ECE 595: Foundations of Blockchain Systems **September 30, 2020**

Lecture 1

Lecturer: Sreeram Kannan *Scribe: Viswa Virinchi Muppirala*

In this lecture, we are going to give an overview of Blockchains, why they are interesting, performances of current systems, a functional decomposition of blockchain systems, and two security models. One of the most popular examples of Blockchain is the cryptocurrency Bitcoin [2]. Bitcoin was launched in 2009 by an unknown entity with a pseudonym Satoshi Nakamoto which provides security in a decentralized manner. While anyone can join and leave, security of blockchain has been proven in the past 11 years as it has accumulated a market of more than 100 Billions, and people are daily using it for trades. Before starting to analyze what makes blockchain secure, we want to step back and explain why blockchain is an interesting subject to study.

1.1 Why Blockchain

We are going to answer this question with three arguments, starting with the evolutionary argument.

1.1.1 Evolutionary Argument

We make use of arguments from the book Sapiens [1] where the author argues that humans, among the millions of species in the world are so successful because

> Humans cooperate flexibly in large numbers

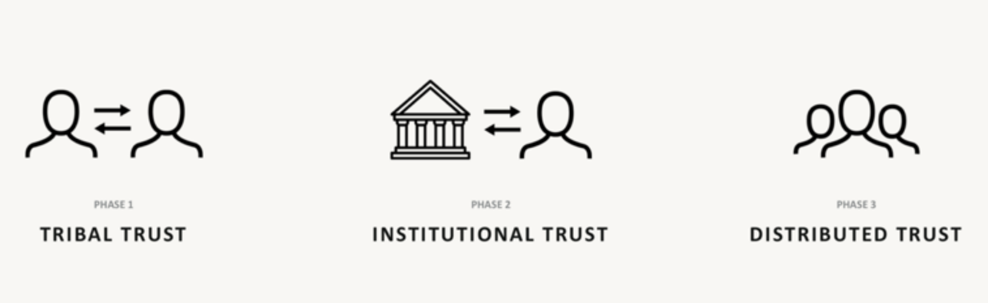


Figure 1.1: Phases of trust

Other species like ants are in large numbers, but they don't cooperate flexibly and their cooperation is rather genetically wired. In contrast, humans can flexibly rearrange their cooperation, thus can easily make their cooperation in large numbers. When the argument is applied to technology, any technology that enables people to cooperate flexibly in large number will have a distinctive evolutionary advantage.

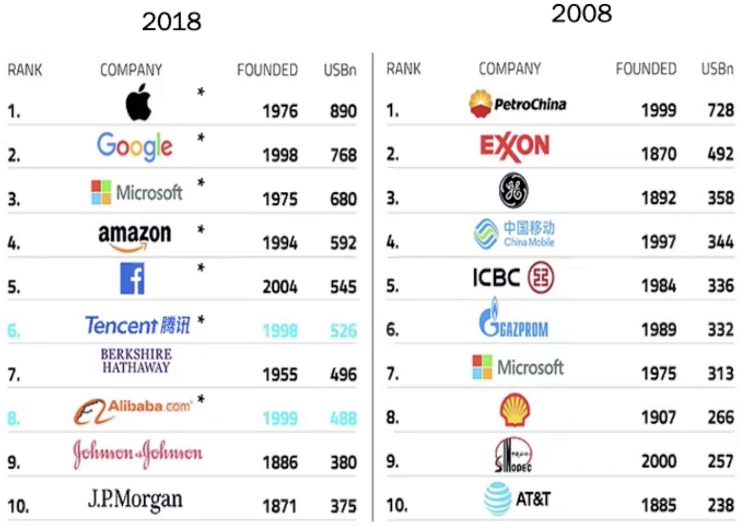
Cooperation between entities banks on trust, and to take the argument further we are going to look at how trust between humans has evolved over the time in three phases in Fig 1.1. In the first phase of human evolution, the trust was tribal where small groups of people cooperated with each other and built trust. In this phase the trust is between each pair of people creating N^2 relations. Although this already gave us the evolutionary advantage but it wasn't enough to carry out feats such as invention of the internet, putting a man on the moon or photographing a black hole.

Only through an institutional trust the mentioned things were possible. Institutional trust brings people in large scale to co-operate with each other through an institution that they trust. An example would be employees who have never met each other working on a product (say, the next Windows launch) at Microsoft. Another example would just be money as an institution where people work together for money. In this phase of trust, the trust flows through the institution making it the bottle neck. If the small group of people through which the trust flows operate in a way other than what is prescribed, then the institutional trust can collapse.

The alternative and the third phase of trust would be distributed trust where the trust emerges out of a large scale of people. Like tribal trust, we do have the N^2 problem where an individual can only keep track of a small scale of people and cannot maintain all N links. If at all such a thing is possible, it will have an immense evolutionary advantage. Blockchain deals with a similar kind of trust model and this argument makes Blockchain interesting.

1.1.2 Economic Argument

One of the most important things in a civilization is commons such as roads, trains and public schools etc. And in a digital setting we have digital commons or digital intermediaries we all inhabit. The examples are Zoom, Twitter, Facebook and Twitter etc. which all intermediate our conversations. Compared to 2008, there are a lot more technology companies in the top 10 list, see Fig 1.2. The economic argument here is that these digital commons are controlled by private companies, and it creates a huge friction when the commons run on their agency's agenda rather than serving a neutral agenda. Those companies have immense values because of locked network effects which is a phenomenon describing the difficulty to jump from an existing big digital platform to a new supposedly better platform. And the value of the company is through this effect. So, this is our economic argument of answering the question ”why blockchain?” because one way of looking at Blockchain is as a digital intermediary without a intermediary.



2018				2008			
RANK	COMPANY	FOUNDED	USBn	RANK	COMPANY	FOUNDED	USBn
1.	*	1976	890	1.	PetroChina	1999	728
2.	Google *	1998	768	2.	EXXON	1870	492
3.	Microsoft *	1975	680	3.	GE	1892	358
4.	amazon *	1994	592	4.	中国移动 China Mobile	1997	344
5.	f *	2004	545	5.	ICBC	1984	336
6.	Tencent 腾讯 *	1998	526	6.	GAZPROM	1989	332
7.	BERKSHIRE HATHAWAY	1955	496	7.	Microsoft	1975	313
8.	Alibaba.com *	1999	488	8.	🐚	1907	266
9.	Johnson & Johnson	1886	380	9.	SINOPEC	2000	257
10.	J.P.Morgan	1871	375	10.	AT&T	1885	238

Figure 1.2: Top companies in 2018 and 2008

1.1.3 Technical Argument

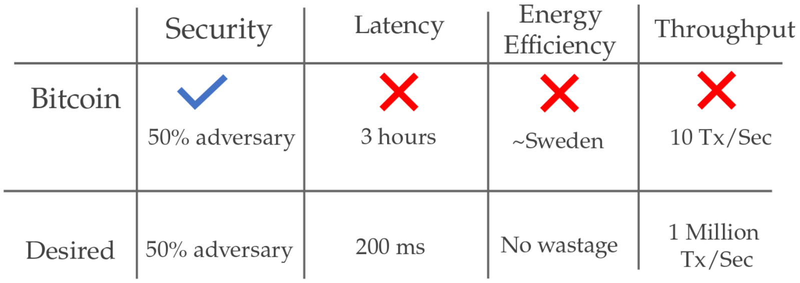
There is a central role for theoretical analysis for security in blockchain. The theoretical analysis are extremely subtle where a protocol transitions from being safe to unsafe through changing of one of the parameters of the protocol. We will look at the analysis and phase transitions in the future lectures. There is also a central role for randomness which is guided by the amount of resource. For example, in Bitcoin the resource is computational power, where the probability of creating a block depends on the computational power. This mechanism is called proof-of-work and similarly there are other mechanisms like proof-of-space and proof-of-stake where the resources are memory and number of coins you own. Along with basic cryptography and distributed computing, ideas from various fields like Information Theory, Coding Theory, Networking, Game Theory, Sharp threshold analysis, P2P design, Incentives, Typicality and Distributed storage are used in Blockchain.

1.2 Background on Blockchain

Blockchain is a new field by itself but historically the field has been called distributed computing. One of the problems people study in this field is to handle crash fail, that the system has to work correctly even if few nodes in the system go off. People also study another model called the adversary resistance or Byzantine fault tolerance where the nodes can not only be faulty but also adversarial. Both systems they considered are permissioned networks with fixed number of nodes. IBM's hyperledger and Facebook's libra are couple of examples where only authorized nodes are allowed to participate the network. Where as, Blockchain is permisionless where anyone can join the network. Examples include Bitcoin, Ethereum and Chia. The application of blockchain goes beyond cryptocurrencies to whatever applications that nodes need to cooperate with each other under a decentralized digital intermediary. In this class, we will focus more on permissionless blockchain.

1.3 Performance of Blockchain and the Blockchain trilemma

Applications like payments (10K Tx/sec), social networking (10K tweets/sec), IoT (100K Tx/sec), prediction market (1000 Tx/sec), gaming assets (1000 Tx/sec) and exchanges (1M Trades/sec) demand a very high throughput. But Blockchain like ethereum only offer a throughput of 20 Tx/sec creating a huge gap that needs to be improved to run those applications. The following table 1.3 analyses the performance of Bitcoin comparing it with desired numbers. Bitcoin has a high latency, low throughput and it spends as much power as an entire country. Bitcoin remains secure as long as less than 50% resource is controlled by the adversarial nodes. We will carefully define what security means in the future lecture; in addition, we will also try to design a blockchain system that achieves all the desired qualities throughout the course by using various ideas from different areas. Another way of summarizing this problem according to the blockchain community is the blockchain trilemma depicted in Fig 1.4. It's a trilemma between Decentralization, Security and Scalability, where one of the aspects must be sacrificed in exchange for the other two properties. Decentralization means the power of each node should be reflected by the amount of resource owned by the node, like memory or Hash power; Scalability means the system should perform better as more nodes join the system; Security is discussed above that the system is secure against less than 50% adversarial power. Bitcoin and Ethereum achieve security and decentralization but do poorly in terms of scalability, so adding more nodes does not improve their system



	Security	Latency	Energy Efficiency	Throughput
Bitcoin	✓ 50% adversary	✗ 3 hours	✗ ~Sweden	✗ 10 Tx/Sec
Desired	50% adversary	200 ms	No wastage	1 Million Tx/Sec

Figure 1.3: Blockchain Today

performance. Protocols like EOS, Tron and Ripple have achieved scalability and security by sacrificing decentralization which roughly operate like Amazon Cloud or Microsoft Azure. Protocols like IOTA and Ghost achieve scalablility and decentralization by sacrificing on security; and traditionally there were other applications in the internet falling into this category. Bittorrent, a peer-to-peer file sharing system that used to occupy half of the internet traffic in 2008 was decentralized and achieved scalability but didn't have security. We'll address this trilemma throughout the course and prove that it's indeed possible to achieve all three in a blockchain.

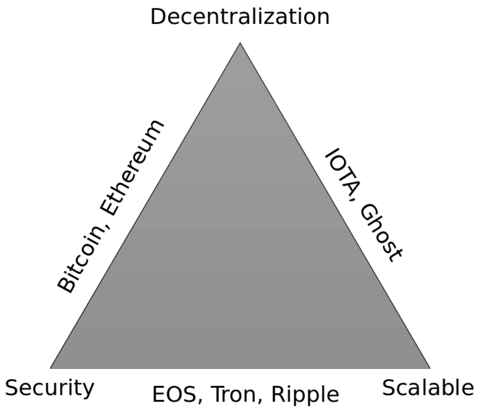


Figure 1.4: Blockchain Trilemma

1.4 Layers of Blockchain

The blockchain consists of different layers, each with its own function. Starting from the bottom, we have them summarised in Fig 1.5.

1.4.1 Peer2Peer

Peer2Peer layer maintains the network connectivity and provides certain primitives, like the ability to broadcast and multicast blocks. The goal of this layer is to optimize latency and bandwidth. Ideas from machine learning, multi-armed bandits and also rumor spreading problem from applied probability come up in the network design.

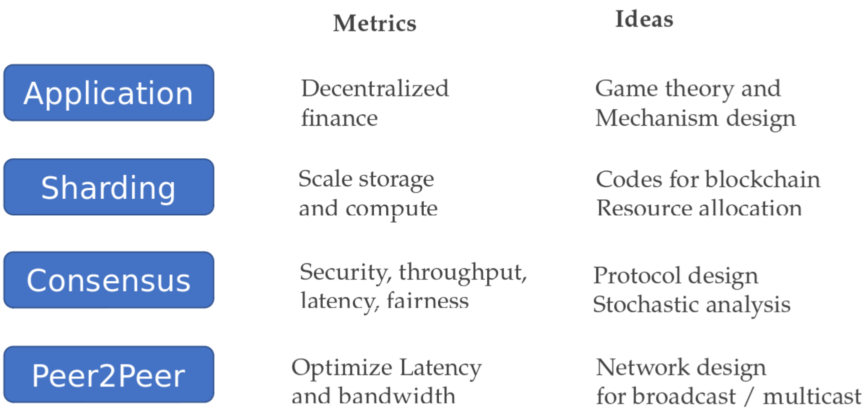

Metrics
Ideas
Application
Decentralized finance
Game theory and Mechanism design
Sharding
Scale storage and compute
Codes for blockchain Resource allocation
Consensus
Security, throughput, latency, fairness
Protocol design Stochastic analysis
Peer2Peer
Optimize Latency and bandwidth
Network design for broadcast / multicast

Figure 1.5: Layers of Blockchain

1.4.2 Consensus

Consensus layer has the details about what the users in the protocol agree on. The goal of the layer is to provide security, throughput, latency and fairness. Ideas from stochastic analysis come up in this layer as the participation of the nodes is stochastically controlled by the amount of resources the nodes have.

1.4.3 Sharding

The sharding(scaling) layer takes care of scalability by resource allocation. The goal of this layer is to provide scalability by dividing the tasks between different nodes. Ideas from coding theory, erasure codes, dynamic game theory are seen here.

1.4.4 Application

One of the examples is decentralized finance, and we will look at more examples throughout the course. For example, Stable coin designs a cryptocurrency by adjusting interest rate and deposit rate in a purely decentralized manner, so that its value does not oscillate.

1.5 Security models in Blockchain

There are two models by which we are going to look at security in blockchain. In an adversarial model, there are two types of nodes: adversarial and honest. Adversarial nodes can do whatever they want, and the rest of node are honest(obedient) that follow the protocol. In the distributed computing context, this model is also called Byzantine Fault tolerant model(BFT model).
In a rational model, each node acts on its self-interest based on the incentives. We asks questions like ”is there a dominant strategy?”, ”is there a Nash-equilibrium?”, ”Do we need to assume that nodes collude in small coalitions?” and ”How large can a collusion be in order for the protocol to be at Nash-equilibrium or to have a dominant strategy?”. In reality, it might be the case that some nodes are adversarial, some are rational and others are honest, but that gets too complicated for analysis, so we will stick to these two models and certify a protocol secure if it is secure in both the models. Even though we are dealing with two models, we will spend more time on adversarial

model, and make connections that sometimes the core of a proof in the adversarial model can be carried over to the rational model.

References

[1] Yuval N. author Harari. *Sapiens : a brief history of humankind*. First U.S. edition. New York : Harper, [2015], [2015]. Includes bibliographical references and index.

[2] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system," http://bitcoin.org/bitcoin.pdf.