Extending Blockchains For Risk Management

Scan this for slides

Raj Jain

Barbara J. and Jerome R. Cox, Jr. Professor

Washington University in St. Louis

jain@wustl.edu

Keynote at the 2nd International Conference on Blockchain Technology (ICBCT), Hawaii, March 13, 2020

Slides and Video Recording at:

http://www.cse.wustl.edu/~jain/talks/pbc_icb.htm

Overview

1. Current Trends in Blockchain
2. Strengths and weaknesses of the current blockchains
3. Extending Blockchains to 4.0: Converting data to knowledge
4. Applications of Probabilistic Blockchains for risk management

Blockchains

- Blockchain is the technology that made Bitcoin secure
- Blockchain was invented by the inventor of Bitcoin
- After Bitcoin became successful, people started looking into the technology behind Bitcoin and found:
  - Blockchain is the key for its success
  - Two complete strangers can complete a transaction/contract a third party

Example of a Contract: Wedding

Example of a Contract: Wedding

Centralized Trust

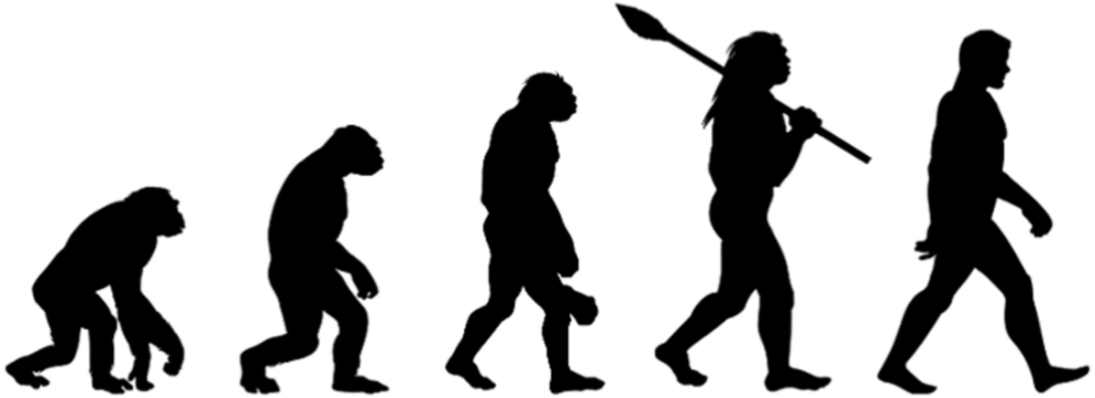
- Centralized registry
- Single point of failure
- Easier to hacked

Decentralized Trust

- Decentralized
- No single point of failure
- Very difficult to hack

Origins of Decentralization

Decentralized



Centralized

- Original humanoids had a decentralized life
- It became centralized only after they started living in villages/cities

Trend: Decentralized – No Central Point of Trust

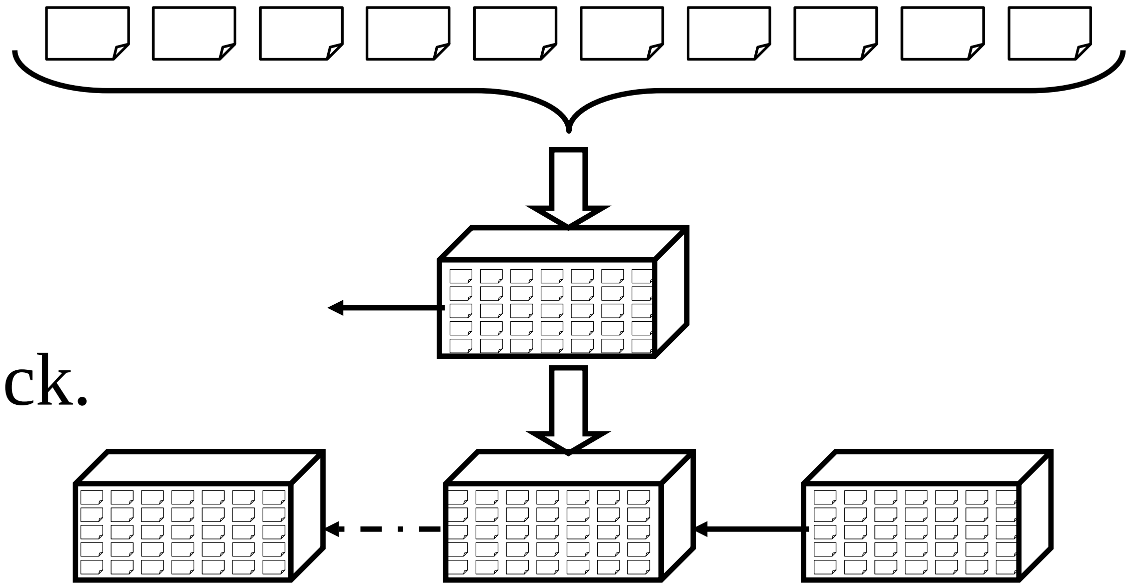
- **Trend**: Make everything decentralized with no central point of trust
- Two perfect strangers can exchange money, make a contract without a trusted third party
- Decentralized systems are
  1. More secure: Attack tolerant
  2. No single bottleneck
  3. More reliable: Fault tolerant
  4. No single point of control ⇒ No monopoly
- Blockchain is one way to do this among **untrusted multi-domain** systems.

Time is a cycle: Decentralized vs. Centralized debate

Examples of Centralized Systems

- **Banks**: Allow money transfer between two accounts
- **City Records**: Wedding registers, Property ownership
- **Networks**: Certificate Authorities, DNS
- In all cases:
  - There is a central third party to be trusted
  - Central party maintains a large database
    ⇒ Attracts Hackers
  - Central party may be hacked ⇒ affects millions
  - Central party is a single point of failure. Can malfunction or be bribed

Blockchain Process



1. **Users** broadcast signed transactions or smart contracts

2. **Mining nodes** validate transactions and create blocks. Point to previous block.

3. **Blockchain nodes** validate blocks and construct a chain

- There are many users, many mining nodes, and many blockchain nodes. More nodes ⇒ Better. Less ⇒ Blockchain not required/useful.

Key Strengths of Blockchains

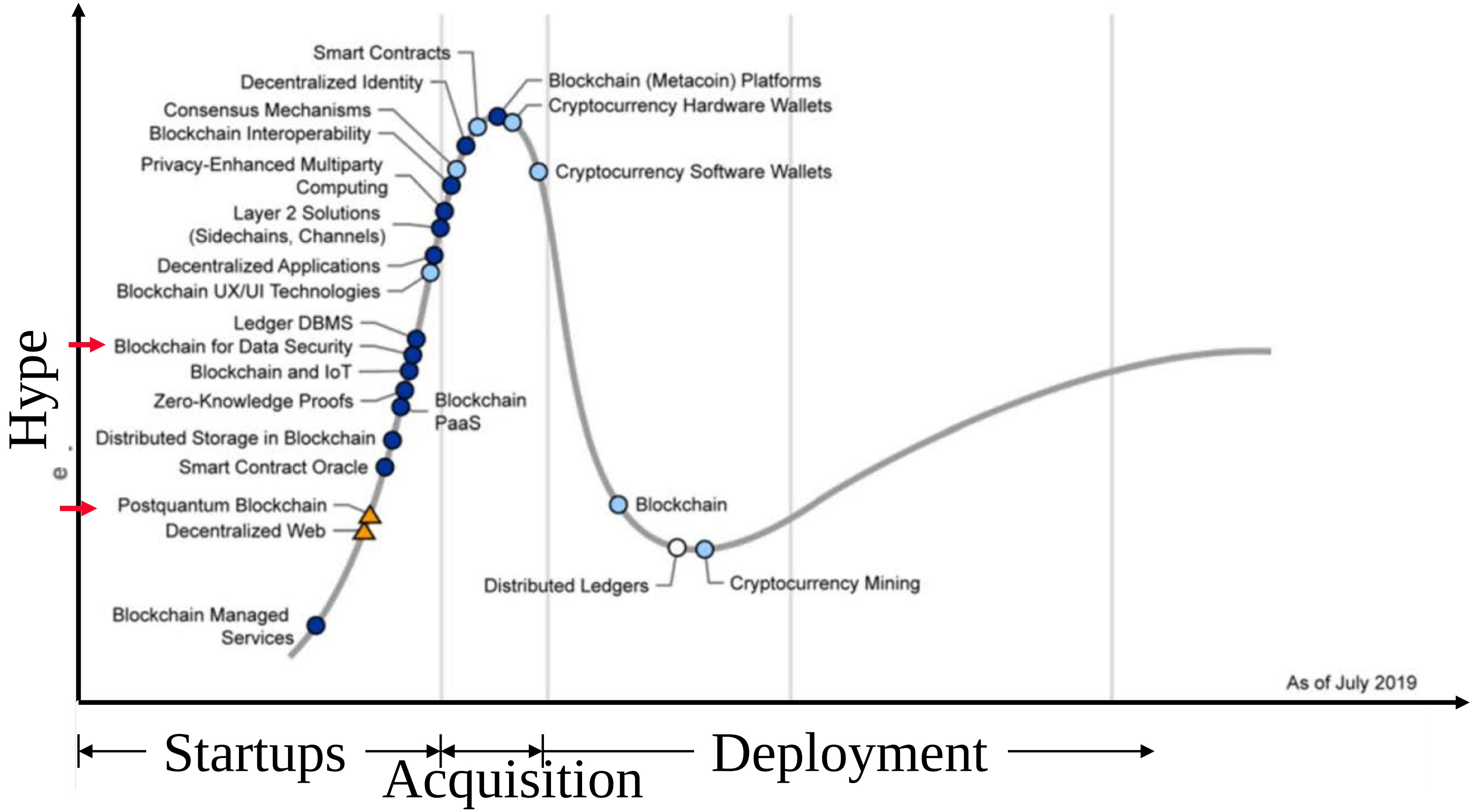
1. **Distributed**: No single point of failure
2. **Decentralized Consensus**: Transactions valid only if agreed by majority
3. **Trustless**: Transacting or processing parties do not need to trust
4. **Cryptographic Security**: Elliptic Curve Cryptography
5. **Non-Repudiation Guarantee**: All transactions are signed

Networking Application 1: PKI

- **Certificate Authorities** (CA): Issue certificates – Heart of SSL
- If a CA is hacked, all certificates issued by it are invalid
- March 2011: A hacker tricked Comodo to issue certificates for Google, Yahoo, Microsoft, …
- Sep 2011: Dutch CA DigiNotar was hacked
- Several fraudulent certificates were issued
- DigiNotar declared bankruptcy
- The hacker claimed he had infiltrated 4 other CAs

Ref: Tara Salman, Maede Zolanvari, Aiman Erbad, Raj Jain, and Mohammed Samaka, "**Security Services Using Blockchains: A State of the Art Survey**" IEEE Communications Surveys and Tutorials, First Quarter 2019, Volume 21, Issue 1, 858-880 pp., http://www.cse.wustl.edu/~jain/papers/bcs.htm

Hype Cycle for Blockchain Technologies, 2019



Ref: A. Litan and A. Leow, "Hype Cycle for Blockchain Technologies, 2019," Gartner Report ID G00383155, July 11, 2019.

Key Strengths of Blockchains

1. **Distributed**: No single point of failure
2. **Decentralized Consensus**: Transactions valid only if agreed by majority
3. **Trustless**: Transacting or processing parties do not need to trust
4. **Cryptographic Security**: Elliptic Curve Cryptography
5. **Non-Repudiation Guarantee**: All transactions are signed

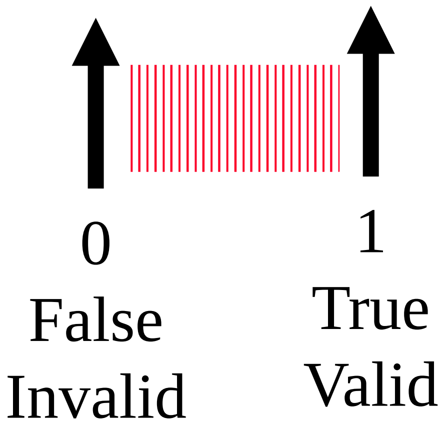
Can the Blockchains be Enhanced?

Limitation 1: Only facts are recorded

- Alice is married to Bob
- Alice gave 20 coins to Bob
- Alice signed a contract with Bob to pay 10 coins for 1 kg of xx.

Limitation 2: Binary Validity

- All transactions recorded on the blocks that are committed are valid
- Those not on the committed blocks and old are invalid
- So the recording is binary: only 0 or 1.

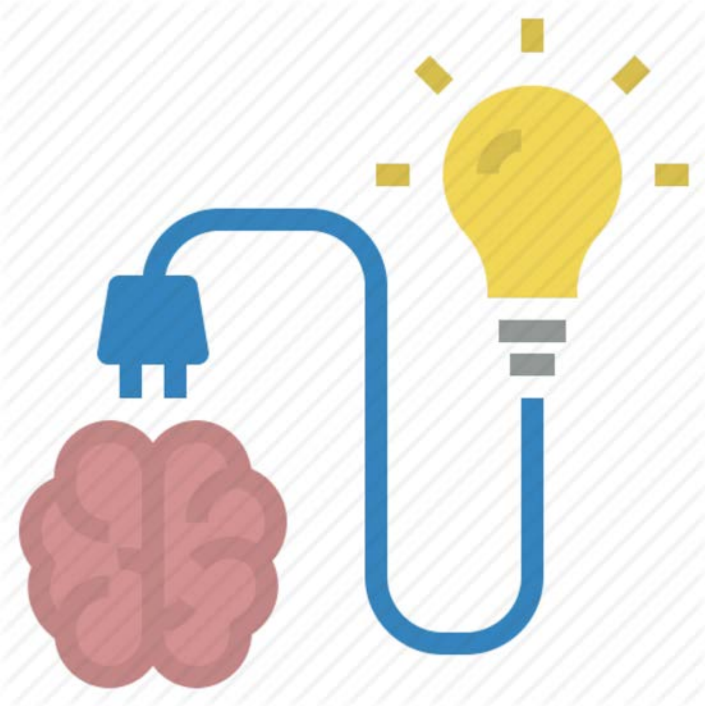


Limitation 3: Deterministic Events only

- Can not record that I am only 90% sure that Alice gave 20 coins to Bob.

Ideas to Enhance Blockchains

- ❑ Blockchain is just a distributed **data storage** of valid transactions
- ❑ All transactions are ***deterministic***
- ❑ What’s Wrong?
  - ❑ Need to convert data to knowledge
  - ❑ We are in big data and machine learning age
  - ❑ Real life is probabilistic
  - ❑ Most to the decisions we make are probabilistic
    ⇒ All decisions have some risk



Risk Propels Progress

- Banks take money from risk-averse savers and give them interest
- Banks invest the money in corporations ⇒ Takes the country forward
- Venture capitalists take risk by investing in half-cooked ideas
- Startups take risk by working in unchartered territories

Decisions with Risk

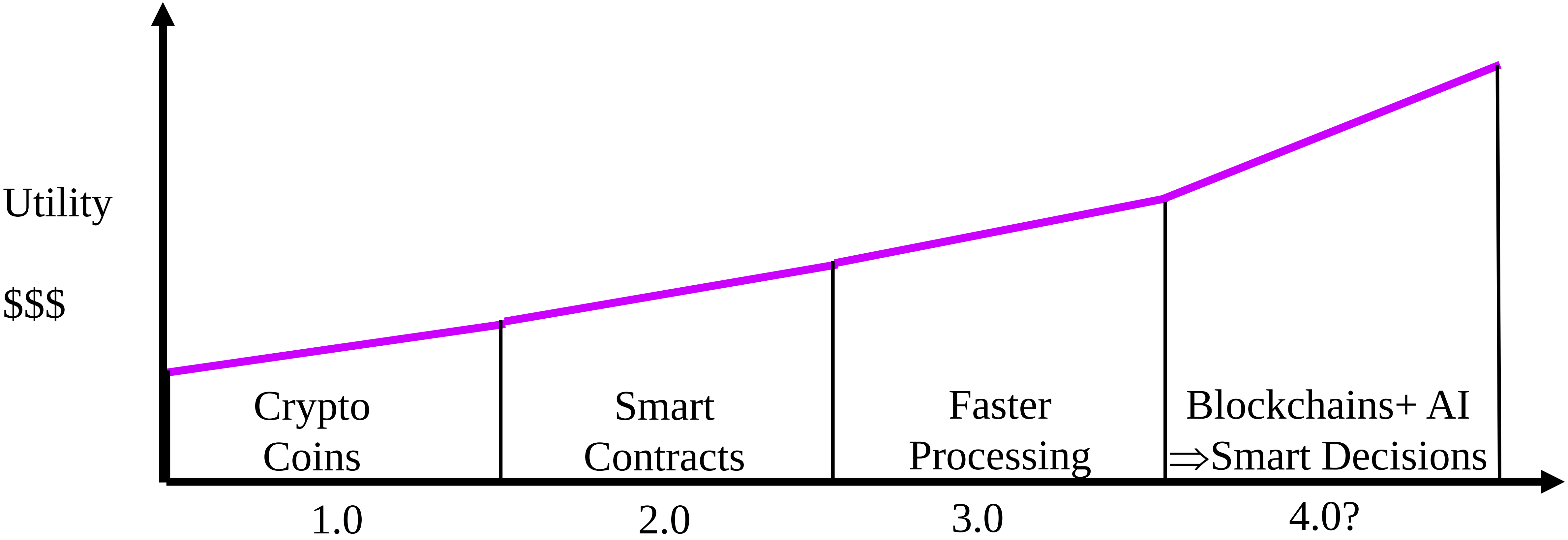
- Sell insurance
- Buy insurance
- Sell a stock
- Buy a stock
- Download a software application on your computer
- Update Windows
- Marry someone
- Travel to some place

Our Goal

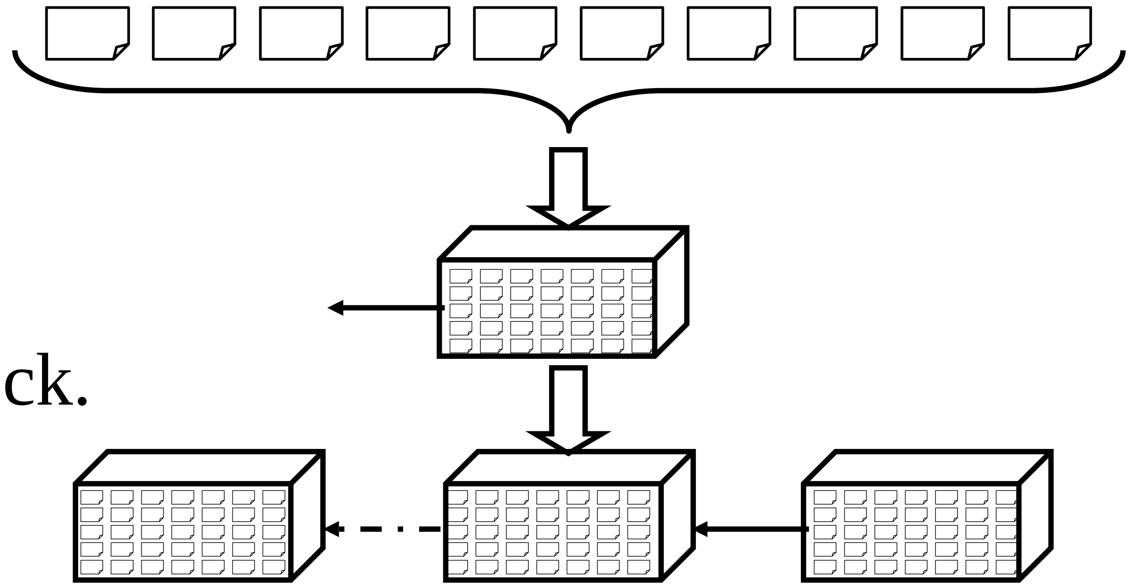
- Moving the chain from deterministic to **probabilistic**
- Moving the chain from storage to **computation**
- Moving the chain from data to **knowledge**
- Moving the chain from information to **decision making**

- Google is moving from “Search” to “Suggest” using AI
- A blockchain that provides knowledge
  – A **knowledge chain** would be more useful

Blockchain Generations



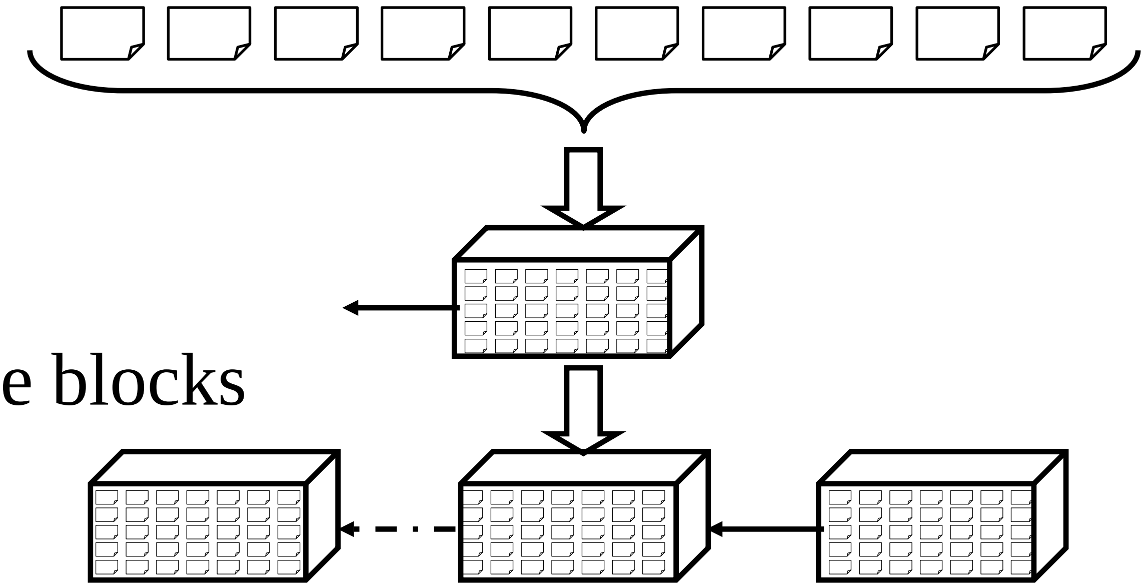
Blockchain Process



1. **Users** broadcast signed transactions or smart contracts
2. **Mining nodes** validate transactions and create blocks. Point to previous block.
3. **Blockchain nodes** validate blocks and construct a chain

- ❑ There are many users, many mining nodes, and many blockchain nodes. More nodes ⇒ Better. Less ⇒ Blockchain not required/useful.

Probabilistic Blockchain Process



1. **Agents** broadcast transactions,
   Transactions = Opinions/decisions

2. **Mining nodes** validate transactions,
   create a knowledge summary and create blocks

3. **Blockchain nodes** validate blocks
   and construct a chain

- Two types of users:
  - **Agent nodes** provide their probabilistic opinions/decisions
  - **Management nodes** that inquire the blockchain and use it for decisions

Probabilistic Blockchain Example

- **Issue**: Whether Cisco stock will go up tomorrow?
- i^{th} Agent says that the probability that it will go up is p_i
- Summary of all opinions related to this issue is:

$$P[\text{Stock will rise}] = G(\{p_1, p_2, \ldots, p_n\})$$

Here, G is the summarizing function

- For Example: Group decision is the first moment of individual decisions

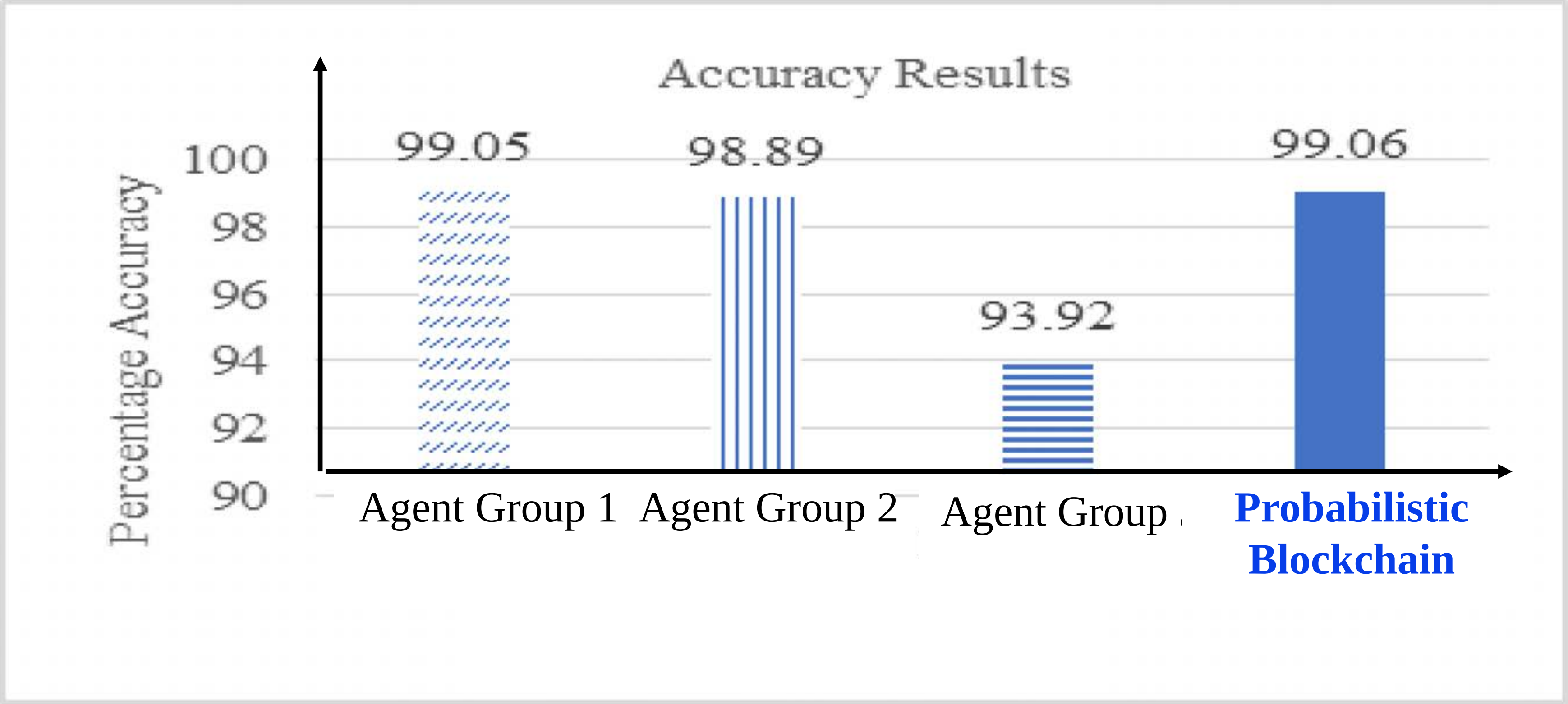
$$P = \frac{1}{n}\sum p_i$$

Ref: T. Salman, R. Jain, and L. Gupta, "**Probabilistic Blockchains: A Blockchain Paradigm for Collaborative Decision-Making**," 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON 2018), New York, NY, Nov. 8-10, 2018, 9 pp., http://www.cse.wustl.edu/~jain/papers/pbc_uem.htm

Empirical Validation

- Issue: Whether a network traffic pattern represents intrusion
- 1000 Agents give their decisions: Yes or No
  - Agents randomly pick one of the 3 machine learning algorithms:
    - Random Forest, Decision Tree, Logistic Regression
- Mining nodes summarize these decisions using the majority function

Results



Distributed decision making is better than any individual decision

Generalizing the Summary Function

- Summary can be any other reasonable function of individual decisions:
  - 90-percentile
  - Median
  - Mode
  - 2nd Moment
- Summary can be a vector: {1st moment, 2nd moment, …, *n*th moment}
- Summary can be the result of any **statistical** algorithm
- Summary can be the result of a **data mining** algorithm
- Summary can be the result of a **machine learning algorithm**

Blockchain 4.0: Database to Knowledge Base

- Blockchain = Distributed database of smart contracts
- Probabilistic blockchain = Knowledge + database
- **Database**: Who bought, who sold, what quantity, what price, what time
- **Knowledge**:
  - Where the market is going?
  - Whether we should buy, sell, or hold?

Knowledge Chain

- Customer query to blockchain network:
  How is the Cisco stock doing today?
- Blockchain to Customer: With 60% confidence, the probability of stock rising is 90%, …
- Ideal for **large** distributed systems with **no national boundaries**, no exchange limitations, no brokers in between
- Crowd-sourced knowledge, crowd source decisions

Application Examples

1. Spam from Email/IP Addresses/Cloud providers/source/public IP
2. Intrusions/attacks from IP Addresses. Anonymously share attack information.
3. Gray domains: Share gray list among agents.
4. Reliability/Issues with recent software updates
5. Error/reliability statistics of network/IoT devices
6. Virus in software

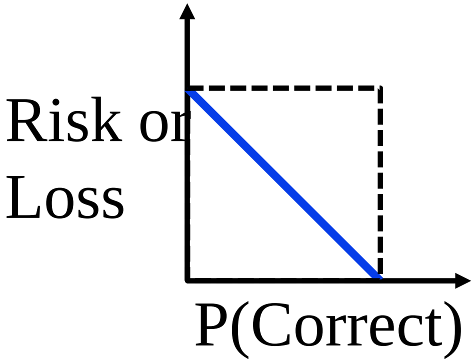
Issues

1. **Summary functions**
2. **Overhead of consensus mechanisms**: Proof of Work, Proof of Stake, …
3. **Reputation of Experts and Bad Actors**:
    - Some agents are more knowledgeable than others
    - Group decisions should give more weight to them
    - How to incentivize better agents?
    - How to penalize bad actors?

Ref: Tara Salman, Raj Jain, Lav Gupta, "**A Reputation Management Framework for Knowledge-Based and Probabilistic Blockchains**," IEEE 1st International Workshop on Advances in Artificial Intelligence for Blockchain (AIChain 2019), held in conjunction with the 2019 IEEE International Conference on Blockchain, Atlanta, July 14, 2019, http://www.cse.wustl.edu/~jain/papers/rpmcewa.htm

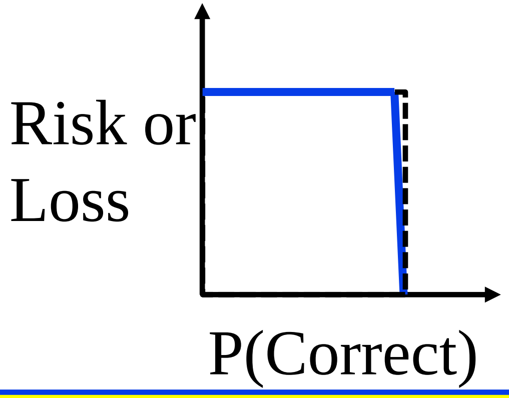
Risk of Incorrect Decisions

1. Identifying objects:



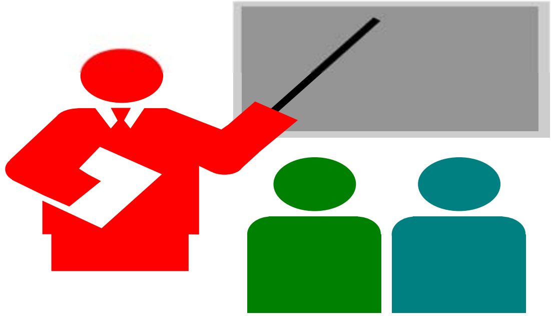
$$P(\text{Correct}) = \text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$$

2. Security:



Summary functions and risk metrics are application dependent

Summary



1. Blockchains provide an immutable, secure, distributed database
2. Three generations: Crypto currency, Assets, Smart contract
3. All three generations are deterministic and **only provide storage**
4. The next generation needs to **connect computation and AI** to make knowledge/decisions in addition to data storage
5. Consensus can be probabilistic result of any statistical algorithm, data mining, or machine learning ⇒ **Knowledge Chain**

Related Papers

- Tara Salman, Raj Jain, Lav Gupta, "**A Reputation Management Framework for Knowledge-Based and Probabilistic Blockchains**," IEEE 1st International Workshop on Advances in Artificial Intelligence for Blockchain (AIChain 2019), held in conjunction with the 2019 IEEE International Conference on Blockchain, Atlanta, July 14, 2019, http://www.cse.wustl.edu/~jain/papers/rpmcewa.htm
- Tara Salman, Raj Jain, and Lav Gupta, "**Probabilistic Blockchains: A Blockchain Paradigm for Collaborative Decision-Making**," 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON 2018), New York, NY, November 8-10, 2018, 9 pp., http://www.cse.wustl.edu/~jain/papers/pbc_uem.htm
- Tara Salman, Maede Zolanvari, Aiman Erbad, Raj Jain, and Mohammed Samaka, "**Security Services Using Blockchains:A State of the Art Survey**" IEEE Communications Surveys and Tutorials, Accepted September 2018, 28 pp., http://www.cse.wustl.edu/~jain/papers/bcs.htm

List of Acronyms

❑ AI	Artificial Intelligence
❑ DNS	Domain Name Service
❑ IEEE	Institution of Electrical and Electronics Engineers
❑ IoT	Internet of Things
❑ IP	Internet Protocol
❑ PKI	Public Key Infrastructure
❑ SSL	Secure Socket Layer

Scan This to Download These Slides

Raj Jain

Jain@wustl.edu

Slides and Video Recording at:

http://www.cse.wustl.edu/~jain/talks/pbc_icb.htm