PhD Research Proposal:

Toufiqur R. Chowdhury | htr.letun@gmail.com

“Testing And Security Vulnerability Detection For Blockchain Based Systems”

Monash University, Australia
Faculty of Information Technology

Supervisor: Professor John Grundy
Co-supervisor: Dr. Zubair Baig

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

1. Introduction

Blockchain, a distributed immutable ledger, was originally proposed in 2008 by Satoshi Nakamoto for the sole purpose of creating a “digital money” what he named Bitcoin. Blockchain has evolved from just being the underpinning technology of Bitcoin to the backbone of many digital currencies with use cases beyond just store of value. With the increase in popularity and the explosion of price of these currencies beyond human expectation, blockchain has attracted many people with ill-intentions to benefit by exploiting its vulnerabilities. In the recent years, most popular cryptocurrencies including Bitcoin and Ethereum have been targeted by attackers. As a result, hundreds of millions of dollars have been stolen from people using these cryptocurrencies. The security vulnerability of blockchain based systems caused not only losing people money, but it is also scaring people and businesses away from using or starting to use cryptocurrencies. This is a huge obstacle for Satoshi Nakamoto’s dream of creating digital money that can replace fiat/printed currency. The aim of the project is to study attacks, discovered vulnerabilities of blockchain based cryptocurrencies and review existing researches on this topic to attempt to come up with one or more ways of fixing any security vulnerabilities and preventing attacks in the future. The primary focus will be on smart contract security and vulnerabilities while taking Ethereum as the platform of choice as it is the most popular blockchain to date for building Decentralised Applications (DApps) using smart contracts.

2. Literature Review

2.1. Introduction

The use of blockchain was first advised in 2008 by Satoshi Nakamoto when he proposed to create a “purely peer-to-peer version of electronic cash (that) would allow online payments to be sent directly from one party to another without going through a financial institution” (Nakamoto, 2008). He named his electronic cash (cryptocurrency), Bitcoin. This literature review will briefly explain what blockchain is, how it works, development stages of blockchain and will review the security risks and vulnerabilities of this technology. After that, the most recent security attack attempts will be briefly reviewed to build a case for the significance of the research that is needed in this area. Bitcoin and Ethereum are the most popular cryptocurrencies as of today and these two symbolises two different development stages of blockchain as well as the most attacked blockchain based systems. Therefore, this literature review will primarily focus on Bitcoin and Ethereum.

2.2. What is Blockchain?

In simplest terms, blockchain means a chain of blocks each of which contains some data. The Bitcoin network uses blockchain to store the historical data of the transactions of the digital asset by the owners. Each of blocks contain a timestamp to record when the transaction occurred, address of the sender, address of the receiver and the amount of BTC (Bitcoin) sent. In other words, blockchain is a decentralised immutable ledger that usually contains records of transfer of assets from one party to another without a central authority.

2.3. Consensus mechanisms

"Being a decentralized system, blockchain systems do not need a third-party trusted authority. Instead, to guarantee the reliability and consistency of the data and transactions, blockchain adopts the decentralized consensus mechanism." - (Li et al., 2017)

The fundamental of the blockchain based systems is a decentralised system where multiple independent nodes form a network. Each of these nodes is required to contain the entire history of transactions stored in a chain of blocks. Whenever a new transaction is to be added to the blockchain each of these nodes must agree upon the terms of the transaction through a consensus mechanism. Below are the most commonly used consensus mechanisms used by popular cryptocurrencies:

2.3.1. Proof of Work (PoW)

The mechanism used by Bitcoin, Ethereum and majority of other cryptocurrencies is called Proof of Work (PoW). The PoW mechanism uses solution of puzzles as a proof of the work to be done by node (AKA block producer). When creating a block, the block producer is required to solve a complex mathematical problem (the puzzle) and submit the result along with the block produced. The result puzzle is usually easy to verify by other nodes in the network. The other nodes in the network come to a consensus that block of transactions is valid by verifying that the result of the puzzle matches exactly with the expected result.

2.3.2. Proof of Stake (PoS)

In the Proof of Stake (PoS) mechanism, the block producer (miner) is required to deposit a certain amount of digital assets. If the produced block is eventually verified by the network as valid, block producer receives the digital asset back as a reward. If verification fails, the producer gets fined and does not get the asset back. PoS is much more attractive than PoW, because it does not require complex puzzles to be solved that takes a lot of computing power and does not waste large amount of electricity doing so.

2.3.3. Delegated Proof of Stake (DPoS)

Delegated Proof of Stake (DPoS) is an improvement to the PoS where miners are not required to stake anything. In DPoS, the coin/digital asset holders are required to vote “delegates”. “DPOS leverages the power of stakeholder approval voting to resolve consensus issues in a fair and democratic way” - (Bitshares.org, 2018). These delegate nodes are responsible for adding transactions, creating blocks and maintaining the network.

2.3.4. Masternode

Masternode system was first used by Dash (previously DarkCoin) cryptocurrency as a consensus mechanism. Masternodes are a specialised nodes in the blockchain network that are required to block a certain amount of digital assets. Unlike in PoS, masternodes are not required to send any digital asset on each block creation and do not risk losing the blocked digital asset. These masternodes are responsible for introducing specific services to the blockchain that cannot be performed by miners on a PoS mechanism. Masternodes in the Dash network provides added features/services such as PrivateSent, InstantSend and so on. It is also used for governance voting in the Dash ecosystem. Due to the lack of requirement of expensive hardware and large amount wasted electricity for computation like in PoW mechanisms and added extra features, masternodes are nowadays more popular and increasingly implemented by many cryptocurrency projects. The Ethereum network is planned to migrate to the masternode systems and eventually eliminate PoW completely.

2.4. Blockchain Development Stages

While it has only been a decade since it was first proposed for Bitcoin in 2008, blockchain has already embraced the second generation of it’s development stage.

2.4.1. Blockchain 1.0

In the first generation, AKA blockchain 1.0, blockchain was primarily used for cryptocurrencies. These cryptocurrencies had only one goal, to become the store of value and rival fiat currencies.

2.4.2. Blockchain 2.0

The second generation of blockchain or blockchain 2.0 is, however, not limited to a single goal. This generation of blockchain basically started with the Ethereum and the implementation of “smart contract” into the Ethereum network. Even though smart contracts were first proposed in 1994 by Nick Szabo, until the era of blockchain it only remained theoretical.

"A smart contract is executable code that runs on the blockchain to facilitate, execute and enforce the terms of an agreement. The main aim of a smart contract is to automatically execute the terms of an agreement once the specified conditions are met.“ - (Moorsel, 2017). In other words, smart contracts are specialized blockchain addresses that are programmable unlike the Bitcoin and other cryptocurrency addresses in blockchain 1.0. The languages used in smart contracts are said to be Turing-Complete. Very similar to computer programs these addresses contain specific codes that are executable when certain conditions are met. Users interact with smart contracts by sending a transaction to it’s address by including certain amount of digital asset or specific data/text or both.

The ability to have cryptocurrency addresses that are programmable opened up whole a new world of possibilities for the use of blockchain. Leveraging the power of smart contracts and blockchain technology the idea of Decentralised Autonomous Applications (DApp) came into existence. The possibility of creating digital assets without an independent network and can serve the purpose of store of value or anything else the developer programmed them to be. These secondary digital assets are called Tokens. According to CoinMarketCap, as of October 2018 there are over 1150 tokens publicly listed (Coinmarketcap.com, 2018).

2.5. Security Vulnerability And Risks Of Blockchain

2.5.1. Private Key Security

Any cryptocurrency address consists of a set of keys: public key and private key. Public key is used as the address or identity of the account that is to be made known and used by any user who wish to send digital assets to the owner of the account. On the other hand, a private key is more like a password of the public key for the owner of the account. It is the owner who is solely responsible in storing the private key securely. Due to the nature of public and private key generation mechanism in blockchain based systems, both of these keys are not human readable and cannot be manually set by the user either. Both in Bitcoin and Ethereum these keys are created using 256 bit hexadecimal number. That simply means to any average user these keys are set of large alphanumeric strings and not possible to remember by average human brains. The users have to store the private key either writing it down in a paper or storing in a storage disk like hard drive or pendrive. This leads to a very simple but critical flaw of losing the private key by its owner. Private keys can also be stolen or forcefully taken from the user by someone else. Due to the decentralised nature of the blockchains, there are no central authority who controls and/or maintains user accounts and the public-private keys. Therefore, losing one’s private key simply means losing the losing the access to the account and all the digital assets it contains and there is not way to recover these fund. According an article published by Aatif Sulleyman in the Independent online news, a British IT worker claimed to have lost 7500 BTC (Bitcoin) due

mistakenly throwing away the hard drive containing the private key to his account. The total value of the lost BTCs was nearly US$80 million at the time of the publication. (Sulleyman, 2017)

Furthermore, ECDSA (Elliptic Curve Digital Signature Algorithm), the algorithm that is used to generate Bitcoin and Ethereum addresses has a vulnerability. According to Hartwig Mayer, an attacker can recover a private key from the public key because it does not generate enough randomness during the signature process. (Mayer, 2018)

2.5.2. Double Spending

The term Double Spending in blockchain means using the same cryptocurrency in multiple transactions. This does not necessarily mean that both of the transactions will be successful. Blockchains with PoW consensus mechanism are most vulnerable to double spending attacks due to the delay between the start and confirmation of a transaction. In a double spending attack, the attacker first creates a transaction that is designed to fail but only after a certain amount of time. The attacker will also create a second transaction with the exact amount of digital asset on the same network, but this time the destination address will be one that the attacker has full control. While the second transaction is being verified the first transaction, usually targeted at cryptocurrency exchanges or vendors online, will be accepted and give the attacker credit and/or allow purchase of other digit goods. By the time the first transaction eventually fails, the attacker will have taken the credits or purchased virtual goods elsewhere. In the end, the attacker gets to keep the original digital assets to him/herself in the destination address of the second transaction and also received the credit and/or digital goods from the vendor.

2.5.3. 51% Vulnerability

“The blockchain relies on the distributed consensus mechanism to establish mutual trust. However, the consensus mechanism itself has 51% vulnerability, which can be exploited by attackers to control the entire blockchain.” - Xiaoqi Li et al. This means in a PoW mechanism if a single miner’s total hashing power accounts for more than 50% of the total network the miner gains the ability to control and manipulate the entire blockchain. The attacking miner can perform malicious actions like Double Spend attack, reverse a transaction, force the entire network to accept the longest chain as the main-chain and so on.

Even after a decade since the birth of blockchain, this core vulnerability has not been resolved. According to an article published on CoinDesk by Alyssa Hertig, more than 5 cryptocurrencies were attacked exploiting the 51% vulnerability during the month May and June 2018. Four of the attacked cryptocurrencies were Verge, Bitcoin Gold, ZenCash and MonaCoin who lost approximately 2.7, 1.86, 0.5 and 0.09 million US dollars accordingly. (Hertig, 2018)

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

The reason for increasing incidents of 51% vulnerability attack is well-explained by the co-founder of ZenCash (now Horizen), “Now we live in a world of ASICs, professional and well capitalized mining farms, and with hash rate for hire services like Nicehash that can be used in lieu of spending a ton of money on your own farm to temporarily hijack a network”. (Viglione, 2018)

2.5.4. Smart Contract Vulnerabilities

While the blockchain 2.0 and smart contracts brought some amazing new capabilities for blockchain based systems, smart contracts themselves came with vulnerabilities and bugs that were never seen before in the blockchain 1.0 era. Smart contracts are programmable and for that each smart contract requires developers to write and test the code just like any other computer programs. While any computer program can have bugs and can be updated with fix, they usually do not cost users a lot of money. Smart contracts, however, due to being based around decentralised cryptocurrencies and if not tested properly can lose people a lot of money. Below are few examples of attacks on the blockchain based systems targeting smart contract vulnerabilities:

Whenever smart contract vulnerability is discussed, the infamous DAO attack on the Ethereum network is brought up. DAO (Decentralised Autonomous Organisation) is a smart contract deployed on the Ethereum chain with the aim to democratize how the Ethereum based projects are funded. The DAO smart contract was attacked only 20 days after it was deployed. In the meantime, it had already raised 3.6 million Ethers (Ethereum) in the ICO (Initial Coin Offering) which was 15% of total circulating Ethers at the time and valued at US$150 million. The attacker stole about US$60 million worth of Ethers.

“*The attacker exploited the reentrancy vulnerability in this case. Firstly, the attacker publishes a malicious smart contract, which includes a withdraw() function call to DAO in its callback function. The withdraw() will send Ether to the callee, which is also in the form of call. Therefore, it will invoke the callback function of the malicious smart contract again. In this way, the attacker is able to steal all the Ether from DAO.*” - (Li et al., 2017)

While it has been over 2 years and much larger ICOs like EOS and Telegram completed without any hiccups and raised about US$4 billion and US$1 billion respectively, smart contract vulnerability attacks did not stop there. In July 2017, the popular Ethereum wallet called Parity was susceptible to attacks multiple times causing loss of millions of dollars. The Parity Freeze bug which was accidentally activated by a user when interacting with the smart contract of the Parity wallet. This incident caused over 0.5 million Ethers (valued over US$100 million) to be locked indefinitely and inaccessible to anyone. Earlier in 2017, Parity’s multisig wallet was attacked exploiting a smart contract vulnerability and over 150 thousand Ethers were lost, valued at US$32 million at the time of the attack.

2.5.5. External Attacks: Attack On MEW By DNS Hack

Attacks on blockchain can be from many angles, even from outside the blockchain itself. The most popular Ethereum web wallet, MEW (MyEtherWallet), was attacked in 24 June, 2018. While the MEW itself was not hacked, external DNS servers were hacked to divert traffic for myetherwallet.com to attackers own fishing website. According to Dmitry Vedenyapin the CTO of cryptocurrency exchange Alluma.io, “The MEW hackers used vulnerabilities in the DNS protocol itself and hijacked the IP address ranges that are resolved by Amazon DNS server. This was possible because some transit providers did not check the announcement before relaying it.” This attack caused a loss of about US$17 million worth Ethers from the users of MEW.

2.6. Conclusion

In the light of the brief discussion above regarding security vulnerabilities and most recent attacks on or around blockchain, it is crystal clear that blockchain based systems are not yet completely bulletproof. The dire need for more in-depth study into this area is not only evident but also urgent. The aim of the research project is to contribute to the ongoing research initiatives and add valuable insights and possibly propose solutions to the security problems in the blockchain industry.

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

3. Research Objectives & Questions

3.1. Research Objectives

I. Background Research
   A. Develop a better understanding of the core concepts of blockchain based systems and implementations by reading up online articles, journals and books.
   B. Search for and read journals and publications by researchers on the topic of blockchain testing, security and vulnerability.
   C. Contact experts on the field with the aim to gain insights into the problems they have faced and solutions they have come up with.

II. Analysis And Deliverables
   A. An analysis of discovered vulnerabilities and previous exploitation attempts on the major blockchain based systems.
   B. Thoroughly investigate any attempts to resolve these vulnerabilities and their outcomes.
   C. Review and analyse literature by other researchers on the topic.
   D. Data analysis from the experiments by implementing proposed solutions
   E. Develop hypothesis on the research questions based on the analysis and experiments.

III. Evaluation
   A. Evaluate the system by developing and implementing test cases
   B. Evaluate any hypothesis based on analysis

3.2. Research Questions

RQ1. Applications of blockchain contracts, vulnerabilities and mis-use.

RQ2. Analysis of erroneous smart contracts deployed on the Ethereum blockchain.

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

4. Research Method

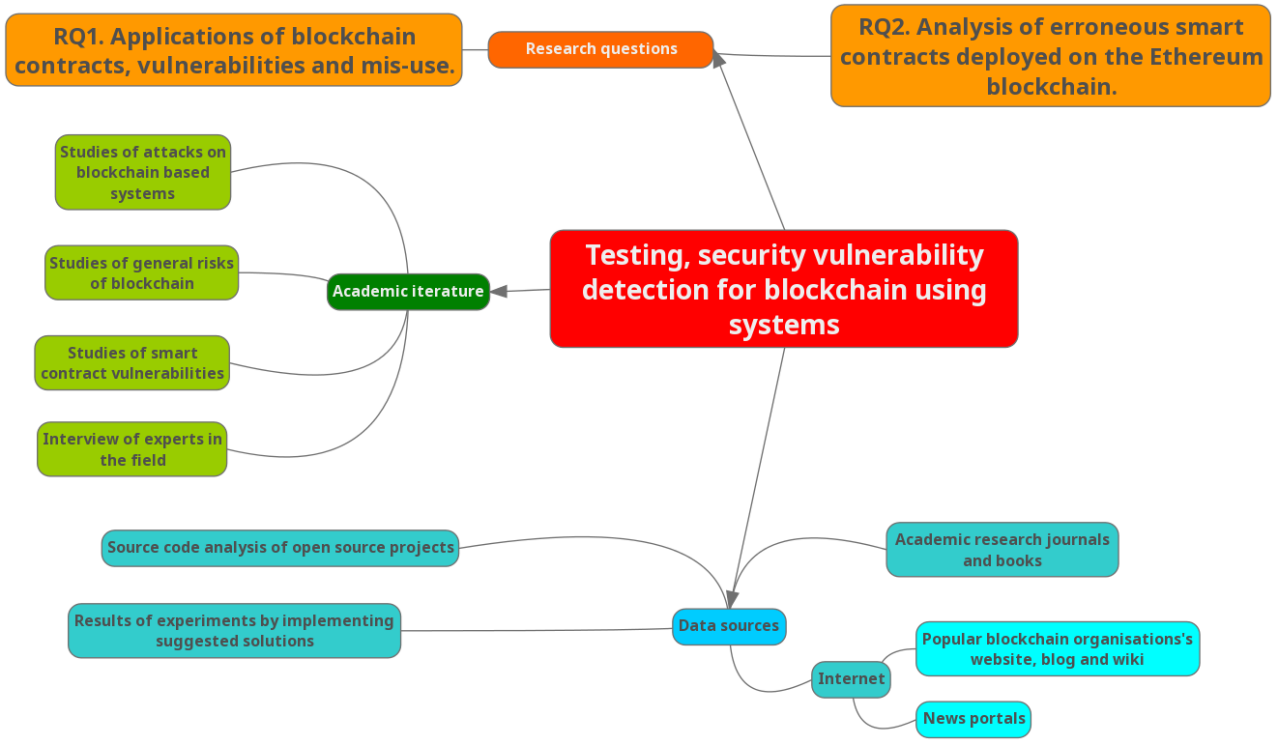

Figure 1: A visual representation identifying key areas and process of the project

The proposed research project will follow a mixed method by combining quantitative and qualitative research methods. The project will first study, review and analyse existing research in this area and the research questions. Any proposed solution to the blockchain contract related security vulnerabilities by existing researchers that have not been implemented by any real world project will be thoroughly analysed and, if deemed necessary, attempts will be taken to implement the solution in a test network environment with the goal to either approve or disapprove feasibility of the proposed solution. If any proposed solution has been already implemented by a real world project or in a test environment by other researchers that is still deemed not to be perfect, will also be studied and analysed using the data collected from existing implementations to validate the effectiveness of the proposed and implemented solution. Furthermore, a significant part of the project will also be to explore the possibilities of solutions to the currently discovered security vulnerabilities that has not been proposed before for blockchain based systems.

The project will primarily revolve around the two research questions mentioned in the section 3.2. In the RQ1, the uses of smart contracts in DApps (Decentralised Applications) will be

researched based on the recent and popular applications deployed on blockchain based systems. By studying and analysing existing literature and attacks on DApps, a systematic mapping of the vulnerabilities and misuses will be created by identifying and categorising them accordingly. Indeed there has already been studies of this kind. For example, in his publication A. V. Moorsel, created a systematic mapping of all studies on smart contracts and classified the issues with smart contracts into four root categories: codifying, security, privacy and performance issues (Moorsel, 2017). The systematic mapping and classification by A. V. Moorsel will be taken as the ground of the research and build up on from there.

Once the research on RQ1 is completed, next phase of the research will focus on the RQ2. This phase will be to analyse erroneous smart contracts. For this purpose, the Ethereum blockchain will be the main focus as it is still the most popular blockchain to date when it comes to developing and deploying public smart contracts. According to Atzei et al, “There are several reasons which make the implementation of smart contracts particularly prone to errors in Ethereum. A significant part of them is related to Solidity, the high-level programming language supported by Ethereum.“ (Atzei, Bartoletti and Cimoli, 2017). Therefore, Solidity will be primary language during the research. While Ethereum and Solidity will be main focus, in order to compare and contrast, the project will also study other competing blockchains and smart contract programming languages as they may distinct and often advanced features.

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

Timeline

Three year research plan:

Dates	Tasks
Mar 2019 - Aug 2019	● Extend knowledge on the field by reading up on existing journals and books ● Refine proposed research questions and execution plan
Sep 2019 - Feb 2020	● Search for and gather existing literature on the specific topic ● Read and record finding and key points from gathered literature ● Rethink and refine thesis topic ● Research and familiarise with new and advanced blockchain based projects ● Deploy one or more popular blockchain in a lab environment and familiarise with the process and procedure
Mar 2020 - Aug 2020	● Continue research ● Find out any developments on the topic from other academics and in the industry. ● Gain better understanding of smart contracts by learning to develop smart contracts and DApps using Solidity and other web technologies
Sep 2020 - Feb 2021	● Map all of the findings from research so far and analyse and re-evaluate all the data ● Prepare a plan for the thesis ● Gain deeper understanding of blockchain and smart contracts by reviewing and studying existing project and open source codes ● Contact experts on the field for interviews or chats
Mar 2021 - Aug 2021	● Prepare a draft of the first chapter of the thesis ● Prepare to implement any idea or proposed solution into a blockchain deployed in the lab. Record findings and analyse data collected from experiments.
Sep 2021 - Feb 2022	● Write draft of second chapter of the thesis ● Revisit the proposal and research questions and refine if necessary ● Implement more ideas and/or proposed solutions into the blockchain deployed in the lab. Record key findings.

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

Expected Outcomes/Impact

Blockchain based systems have been around for only a decade. This is a very short time for a new technology to mature. At its core, blockchain based systems mostly fall in the field of FinTech (Financial Technology). The field FinTech itself is very sensitive in nature due to its involvement of money. Blockchain based projects are even more sensitive because this in an emerging technology that involves money. The success of these projects may very well be defined by how security handled in them as mass adoption is the only way projects like Bitcoin and Ethereum can succeed. Due to the recent attacks on Ethereum and loss of millions of dollars caused panic among the average users which is reflected on the price of Ethereum today compared to the price it had end of last year.

This project aims to further the knowledge in the area of blockchain security in the following ways. The existence of research in this field cannot be disregarded. However, this project will investigate the topic through an alternative analytical lens. Existing research will be reviewed and analysed as well as be taken as a starting point when exploring new ideas and solutions. This project will not only research and build hypothesis but also develop and implement concepts and ideas into code for test and analytical purposes. This will provide a combination of theoretical and practical knowledge which will help future researcher gain further knowledge and deeper insights into specific problems.

Testing And Security Vulnerability Detection For Blockchain Based Systems

Proposal by Toufiqur R. Chowdhury

References

Atzei, N., Bartoletti, M. and Cimoli, T. (2017). A survey of attacks on Ethereum smart contracts.

Bitshares.org. (2018). Delegated Proof-of-Stake Consensus - BitShares. [online] Available at: https://bitshares.org/technology/delegated-proof-of-stake-consensus/ [Accessed Oct. 2018].

Coinmarketcap.com. (2018). All Tokens | CoinMarketCap. [online] Available at: https://coinmarketcap.com/tokens/views/all [Accessed Oct. 2018].

Duffield, E. and Hagan, K. (2014). Darkcoin: Peer-to-Peer CryptoCurrency with Anonymous Blockchain Transactions and an Improved Proof-of-Work System.

Hertig, A. (2018). Blockchain's Once-Feared 51% Attack Is Now Becoming Regular - CoinDesk. [online] CoinDesk. Available at: https://www.coindesk.com/blockchains-feared-51-attack-now-becoming-regular/ [Accessed Oct. 2018].

Jiang, B., Liu, Y. and Chan, W. (2018). ContractFuzzer: Fuzzing Smart Contracts for Vulnerability Detection.

Li, X., Jiang, P., Chen, T., Luo, X. and Wen, Q. (2017). A survey on the security of blockchain systems. Future Generation Computer Systems.

Mayer, H. (2018). ECDSA Security in Bitcoin and Ethereum: a Research Survey.

Moorsel, A. (2017). Blockchain Based Smart Contracts : A Systematic Mapping Study.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System.

Sulleyman, A. (2017). Man wants to dig up landfill site after he ‘threw away’ bitcoin haul now worth over $80m. [online] The Independent. Available at: https://www.independent.co.uk/life-style/gadgets-and-tech/news/bitcoin-value-james-howells-newport-landfill-hard-drive-campbell-simpson-laszlo-hanyecz-a8091371.html [Accessed Oct. 2018].

Szabo, N. (1997). Formalizing and Securing Relationships on Public Networks. First Monday, 2(9).

Viglione, R. (2018). Zen is Antifragile: Beyond a 51% Attack - Horizen. [online] Horizen. Available at: https://blog.zencash.com/zen-is-antifragile-beyond-a-51-attack/ [Accessed Oct. 2018].