	<b>Shared data</b> enables near real-time updates across the network to all parties.
-	<b>Limited Access to Population Health Data,</b> as HIE is one of the few sources of integrated records.	<b>Distributed, secure access</b> to patient longitudinal health data across the distributed ledger.
Yn.	<b>Inconsistent Rules and Permissions</b> inhibit the right health organization from accessing the right patient data at the right time.	<b>Smart Contracts</b> create a consistent, rule-based method for accessing patient data that can be permissioned to selected health organizations.

Ē

### Illustrative Healthcare Blockchain Ecosystem



### Blockchain as an Enabler of Nationwide Interoperability

	On Chain Data	Off Chain Data		
Data Types	• Standardized data fields containing summary information in text form (e.g. age, gender)	• Expansive medical details (e.g. notes) and abstract data types (e.g. MRI images, human genome)		
Pros	• Data is immediately visible and ingestible to all connected organizations, making blockchain the single source of truth	• Storage of any format and size of data		
Cons	• Constrained in the type and size of data that can be stored	<ul> <li>Data is not immediately visible or ingestible, requiring access to each health care organization's source system for each record</li> <li>Requires Off-Chain micro-services and additional integration layers</li> <li>Potential for information decay on the blockchain</li> </ul>		

### **Challenges and Considerations**

- Scalability constraints
  - tradeoffs between transaction volumes and available computing power
- Data Standardization and Scope
- Adoption and Incentives for Participation
- Costs of Operating Blockchain Technology
- Regulatory Considerations

# Open-source landscape map for healthcare-related blockchains

#### Healthcare-related blockchain projects



Have an update for the map? DM @andreacoravos or submit a pull request via our GitHub repo.

### Blockchain and Health IT: Algorithms, Privacy, and Data

- A new system architecture in development at MIT
  - the OPAL/Enigma project3
  - which creates a peer-to-peer network
    - enables parties to jointly store and analyze data with complete privacy
    - based on highly optimized version of multi-party computation with a secret-sharing
  - an auditable, tamper-proof distributed ledger (a permissioned blockchain) records and controls access through smart contracts and digital identities.

#### Blockchain and Health IT: Algorithms, Privacy, and Data



#### Blockchain and Health IT: Algorithms, Privacy, and Data

- OPAL/Enigma applied to precision medicine
  - Precision Medicine Initiative (PMI) is to set "the foundation for a new way of doing research that fosters open, responsible data sharing with the highest regard to participant privacy."
    - the data repository nodes at each of the clinical sites (clinical sites function as the data owner) would require a local solution to abstract the clinical data in real time from electronic health records (EHR)
    - the data could queried securely by others through permissioned queries; the data remains always encrypted at the clinical center/data repository node which generated it.
    - This creates a new paradigm for broadening transparency and data sharing of clinical trial data

### A Case Study for Blockchain in Healthcare: "MedRec"

- MIT Media Lab, Beth Israel Deaconess Medical Center
- Problem Domain and Solution
  - Distributed Consent, Record Locator Service
  - Use Blockchain to store PPR and encounter info to create a "patient history" on the chain, and act as a notification directory by determining the list of people that have touched a patient.
- Target Industry
  - Care organizations
- Primary Participants
  - Providers, Patients, Researchers

Source: ONC/CCC Use of Blockchain in Health IT and Health-Related Research Challenge (2016) A Case Study for Blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data, Ariel Ekblaw, et al

### A Case Study for Blockchain in Healthcare: "MedRec"

Blockchain: Ethereum

Content Stored on Chain:

Registration Contract - MRN to Ethereum address mapping;

Patient-Provider Contract (PPC) - hash of a subset of clinical data contained at provider, along with provider endpoint info, Summary Contract (SC)

Rewards for Participation: Clinical Research orgs provide compute, and can submit queries as reward mechanisms to participating organizations (eg. "Give me a count of patients currently taking XYX medications")



Figure 1. MedRec smart contracts on the left, showing data content for each contract type. Sample relationship graph between contracts and network nodes on the right.

### Lab Overview

### Lab Overview

- Two Labs
  - Decentralized Application using Smart Contracts and IPFS
  - Access Control and Security Issues in Smart Contracts

### Thanks.