Module 4 – Introduction to Blockchains

Takashi Sugimura, PhD
University of Hawaii Maui College

taskasug@Hawaii.edu
@uhmcbait
maui.hawaii.edu/cybersecurity
What is Blockchain .... Bitcoin ?
Who is Satoshi Nakamoto ?

- Bitcoin revolutionized our concepts of currency and money. ... 2008 Height of the banking crisis
- Like the currency he created, the identity of Bitcoin’s creator Satoshi Nakamoto is virtual, existing only online.
Blockchain in some form will revolutionize transaction systems.
Software, Game Theory and Cryptography

Separately these fields have existed for a long time, but for the first time they have intersected and morphed into *Blockchain Technology*.

*Game Theory* is the “study of mathematical models of conflict and cooperation between intelligent rational decision makers”.

Is related to Blockchain because the Bitcoin Blockchain (public) originally conceived by Satoshi Nakamoto had to solve a known game theory conundrum called the *Byzantine Generals Problem*.
The Byzantine Generals Problem (1982)

LESLEY LAMPORT, ROBERT SHOSTAK, and MARSHALL PEASE
SRI International

Reliable computer systems must handle malfunctioning components that give conflicting information to different parts of the system. This situation can be expressed abstractly in terms of a group of generals of the Byzantine army camped with their troops around an enemy city. Communicating only by messenger, the generals must agree upon a common battle plan. However, one or more of them may be traitors who will try to confuse the others. The problem is to find an algorithm to ensure that the loyal generals will reach agreement. It is shown that, using only oral messages, this problem is solvable if and only if more than two-thirds of the generals are loyal; so a single traitor can confound two loyal generals. With unforgeable written messages, the problem is solvable for any number of generals and possible traitors. Applications of the solutions to reliable computer systems are then discussed.

Categories and Subject Descriptors: C.2.4. [Computer-Communication Networks]: Distributed Systems—network operating systems; D.4.4 [Operating Systems]: Communications Management—network communication; D.4.5 [Operating Systems]: Reliability—fault tolerance

This research was supported in part by the National Aeronautics and Space Administration under contract NAS1-15428 Mod. 3, the Ballistic Missile Defense Systems Command under contract DASG60-78-C-0046, and the Army Research Office under contract DAAG29-79-C-0102. Authors’ address: Computer Science Laboratory, SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025.
Cryptography Science is used in multiple places to provide security for a Blockchain network and rests on three basic concepts:

- Hashing
- Keys
- Digital signatures

Software has been developed that combines Cryptography with Game Theory to produce the overall constructs of Blockchain, where seeming uncertainty is mitigated with overwhelming mathematical certainty.
Blockchains Fit in the Technology Revolution

• Pre 1993 .... Internet browser introduced
  • Database (IBM 1970) applications/ data storage, analytics
  • Centralized networks / client –server systems

• Internet Era (1993-2010)
  • email / e- Commerce
  • Decentralized networks
  • Social networks

• 2010 - Present
  • High performance computing/ GPU
  • Distributed networks / cloud computing
  • Big Data/ AI and Machine Learning
  • Blockchain/ Bitcoin (10 years ago)
IBM’s database model stood unchanged until about 10 years ago, when the Blockchain came into this space with a radical new proposition:

- **What if your database worked like a network** — a network that’s shared with everybody in the world, where anyone and anything can connect to it?

- Blockchain calls this “decentralization.”

  - offers the promise of nearly friction-free cooperation between members of complex networks
  - adds value to each other by enabling collaboration without central authorities and middle men.
networks

<table>
<thead>
<tr>
<th>Centralized Systems</th>
<th>Decentralized Systems</th>
<th>Distributed Systems</th>
</tr>
</thead>
</table>

No central nodes, peer-to-peer

Legend
- Blue circle: Centralized Server
- Green circle: Server connected to main central server
- Red circle: Individual nodes or terminals
Blockchain Process... Decentralization

No central nodes .... All the nodes are not connected to each other

All nodes could be miners

> The blockchain network is a peer-to-peer network of independent nodes communicating together by message broadcasting.
Distributed Consensus

At any given time:

- All nodes have a sequence of blocks of transactions they’ve reached consensus on.
- Each node has a set of outstanding transactions it’s heard about.

(Not all nodes have all the information at the same time (Latency))

Why consensus is hard:

Nodes may crash
Nodes may be malicious

Network is imperfect
- Not all pairs of nodes connected
- Faults in network
- Latency
- No notion of global time
Bitcoin consensus works better in practice than the Byzantine Generals Problem (BGP theory)

• The theory is important because it can help predict unforeseen attacks
• **Consensus works better for Bitcoin due to the additional conditions**
  • Offers incentives
  • Embraces randomness
    • Doesn’t have a specific endpoint
    • Consensus happens over a longer time scale..... 1 hour
How Blockchain Works
Here are five basic principles underlying the technology.

1. Distributed Database
   • Each party on a blockchain has access to the entire database and its complete history.
   • No single party controls the data or the information. Every party can verify the records of its transaction partners directly, without an intermediary.

2. Peer-to-Peer Transmission
   • Communication occurs directly between peers instead of through a central node.
   • Each node stores and forwards information to all other nodes.
3. Transparency with Pseudonymity
• Every transaction and its associated value are visible to anyone with access to the system. (public key)
• Each node, or user, on a blockchain has a unique 30-plus-character alphanumeric address that identifies it. (private key)
• Users can choose to remain anonymous or provide proof of their identity to others. (signatures) Transactions occur between blockchain addresses.

4. Irreversibility of Records
• Once a transaction is entered in the database and the accounts are updated, the records cannot be altered, because they’re linked to every transaction record that came before them (hence the term “chain”).
• Various computational algorithms and approaches are deployed to ensure that the recording on the database is permanent, chronologically ordered, and available to all others on the network.
5. Computational Logic

- The digital nature of the ledger means that blockchain transactions can be tied to computational logic and in essence programmed.
- **users can set up algorithms and rules that automatically trigger transactions between nodes.**

  - Data Security
    - Keys
    - Signatures
    - Hashing
    - Redundancy
    - Improved workflow
signature = f(message, private key)  ....unique for every transaction

Verify = f( signature, message, public key) = { true or false}
Hash function

• Takes any string as input
• Creates a fixed size output (256 bits for Bitcoin)
• Efficiency Computable

• Security properties
  1. Collision – free
  2. Puzzle- friendly.... consensus
Hash property 1: Collision-free

Nobody can find x and y such that
\[ x \neq y \text{ and } H(x) = H(y) \]

If \( H(x) = H(y) \) then \( x = y \)

compare hashes instead of long messages
Collisions do exist ...

possible inputs → possible outputs

How to find a collision:

- try $2^{130}$ randomly chosen inputs
- 99.8% chance that two of them will collide

This works no matter what $H$ is ...

... but it takes too long to matter
The SHA (Secure Hash Algorithm) is one of a number of cryptographic hash functions. ... - algorithm that generates an almost-unique, fixed size 256-bit (32-byte) hash. (a “bit” is the smallest unit of storage 1 or 0/ a byte is 8 bits & encodes a single character \(256/8 = 32\))

<table>
<thead>
<tr>
<th>Data</th>
<th>SHA-256 hash</th>
</tr>
</thead>
<tbody>
<tr>
<td>tak</td>
<td>ddeb49eb1ed10bb3bf2c3e6070c56b3e3755701d0c0c72c5e63d15a314f3e29</td>
</tr>
<tr>
<td>tak sugimura</td>
<td>ec168a4a3ae4f63716ba4e78f2c26573ec584babbd25ad46c091e10224b0f19</td>
</tr>
</tbody>
</table>

- Creates a unique character set of fixed length / hash
- Not encryption... cannot be decrypted to original text
- The entire Library of Congress can be “hashed”

Regardless of the length of the data (entire Library of Congress).... The length of the hash value is fixed.
Blockchain Demo: Interactive:  https://anders.com/blockchain/blockchain.html

Block 1: Genesis Block
req: start with “0000”
In the demo there is an example of 3 “Peers” and a demonstration of consensus management.
The hash output, or fingerprint, is actually what’s used as the “previous block” reference.

• One result of this is that there’s no way to switch out a block in the middle of the chain, because the hash value for the new block would be different, and the next’s block reference would no longer point to it.

• Even more importantly, a block cannot be solved before the previous block is solved. The previous block reference is part of the text that goes through the hash function, so any changes to it would require resolving.
solving a block involves trying to get the cryptographic hash of the block to be below a certain value, and you do that by trying different random numbers. Once solved, the hash output is like a fingerprint that uniquely identifies that block. If even a single character in the block is changed, the block’s hash would be completely different.
What Do Bitcoins “Look” Like?

1454A2geTxaJwF8eqry70LECdomgDSj6Zx

Public Key (“Address”)
34 characters starting with 1 or 3
Represents a possible destination for payment

5JHkYd4mYkTsCsF5axnFj573PG6tqpeJ39Rz2M33vwBka4S1hu6

Private Key
51 characters starting with 5
Required to transfer value from the address
Miners are nodes with high performance computing resources that resolves consensus.

**China: 60%**
- F2Pool
- AntPool
- BTCC
- BTC
- Bixin
- BW

**Other Miners**
- BTCC Pool: 7.2%
- F2Pool: 7.1%
- Rixin: 7.7%
- BW.COM: 3.9%
- SlushPool: 3.5%
- BitClub Network: 4.5%
- BitFury: 5.7%
- ViaBTC: 6.1%
- BTC.com: 10.4%
- BTC.TOP: 11.4%
China mines the most bitcoins of any nation and therefore also exports the largest number of bitcoins. This has been driven by the following:

- **China has a weak banking system** and has been fueled by a cash economy.

- China has the largest supercomputers in the world and pioneered the use of "graphic" chips in parallel processing supercomputing which are ideal for processing Bitcoins. They are producing a special chip for Bitcoin processing. **Cheap electricity**

- **Most of the world’s largest, mining pools are based in China.** Chinese Bitcoin miners have gained an advantage by capturing a large percentage of Bitcoin’s hash power.
Global Supply Chain: A larger Transaction

The “life-cycle” of a supply chain for a retail business

Physical “nodes”
- Raw Materials
- Producer
- Distributor
  - Shipping/customs
  - Freight Forwarder
  - Warehouse
- Taxes
- Retail Facility
- Customer
Example: Financial services....Banking

• The banking industry is filled with shared resources. Consider ATM machines: each machine is owned by a single institution, but accepts cards from a huge network. This sharing requires a complicated management apparatus mostly managed by VISA

• The four major credit card networks in the United States are Visa, Mastercard, American Express and Discover. Visa is the largest.
  • In 2014, Visa's U.S. credit purchase volume was $1.2 trillion, $1.1 trillion in 2013. Total credit card volume was $4 trillion
  • There were 304 million Visa credit cards in circulation in the United States and 545 million Visa credit cards in circulation outside of the United States

• If ATMs were created today Blockchain would be used
Why Blockchain is currently not used in Financial Transactions:

• There are over **700 cryptocurrencies** being traded on various online exchanges
• Bitcoin transaction processing is restricted to **7 transactions per second (tps)**, compare this with Visa processing speeds which has a peak capacity of around **56,000 transactions per second**
• Most people don’t use cryptocurrencies to move value around because the infrastructure necessary for such **trustless transactions don’t scale well in their current form**.

• **Example: pay for your Starbucks using bitcoins.**
  • Once you made the payment you will have to wait until your transaction is validated by bitcoin’s blockchain network @ ~7 tps.
  • Your wait times are positively correlated with how many others are paying with bitcoins at that very moment as well as the number of active public nodes validating the transactions
  • you could be waiting anywhere from **20 minutes to 2 hours**. (a block is processed every 10 minutes)
  • **Your coffee will go cold before your payment is confirmed.**
**Hashgraph** is solving for the scaling problem through their DAG (Direct Acylic Graph), which works by combining the gossip protocol with a voting algorithm that enables Hashgraph to reach consensus quickly and securely **without proof of work**.

• Instead of the **3–7 transactions per second (TPS)** that the Bitcoin blockchain enables, or the **10–20 TPS** that Ethereum can handle, **Hashgraph promises to process hundreds of thousands of TPS**, surpassing Visa’s current network transaction speeds.

• Major credit cards like Visa average about 2,000 transactions a second, with peaks reaching 50,000 transactions a second, while maintaining service standards of 7 seconds max.

• **Hashgraph** has achieved **100,000 transactions a second**, across 8 regions, with a latency of just 3.3 seconds.

*Those speeds will revolutionize the types of applications that will be able to run on a Blockchain.*
Major Blockchain Platforms

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Ethereum</th>
<th>Hyperledger Fabric</th>
<th>R3 Corda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description of platform</td>
<td>Generic blockchain platform</td>
<td>Modular blockchain platform</td>
<td>Specialized distributed ledger platform for financial industry</td>
</tr>
<tr>
<td>Governance</td>
<td>Ethereum developers</td>
<td>Linux Foundation</td>
<td>R3</td>
</tr>
<tr>
<td>Mode of operation</td>
<td>Permissionless, public or private</td>
<td>Permissioned, private</td>
<td>Permissioned, private</td>
</tr>
<tr>
<td>Consensus</td>
<td>Mining based on proof-of-work (PoW), Ledger level</td>
<td>Broad understanding of consensus that allows multiple approaches (i.e., notary nodes), Transaction level</td>
<td>Specific understanding of consensus (legal prose), Transaction level</td>
</tr>
<tr>
<td>Smart contracts</td>
<td>Smart contract code (e.g., Solidity)</td>
<td>Smart contract code (e.g., Go, Java)</td>
<td>Smart contract code (e.g., Kotlin, Java)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smart legal contract (legal prose)</td>
</tr>
<tr>
<td>Currency</td>
<td>Ether, Tokens via smart contract</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Currency and tokens via chaincode</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Comparison of Ethereum, Hyperledger Fabric and Corda
Amazon Web Services launches ‘blockchain templates’

- Amazon is the world cloud leader (distributed computing) with AWS
- Amazon launched AWS Blockchain Template, on **April 19th, 2018**
- Amazon wants to make it easier for developers to create projects based on the blockchain technology underlying bitcoin and other cryptocurrencies, with “**blockchain templates**”.
- The product gives users pre-set blockchain frameworks that support two versions of the technology: **Ethereum** and the **Linux Foundation’s Hyperledger Fabric**.
- The templates allow users to create and launch secure blockchain network using open source frameworks in a fast and easy way.

**Google is working on blockchain-related technology** to support its cloud business and head off competition from emerging startups that use the heavily-hyped technology to operate online in new ways, according to people familiar with the situation.
Business Model: Amazon, Google, Microsoft IBM

• **The technologies that is essential for Blockchain applications**..... Supercomputers/ GPUs.... Cloud computing/ Distributed systems..... AI/ Machine learning

• Just as the model of Blockchain developments being driven by “open source” platforms such as IBM/Hyperledger, AI machine learning software and algorithms that are powering all of Google’s AI activity — TensorFlow — is now being given away for free. TensorFlow is now an open source software project available to the world.

• **Why?...** There are **20 million** organizations in the world that could benefit from machine learning today. If millions of companies use this best in class free AI software then they are likely to need lots of computing power.....

• **Who?...** Google Cloud is of course optimized for TensorFlow and related AI services .... Amazon AWS once you become reliant on their software and their cloud you become a very sticky customer

• **Amazon, Google, Microsoft, IBM** are competing for distributed computing applications .....**Blockchain** may not be as important as AI but exists in the same space
Who Will Build the Health-Care Blockchain?

Decentralized databases promise to revolutionize medical records, but not until the health-care industry buys in to the idea and gets to work.

MIT Technology Review September 2017
Observations:

• 26 different electronic medical records systems used in the city of Boston, each with its own language for representing and sharing data.

• Critical information is often scattered across multiple facilities, and sometimes it isn’t accessible when it is needed most—a situation that plays out every day around the U.S., costing money and sometimes even lives. …tailor-made for a blockchain to solve

• Instead of payments, this blockchain would record critical medical information in a virtually incorruptible cryptographic database, maintained by a network of computers, that is accessible to anyone running the software
Researchers at the MIT Media Lab have developed a prototype system called using a private blockchain.

It automatically keeps track of who has permission to view and change a record of medications a person is taking. It also solves a key issue facing just about anyone who wants to take blockchain outside the realm of digital currency: miners.

- With Bitcoin and other cryptocurrencies, miners use computers to perform calculations that verify data on the blockchain—a crucial service that keeps the system functioning. In turn, they’re rewarded with some of that currency.

- The system incentivizes miners—generally medical researchers and health-care professionals—to perform the same work by rewarding them with access to aggregated, anonymized data from patients’ records that can be used for medical studies (as long as patients consent).

- Data is the most important commodities in AI applications and insurance, pharmas and other companies would pay for the use of the data…. In a similar way that Facebook, Google, financial institutions do today.
Blockchain Governance:
• On-Chain
• Off-Chain