BLOCKCHAIN TECHNOLOGY

Jamie Munro
University of St Andrews
School of Computer Science

***Abstract* – After providing an overview of blockchain technology and its applications, security considerations are discussed and a critical evaluation is provided.**

***Keywords* – Blockchain, Smart Contracts, Security, Cryptocurrency.**

***Word Count (main body)*: 3490**

1. INTRODUCTION

In its purest form, a blockchain is a distributed and verifiable database where all participants of the chain can be sure that they have identical data. Every transaction that occurs on the blockchain is executed and shared by every participant. Once data has been recorded on the blockchain, it can never be erased. Every transaction that has ever occurred on the blockchain is verifiably recorded [1].

The concept of blockchain was invented by an unknown individual (or group) known as Satoshi Nakamoto in 2008. The motivation behind the creation of blockchain was to create a digital currency. Previous digital currencies were plagued by a problem known as "double spending" whereby an attacker can exploit vulnerabilities in the system to spend a unit of currency more than once. Previously the only available solution was to resort to some central authority for verification. Nakamoto rejected this solution and wanted to create a decentralised and trust-less digital currency that did not rely on central authorities (governments, banks). To this end, Nakamoto presented bitcoin and blockchain in their 2008 paper [2]. Bitcoin remains the most popular application of blockchain technology to date.

In this paper, the concept and workings of the blockchain will be explored in Section 2. Following this, Section 3 will discuss the various applications of blockchain technology. Section 4 will discuss the security considerations that arise from blockchain and finally Section 5 will evaluate the technology.

2. BLOCKCHAIN CONCEPTS

Blockchain is an attempt to solve the traditional database synchronisation problem across a peer-to-peer network using a distributed consensus algorithm. Mathematics and cryptography are relied upon to secure the network [3].

2.1. DESIGN PRINCIPLES

Decentralisation – The most fundamental concept of the blockchain. The blockchain should not rely on any single centralised authority.

Transparency – All data and operations that occur on the blockchain should be visible and auditable to all users.

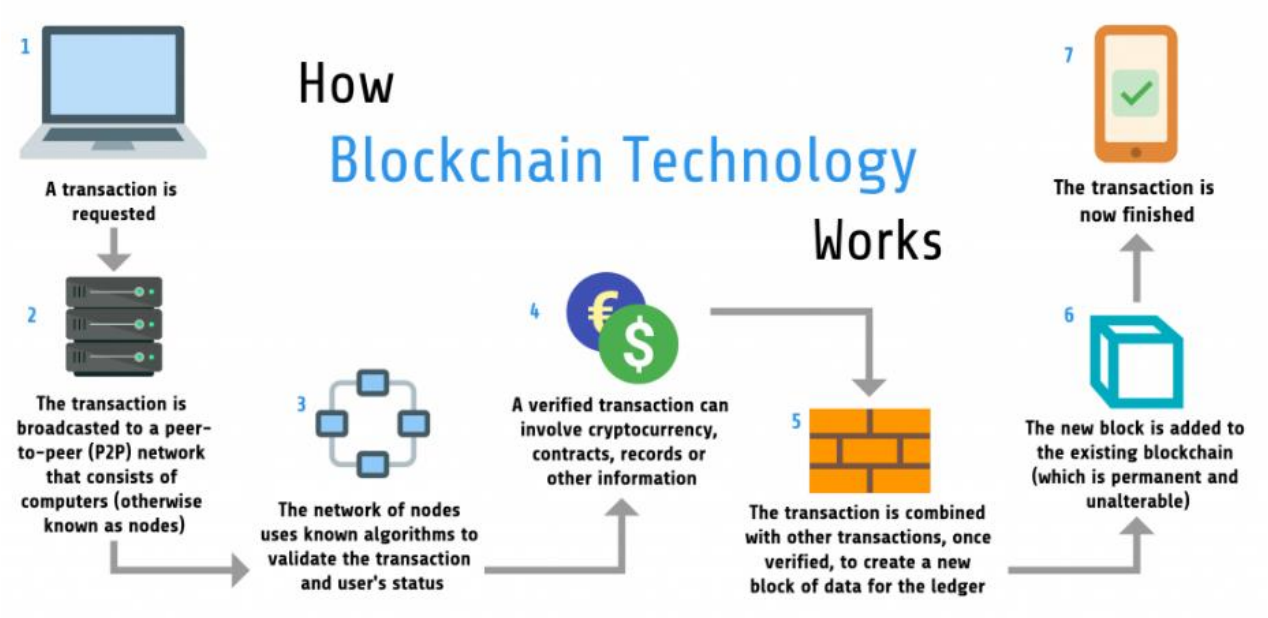

FIGURE 1 SIMPLIFIED BLOCKCHAIN MECHANICS [10]

Trust-less – Users should not need to trust any central authority or individual.

Immutability – Once data is stored on the blockchain it should not be possible to erase it and any changes to the data should be recorded so that there is an audit trail.

Anonymity – Users should not need to reveal their real world identities and can be identified simply through a unique alphanumeric identifier.

Open-Source – Users should be able to inspect the code that the blockchain runs on so that it can be verified. Anyone should be able to build a technology or application on top of the blockchain.

[3]

2.2. Blockchain Mechanics

New records are added to the blockchain through the following process:

1. Senders broadcasts a new record to the network
2. The record is added to a pool of pending transactions
3. Verifiers select a number of records from the pool, validates them and adds them to their local block
4. Each verifier competes for their block to be accepted by the network. There are numerous algorithms that can be used to settle the contest
5. Every node in the network uses a consensus algorithm to decide which block to confirm
6. Every node in the networks adds the confirmed block to their chain

2.3. Blockchain Structure

The blockchain is made up of three fundamental elements: blocks, nodes and verifiers.

2.3.1. BLOCKS

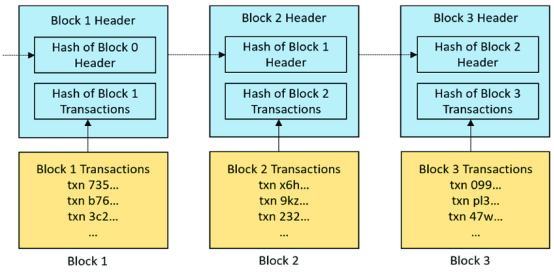

FIGURE 2 CHAINING BLOCKS TOGETHER [11]

The blockchain is composed of a number of sequential "blocks". Each block is linked to the preceding block forming a chain, hence the name blockchain.

Each block consists of a batch of data and a header containing a hash of the previous block, a hash of the current block and a timestamp. This information can be supplemented with additional fields depending on what the blockchain is being used for [3].

The batch of data will also vary on the application of the blockchain, for example the data in bitcoin's blockchain would be a set of transactions.

In computer science, a hash is a function that takes an input of arbitrary size and produces an output of fixed size. In cryptography this definition is extended to include three key principles:

1. Given any input, it is computationally easy to produce a hash
2. Given any output, it is computationally infeasible to compute the input to the hash function
3. It is computationally infeasible to find any two inputs to the function that produce the same output

There are many hash functions available that can satisfy these requirements, however finding new hash functions and breaking existing ones remains an area of research [4].

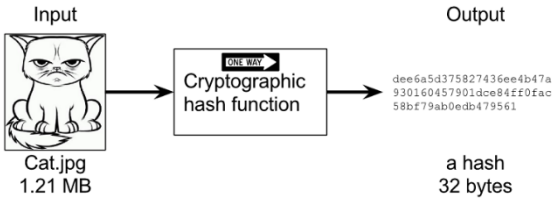

FIGURE 3 CRYPTOGRAPHIC HASH FUNCTION [12]

Each block contains a hash of both the current block, and the preceding block. These hashes are effectively a summary or "fingerprint" of the data stored inside the block and form the information necessary to verify the blockchain.

The chain can be verified by hashing the data inside the current block to ensure that the hash of the current block stored in the block header is accurate. Next the data from the previous block can be hashed to ensure that the hash of the previous block in the current block header is accurate. This process can be repeated by moving further back along the chain until eventually the first block is encountered and the blockchain can be considered verified.

Under principle 3 of a good hash function, no two inputs should produce the same output. This means that if any parts of the data in a block have been tampered with or corrupted, the hash that is produced will not match the expected values and verification will fail.

[8]

2.3.2. NODES

Of fundamental importance in blockchain technology is decentralisation. No single device or organisation can own the whole chain. Instead there must be numerous machines each maintaining their own copy of the blockchain. A node can be any machine capable of doing this.

As each node maintains their own copy of the blockchain, there must be consensus across the whole (or a majority) of the network in

order to update the chain. Once the nodes reach consensus, a new block can be added to the blockchain. Once this has happened, any nodes that have not received the new block or are unable to verify the block due to corruption or tampering, can obtain the block from other nodes.

Note that it is possible to use the blockchain without being a "full node". In this case users obtain the most recent block from the longest chain in the network and verify it with a number of nodes in the network. This is not as secure as maintaining a full-node as the user has not done the verification themselves [2].

[8]

2.3.3. VERIFIERS

Verifiers have the job of producing new blocks to be added to the network. Verifiers select a number of records from the pending transaction pool and include them in a block.

Before including a record in the block, verifiers must check that they are valid. The exact process for this will vary based on the application but common examples would be ensuring that the sender of each transaction has signed the transaction with their private key or ensuring that the sender of a currency has a high enough balance.

Once the verifiers have completed a block, they compete with other verifiers on the network for their block to become accepted to the blockchain. The form of this competition is prescribed by the specific blockchain, but the two most popular methods, which will be discussed in detail below, are Proof of Work (PoW) and Proof of Stake (PoS). Verifiers are usually rewarded financially when their block is adopted which acts as an incentive for them to provide their services.

Note that in many blockchains, the machines that act as (full) nodes are also verifiers and vice versa.

[8]

2.4. PROOF OF WORK (POW)

Proof of Work was the algorithm employed by Nakamoto for the bitcoin cryptocurrency. PoW involves finding a piece of data that is computationally difficult to produce, but easy for others to verify. This data is usually a value that can only be obtained through trial and error.

The process of calculating PoW is known as "mining" and the verifiers are called "miners". Blockchains that employ this method store an additional value in the block header called the "Nonce". The task of a miner is too adjust this value so that the hash of the block header is less than a certain value known as a "difficulty target". This value is pre-set by the network and can be adjusted [3].

Due to the low probability of finding a suitable value for the Nonce, it is not possible to predict which miner will find the correct value. Nakamoto described PoW as "One-CPU-one-vote". The longest chain in the network will represent the majority decision as it has the most CPU time invested into it. For a past block to be modified an attacker would have to be able to redo the proof-of-work algorithm for all blocks after the modified block, and then get their final block accepted by the network. The probability of this happening continues to approach zero as more blocks are added to the network [2].

2.5. PROOF OF STAKE (POS)

Proof of Stake is a method that is able to verify the network without the huge CPU requirement of PoW. Under this algorithm, the probability of a verifier's block being accepted

by the network is based on how big of a stake in the network they have. In digital currencies, this would be how much of the cryptocurrency they hold. Under this system, two principles protect the network:

1. An attacker would need to control a large amount of the currency in order to execute an attack. This would be very expensive.
2. The reputational damage to a currency from being attacked would likely drastically reduce the value of the currency. As an attacker would need to own a vast quantity of the currency, the incentive to perform such an attack would be reduced as they would be de-valuing their own assets.

[3]

Cardano is an example of a PoS cryptocurrency [9].

3. Blockchain Applications

Blockchain has been employed in a wide range of settings. Just two of these applications will be discussed below.

3.1. Digital Currencies

Digital currencies also known as cryptocurrencies were the original application for blockchain technology. As the first cryptocurrency, bitcoin remains the most popular and at the time of writing boasts a market capitalisation of over US $700 billion [5]. Bitcoin is a global currency and payment system that enables users to transfer value without needing to use intermediaries such as banks. Each user is assigned a public key which is used as an address, and maintains a private key which is used to secure access to their "wallet". Bitcoin uses PoW to secure the network and adjusts the difficulty target to ensure that a new block is produced every 10 minutes. The miner that finds the correct Nonce is rewarded with a number of bitcoins.

3.2. Smart Contracts

Smart contracts are computer programs that are able to automatically execute, control or document legal events according to the terms of a contract [6]. In 2013, Vitalik Buterin proposed the use of the blockchain as a decentralised distributed global computer that could run a variety of applications, services or contracts [7]. Ethereum specifies the Ethereum Virtual Machine (EVM) and a cryptocurrency called Ether which is used to pay "gas". The computational difficulty of a smart contract is described in units called gas and users decide how much Ether they want to offer per unit of gas. The verifiers on the Ethereum network perform the actual computation and update the state of the virtual machine, which is recorded in the blockchain, receiving the gas as a reward. Once a contract has been "deployed" to the Ethereum network, it is immutable and the code can be inspected by anyone. Today numerous programming languages can compile down to EVM instructions and there are thousands of Decentralised Applications (Dapps) running on the Ethereum network, ranging from the simple savings account which enables you to deposit currency to be returned to you at a specified date, to more advanced ideas such as online casinos.

4. Security Considerations

Blockchain has been designed from the ground up with security in mind. Despite this there are still many security concerns.

4.1. 51% Attack

In PoW based blockchains, the probability of finding the correct Nonce value increases in proportion to the CPU power at the Miner's disposal. If an individual or group are able to

control 51% of the CPU power of the network, they will be able to find the Nonce quicker than the rest of the network. This means they can decide which block should be adopted enabling them to:

1. Modify transaction data, possibly enabling a double-spending attack
2. Stop a valid transaction from being verified
3. Stop a miner from mining a block

51% attacks become less likely as more total power joins the network as an attacker would need to control a greater amount of computational power.

4.2. Selfish Mining

Selfish mining or withholding blocks is a malicious technique where a group of miners form a cartel to improve their share of mining rewards.

In 2013 Sirer and Eyal [13] showed that a cartel could improve their mining rewards by revealing a newly mined block to other members of the cartel, but not the public network. This means that members of the cartel can gain a head start on mining the next block, whilst honest miners continue to work on the public blockchain. This process is continued for an arbitrary length of time until the cartel's blockchain is longer than the public blockchain. At this point, the cartel will reveal the longer chain to the public, forcing miners currently working on the public chain to abandon it and join the cartel's chain. This works because the longest chain is always considered the valid one.

4.3. Denial of Service

Whilst the blockchain's distributed nature makes it more resilient to Distributed Denial of Service (DDOS) attacks, these attacks are still possible on some aspects of the network.

Generally DDOS attacks will be directed against components that have single points of failures like centralised cryptocurrency exchanges, mining pool servers and online wallets [18].

4.4. Malleability Attacks

Transaction malleability attacks are techniques that are used to trick users into sending a transaction more than once. In most blockchains, every transaction has a transaction ID. If an attacker can alter this ID and have it confirmed before the original transaction, it will appear to the victim that their transaction has failed, when in reality it has succeeded under a different ID. This may lead them to repeat the transaction.

Mt. Gox, a major exchange during the early days of bitcoin was famously bankrupted as a result of a malleability attack in 2014 [14].

Bitcoin has now solved this issue using a process called Segregated Witness or SegWit, which replaced the vulnerable component with a non-malleable hash [16].

4.5. Routing Attacks

Routing attacks attempt to tamper with transactions before they reach the wider network. Routing attacks are typically very difficult to detect as the attacker aims to divide the network into partitions that cannot communicate with each other in order to control the flow of information in and out of that partition. This is also called a partition attack [16].

4.6. Sybil Attacks

Sybil attacks occur when an attacker in control of a small number of nodes in a peer-to-peer network, creates a large number of identities for their nodes. This enables the attacker to manipulate the network such that they have more control than the hardware they have

available would entitle them too. For smaller attacks, this enables an attacker to surround a victim with fake nodes, similar to a partition attack, see figure 4. For larger attacks, if an attacker is able to occupy the majority of the network with Sybil nodes, this could enable them to mount a 51% attack. The only ways to combat this technique are increasing the costs of creating a new identity [16, 15].

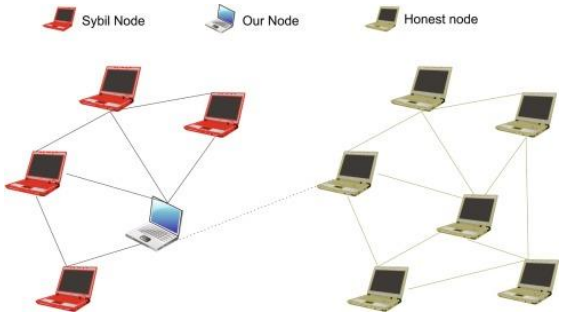

FIGURE 4 SURRONDED BY SYBIL NODES [15]

4.7. Eclipse Attacks

Eclipse attacks are similar to Routing and Sybil attacks. In this type of attack the attacker must control a large number of IP addresses. The attacker attempts to redirect all of the connections from a victim node to IP addresses that they control. Once again this enables them to restrict the flow of information to the victim node and manipulate it in various ways [17].

4.8. Timejacking

In a Timejacking attack, an attacker can trick a node into accepting an alternative blockchain by manipulating the timestamp on the blocks. This can be achieved if the attacker creates multiple malicious nodes that are connected to the victim node all using inaccurate timestamps. These attacks are however fairly easy to prevent by using the node system time or restrictive time ranges [16].

4.9. Race Attacks

In a race attack, an attacker sends two conflicting transactions at the same time. This can result in double spending if a victim accepts the payment (and ships a product) before the first transaction is eventually invalidated by the second [16].

4.10. Alternative History Attack

Alternative history or blockchain reorganization attacks can occur when an attacker sends a transaction but at the same time mines an alternative fork which does not include the transaction. The attacker then waits until the product has been shipped before releasing the alternative chain and recovering their funds. These attacks are only possible when an attacker has access to vast quantities of computing power.

4.11. User Wallet Attacks

As with most systems, the greatest security threat comes from the users of blockchain. There are numerous attack vectors against individual users.

Phishing attacks are very common in the cryptocurrency world where attackers attempt to trick victims into revealing sensitive information such as wallet keys or passwords. Similarly attackers may try dictionary attacks against poorly defended wallets.

4.12. Smart Contract Vulnerabilities

For blockchain technologies that offer smart contract functionality, there is always the risk that errors in the smart contract code could open vulnerabilities.

It is also possible that the underlying virtual machine implementation contains errors that expose vulnerabilities.

5. EVALUATION

5.1. ADVANTAGES

The main advantage of blockchain is the ability to share a database with others without having to trust any central authority. This has the added bonus of cutting out any fees the intermediatory could charge.

The blockchain provides transparency, verification and audit trails as standard. The blockchain provides stability, once something has been recorded, it is extremely difficult to reverse it.

The distributed nature of the blockchain makes it very resilient and the network could sustain even a very large percentage of the network becoming unavailable. Fault recovery is also built in, once these nodes become available again they can recover the blocks they missed by querying other nodes.

Blockchains lends themselves well to (pseudo) anonymity. This makes them of benefit to users in oppressive totalitarian regimes and users concerned about privacy.

[18]

5.2. DISADVANTAGES

Blockchains, especially those that use PoW, are highly inefficient. Massive quantities of CPU time are expended trying random numbers simply to secure the network. Additionally, only one miner can win and the work of all the other miners is wasted. This also makes blockchain a huge consumer of electricity and these issues are exasperated as the chain grows longer.

Storage is also a major cost for blockchain technologies. The storage requirement for downloading the complete ledger continues to grow as more data is recorded in the chain. As chains become bigger than average disk sizes, a new barrier for entry will exist for users. At the time of writing, the Ethereum ledger is already approaching 1 TB [19].

Whilst the immutability of the blockchain is considered an advantage, changing historical data or code usually requires a hard fork where a new chain is created and the old abandoned.

Although mentioned previously as an advantage, the anonymity of blockchain can make it useful for criminal activity and make law, tax and regulatory enforcement challenging.

A final concern for blockchain is scalability. As the number of users grows, congestion becomes an issue and it can take longer for transactions to be processed. These issues are usually the result of verifiers bottlenecking the process as they are unable to verify transactions quicker than they are added to the network [20].

[18]

5.3. CONCLUSION

Blockchain is a powerful technology that has already changed the world and will continue to do so. Like all technologies, it has its disadvantages, but the community continues to develop new solutions. Blockchain is still young and has not yet seen mass adoption, many of its most exciting applications may still be in the future.

REFERENCES

[1] Michael Crosby, Nachiappan, Pradan Pattanayak, Sanjeev Verma, Vignesh Kalyanaraman 2016 “Blockchain Technology: Beyond Bitcoin”

[2] Satoshi Nakamoto 2008 “Bitcoin: A Peer-To-Peer Electronic Cash System”

[3] Iuon-Chang Lin, Tzu-Chun Liao 2017 "A survey of Blockchain Security Lessons and Challenges"
[4] Rajeev Sobti, G. Geetha 2012 "Cryptographic Hash Functions: A Review"
[5] CoinMarketCap, https://coinmarketcap.com/
[6] Shuai Wang, Yong Yuan, Xiao Wang, Juanjuan Li, Rui Qin, Fei-Yue Wang 2018 "An Overview of Smart Contract: Architecture, Applications, and Future Trends"
[7] Vitalik Buterin 2013 "Ethereum Whitepaper"
[8] Dylan Yaga, Peter Mell, Nik Roby, Karen Scarfone 2019 "Blockchain Technology Overview"
[9] Charles Hoskinson 2017 "Why we are building Cardano"
[10] IP Specialist 2019 "How Blockchain Technology Works" https://medium.com/@ipspecialist/how-blockchain-technology-works-e6109c033034
[11] Cornelius C. Agbo, Qusay H. Mahmoud, J. Mikael Eklund 2019 "Blockchain Technology in Healthcare: A Systematic Review"
[12] Kalle Rosenbaum 2019 "Grokking Bitcoin"
[13] Ittay Eyal, Emin Gün Sirer 2013 "Majority is not Enough: Bitcoin Mining is Vulnerable"
[14] Christian Decker, Roger Wattenhofer 2014 "Bitcoin Transaction Malleability and MtGox"
[15] Shubhani Aggarwal, Neeraj Kumar 2020 Advances in Computers, "Chapter 20 – Attacks on blockchain"
[16] Anna Katrenko, Mihail S. 2020 Apriorit "Blockchain Attack Vectors" https://www.apriorit.com/dev-blog/578-blockchain-attack-vectors
[17] Ethan Heilman, Alison Kendler, Aviv Zohar, Sharon Goldberg 2015 "Eclipse Attacks on Bitcoin's Peer-to-Peer Network"
[18] Julija Strebko, Andrejs Romanovs 2018 "The Advantages and Disadvantages of the Blockchain Technology"
[19] Etherscan 2021 https://etherscan.io/chartsync/chaindefault
[20] Anamika Chauhan, Om Prakash Malviya, Madhav Verma, Tejinder Singh Mor 2018 "Blockchain and Scalability"

TABLE OF FIGURES