2

Breaking Bitcoin at any Cost

Irrational Attacks on Proof-of-Work Blockchain

Ashish Rajendra Sai

1. Introduction

The unrest caused by the 2009 economic crisis is thought to have influenced the launch of Bitcoin, the first decentralized digital currency [24]. Bitcoin proposes a method of value transfer without the use of a trusted third party in an untrustworthy environment [12]. Bitcoin accomplishes this by employing a peer-to-peer network that maintains a global ledger of all transactions. A consensus algorithm is used to reach an agreement on the ledger's view of the data, in which the majority of network participants agree on a single view of the data.

The novel contribution of Bitcoin in the field of digital currencies is the application of a clever incentive mechanism for participants of the network. The participants of the Bitcoin network are anticipated to solve a computationally expensive trivial mathematical problem. The first participant to solve the mathematical problem acquires a reward for the computational cycles spent on the solution. This numerical solution is frequently referred to as Proof-of-Work. The reward mechanism serves as an incentive for honest behavior in the network. Dishonest behavior is de-incentivized in the form of lost rewards if half of the network dissents with the attackers' view of the data. It is presupposed based on the economic barriers that half of the network participants are honest.

University of Amsterdam.
Email: a.r.k.sai@uva.nl

Bitcoin's underlying technology Blockchain has seen numerous applications including supply chain. The utility of Blockchain in the supply chain relies on the secure operation of the underlying consensus mechanism.

Because the financial asset is entangled in the consensus, it is assumed that an attacker would try to maximize the profit from the attack. This assumption of a rational miner is employed by numerous Bitcoin security evaluation frameworks [32], [15], [35], [18], [14], [41], [31]. We believe that the assumption may limit the applicability of current attacking strategies that do not account for extreme scenarios in which the adversary is unconcerned about any financial gain or loss incurred during the attack.

In this chapter, we aim to model an irrational attacker. We describe the irrationality as the intent of harming the network, notwithstanding the cost (remunerative or reputational). The chapter intends to answer the following research question:

What attack strategies could an irrational attacker use to put the bitcoin network in jeopardy?

To answer the research questions, we devise four dimensions that an irrational attack can deploy attack vectors on to exploit the network. The design of Bitcoin determines the choice of these dimensions. The bitcoin core protocol is an implementation guideline for the participants of the network. These guidelines are realized in the core client, which is deployed over a peer-to-peer network of participants. We describe the implementation of guideline oriented attack vectors as Protocol Layer Attacks. The protocol by its design induces a race condition for the proof-of-work in the network to establish consensus on the view of the data.

The attack vectors associated with the race conditions are defined as Consensus Layer Attacks. The consensus pivots copiously on the transmission of information between the participants of the network. The attack vectors that strive to exploit the transmission of information are categorized as Network Layer Attacks.

The classifications listed above are imperative for accurate computational execution of the implementation guidelines, but one riveting factor that affects the network is the economics involved in the incentives for the participation. We categorize attacks on the incentives as Economic Layer Attacks. We can recapitulate the attacks modelled in this chapter in four categories: Protocol, Consensus, Network, and Economic layer attacks. The chapter makes the following contributions:

1. A novel investigation of attack vectors for an irrational attacker on the Bitcoin network (Section 3).

2. The chapter proposes new attack strategies on the Protocol layer including the Nonce Attack (Section 4.1) and Difficulty Recalibration Attack (Section 4.1).
3. We propose and appraise the possibility of an Extortion Attack on the Consensus Layer of the network (Section 4.2).
4. We introduce novel Denial of Service attack vectors for an Irrational Attacker on the Network Layer of Bitcoin (Section 4.3).
5. Provided a wealthy attacker; we model the feasibility of an Economic attack on the Bitcoin network by modelling artificially fabricated panic selling and a novel Griefer attack (Section 4.4).

2. Background

In this section, we first probe the structure of a blockchain followed by how Bitcoin implements a blockchain data-structure on peer-to-peer distributed systems. We also list the prominent attacks on the Proof-of-Work (PoW) based blockchains.

2.1 Structure of Bitcoin

The term Blockchain is often used as an umbrella term to refer to the broad field of Distributed Ledger Technology [33]. The term Blockchain was first used by the creator of Bitcoin Satoshi Nakamoto in a GitHub commit to note the data structure used by Bitcoin [49]. Bitcoin uses a cryptographically linked structure of blocks that accommodate transactions. Each new block of the transaction is connected to the previous block cryptographically developing a chain of blocks. The participants of the bitcoin network are tasked with the accordance on a single view of this append-only structure. To accomplish this consistency of data in an unconstrained distributed environment, Bitcoin utilizes a peer-to-peer distributed system with cleaver Proof-of-Work based incentive mechanism. The implementational details of such a system are beyond the scope of this chapter.

We cohere to the structure proposed by [25], in which the authors identify three abstract technical components to describe the complex functioning of the bitcoin network. These components are the consensus protocol, the communication network, and the transactions (including scripts). The consensus protocol is used to elect a leader via a race over the solution of a mathematically hard computational problem. The communication network component enables the participants (also referred to as Nodes) to exchange messages concerning new transactions or blocks on the network amongst other protocol level messages. The third component identified by [25] is the transactions component; in this component, the authors accumulate the protocol specifications correlated with the transactions on the bitcoin networking, including the script execution.

Even though the proposed abstracted categorization provides [25] an insightful overview of the functioning of the bitcoin network, we extend the categorization with minor alterations in the study. We design a layered architecture for the technical components of the system. The lowest layer in the proposed architecture is Protocol Layer. In the protocol layer, we define the structure of the system, including how transactions are stored and processed encompassing the transactions component proposed by [25].

Other protocol layer components include the blockchain structure used by the Bitcoin network along with the cryptographic primitives employed on the protocol level. The network participants must adhere to the protocol layer specifications to participate in the network. This adherence is often accomplished by deploying the Bitcoin-core client on the nodes.

The protocol layer specifies the behavior of the network over a distributed network of nodes (network participants). This network connection establishment and intercommunication between nodes are captured in Network Layer of the architecture. The network layer is responsible for the discovery of nodes that deploy the adhering protocol client. The network layer is also responsible for efficient communication between the nodes present in the network. The Network Layer serves as the information dissemination mechanism of the system.

The main aim of the Bitcoin network is to deterministically agree on a single view of the data under certain assumptions (the primary assumption being the requirement of the majority of nodes being honest) [47]. We describe the consensus Layer that ensures that the network reaches a consensus with some degree of assurance. Agreement between nodes over the view of data is attained through a cleaver Proof-of-Work incentive mechanism. This incentive mechanism is a part of the consensus layer in the proposed architecture.

As the article endeavors to analyze the behavior of an irrational attack, we extend the proposed technical architecture to embody the economic aspects of the system additionally. We posit that as the Bitcoin network inherently depends on the strength of the incentives to ensure that the majority of the participants are honest, it is imperative to capture the economics while analyzing the attacks. Thus we suggest a fourth layer in the layered architecture, the Economic Layer, which incorporates the economic aspects of the Bitcoin System.

2.2 Security Attacks on Bitcoin

Conventionally, payment systems have relied on a central authority to guarantee that the system is secure by validating all the transactions processed by the system. Unlike the conventional systems, bitcoin does not have a central authority, but the majority of the network is tasked with the honest validation of the transactions in the network [46]. The distributed and decentralized nature of Bitcoin introduces new attack vectors that are not present in the traditional payment systems. In this subsection, we analyze the most prominently researched security attacks on the Bitcoin network.

2.2.1 Double Spending Attack

The paramount impediment in the development of a decentralized payment system before bitcoin was the plight of a Double Spend. A double spend is defined as the payment originated with the same currency unit twice, i.e., using the same coin to pay for two transactions. Traditionally, this is solved by the central authority which maintains a ledger and can verify against the ledger if the coin has already been spent. Bitcoin provides a probabilistic guarantee that a double spending attack will not succeed with the assumption that the majority of the network is honest. One way of achieving a double spend on the Bitcoin network is by engaging a Finney Attack [55].

2.2.2 Finney Attack

An attacker who intends on double-spending transaction t_d will send the subjected transaction to a recipient simultaneously the attacker creates a conflicting transaction with a different recipient but keeps this conflicting transaction in a private chain. Once the transaction t_d is appended to a block in the Blockchain, the recipient releases the service or product to the attacker. At this point, the attacker decides to publish the block with the conflicting t_d to the Bitcoin network resulting in a fork. If the majority of the participants decide to adopt the conflicting block, the attacker has successfully spent the same bitcoin twice. The prospect of such an attack can be minimized substantially by the recipient if the recipient decides to wait for a reasonable number of block confirmation (k). The de-facto standard value of k for the Bitcoin network is considered to be 6.

In a Finney attack, the attacker exploits a private blockchain to execute the attack. These types of attacks are identified as Block Withholding Attacks [35]. Supplementary sophisticated attacks have been proposed that utilize a similar strategy of performing the mathematically intense Proof-of Work operations (also referred to as mining) in private.

2.2.3 Selfish Mining

In Selfish mining [41], the attacker aims to augment an unfair share of reward by causing harm to the other participants of the network. The attacker achieves this by mining on a private chain and publishing the private chain strategically to ensure the loss of the reward for other participants. The attacker does this with the anticipation that other participants will join the attackers' coalition.

The attackers' success probability is profoundly reliant on the direct network connection that the attacker has [41]. It is reported that network connectivity may be a decisive factor. Other attacks that exploit the network connectivity include the Eclipse Attack [34].

2.2.4 Eclipse Attack

In an Eclipse mining [26], the attacker partitions the victim from the rest of the network. The view of the data that the victim sees is a monopolized version of the actual data based on the will of the attacker. The attacker in this attack manipulates all the incoming and outgoing network connections of the victim. The attacker can exploit the eclipse miners computational power for their benefit.

These types of attacks aim to accomplish monetary gain by economically harming other entities (e.g., lost block reward). The economic attacks on the bitcoin network may prove to be most likely as the reliance on monetary value for security is a significant limitation. Some other attacks that exploit the economic properties of bitcoin include Whale Transaction [38] to reduce the growth rate of the network by promoting more forks.

2.2.5 Whale Transaction

A Whale transaction is defined as a transaction with very high monetary value as the transaction fee, which may promote forks in the network [38], [48]. All the miners are incentivized to include the whale transaction in their block to receive the significantly high transaction fee in their block. This race to include the whale transaction in the block limits the chain growth rate as most miners are indulged in the race to win the transaction fee reward.

In this attack, the authors introduced the idea of miners favoring a chain fork based on incentives. Another attack that exploits the same economic gain incentive is a Bribery attack [28].

2.2.6 Bribery Attack

In a Bribery attack [28], the attacker intends to double spend a transaction t_d by incentivizing other miners to favor the fork that includes the conflicting transaction. Combining a whale transaction to favor the fork that includes a conflicting transaction may allow for a successful double spend attack.

The attacks listed above are some of the most prominently reported attacks on the Bitcoin network in the literature. These attacks assume that the attacker seeks to attain a monetary gain by performing the attack. This assumption may seem reasonable as the cost involved in the successful execution of any of these attacks is high. However, we argue that if a largescale organization is motivated to compromise the network, they may not be incentivized by the possible gain from the attack. In this chapter, we identify more attack vectors in the architectural layers identified above for an irrational attacker. The following section provides an overview of the study design.

3. Study Design

As the Bitcoins network is appreciably intricate, an exhaustive search of all possible attack strategies is considered injudicious. We abstain from the exhaustive search by conducting a review of the architectural layers (Section 2) from the attackers' perspective. The review results in a list of potential attacks. We adhere to the reflective action science research method [3]. We do not validate if the attacks are realistic, practical or feasible, we leave this up to future research.

The reflective action science approach details a five-phase cyclic process [7], [1] consisting of Diagnosis, Action planning, Action taking, Evaluating, and Specifying Learning. In the following section, we will review the action plan in detail.

Prior to the initialization of reflective action research, a research environment must be established. In our case, the research environment is the aforementioned four-layered architecture. After the establishment of the research environment, the cyclic process is initiated with the diagnosis of the problem with the information system.

In the diagnosis phase, we aim to identify the possible attack strategies for the irrational attacker. In line with the current research in action science, the diagnosis entails self-interpretation of the problem, not by reduction and simplification, but preferably in a holistic fashion [7]. The diagnosis also assists us in the development of certain theoretical assumptions about the characteristics of the problem [9].

The diagnosis is followed by action planning phase in which researchers and practitioners collaborate to determine the methods of validation of previously identified problems. For this study, the action planning phase involves the identification of evaluation techniques for the identified problems. Theoretical frameworks traditionally drive these methods.

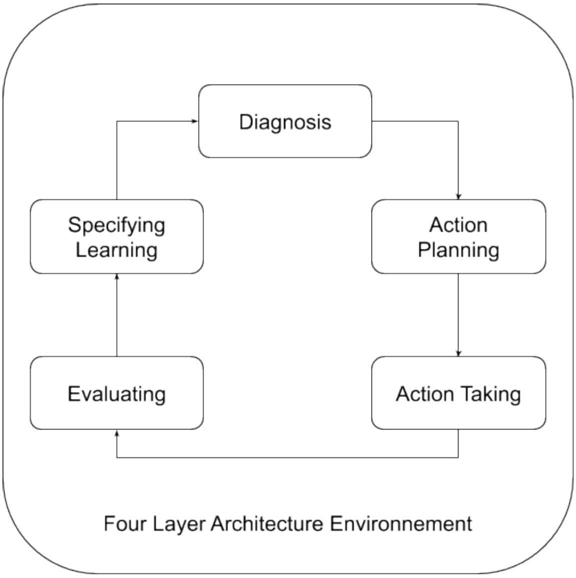

Figure 1. Action research plan.

In the next phase known as Action Taking, the identified methods of validations are implemented. After the action implementation phase, the results are evaluated by the researchers and practitioners in the Evaluation phase. The evaluation phase is followed by the Specifying Learning phase in which the knowledge gained by the action research is used to cultivate the attack strategy until satisfactory evaluation results have been obtained.[1] The reflective action science methodology, followed by the research, is illustrated in Figure 1.

4. Diagnosis

We define the problem in our diagnosis as the possible attack strategies for an irrational attacker. To identify new potential attack strategies, we rely on a systematic review of Bitcoins design choices on the previously identified architectural layers (Section 2). For each layer, we define the goal and assumptions of our irrational attacker and review the layer's design choices. The design choices analyzed are only a subset of all properties associated with the layer; we restrict our review to only the components that have already been identified by the previous research as security

[1] As stated earlier, this is beyond the scope of this chapter, we envision that future work would assess the feasibilities of these attacks.

critical. This study's design choice restricts the complexity of reviewing a broad set of the Bitcoin design component. The rest of this section reviews each of the layers and state the goal and assumptions of the irrational attacker.

4.1 Protocol Layer

The protocol layer is considered the core of the Bitcoin network as the rules defined in the protocol drive the Bitcoin system. During the inception of Bitcoin, the protocol design was largely driven by the community of a few developers, and the choices made during the early days of Bitcoin persist in most cases. These decisions may not have considered all the attack vectors. Previous research studies have reported several issues related to the protocol layer design choices [53], [51], [50], [52]. Our study aims to identify the protocol layer decisions that may prompt security threats. We define the **Goal** and **Assumption** of irrational miner on protocol layer as follows:

Goal: *Identification and exploitation of design choices that may be prone to alteration to cause harm to the network.*

Assumption: *We assume that our irrational attacker is unable to change the protocol implementation for the participating nodes.*

4.2 Consensus Layer

The novel idea of bitcoin is the clever use of incentive based consensus algorithm [37]. As demonstrated in Section 2, most of the attackers theorized are consensus mechanism oriented such as the double spending attack. The security of the bitcoin network inherently depends on the quality of the consensus mechanism. This dependency may prove it crucial for an irrational attacker to attempt and exploit the consensus layer constructs. The attacker has the following goal and works under the following assumption:

Goal: *Identification and exploitation of design choices that may be prone to alteration by external factors such as computational power of the adversary.*

Assumption : *We also assume that the attacker has less than 50 % of the computational power of the network i.e.,* $\alpha < 50\%$. *We also reduce the complexity of analysis by not considering the network latencies involved in propagating information from the attacker to the rest of the network.*

4.3 Network Layer

Any computing system that is connected to a network may possess one of many identified network vulnerabilities [10]. As the bitcoin system is a peer to-peer system, most of the possible actions in the system execute over a networking environment. Due to the open nature of the bitcoin system, the networking component of the system is exposed to the public domain where any computing node can become a part of the network of bitcoin. This dependence and open nature make bitcoin a very lucrative target for networking attacks such as DOS (Denial of Service) [23]. We assume that the irrational attacker functions under the following goal and assumption:

Goal: *Identification and exploitation of networking attacks to reduce the reliability of the network by inducing latency.*

Assumptions : *We assume that our irrational attacker can directly establish a connection to a particular portion of the network.*

4.4 Economic Layer

The consensus mechanism of bitcoin assumes that most participants are rational and will be honest to network as the reward they receive for honest behavior is more significant than dishonest behavior. The reward in bitcoin is disseminated to the participant in the form of BTC. The intrinsic value of BTC has been argued upon in the research [30], [17], [16]. This argument further strengthens the dependence of Bitcoin on the economics associated with participation [36]. Prior research in the economics of Bitcoin has reported that it is indeed possible to manipulate the intrinsic value of the BTC [43]. The illustrated irrational attacker aims to exploit the economic vulnerabilities of the Bitcoin system with the following goal and assumptions:

Goal: *Identification of economic vulnerabilities that may be exploited given an irrational motive.*

Assumptions : *We assume that our irrational attacker has an unconstrained supply of monetary assets that the attacker may use to manipulate the network.*

Based on the goals and assumptions listed above using self-interpretation of the system layers result in the attacks identified in Table 1. Table 1 contains a non-exhaustive list of possible attacks by the irrational attacker. As demonstrated in numerous action research studies, the diagnosis of these attack vectors is mostly driven by self-interpretation of the problem [7], [4], [8]. The validity of these attacks is further evaluated in the subsequent phases of the action research plan (Section 3). The remainder of this section outlines the identified attacks in detail.

Table 1. Irrational attacks.

Layer	**Identified Attacks**
Protocol Layer	Nonce Attack and Difficulty Re-calibration Attack
Consensus Layer	Extortion Attack
Network Layer	Denial of Service Attack
Economic Layer	Panic Selling and Griefer Attack

4.5 *Protocol Layer Attacks*

4.5.1 Nonce Attack

Every block in the bitcoin's blockchain contains a 32-bit field called **Nonce**.

The value of this field is varied by the miners to find the hash of the block that fulfils the target requirement of the network. This calculation is performed by specialized mining equipment. The successful calculation may result in a reward. Even though the irrational attackers' aim is not to attain higher profit, but if the attacker can gain a profit greater than the hashing power proportion, it may impact the profit of others. We observe that the honest portion of the network is probabilistic able to mine the following BTC as a reward:

$$Reward_{Honest} = B_r * (100 - \alpha)/100 \tag{1}$$

If the attacker can successfully mine more blocks without increasing the α, it may reduce the value of $Reward_{Honest}$. The loss in the reward due to the attacker's ability to mine more blocks without a significant increase in α causes harm to honest participants. The proposed nonce attack is an attempt to exploit this in order to cause harm to the network.

We speculate that due to the endianness of the mining hardware, one type of hardware may favor finding odd values of nonce than even. If the attacker can establish a pattern of nonce value, it may significantly impact the attackers' possibility of finding the actual value of nonce, thus attaining a higher reward leading to the loss of others. **Attack Strategy:**

- *Step 1*: Retrieve historical nonce data from all blocks on the blockchain.
- *Step 2*: Observe the odd and even ratio of nonce over time to devise a new mining strategy to maximize the chances of mining the next block. The attacker may also examine other patterns with a nonce to observe any other numerical biases induced by the hardware used.
- *Step 3*: Test the new mining strategy in a private bitcoin deployment to observe any improvements.
- *Step 4*: Based on the results from Step 3, the attacker may alter the strategy devised in Step 2.

- *Step 5*: The new strategy is deployed on the actual network in an attempt to increase α, consequently resulting in a lower $Revenue_{Honest}$.

4.5.2 Difficulty Recalibration Attack

The Bitcoin network aims to, on average, generate one block every 10 minutes. To maintain this block creation time, the network periodically makes it harder or easier for the rest of the network to find a solution for the nontrivial mathematical problem. This change in the difficulty of mining is referred to as the difficulty of recalibration. In Bitcoin, the difficulty recalibration is performed after 2015 blocks [54].

The block creation time is often referred to as a bottleneck for the performance of the blockchain, which is often measured in transactions per second (TPS) [44], [31]. In contrast to the traditional payment systems, Bitcoin exhibits a low TPS speed. The low TPS has been speculated to be a barrier in the mainstream adoption of bitcoin as a payment system [40].

The irrational attacker can further slow the adoption by artificially manipulating the TPS of Bitcoin. We speculate that the time difference between two difficult recalibrations may be exploited to harm the network. We devise a novel strategy in which the attacker will first attain a significant α followed by sudden withdrawal from the network. We have illustrated the attack flow in Figure 2.

The attack is significantly limited by the security provision in the Bitcoin core client, which limits the maximum rise or fall in the value of the difficulty after every recalibration [22]. However, this limit does not omit the possibility of an attack over a more extended period. The abrupt change in the total hash power induced by the irrational attacker will cause the $Revenue_{honest}$ to fluctuate significantly over time, making it less reliable to mine. When the difficulty decreases after the attacker leaves the network, the block creation time will increase significantly due to the high difficulty. This increase in block creation time will subsequently

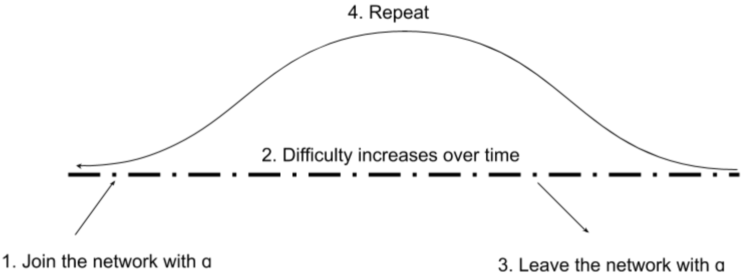
4. Repeat
2. Difficulty increases over time
1. Join the network with α
3. Leave the network with α

Figure 2. Difficulty recalibration attack.

result in a low TSP. As critiqued above, low TSP is highly undesirable as it causes direct harm to the adoption of Bitcoin as a payment system. **Attack Strategy:**

- *Step 1*: The attacker with hash power α joins the network with hash power (H_{total}) resulting in a new increased hash power ($H_{network} = H_t otal + \alpha$).
- *Step 2*: Assuming that the value of α is a significant portion of $H_{network}$, the sudden growth with result in a significantly low B_t due to a D that can only be changed after 2016 blocks. The attacker will wait till the difficulty is raised by a factor proportional to α.
- *Step 3*: Once the value of D has gradually increased to maintain a B_t of 10 minutes, the irrational attacker will decide to withdraw from the network.
- *Step 4*: The withdrawal of */alpha* from the network will result in a significantly lower $H_{network}$ but with the same D for the first recalibration period. This will result in a high B_t, thus a low TPS.
- *Step 5*: Once the network re-calibrates the difficulty to attain a B_t of 10 minutes, the attacker will start the process of B_t manipulation by repeating the attack from Step 1. The practicality of such an attack is analyzed in subsequent phases of the action research plan.

Theoretically, the attack can be carried out for an unbounded period leading to a very unreliable payment system with high variance in block creation. This may significantly harm network adoption.

4.6 Consensus Layer Attack

4.6.1 Extortion Attack

Gaining the majority in the Bitcoin network is considered hard due to the increasing cost of attaining mining hardware. One way of attaining a higher reward than the attackers hashing power proportion is Selfish Mining [32]. As reported in Section 2, in selfish mining, the attacker aims to get other miners to join his coalition to lose less reward. The rationale behind other miners joining the attackers' collocation is reduced loss in rewards. We propose another technique that aims to increase the profit of others while slowing down the network growth rate at the same time.

It has been reported that the transaction fee (T_f) only constitutes a tiny portion of total block reward [21], [19]. We propose exploiting this high dependence of consensus mechanism on the Block Reward (B_r) as the prime mean of reward dissemination. In line with previous attacks that have proposed using a high-value transaction to induce more forks in the network [38], we propose imposing a lower bound value on T_f for each block. The attacker may attempt to censor the network by only accepting transactions with $T_f > 1BTC$. This condition will result in only a tiny portion of the transactions being included in the block mined by the α proportion of the networks hashing power. Based on the value of α, the transactions backlogged may be significant, resulting in a low TPS rate without manipulating B_t.

We also speculate that other miners would be incentivized to join the attacker coalition as the attacker on average will earn more BTC from each block than honest miners as the lower bound value of T_f of the censored blocks will likely be higher than the honest blocks. The more significant the proportion of the network that participates in the attack, the more reward the participants will earn. This lucrative opportunity to earn more reward may incentivize the honest miners to join the attacker. **Attack Strategy:**

- *Step 1*: The attacker with hash power α will publically announce that the attacker will only include transactions that have a T_f higher than the lower bound defined by the attacker. This lower bound is higher than the moving average total T_f of past 100 blocks.
- *Step 2*: The attackers initializes the attack with its mining power.
- *Step 3*: The attacker will wait for others to join his attack. The public announcement may prompt rational miners to participate in the attack to earn more profit. As more rational honest miners join the coalition, the attackers α increases.
- *Step 4*: The attacker continues the attacks indefinitely causing the network to suffer a significant backlog of transactions slowing down the TPS, causing the Bitcoin network significant harm.

4.7 Network Layer Attack

Information system security is often evaluated based on the CIA triad model (Confidentiality, Integrity, and Availability) [11]. The desired security properties include confidentiality, i.e., the data processed or transmitted from the information system should not be exposed to unauthorized actors. The confidentiality in Bitcoin in ensured by the use of known secure cryptographic algorithms such as Elliptic-curve cryptography [2]. Another curial security property is the integrity of the data, i.e., the data transmitted or stored in the information system should not be altered and reminded in the intended form. This property is especially important in reference to Bitcoin as it aims to provide an immutable ledger of transactions. To attain this property, Bitcoin relies intensely on secure hashing function such as SHA256.

As Bitcoin is primarily a peer-to-peer networking system, the last of CIA triad property plays vital importance. Availability in a security context refers to the ability of the users to access the information and services provided by the information system. This property can easily be void in networking systems by inducing significant traffic to the networking device. The attacks on this property of the information system are known as Denial of Service attack, where the attacker aims to obstruct the service for a period of time [6].

As reported by [23], Bitcoin has implemented several safeguards to ensure that traditional DOS attacks do not result in loss of availability. This includes the blacklisting of bad actors in the system that sends out garbage block data [56]. We propose a new technique that aims to circumvent this limitation and exploit the node discovery protocol to induce noisy data in the network with the aim of slowing down the network. The node discovery protocol allows new nodes in the network to connect with existing nodes [27].

We aim to circumvent two security provisions against DOS implemented in Bitcoin Core Client [45]. The first provision aims to ensure that the connection to the DOS node is disconnected based on DoS score calculated by the client application for every connected node. The second provision bans the misbehaving nodes for 24 hours to ensure that the DoS attack is not successful. In our attack, we propose deploying lightweight bitcoin clients with just networking capabilities in a virtual environment. These virtual nodes actively connect with as many nodes in the network as possible and send garbage data to clog the networking link to the node. Once the virtual node is blocked by one of the peers, a new virtual node is deployed to connect to the peer with a different IP address and node address.

By virtualizing the process of node deployment, we can increase the number of full nodes on the network and attempt to connect to as many peers as possible on the network. These nodes can be deployed similarly to that of the one proposed in [45]. **Attack Strategy 1:**

- *Step 1*: If the total discoverable nodes present in the bitcoin network is *N*, the attacker will deploy greater than *N* virtual full nodes. Each virtual node will contain a list of all other virtual nodes to avoid a high degree of intercommunication.
- *Step 2*: After the deployment, the attacker′s node will act as an honest node which only relays the messages between honest nodes while still connecting to as many nodes as possible.
- *Step 3*: After a predefined period, the attacker will start sending garbage data periodically to the connected peers.
- *Step 4*: As honest peers start blocking or disconnecting from the virtual nodes, the attacker will generate more virtual nodes that will attempt to repeat the process.

We speculate that such an attack may cause significant harm for its first iteration but will most likely be patched and does not constitute a longitudinal attack. To cause more harm to the network by DoS, the attacker may wish to exploit the vulnerabilities in Mining pools. The following text explores an alternate DoS strategy in which the target is Mining pools rather than the Bitcoin network itself.

As reported in [20], even deliberate DoS attack on the mining pool may yield a profit. A well-orchestrated strategy of DoS against mining pools by other mining pools may be profitable. They report that smaller mining pools have a greater possibility of earning a profit if they attack larger pools. We propose a variant of the deliberate attack to reflect the irrationality of the attacker. The irrational attacker is not concerned about the profitability and aims to cause harm to the network. We speculate that the attacker may reduce the total hashing power of the network significantly by reducing the effectiveness of several large mining pools. We propose the following attack strategy against mining pools:

- *Step 1*: The attacker may start by shortlisting target mining pools by identifying associated IP addresses or node addresses to communicate with. This can be done by maliciously joining the mining pool and connecting to as many honest miners in the mining pool as possible.
- *Step 2*: In this step, the attacker will attempt to congest the mining pool network by disseminating a large quantity of garbage data. The attacker can repeat this process until the mining pool blocks attackers node. In the case of being blocked, the attacker may similarly morph a new node to that of Attack Strategy 1.

4.8 Economic Attacks

4.8.1 Panic Selling

As reported in [31], the security of Bitcoin is largely dependent on the reward for mining. We observe that the real world value of the reward is a significant factor that provides intrinsic value to Bitcoin [39]. This dependence on the exchange rate of Bitcoin to fiat currencies renders it vulnerable to external manipulation. The price manipulation of Bitcoin has been studied in [42]. We now attempt to analyze it by assuming an irrational actor with the intent of harming the system. By reducing the exchange value of Bitcoin significantly, the attacker may compromise the honest majority because of a lack of incentive. We propose Panic Selling as a mechanism to harm the network, given that the attacker has substantial monetary assets. The proposed Panic Selling attack consists of the following steps:

- *Step 1*: The irrational attacker gradually accumulates a large amount of Bitcoin from numerous exchanges inducing an artificial need of Bitcoins. As bitcoin's market price is driven by supply and demand model [29], the attacker may increase the price significantly during the acquisition period.

- *Step 2*: Once the attacker has accumulated a substantial proportion of the Bitcoins in supply, a selling operation is performed. By suddenly dumping a large sum of Bitcoins, the attacker can reduce the exchange rate significantly. Depending on the acquisition power of the irrational attacker, the dumping may cause bitcoin to lack any incentive for participation.
- *Step 3*: The sudden drop in the price of Bitcoin may induce fear in stakeholder leading a panic selling, the attack may benefit from such a voluntary selling increasing the effectiveness of the attack.

Panic selling attack assumes that the attacker will sell the assets at the present market price and aims to drop the exchange rate by excessive selling. We propose another economic attack, in which the attacker sells Bitcoin at an irregular price to annoy other sellers and eventually reduce the value of Bitcoin.

4.8.2 Griefer Attack

In ludology, a griefer is defined as an action that acts irrationally to antagonize other participants [13]. We propose that the irrational attacker follow a Griefer pattern when conducting an economical attack to induce uncertainty in the environment. Uncertainty and its relation with the price have been widely studied [5] suggesting a strong reliance on price on the certainty of return. The exchange rate of Bitcoin to more traditional fiat currencies is largely decided by the exchange platform. These exchange platforms often tend to have some variance in the exchange rate of Bitcoin. We propose to exploit this delegation of exchange rate determination to the third party. In the speculated attack, the attacker will arbitrarily choose the price of Bitcoin to confuse the market of resell value. The attacker may follow the following attack strategy:

- *Step 1*: Similar to Panic Selling, the attacker must gather a significant amount of Bitcoin before attacking the network.
- *Step 2*: Based on the accumulated amount of Bitcoin, the attacker may wish to either sell the Bitcoin at an arbitrary price to buyers on buying-selling platforms or establish a new exchange that offers Bitcoins at a significantly low price.
- *Step 3*: The price difference may prompt users to buy Bitcoin from the newly established exchange rather than more traditional exchanges. This migration of potential buyers may lead to most exchanges adopting the low exchange rate dictated by the malicious attackers' exchange.

5. Conclusion

The security provisions of Bitcoin have been subjected to academic scrutiny since its conception in 2009. Numerous research articles have investigated the security characteristics of decentralized cryptocurrency. Due to the economics involved in the cryptocurrencies, numerous of these articles assume a rationale attacker with the scheme to maximize the profit. This conjecture omits the possibility of an irrational attacker with the purpose of harming the cryptocurrency irrespective of the cost.

This chapter presents various novel attack vectors that an irrational attacker may exploit to jeopardize the bitcoin network. The paper manifests attacks on four facets: Consensus layer attacks, Network layer attacks, Protocol layer attacks, and Economic attacks. By modelling an irrational attacker, we can investigate the plausibility of a potential large-scale organizational or governmental attack on the Bitcoin network.

We have identified nonce and difficulty recalibration attacks as two potential sources of irrational attacks on the protocol layer. Both of these strategies have not yet been seen as an attack vector. Both of the attacks seem theoretically viable however it would be worthwhile to explore these attack vectors as these can likely be fixed by a protocol level fix.

Another type of attack that we have identified relies on the consensus protocol of the network. Other attacks that can induce instability in the network include network and economic attacks. For instance, panic selling can induce a drop in the exchange rate of Bitcoin that could potentially impact the profitability of operating in the network.

5.1 Future Work

The purpose of this chapter is to illustrate the possibility of irrational attacks on Bitcoin and other cryptocurrencies. We believe that these strategies need to be examined further to rule out any potential threat these attack strategies may pose to the network.

We intend on constructing an experimental setup to test the feasibility of these attack vectors. Specifically, examining the possibility of economic attacks using reinforcement learning may allow us to understand the economic dynamics present in the crypto-economies.

Glossary

Bribery attack is an attack in which the attacker intends to double spend a transaction t_d by incentivizing other miners to favor the fork that includes the conflicting transaction.. 7

Eclipse mining is an attack in which the attacker partitions the victim from the rest of the network.. 7

PoW Proof of work (PoW) is a method of securing a cryptocurrency network and confirming transactions by having computer networks verify large amounts of data.. 4

Selfish mining is an attack in which the attacker aims to collect an unfair share of reward by causing harm to the other participants of the network.. 6

Whale transaction is defined as a transaction with very high monetary value as the transaction fee, which may promote forks in the network.. 7

References

[1] Roger, D. Evered. 1978. An assessment of the scientific merits of action research Gerald 1. Susman and. In: *Administrative Science Quarterly,* 23.4: 582–603.

[2] Neal Koblitz. 1987. Elliptic curve cryptosystems. In: *Mathematics of Computation,* 48.177: 203–209.

[3] Harold, G. Levine and Don Rossmoore. 1993. Diagnosing the human threats to information technology implementation: A missing factor in systems analysis illustrated in a case study. In: *Journal of Management Information Systems,* 10.2: 55–73.

[4] Richard, L. Baskerville and Trevor A. Wood-Harper. 1996. A critical perspective on action research as a method for information systems research. In: *Journal of Information Technology,* 11.3: 235–246.

[5] Raymond Deneckere, Howard P. Marvel and James Peck. 1997. Demand uncertainty and price maintenance: Markdowns as destructive competition. In: *The American Economic Review*, pp. 619–641.

[6] Christoph, L. Schuba et al. 1997. Analysis of a denial of service attack on TCP. In: *Proceedings. 1997 IEEE Symposium on Security and Privacy (Cat. No. 97CB36097).* IEEE., pp. 208–223.

[7] Richard, L. Baskerville. 1999. Investigating information systems with action research. In: *Communications of the Association for Information Systems* 2.1: 19.

[8] Richard, Baskerville and Michael D. Myers. 2004. Special issue on action research in information systems: Making IS research relevant to practice: Foreword. In: *MIS Quarterly*, pp. 329–335.

[9] Rikard Lindgren, Ola Henfridsson and Ulrike Schultze. 2004. Design principles for competence management systems: a synthesis of an action research study. In: *MIS Quarterly*, pp. 435–472.

[10] Simon Hansman and Ray Hunt. 2005. A taxonomy of network and computer attacks. In: *Computers & Security* 24.1: 31–43.

[11] Sattarova, Y. Feruza and Tao-hoon Kim. 2007. IT security review: Privacy, protection, access control, assurance and system security. In: *International Journal of Multimedia and Ubiquitous Engineering,* 2.2: 17–32.

[12] Satoshi Nakamoto et al. 2008. Bitcoin: A peer-to-peer electronic cash system. In.

[13] Don Gotterbarn and James Moor. 2009. Virtual decisions: Video game ethics, Just Consequentialism, and ethics on the fly. In: *ACM SIGCAS Computers and Society* 39.3: 27–42.

[14] Ghassan Karame, Elli Androulaki and Srdjan Capkun. 2012. Two bitcoins at the price of one? Double-spending attacks on fast payments in bitcoin. *In*: *IACR Cryptology ePrint Archive,* 2012.248.

[15] Ghassan, O. Karame, Elli Androulaki and Srdjan Capkun. 2012. Doublespending fast payments in bitcoin. pp. 906–917. In: *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM.

[16] Fran Velde et al. 2013. Bitcoin: A primer. In.

[17] Adrian Blundell-Wignall. 2014. The Bitcoin Question. In.

[18] Nicolas, T. Courtois and Lear Bahack. 2014. On subversive miner strategies and block withholding attack in bitcoin digital currency. In: *arXiv preprint arXiv:1402.1718*.

[19] Nicolas Houy. 2014. The economics of Bitcoin transaction fees. In: *GATE WP* 1407.

[20] Benjamin Johnson et al. 2014. Game-theoretic analysis of DDoS attacks against Bitcoin mining pools. In: *International Conference on Financial Cryptography and Data Security*. Springer, pp. 72–86.

[21] Kerem Kaskaloglu. 2014. Near zero Bitcoin transaction fees cannot last forever. In.

[22] Karl, J. O'Dwyer and David Malone. 2014. Bitcoin mining and its energy footprint. In.

[23] Marie Vasek, Micah Thornton and Tyler Moore. 2014. Empirical analysis of denial-of-service attacks in the Bitcoin ecosystem. pp. 57–71. In: *International conference on financial cryptography and data security*. Springer.

[24] Beat Weber. 2014. Bitcoin and the legitimacy crisis of money. *In*: *Cambridge Journal of Economics* 40.1: 17–41.

[25] Joseph Bonneau et al. 2015. Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. pp. 104–121. In: *2015 IEEE Symposium on Security and Privacy*. IEEE.

[26] Ethan Heilman et al. 2015. Eclipse attacks on bitcoins peer-to-peer network. In: *24th {USENIX} Security Symposium ({USENIX} Security 15)*, pp. 129–144.

[27] Andrew Miller et al. 2015. Discovering bitcoins public topology and influential nodes. In: *et al.*

[28] Joseph Bonneau et al. 2016. Why buy when you can rent? bribery attacks on bitcoin consensus. In.

[29] Pavel Ciaian, Miroslava Rajcaniova and dArtis Kancs. 2016. The economics of BitCoin price formation. In: *Applied Economics* 48.19: 1799–1815.

[30] Anne Haubo Dyhrberg. 2016. Bitcoin, gold and the dollar–A GARCH volatility analysis. In: *Finance Research Letters* 16: 85–92.

[31] Arthur Gervais et al. 2016. On the security and performance of proof of work blockchains". In: *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. ACM, pp. 3–16.

[32] Ayelet Sapirshtein, Yonatan Sompolinsky and Aviv Zohar. 2016. Optimal selfish mining strategies in bitcoin. In: *International Conference on Financial Cryptography and Data Security*. Springer, pp. 515– 532.

[33] Mark Walport. 2016. *Distributed ledger technology: beyond blockchain. UK Government Office for Science*. Tech. rep. Tech. Rep.

[34] Karl Wu¨st and Arthur Gervais. 2016. *Ethereum eclipse attacks*. Tech. rep. ETH Zurich.

[35] Samiran Bag, Sushmita Ruj and Kouichi Sakurai. 2017. Bitcoin block withholding attack: Analysis and mitigation. In: *IEEE Transactions on Information Forensics and Security* 12.8: 1967–1978.

[36] Jonathan Chiu and Thorsten V. Koeppl. 2017. The economics of cryptocurrencies–bitcoin and beyond. In: *Available at SSRN 3048124*.

[37] Nicola Dimitri. 2017. Bitcoin mining as a contest. In: *Ledger* 2: 31–37.

[38] Kevin Liao and Jonathan Katz. 2017. Incentivizing blockchain forks via whale transactions. In: *International Conference on Financial Cryptography and Data Security*. Springer, pp. 264–279.

[39] Dirk, G. Baur, Kihoon Hong and Adrian D. Lee. 2018. Bitcoin: Medium of exchange or speculative assets? In: *Journal of International Financial Markets, Institutions and Money,* 54: 177–189.

[40] Mauro Conti et al. 2018. A survey on security and privacy issues of bitcoin. In: *IEEE Communications Surveys & Tutorials,* 20.4: 3416– 3452.

[41] Ittay Eyal and Emin Gu¨n Sirer. 2018. Majority is not enough: Bitcoin mining is vulnerable. In: *Communications of the ACM,* 61.7: 95–102.

[42] Neil Gandal et al. 2018. Price manipulation in the Bitcoin ecosystem. In: *Journal of Monetary Economics,* 95: 86–96.

[43] John, M. Griffin and Amin Shams. 2018. Is bitcoin really un-tethered? In.

[44] Ashish Rajendra Sai, Jim Buckley and Andrew Le Gear. 2018. Optimal block time for proof of work blockchains. In.

[45] Bitcoin. *Bitcoin Core Implementation Github*. 2019. url: https:// github.com/bitcoin/bitcoin.

[46] Ashish Rajendra Sai, Jim Buckley and Andrew Le Gear. 2019. Assessing the security implication of bitcoin exchange rates. In: *Computers & Security,* 86: 206–222.

[47] Ashish Rajendra Sai, Andrew Le Gear and Jim Buckley. 2019. Centralization threat metric. In.

[48] Ashish Rajendra Sai. 2021. Towards a holistic assessment of centralization in distributed ledgers. In.

[49] Ashish Rajendra Sai et al. 2021. Taxonomy of centralization in public blockchain systems: A systematic literature review. In: *Information Processing & Management,* 58.4: 102584.

[50] Gavin Andresen. *Block v2, Height in Coinbase*. url: https://github.com/bitcoin/bips/blob/master/bip-0034.mediawiki.

[51] Gavin Andresen. *March 2013 Chain Fork Post-Mortem*. url: https: //github.com/bitcoin/bips/blob/master/bip-0050.mediawiki.

[52] *CVE-2013-2273*. url: https://nvd.nist.gov/vuln/detail/CVE2013-2273.

[53] *CVE-2013-5700*. url: https://nvd.nist.gov/vuln/detail/CVE2013-5700.

[54] *Difficulty*. url: https://en.bitcoin.it/wiki/Difficulty.

[55] Finney, H. *The Finney attack(the Bitcoin Talk forum).*

[56] *Weaknesses*. url: https://en.bitcoin.it/wiki/Weaknesses\ #Denial_of_Service_.28DoS.29_attacks.