University of Pisa

Department of Computer Science

Ph.D. Thesis Proposal

Security Risk Analysis of Blockchain Technology for Electronic Voting Purposes

Ricardo Chica Cepeda

Supervisor
Prof. Fabrizio Baiardi

October 2017

Contents

Chapter 1.

1. Introduction

The constant growth and development of information technology in all fields of society have enabled a substantial improvement in activities related to the electronic government and the way in which the public sector connects with the citizens and improves its own services.

Voting is the basis of any democratic system, either to elect representatives, to take decisions (referendum) or to reach a large-scale agreement. REV (Remote Electronic Voting), the citizens will have the possibility with the use of electronic devices like personal computers or smartphones connected to the internet, to record and transmit their votes during a specific time, set by the authorities of the election.

The daily activities, the geography and the disposition of the resources used for traditional voting, make that in the majority of cases, the eligible citizens do not participate in the elections, which is harmful to democracy and in some cases, affect the results when not counting the minimum number of participants, cases like Colombian referendum that was made in 2016, to approve or deny the negotiation between the government and the guerillas group known as FARC, to end a fifty years arm conflict had a 62% of abstention [1], or in 2016 the United Kingdom Brexit election, which decided if the country should remain or leave the European Union, had more than 28% of abstention as well. [2]

With the rise and massification of information and communication technologies, new forms have been developed in recent years to improve electoral processes, including internet voting, which has already been carried out in countries such as Estonia and Switzerland on a large scale, and some North American and Latin American cities as Santa Catarina, Brazil and Santo Domingo de Los Colorados in Ecuador, as a pilot test.

The blockchain technology allows the communities to redesign their interactions in different fields like government, economy, business and much more in a large scale, based on automated and trustless transactions. This situation had been changing the idea of the traditional role of the State and centralized Institutions. Indeed, the blockchain followers claim that the civil society could organize itself and protect its own interest more effectively, by replacing the traditional functions of State with blockchain based services and decentralized. This revolution has been creating a new system of governance, in which centralization, coercion, and hierarchies are replaced by a mechanism of distributed consensus. [3]

Cryptocurrency and its underlying technologies have been gaining popularity for transaction management beyond financial transactions. Transaction information is maintained in the blockchain, which can be used to audit the integrity of the transaction. [4]

A blockchain is a transaction database shared by all nodes participating in a system based on the Bitcoin protocol. A full copy of a currency's blockchain contains every transaction ever executed in the currency. With this information, we can find out how much value belonged to each address at any point in history.

Every block contains a hash of the previous block. This has the effect of creating a chain of blocks from the genesis block to the current block. Each block is guaranteed to come after the previous block chronologically because the previous block's hash would otherwise not be known. Each block is also computationally impractical to modify once it has been in the chain for a while because every block after it would also have to be regenerated. These properties are what makes blockchain transactions irreversible.

Since the information stored in blockchain is not related to personally identifiable information, it has the characteristics of anonymity. Also, the blockchain allows for transparent transaction verification, since all information in the block is open to the public.

To be distributed, an electronic voting system needs to decentralize one of the main operation, the validation of transactions exchanging value, this problem is usually solved with cryptography, and the whole idea is to shift the user's trust from a human-controlled central entity to a few reliable cryptography functions.

During the Ph.D. studies, we will work on the Blockchain technology, focus on its functionality for e-voting purposes, especially on user's privacy protection.

The present Thesis proposal is organized in four chapters. The first one is an introduction of the E-voting and Blockchain Technology, describing their components, vulnerabilities, and risk. The second Chapter presents the state of the art for this technology, the next chapter will describe the proposal and chapter fourth will describe the second year working plan.

Our main goal is an approach to manage the risks of this technology focus on electronic voting. To reach this objective we will analyze the protocol in details and describe the vulnerabilities in the components, the attacks they enable and the countermeasures that can be taken.

1.1. E-voting (Electronic voting)

E-voting is a technology where eligible citizens can vote using electronic devices such a laptop or smartphone, through an internet connection, while ensuring privacy and integrity of the results in a way to improve accessibility, as well as alternative method to traditional on-site elections, without losing sight of the main fundamental objectives:

- Ensure universal, free, equal, secret and direct vote.
- Achieve greater citizen participation.
- Ensure the transparency of the electoral process

There are two types of internet voting: On-site, which is conducted at controlled places, where election officials can authenticate eligible voters and the electronic infrastructure that must be used. The second type allows voters to transmit their votes from any internet connection to which they have access using a computer or smartphone.

When casting votes, the system gives a unique digital identification number (PIN) to the citizens that allow them to access the screens where the choice is made. Once the voter enters the site he/she can select the candidate of his preference and send the choice instantly. Voting

is transmitted through a network of communications, either in a centralized or decentralized protocol, from the place where it has been issued up to a remote digital urn or central server.

1.2. Requirements for Internet Voting Protocol

All the voting protocols tend to meet the same set of security requirements, the privacy of all the voters is the main one, and the result must be totally secret until the election is completed and verifiable. That provides the user the confidence that their votes had been treated correctly.

SECURITY REQUIREMENT	DESCRIPTION
Privacy	Is not revealed to anyone the way an eligible user voted
Authentication of voter's	To ensure that only eligible voters can vote and only one vote per person is counted.
Accuracy	Valid votes cannot be removed or manipulated. No invalid votes can be added
Secrecy of intermediate results	All results are kept secret until the election is completed.
No-coercion	The system must not enable the selling of votes or the coercion of voters.
Verifiability	Voters must be assured of the correct treatment of their votes, and have means to irrefutably prove of any fraud.

Table 1. General Security requirement for electronic voting protocols

1.3. Features and functionalities of electronic voting system

The basic features and the end-user functionalities that systems should offer to both voters and election officials are the following:

- **Universal:** The voting system must be available for all eligible voters, without requiring special knowledge, and be easy to navigate, including graphics and sounds mechanism for people with disabilities.

- **Availability:** Must never enter an undefined state, and have a backup mechanism to recover the system in case of an emergency.

- **Free:** Voters should make their choice without any interference or influence of anybody, as well they must not be paid or get paid for it.

- **Equal:** Voters should authenticate themselves to prevent unauthorized access, and each person can only vote once, each ballot is counted exactly once within the result. All ballots have the same influence on the result. [5]

Basic software components:

- **I-voting client application.** This user application allows voters to cast e-votes from a wide range of platforms. It can be customized to support any kind of election.

- **I-voting system.** It is comprised of a group of protected servers that collect, store, tabulate votes and create reports for election management. All these servers are controlled by the election commission.

- **I-voting verification application.** Because every voter should be certain that their vote is counted as intended, this mobile app allows voters to confirm that their vote was registered appropriately. [1]

1.4. Blockchain as a Decentralized Protocol

A protocol defines all the rules and conventions for communication between network devices. Network protocols include mechanisms for devices to identify and make connections with each other, as well as formatting rules that specify how data is packaged into messages sent and received. Some protocols also support message acknowledgment and data compression designed for reliable and high-performance network communication.

A decentralized protocol is based on the concept of 'client' and 'host' nodes, combining to create a general network. Both types of nodes should be supported by any piece of software for the protocol. The network is supported by a 'backbone' of host nodes, which are all connected together, and each provide a 'gateway' to the network for a number of client nodes; The hosts pass messages on from any of their clients that send them to all the other hosts in the network, and messages they receive from other hosts to all the clients they support

Cryptography is a method of using advanced mathematical principles in storing and transmitting data in a particular form so that only those, for whom it is intended, can read and process it. One of the newest cryptographic decentralized protocols is the blockchain, which has been used for the crypto-currency known as Bitcoin, but it also can be used for a wide range of services (Government, Private Companies, E-commerce, etc)

A blockchain is an audit trail for a database which is managed by a network of computers where no single computer is responsible for storing or maintaining the database, and any computer may enter or leave this network at any time without jeopardizing the integrity or availability of the database. Any computer can rebuild the database from scratch by downloading the blockchain and processing the audit trail.

The obvious way to ensure that no single entity can manipulate the database is to make the database public, and allow anyone to store a redundant copy of the database. In this way, everyone can be assured that their copy of the database is intact, simply by comparing it with everyone else's. [7]

Taking in consideration that in decentralized protocol there are no authorities or trusted parties, all voters operate independently with equal mutual suspicion. All traffic is performed on regular communication channels. The protocol is also accurate in that cheating is discovered immediately and in some cases, the perpetrator may be identified.

The system used in the decentralized protocol based on blockchain, offers a transparent public ledger which is a collection of accounting entries that is not centrally controlled by an individual or organization and the ledger entries only get confirmed as correct and officially enter into the ledger once they have been mathematically verified by the blockchain. At the same time, the ledger is completely public.

The most prominent concern about an implementation of blockchain voting system is the lack of experimental evidence that such a system could hold up to a large scale use, for example in a national election. Another important issue is regarding the use of a cryptographic key in which a verified voter can cast their ballot, and in some cases, can be difficult to deal with this aspect as well making the attackers to compromising the voter's key instead of the system. [8]

Blockchain uses security methods like asymmetric cryptographic keys which utilize two keys (public key and a private key) to encrypt and decrypt a particular data. The public key that may be disseminated widely, and private key which is known only by the owner, this accomplishes two functions: Authentication when the public key is used to verify that a holder of the paired private cast the vote, and encryption, whereby only the holder of the paired private key can decrypt the message encrypted with the public key.

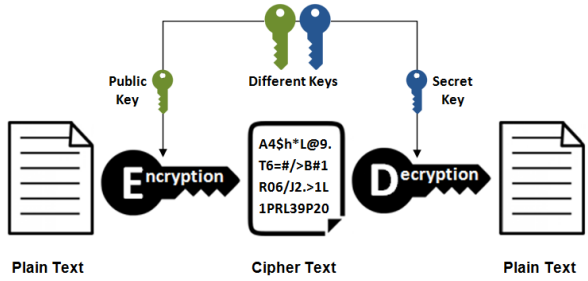

Figure 1: Asymmetric encryption

In the case of Blockchain when a legitimate user cast his vote, what the system does is broadcast a transaction to all the nodes that compromise the peer–to–peer network.

Giving that a variety of users are broadcasting the transaction to the network, the nodes must agree on exactly which transaction was broadcast and the order in which these transactions happened. This will result in a single, global ledger for the system.

So, at any given point, all the nodes in the peer-to-peer network have a ledger consisting of a sequence of blocks, each containing a list of transactions, that they've reached consensus on [10]

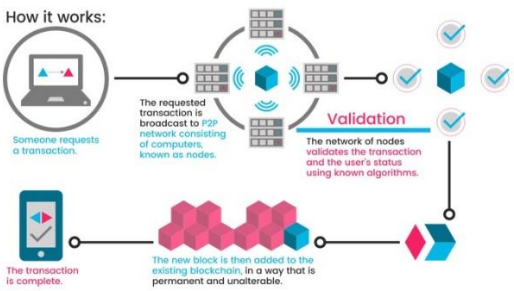

Figure 2. Blockchain Working Scheme. [2]

Chapter 2

2. State of the Art

The present chapter describes the state of the art based on the following topics:

1. Internet voting systems
2. Risks of internet voting systems
3. Blockchain as a decentralized protocol

2.1. Internet-Voting Systems:

Systems for electronic voting has been increased in many countries, recent examples are US, Estonia, Switzerland, Brazil, and Venezuela, one of the first ones was Estonia in 2005, The base of this system is the national ID card that all Estonian citizens are given. These cards contain encrypted files that identify the owner and allows the owner to carry out a number of online and electronic activities including online banking services, digitally signing documents, access their information on government databases and i-voting. [11].

In Venezuela, a company name Smartmatic has been the main technology supplier and system integrator for election since 2004, with touch-screen electronic voting machine designed to support multiple choice scenarios and real-time electoral information systems. [6]

In 2014 the company launched a remote voting platform that offers digitally verifiability, integrity of the vote, from the point of casting to counting.

Some other countries in Europe are exploring as well possibilities with blockchain, Switzerland, United Kingdom and Germany are working with private companies to build a robust blockchain e-voting infrastructure.

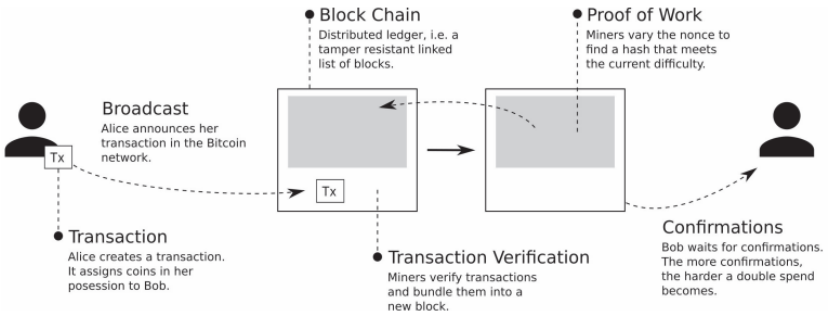

Figure 3: Blocks Transaction

In 2016 the Spain Company called Scytl, offered its technology, over a multilayer security and verifiable framework with the following privacy, security, integrity and auditability characteristics:

- End to end encryption: The votes are encrypted at the voting device guaranteeing they cannot be tampered during transmission or when received by the voting server.

- Encrypted votes are digitally signed on the voter's device using key roaming mechanisms ensuring voter eligibility and protecting vote integrity.
- Voting servers verify if the contents of the encrypted votes are valid without needing to decrypt them. This allows verifying that invalid votes are not cast accidentally, all without breaking vote secrecy.

- Votes are decrypted using a verifiable mix-net and secret sharing scheme, this ensures that voter anonymity is maintained when vote contents are counted.

- Return codes allow voters to check if the encrypted votes received by the server contain the selected voting options, without breaking vote secrecy.

- Receipts allow voters to verify if their votes have been stored in the ballot box by checking on a public bulletin board.

- Ballot box integrity is verified by using the public bulletin board, digital signatures and mathematical proofs of the votes.

- The mathematical proofs produced by the mix-net and decryption processes provide evidence of accuracy of the counting process and therefore, the accuracy of the results.

- To facilitate public audits, mathematical proofs produced by the mix-net and decryption processes are published in an Immutable bulletin board. [3]

Any online voting protocols should follow the following concepts to ensure secrecy of the ballot and eligibility:

a. Trusted Authorities: Voters transmit their ballot to the authorities using an authenticated channel that allows to verify the eligibility. Authorities are trusted to keep ballots confidential and to produce the correct tally.

b. Anonymous Voting: Using an authenticated channel, voters acquire from one authority an eligibility token.

c. Random Perturbation: Voters send encrypted ballots to a group of authorities that shuffle one after each other the set of all ballots.

d. Homomorphic Encryption: encrypted ballots are tallied and decrypted only afterwards. The use of cryptography implies technological trust and trust in authorities.

e. Balancing Verifiability and Secrecy of the Ballot: Online voting protocols offering end-to-end verifiability E2E-V allow voters to verify the online voting outcome using cryptographic proofs. [4]

2.2. Risks of E-voting Systems

ELECTION		
	Authentication	- There is not a physical probe that the person voting is really the authorized voter. - Possibility of stolen voter packages or identification cards - Misuse of elector's ID card and personal information voting by others without the knowledge of the elector
	Voting	- Unable access to election website - Network Saturation - Internet signal cut off - Dissociation of the instructions for user verification and voting options - Phishing - Malware
	Validation	- Internet signal cut off - Attacking the web application
	Storage	- Hacker - Manipulation of the algorithm of the voting counting program in the server (*The company that installed can decide also who win*) - Replacement of the voting counting software
	Decryption	- Remove or replace de cryptography parameters

The 2013 Estonia local election used REV (Remote Electronic Voting) and there were identified many potential security risks, like malware on the client side machine, that monitors the user while placing his vote and then later changing the vote to a different candidate. Another weakness was regarding the HTTP. If a client sends a request containing unexpected header fields, the server logs the field names to disk, by sending many specially crafted requests containing fields with very long names, an attacker can exhaust the server's log storage, after which it will fail to accept any new votes.

The encrypted ballots are separated from the signatures and copied to an isolated machine before being decrypted and counted, an attacker who can smuggle this information out through a covert channel can compromise every voter's secret ballot.

The counting server malware can sort the encrypted ballots and leak the voter choices corresponding to each as a sequence of integers in the same order.

Another possible risk has infected the server through malware being placed on the DVD´s used to set up the servers and transfer the votes. [13]

Estonia's system also fails to provide compelling proof that election outcomes were correct. The tabulation process at the end of the election was also concerning, because after the votes were decrypted on the counting server, an unknown technical glitch prevented workers from writing the official counts and log files on a server DVD, and transfer them to a computer where they sign the results officially, instead the electoral authorities decided to use a regular personal USB to transfer those files, that might add a multiple potential attack vectors. [14]

2.3. Blockchain Decentralized Public Ledger Platforms

Most of the people use the term of Blockchain for bitcoin or any other cryptocurrency, but this technology has gained more uses in the past few years and is working for many other functionalities, like smart contracts or voting.

Early value transfer systems embodied the concepts of value storage, encryption, and cryptographic public/private key pairing, at the heart of modern crypto-currencies. The main difference between blockchain technology and predecessors is the level of decentralization of the network.

- **Internet-based Value Containers:** The blockchain, does not encompass all the "distributed application/ledger" blockchain set, is a modern value transfer system, namely a protocol for sending, receiving and recording value on a public ledger.

- **Incentives for Collaborative Effort:** Decentralized public ledgers are intensive in labor and computer processing time, thereby reflecting how miners are rewarded

- **Open Source Licenses and Governance Mechanisms:** The licensing model for enabling changes to the software of either the public ledger currency platform or a token-free blockchain application.

- **Immutability of the System:** The ability to declare a truth, globally and without a Center of authority [5]

The common themes seem to be a data store which:

- Contains financial transactions.
- Is replicated across a number of systems in almost real-time.
- Works over a peer-to-peer network.
- Uses cryptography and digital signatures to prove identity, authenticity and enforce read/write access rights.
- Can be written by certain participants.
- Can be read by certain participants, a wider audience.

- Has mechanisms to make it hard to change historical records, or at least make it easy to detect when someone is trying to do so. [6]

2.3.1. Addresses

The structure of a blockchain is that a block of multiple transactions is connected to a previous block in chain-like form. Each block within the blockchain is identified by a hash, generated using the SHA256 cryptographic hash algorithms in the header of the block. Each block also references a previous block, known as the parent block, through the "previous block hash" field in the blocker header. Which means each block contains the hash of its parent inside its own header. The sequence of hashes linking each block to its parent creates a chain going back all the way to the first block ever created, known as the genesis block.

The block as a container data structure aggregates transactions for inclusion in the public ledger, the blockchain. The header of the block contains metadata, following by a long list of transactions that make up the bulk of its size. The block header is 80 bytes, whereas the average transaction is at least 250 bytes, and the average block contains more than 500 transactions.

The block header consists of three sets of block metadata. First, there is a reference to a previous block hash, which connects this block to the previous block in the blockchain. The second set of metadata, namely the difficulty, timestamp, and nonce, relate to the mining competition and the third piece of metadata is the Merkle tree root, a data structure used to efficiently summarize all the transactions in the block. [7]

Blockchain uses cryptographic proof instead of the trust-in-the-third-party mechanism for two willing parties to execute an online transaction over the Internet. Each transaction is protected through a digital signature, is sent to the "public key" of the receiver, and is digitally signed using the "private key" of the sender.

In the case of bitcoin for the generation of user keys, it uses the algorithm ECDSA based on elliptic curve cryptography. i.e In bitcoin, the user who wants to receive a payment can send the payer its public key, and it links the payment to this public key so that only the user who has the private key can access the payment and therefore to those funds.

Size	Field	Description
4 bytes	Block Size	The size of the block, in bytes, following this field
80 bytes	Block Header	Several fields form the block header
1-9 bytes (VarInt)	Transaction Counter	How many transactions follow
Variable	Transactions	The transactions recorded in this block

Figure 4: Structure of a block

Size	Field	Description
4 bytes	Version	A version number to track software/protocol upgrades
32 bytes	Previous Block Hash	A reference to the hash of the previous (parent) block in the chain
32 bytes	Merkle Root	A hash of the root of the merkle tree of this block's transactions
4 bytes	Timestamp	The approximate creation time of this block (seconds from Unix Epoch)
4 bytes	Difficulty Target	The proof-of-work algorithm difficulty target for this block
4 bytes	Nonce	A counter used for the proof-of-work algorithm

Figure 5. Structure of the block header

2.3.2. Transactions:

The key elements of a transaction are the hash value as the transaction identifier (TXID) and a list of inputs and outputs.

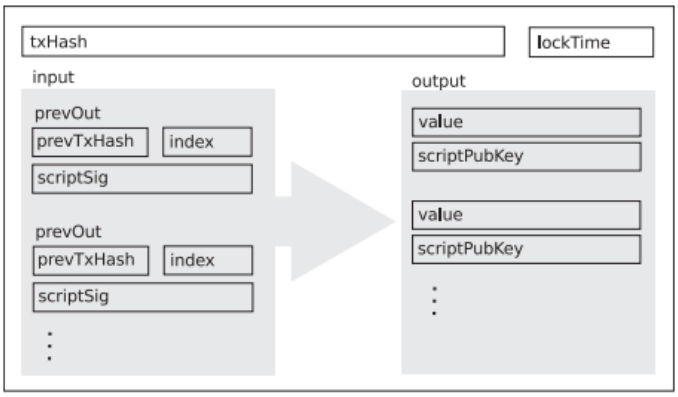

Figure 6: Transaction

Transaction input does not specify how many coins are spent because each output of a previous transaction can be used only once, inputs necessarily always use all the coins from the referenced output.

A transaction is a funds exchange between addresses. Transactions are multi-input, multi-output, it means that a transaction may have more than one input (address from which funds are withdrawn) and more than one output (address in which funds are stored). Each transaction completely transfers funds from the inputs to the outputs (no change is left in the input addresses).

The transaction hashes and the reference to the previous transactions take the role of serial numbers as they are used in cryptocurrency systems like bitcoin, in this case, eliminates the need for a third party, for example, a bank to issuing serial numbers.

2.3.3. Scripts:

Through scripting, there is a certain degree of programmability what exactly a transaction does. Scripting is realized by a simple stack-based language. It is intentionally designed not to be Turing complete so that it is easier to handle and unintended side effects can be avoided.

Each output in a Bitcoin transaction is described by a script. The operations to be performed, potentially along with constants, constitute the so-called scriptPubKey. A script expects a number of "arguments", the scriptSig. An input which connects to an output must provide the scriptSigthe respective script. The connection is considered valid when the output's script evaluates to true given the scriptSig provided in the connecting input. The probably most essential and most common script of all is "Pay-to-PubKeyHash" (P2PKH). Semantically, a transaction employing P2PKH transfers coins from one or more origin addresses to a destination address. The key idea is to have a script at the output which checks whether the connecting input has been signed with the public key "owning" the coins at the output. Script 1 provides the generic P2PKH script template. Bitcoin scripts are processed from left to right. In its scriptSig, P2PKH requires a public key (pubKey) which hashes to the specified Bitcoin address (pubKeyHash) and a signature (sig) proving the possession of the respective private key.

The hash is included in a generic output script and can be redeemed as specified in Script 2. In principle, the redeem script can be any script, but the transaction is considered as a standard

transaction only if the redeem script follows one of the standard "pay-to-x" scripts, e.g., a P2PKH. Script 3 depicts the generic script template for an m-ofn multi-signature transaction.

After pushing the constants to the stack, OP_CHECKMULTISIG takes the integer n first (because after pushing it is the topmost entry), then n pubKey items, subsequently the integer m, and finally m sig items off the stack. Now, in essence, OP_CHECKMULTISIG iterates over public key/signature pairs and executes OP_CHECKSIG.

```
scriptPubKey:OP_DUP OP_HASH160 <pubKeyHash>
  OP_EQUALVERIFY OP_CHECKSIG

scriptSig: <sig> <pubKey>
```

Script 1. "Pay-to-PubKeyHash" (P2PKH) script template

```
scriptPubKey: OP_HASH160 <redeemScriptHash>
  OP_EQUAL

scriptSig: [<sig> ...] <redeemScript>
```

Script 2. "Pay-to-PubKeyHash" (P2SH) script template

```
scriptPubKey: <m> <pubKey> [<pubkey> ...] <n>
  OP_CHECKMULTISIG

scriptSig: 0 [<sig> ...]
```

Script 3. *m*-of-*n* multi-signature transaction script template

Figure 7. Scrip on blockchain

2.3.4. Consensus:

Blockchains use distributed ledgers to record information primarily about the balance of every address for value transfer platforms (like bitcoin and most cryptocurrencies), though the approach can be extended to any kind of information. Key to the operation of the blockchain is that the network should collectively agree on the contents of the ledger: instead of authority for keeping accounts being centralized.

This requires that the network maintains consensus around the information recorded on the blockchain, the way this consensus is achieved impacts the security and economic parameters of the protocol. One the most common ways to solve this problem is the proof of work.

Proof of Work (PoW):

Proof of work is a distributed consensus mechanism for blockchain. In this mechanism the miners, which are the ones that produce the consensus by solving a puzzle consisting of a mathematical function process call hash (takes input data of any size, performs an operation on it, and returns output data of a fixed size).

PoW operates on the principle that it is expensive to add a tranche of new transactions to the blockchain but very easy to check if the transactions are valid due to the transparent nature of the ledger. Miners collectively verify the entire blockchain, and transactions aren't considered to be fully 'confirmed' until several new blocks have been added on top of them.

If a malicious actor tries to spend coins fraudulently, those transactions will be ignored by the rest of the network. [20] Distributed storage of multiple copies of the blockchain opens up new possibilities for cheating. In particular, it may be possible to issue two separate

transactions to two different receivers, transferring the same data (coins, votes). This is called double spending. If two persons verified and accepted the transactions independently (based on their respective local copy of the blockchain), this would drive the blockchain into an inconsistent state.

To reference a specific block, its header is hashed twice with the SHA-256 function [5]; the resulting integer value belongs to the interval $[0,2^{256}-1]$. In common blockchain implementations the generic hashing function hash(·) with a variable number of arguments and range [0, M]. The arguments of the function can be treated as binary strings and merged together to form a single argument that can be passed to the SHA-256 hashing function. The block reference is used in the proof of work protocol. $hash\ (b) \leq M/D$, where $D \in [1, M]$ is the target difficulty. There is no known way to find B satisfying (1) other than iterating through all possible variables in the block header repeatedly. The higher the value of D, the more iterations are needed to find a valid block; the expected number of operations is exactly D. The time period $T(r)$ for a miner with hardware capable of performing r operations per second to find a valid block is distributed exponentially with the rate r/D. $P\{T(r) \leq t\} = 1 - \exp(-rt/D)$ [8]. If we take into consideration a n bitcoin miners with hashes rates $(r_1, r_2, \dots, r_n)$. The period of time to find a block T is equal to the minimum value of random variables Tr_i and the miner T publishes a found block and it reaches other miners immediately. T is also distributed exponentially:

$$P\left\{T \overset{def}{=} \min\left(T_{1,\dots\dots,}T_N\right) \leq t\right\} = 1 - \exp\left(-\frac{t}{D}\sum_{i=1}^{n} r_1\right);$$

$$P\{T = T_1\} = \frac{r_1}{\sum_{j=1}^{n} r_j}$$

A miner with a share of mining power p has the same probability p to solve a block before other miners.

Forks:

Is a technical event that occurs when there is a change to the blockchain software and creates two separate versions of the blockchain with a shared history, it can be temporary, lasting for a few minutes, or permanent. A blockchain diverges into two potential paths forward either with regard to a network's transaction history or a new rule in deciding what makes a transaction valid, because diverse participants need to agree on common rules.

Hard fork: Is a software upgrade that introduces a new rule to the network that isn't compatible with the older software. Nodes that continue running the old version of the software will see the new transactions as invalid. So, to switch over to the new chain and to continue to mine valid blocks, all of the nodes in the network need to upgrade to the new rules. The problem comes when some sort of political impasse arises, and a portion of the community decides to stick by the old rules no matter what. The hash rate, or network computing power, behind the old chain, is irrelevant.

Soft Fork: Is any change that's backward compatible. For example of 1MB blocks, a new rule might only allow 500K blocks. In this case, non-upgraded nodes will still see the new

transactions as valid (500k is less than 1MB in this example). However, if non-upgraded nodes continue to mine blocks, the blocks they mine will be rejected by the upgraded nodes.

2.3.5. Transaction Malleability:

Refers to a bug in the Bitcoin protocol, which makes it possible to change the TXID without invalidating the transaction. It includes references to previous transactions (inputs) with a respective redeem script (scriptSig) and specifies one or multiple destinations (outputs). Each transaction can be uniquely identified by its TXID, which is a hash of the transaction data, including the redeem script(s)

A transaction malleability attack against an exchange proceeds as follows: (i) the attacker withdraws coins from an exchange and (ii) as soon as the attacker receives the respective withdrawing transaction issued by the exchange, he rebroadcasts the altered version of this transaction with a different TXID. One of the two transactions eventually makes it into the blockchain. Due to propagation delays and precautions the attacker can take, there is a chance that the modified transaction wins over the original withdrawal. If the exchange relies on TXIDs only, it will not find the withdrawal transaction in the blockchain and believe the withdrawal has failed. As consequence, the attacker may withdraw again.

The transaction malleability attack can be thought of as a variant of a double spending [21]. In contrast to a typical double spend, however, the attacker is the receiving and not the spending party. The success of the attack depends on a number of constraints, i.e., the malleable transaction must be confirmed and the exchange must check for TXIDs only. [22]

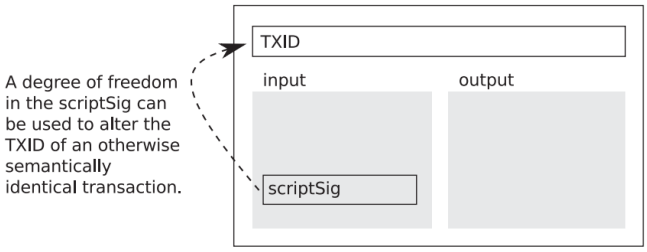

Figure 8. Transaction Malleability

2.3.6. Scalability

The main objective of the peer-to-peer network in blockchain is to quickly distribute the information into every part of the network. Variations in the propagation mechanisms directly affect the formation of the distributed consensus and thus the security. In general, inconsistent states, i.e., blockchain forks, are undesirable, because they facilitate double spending. However, the network is faced with scalability issues. Especially network bandwidth, network size and storage requirements pose challenges.

The protocol is capable of much more than the current transaction rate and is thus able to scale to higher demands. Currently, in the case of Bitcoin has an artificial maximum block size of 1 MiB, which limits the number of transactions per block and therefore also the growth rate of the blockchain. This limit is enforced to prevent from inflating the blockchain before the Bitcoin protocol is capable of handling more transactions. [22]

2.3.7. UTXO (Unspent Transaction Output)

A UTXO is an unspent transaction output. In an accepted transaction in a valid blockchain payment system (such as Bitcoin), only unspent outputs can be used as inputs to a transaction. When a transaction takes place, inputs are deleted and outputs are created as new UTXOs that may then be consumed in future transactions.

In the Bitcoin network, which uses this model, a UTXO is the amount that is transferred to a Bitcoin address (along with information required to unlock the output amount*) during a transaction. Received amounts (UTXOs) are used individually during a transaction and new outputs are created – one for the receiver and, if applicable, one for the amount that is left over (change output). The amount sent to the recipient becomes a new UTXO in the recipient's address while the change output becomes a new UTXO in the sender's address that may be used in a future transaction.

2.3.8. Merkle Tree

A Merkle tree, also known as a binary hash tree, is a data structure used for efficiently summarizing and verifying the integrity of large sets of data. Merkle trees are binary trees containing cryptographic hashes. Merkle trees are used in bitcoin to summarize all the transactions in a block, producing an overall digital fingerprint of the entire set of transactions, providing a very efficient process to verify whether a transaction is included in a block. A Merkle tree is constructed by recursively hashing pairs of nodes until there is only one hash, called the root, or Merkle root. The cryptographic hash algorithm is SHA256 applied twice, also known as double-SHA256.
Because the Merkle tree is a binary tree, it needs an even number of leaf nodes. If there is an odd number of transactions to summarize, the last transaction hash will be duplicated to create an even number of leaf nodes, also known as a balanced tree. [23]

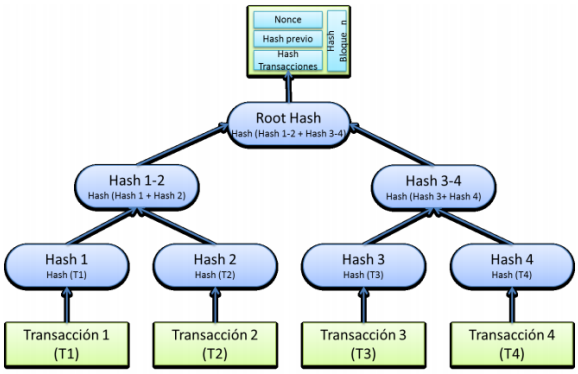

Figure 7: Merkle tree

2.3.9. Public and Private blockchains

Blockchains such as the Bitcoin have also been referred to as public blockchains, in the sense that the networks are open to the general public to join as users or serve in as nodes, but also in the sense that the blockchain data is publicly transparent.

But when we talk about the use of this technology for purposes that can be regulated by a specific entity like a corporation, country or city for example for voting purposes, it may generate a public discussion in the way that even known is a decentralized technology and open to the general public to use as a user or serve in as nodes, also can become a private blockchain and generate some constraints. The BitFury Group [8] proposed the following classification of permission:

1. A public blockchain is a blockchain, in which there are no restrictions on reading data and submitting transactions for inclusion into the blockchain
2. A private blockchain is a blockchain, in which direct access to blockchain data and submitting transactions is limited to a predefined list of entities.
3. A permissionless blockchain is a blockchain, in which there are no restrictions on identities of transaction processors
4. A permissioned blockchain is a blockchain, in which transaction processing is performed by a predefined list of subjects with known identities.

In the case of voting purposes, we can define this type of blockchain as a private one, since it has to be regulated by some entity, that specifies who can use it, regarding their own conditions (Age, city, country, etc).

#	Parameter	Blockchain for Bitcoin	Local Blockchain	Shared Blockchain
1	**BC Visibility**	Public	Secure/Private	Secure/Private
2	**Transaction Chainging**	Input / Output	Previous T of the same D	Previous T of the same D
3	**Transaction mining**	All Ts	All Ts	All Ts
4	**Mining Requirement**	Proof of work	None	None
5	**Forking**	Not Allowed	Allowed	Allowed
6	**Double Speding**	Prohibited	Not Applicable	Not Applicable
7	**Transaction verification**	Signature	Not Verification	Not Verification
8	**Transaction parameters**	Input, output	Block-number, hash of data, time, output, PK, policy rules.	Block-number, hash of data, time, output, PK, policy rules.
9	**Transaction Dissemination**	Broadcast	Unicast	Unicast
10	**Deference in block header**	Puzzel	Policies	Policies
11	**Block stored by miner**	All blocks	All blocks	All blocks
12	**New block verification**	Blocks and Ts in blocks	No verification	No verification
13	**BC control**	No one	Owner	Owner
14	**Miner Checks**	No one	No one	No one
15	**How many blocks each T is store in?**	One block	One block	
16	**Miner joining overhead**	Download all blocks in BC	Download all blocks in BC	Download all blocks in BC

17	**Miner selection**	Self-selection	Owner chooses the miner	Owner chooses the miner
18	**Miner reward**	Coins	Nothing	Nothing
20	**Malicious Miner**	Allowed to join	not possible	not possible
21	**Effects of 51% attack**	Double spending	not possible	not possible
22	**Encryption Method**	Public / Private Keys	Optional	Optional

T: Transaction
D: Device
Figure 8: Comparison of bitcoin blockchain and local and shared the blockchain

2.3.10. Side Chains:

Side chains are the concept of parallel blockchains that allow assets from one blockchain to be transferred into another. [9] The intention of sidechains is to enable the transfer of data to other blockchains. Sidechains can extend the functionality of a parent blockchain by introducing new features on the sidechain, i.e, sidechains can potentially bring to permissionless networks are things such as supporting multiple asset types from different blockchains to exist on a mutual sidechain, smart contracts and prediction markets [10]

2.3.11. Smart contracts

Smart Contracts describes a type of cryptographic contract in which the contractual obligations of it are executed through self-enforced computer code. Such contracts need to be resolved by an unbiased mediator. This makes distributed consensus-oriented ledger networks ideal platforms for this purpose since distributed ledger networks can apply theoretical motives for the mediator network to act honestly. It is the decentralized verification processes that make distributed ledgers suitable for smart contracts. Smart contracts are verified in the same way as Bitcoin's regular script-based transactions, which are verified by every node in the network. This means that every node has to run every contract and thus execute every contracts code placed on the blockchain. The security can be viewed as a contract between an issuer and the holder of the certificate, encoded the contractual clauses into self-enforcing computer code, and give the holder some rights, for example for voting. [11]

In order to develop or utilize a smart contract, the terms and conditions of an agreement need to be translated into code. In this code structure, all potential output capabilities need to be predefined and be agreed on. The application of smart contracts reaches from simple account transfers of cryptocurrency to applications within the Internet of Things (IoT) and smart property use cases.

2.3.12. Smart Government

Many Government institutions will benefit considerably by having simultaneous access to a distributed database that stores public records. For example, identity management passports or drivers' licenses can be placed on the Blockchain, enabling multiple agencies to share, access and verify identification in real time. Some Latin American countries like Colombia and Ecuador are exploring the use of this technology for further elections. In Ecuador on 2014 the city of Santo

Domingo, made the first electronic election in the country with a centralized protocol, now local authorities are following close this technology for 2022 elections.

Chapter 3

3. Thesis Proposal

Our main goal is an approach to manage the risks of this technology. We will analyze the protocol in details and describe the vulnerabilities in all the components, the attacks they enable and the countermeasures that can be taken.

Many of the concerns about BEV (Blockchain electronic Voting) are related to anonymity, data integrity, coercion and security attributes that our research will focus.

- We will start introducing the problems with current voting practices, description of present day's deployments of BEV, and then a breakdown of this model to analyze potential flaws and threats.

- Evaluate how the security standards for e-voting on decentralized protocol are defined in order to verify later if those properties are fulfilled on blockchain technology, taking in consideration that this protocol that started for cryptocurrency uses like bitcoin, has gained more uses nowadays, but in case of voting, still at early stage.

- Outline possible routes of adoption of blockchain technologies in the securities voting while addressing the challenges ahead.

- Research how different version of this technology helps minimize the risks of voting over Internet.

- Define a methodology to model Blockchain vulnerabilities.

4. Working plan

1.1. First year work

During the first year as a Ph.D. student, I focus on the main topics regarding decentralized protocols, and the blockchain technology, especially for cryptocurrency (Bitcoin).

We also researched some of the electronic voting platform used in countries like Estonia and Venezuela to familiarize with their implementations, and discover some of the risk and problems they faced.

I spend a few months in Ecuador at the Universidad Laica Eloy Alfaro de Manabí, at the Computer Systems Department, researching about the Weaknesses of centralized and Decentralized protocols.

Participated to the following courses and seminars.

Courses	
	Cryptography
	Introduction to Network Science
	Bertinoro International Spring School (*Approximation Algorithms, Probabilistic Graphical Model, Kleene Algebra with Test and Application to Network Programming*)
	The internet of everything everywhere: Methods and technologies for internet working land, air and sea.
	Design and analysis of secure systems
	Scientific writing and elaboration of papers (Ecuador)
Seminars	
	Mauriana Pesaresi Seminar
	Service, cloud and fog computer
	Research innovation and future of ICT

1.2. Second year working plan:

The main goal of the second year is to explore the potential of this technology from a government perspective, taking in consideration that an election involves its own uniquely difficult set of design aspects that will be research:

1. Analyses some of the most prominent projects that already implement this technology. (*Follow my vote*, *Vote Watcher, Bit Congress*), and find out the weaknesses.
2. How this technology ensures that an individual can check to see if her own vote was counted, and not be able to discern about how other people voted.
3. If there is any way that the system does not enable coerced voting.
4. Depending on the rules of the election, how the system produces or obscure interim results as desired.
5. How blockchain manages audit parameters and the level of security regarding this.
6. Analyzes recounting latency, and the lack of transparency

I will spend a few months abroad in Ecuador at the ULEAM (*Universidad Laica Eloy Alfaro de Manabí*), institution in which I have worked the last few years at the Department of Computer Science. The faculty would begin researching in blockchain technology, I will work with them in that matter during my time abroad and participate in the courses they will offer regarding that matter.

We will also participate on the 6th (ICEDEG) International Conference on Democracy and eGovernement conference held in Ecuador, also Conferences and seminars on Italy.

5. Bibliography

[1] Smartmatic, «Estonia Election,» [En línea]. Available: http://www.smartmatic.com/uploads/tx_news/CS_Estonia_elections_2014_2015.pdf.

[2] Blockgeeks. [En línea]. Available: http://blockgeeks.com/guides/what-is-blockchain-technology.

[3] SCYTL, «https://www.scytl.com/,» [En línea]. Available: https://www.scytl.com/en/online-voting-technology-security/#. [Último acceso: 15 09 2017].

[4] S. G. Robert Riemann, *Distributed Protocols at the Rescue for Trustworthy Online Voting,* 2017.

[5] M. Pilkington, *Blockchain Technology: Principles and Applications,* 2016.

[6] A. Lewis, *A Gentle Introduction To Blockchain Technology,* BRAVE NEWCOIN.

[7] A. M. Antonopoulos, *Mastering Bitcoin,* O.REILLY, 2014.

[8] B. Mundo, «bbc.com,» [En línea]. Available: http://www.bbc.com/mundo/noticias-america-latina-37539590.

[9] E. R. Results, «Electoral Commission,» 2016. [En línea]. Available: http://www.electoralcommission.org.uk/find-information-by-subject/elections-and-referendums/past-elections-and-referendums/eu-referendum/electorate-and-count-information.

[10] J. I. J. T. G. E. H. J. K. Kibin Lee, «ELECTRONIC VOTING SERVICE USING,» *Journal of Digital Forensics, Security and Law,* vol. 11, nº 2, 2016.

[11] P. Paper, «Introducing Electronic Voting - Esential Considerations,» Diciembre 2011. [En línea]. Available: http://www.eods.eu/library/IDEA.Introducing-Electronic-Voting-Essential-Considerations.pdf.

[12] S. d. A. P. OEA, «“TECNOLOGÍAS APLICADAS AL CICLO ELECTORA,» 08 2014. [En línea]. Available: https://www.oas.org/es/sap/docs/deco/Tecnologias_s.pdf.

[13] Followmyvote, «Online Voting Technology,» [En línea]. Available: https://followmyvote.com/online-voting-technology/blockchain-technology/.

[14] Francesca Caiazzo, «The Benefits and Risks of Block-Chain Voting,» 14 12 2016. [En línea]. Available: http://www.cs.tufts.edu/comp/116/archive/fall2016/fcaiazzo.pdf.

[15] J. B. E. F. Arvind Narayanan, «Bitcoin and Cryptocurrency Technologies,» 02 2016. [En línea]. Available: https://d28rh4a8wq0iu5.cloudfront.net/bitcointech/readings/princeton_bitcoin_book.pdf?a=1.

[16] https://followmyvote.com/, «https://followmyvote.com/,» [En línea]. Available: https://followmyvote.com/ .

[17] C. B. a. T. P. Team Plymouth Pioneers – Plymouth University. Andrew Barnes, «Digital Voting with the use of Blockchain Technology,» [En línea]. Available: https://www.economist.com/sites/default/files/plymouth.pdf. [Último acceso: 15 09 2017].

[18] Smartmatic, «www.smartmatic.com,» [En línea]. Available: https://www.smartmatic.com/uploads/tx_news/CS_Venezuela_2004_2015_carta_v9.1.pdf. [Último acceso: 15 09 2017].

[19] C. B. a. T. P. Andrew Barnes, «Digital Voting with the use of Blockchain Technology,» Team Plymouth Pioneers – Plymouth University.

[20] T. F. Z. D. H. H. M. M. J. A. H. Drew Springall, «Security Analysis of the Estonian Internet Voting System,» [En línea]. Available: https://jhalderm.com/pub/papers/ivoting-ccs14.pdf.

[21] S. factsheet, «Smartmatic factsheet,» [En línea]. Available: http://www.smartmatic.com/fileadmin/user_upload/Factsheet_TIVI.pdf.

[22] SCYTL, «SCYTL,» [En línea]. Available: https://www.scytl.com/es/online-voting-benefits/.

[23] S. Matsuo, «https://cyber.stanford.edu/sites/default/files/shinichiromatsuo.pdf,» de *How Formal Analysis and Verification add Security to Blockchain Based Systems*, 2017.

[24] M. Atzori, «Blockchain Technology and Decentralized,» 12 2015. [En línea]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2709713. [Último acceso: 09 2017].

[25] M. Swan, «Blockchain Blueprint for a New Economy,» O'Reilly Media, 2015, p. 54, p. 54.

[26] S. Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System,* 2014.

[27] D. E. I. a. T. Hansen, *US Secure Hash Algorithms (SHA and SHA-Based HMAC and HKDF),* 2011.

[28] J. A. H. N. H. J. M. M. N. J. W. Bos, *Elliptic curve cryptography in practice,* 2014.

[29] D. R. a. A. Shamir, *Quantitative analysis of the full bitcoin transaction.*

[30] M. K. a. S. P. M. Fleder, *Bitcoin transaction graph analysis,* 2013.

[31] G. Kostarev, *Review of blockchain consensus mechanisms.*

[32] C. D. a. R. Wattenhofer, *Bitcoin transaction malleability and MtGox,* 2014.

[33] F. T. a. B. Scheuermann, *Bitcoin and Beyond: A Technical Survey on,* 2016.

[34] P. Franco, *Understanding Bitcoin - Cryptography, Engineering and Economics,* Wiley, 2014.

[35] National Institute of Standards and Technology, *Digital Signature Standard (DSS),* 2013.

[36] M. V. Hankerson, *Guide to Elliptic Curve Cryptography*, Springer, 2004.

[37] bitcoinWiki, *Blockchain.*