Selfish Mining and Networking Effects

Thesis zur Erlangung des Grades
Master of Science (M. Sc.)
im Studiengang Computer Science

Alexander Wagner
wagner.2@campus.tu-berlin.de

Distributed Security Infrastructures
Institut für Softwaretechnik und Theoretische Informatik
Fakultät Elektrotechnik und Informatik
Technische Universität Berlin

Gutachter:
Prof. Dr. Florian Tschorsch
TBA: Second supervisor

eingereicht am: TBA

Hiermit erkläre ich an Eides statt, dass die vorliegende, dieser Erklärung angefügte Arbeit selbstständig und nur unter Zuhilfenahme der im Literaturverzeichnis genannten Quellen und Hilfsmittel angefertigt wurde. Alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind kenntlich gemacht. Ich reiche die Arbeit erstmals als Prüfungsleistung ein.

Berlin, TBA

.....

(Alexander Wagner)

Zusammenfassung

Abstract

Contents

Chapter 1

Introduction

Bitcoin is the most prominent example of a decentralized cryptocurrency.
noch mehr über bitcoin allgemein? "trivia" wie groß etc.
It utilizes proof-of-work blockchain as a distributed ledger technology. It includes transactions into so called blocks. Blocks possess a unique ID and reference a previous block Tschorsch and Scheuermann. This construct builds a directed acyclic graph. The root of this tree is also called genesis block. Thus, every block directly or indirectly references the genesis block.
A correct block includes a nonce, which solves a cryptographic puzzle. The challenge is to alter the nonce until the hash of the set of transactions, the hash of the previous block and the nonce produce a partial hash collision. Essentially, the hash has to be smaller than some threshold value, which is also referred to as difficulty Tschorsch and Scheuermann. Thus, Bitcoin binds block creation to the computational ressources a peer possesses, since the partial hash collision can only be solved through trial and error. The correctness of the block is easily verifiable through third parties. Thus, Bitcoin ensures a fair leader election through this process.
Bitcoin uses a peer-to-peer network to propagate the mined blocks in the system. The network is unstructured as every peer tries to maintain a minimum of eight connections and performs neighbor discovery over DNS, IRC and asking neighbors Tschorsch and Scheuermann. Blocks are propagated over the peer-to-peer layer through flooding.
Once a miner mines a block through solving a cryptographic puzzle, he can publish the block and receives rewards through transaction fees and mining rewards. This provides an incentive to the miner to generate as many correct blocks as possible Barber et al..
Consensus is established over the longest chain rule Barber et al.. This means that the block ending the longest chain determines the state of the blockchain. This also implies that a miner only receives rewards, if his mined blocks are included in the main chain. Thus, a miner wants to produce as many correct blocks, that are part of the main chain, as possible. A protocol maximizing reward gain is thus incentive compatible. A miner produces a relative share of blocks proportionally to his relative share of computational power of the whole network. Thus, a miner should produce a relative share of the main chain proportional to his relative share of computational power.
The original protocol, also called honest mining, assumes publishing blocks immediately af-

ter mining. Honest mining is assumed to be incentive compatible. It follows that no miner can earn disproportionate rewards by deviating from the protocol. Consequently, earning disproportionate rewards through deviation from the honest mining protocol, would disprove Bitcoin's incentive compatability claim.

One protocol deviation is selfish mining, which was first introduced by Eyal and Sirer. Selfish Mining is a vulnerability, which aims at increasing revenue through block withholding. The selfish miner aims at producing a greater relative share of blocks of the main chain, than the relative share of computational power of the network. Therefore, selfish mining violates Bitcoins incentive compatibility claim, as it offers a more profitable mining protocol than honest mining. This is problematic, since it not only breaks fair leader election, but also results in potentially longer confirmation times for transactions of users.

Selfish mining is a statistical attack. To analyze profitibality it is therefore beneficial to analytically model selfish mining. In order to study the impact of deviating mining strategies it is very important to represent the blockchain network as close to reality as possible, to estimate realistic results.

Blockchain Mining is commonly modelled through markov decision processes. Thus, revenue gains can be analytically estimated. Eyal and Sirer first described a selfish mining model. Sapirshtein et al. further extended the model to consider all possible actions a selfish miner can perceive. This markovian model was widely used and adopted in other research directions studying other aspects of selfish mining. Bai et al. extended the model even further to analyze multiple selfish miners.

The block creation interval is much bigger than the block propagation time and is therefore modelled as instantaneous Sapirshtein et al. [2015], disregarding block propagation effects. Since most research, which is concerned with selfish mining, builts on top of the model presented by Sapirshtein et al. networking factors remain unassessed. Assuming that the underlying network does influence the system built on top, this master thesis aims to analyze the impact of networking effects on selfish mining.

Chapter 2

Model

In order to model networking effects and selfish mining, it is essential to capture network properties in an analytic model. The model can then be used to estimate selfish mining profitability. Gopalan et al. have introduced a new blockchain model, which captures network properties.

Gopalan Model The model of Gopalan et al. consists of a set of peers P connected through a peer-to-peer network. Peers add blocks to the blockchain through a process called mining. The peer-to-peer network is modelled as an undirected Graph $H = (V, E)$. An edge $(i, j) \in E$ represents communication possibilities between $v_i \in V$ and $v_j \in V$. The set of vertices is finite, such that $|V| = N \in \mathbb{N}$. Vertices are associated with peers, such that v_i represents peer $p_i \in P$. Additionally, a directed acyclic graph $G_{p_i}(t) = (B_{G_{p_i}}(t), E_{G_{p_i}}(t))$ is associated with each peer p_i, at each point in time $t \in \mathbb{R}+$. The vertex set $B_{G_{p_i}}(t) \subset \mathbb{N}$ represents the blocks known of peer p_i at time t. The associated edge set of $E_{G_{p_i}}(t)$ represents references between blocks. The following holds true for shorter notations: $B_G(t) = \cup_{i=1}^{N} B_{G_{p_i}}(t)$ and $E_G(t) = \cup_{i=1}^{N} E_{G_{p_i}}(t)$
aber warum eigentlich? macht die zeitabhängigkeit das ganze nicht kaputt?
Furthermore, the following equations hold for the principle of blockchains:

$$\forall p \in P : G_{p_i}(0) = (\{0\}, \emptyset) \tag{2.1}$$

$$t_1 < t_2 \rightarrow B_{G_{p_i}}(t_1) \subseteq B_{G_{p_i}}(t_2) \tag{2.2}$$

$$t_1 < t_2 \rightarrow E_{G_{p_i}}(t_1) \subseteq E_{G_{p_i}}(t_2) \tag{2.3}$$

Note that in this representation 0 denotes the genesis block described in equation 2.1.
$G_{p_i}(t)$ evolves over time. Blocks arrive over continuous time according to a stationary point process A with intensity λ. Each block $b \in \mathbb{N}$ arrives at a random peer p_i. This models peer p_i mining block b at time t and that at this time the block is also added to $B_{G_{p_i}}(t)$.
References are added to $E_{G_{p_i}}(t)$ according to policy and depending on $G_{p_i}(t^-)$, where t^- is a moment in time infinitesimally before t. O_i denotes the set of outgoing neighbors of block i.
The communication is modelled as a marked point process T_{p_i}. Each mark corresponds to another peer $p_j \in P \backslash \{p_i\}$. In an epoch peer p_i contacts p_j and thus, adds the lowest numbered

block of $B_{p_i}(t) \backslash B_{p_j}(t)$ to the set of Vertices B_{p_j}. If $B_{p_i}(t) \backslash B_{p_j}(t)$ is not empty, E_{p_j} is also updated accordingly.
The peer-to-peer network dynamics are modelled as a continuous time rumor-spreading process with exogenous arrivals [Gopalan et al., 2020]. Since communication is bound to the process T_{p_i}, the block dissemination is bandwidth limited. Reference selection and thus O_{p_i} is chosen accordant to longest chain policies [Gopalan et al., 2020].
Let $L_{p_i}(t)$ denote the set of nodes farthest away from the genesis block 0, known to peer p_i at time t.

$$L_{p_i}(t) := \{j \in B_{p_i}(t) : d(j, 0) \geq d(j', 0), \forall j' \in B_{p_i}(t)\} \tag{2.4}$$

Note that the set $O_{p_i} \cap L_{p_i}(t)$ is non empty. This constructs a simple directed acyclic graph. The Tree Policy [Gopalan et al., 2020] can then be determined as $|O_{p_i}| = 1$ and establishes the following relationship:

$$|E_{G_{p_i}}(t)| = |B_{G_{p_i}}(t)| - 1 \tag{2.5}$$

Every block will have exactly one outgoing reference, according to some deterministic rule [Gopalan et al., 2020]. Gopalan et al. assume that the block with the lower index number will be chosen.

model extension – selfish mining inclusion The selfish mining attack is described as a peer executing a protocol deviant from honest mining [Eyal and Sirer, 2013]. Therefore a selfish miner can be modelled according to the model described in 2 through altering the reference selection and communication process. The reference selection process is policy driven, and can thus be modified by providing a new selfish policy.
Peer $SM \in P$ has an associated policy slightly different to 2.4. Note that to follow the Tree Policy [Gopalan et al., 2020], a deterministic rule has to be established for the case that $|O_{SM} \cap L_{SM}(t)| > 1$.
Assume that SM has the knowledge of the set of blocks mined through him, $M_{SM}(t) \subset B_{G_{SM}}(t)$. SM will set

$$(L_{SM}(t) \cap M_{SM}(t)) \neq \emptyset \rightarrow L'_{SM}(t) \subset (L_{SM}(t) \cap M_{SM}(t)) \tag{2.6}$$

It then follows that $|L'_{SM}(t)| = 1$. This modified tree policy sets references according to the original selfish mining protocol described by Eyal and Sirer.
The second aspect to be modified is the communication process. It will be shown that through modification of the communication process the number of cases, where $|O_{p_i} \cap L_{SM}(t)| > 1$ increases, thus, making 2.6 impactful.
Key idea of selfish mining is block withholding. As such the selfish miner not only possesses a blockchain representation of the form $G_{p_i}(t) = (B_{G_{p_i}}(t), E_{G_{p_i}}(t))$, but rather $G_{p_i}(t) = G_{SM_{public}}(t) \subseteq G_{SM_{private}}(t)$, where $G_{SM_{private}}(t) \backslash G_{SM_{public}}(t)$ represents blocks mined but unpublished by the selfish miner.
The concept has been visualized in
$G_{SM_{comm}}(t)$, a third representation, is used to update other peers just as described in 2. $G_{SM_{private}}(t)$ interacts with $G_{SM_{public}}(t)$ through an instantaneous update process U. The first task of U is to ensure that $G_{SM_{public}}(t) \subseteq G_{SM_{private}}(t)$ holds true, meaning U updates $G_{SM_{private}}(t)$, when new blocks arrive to $G_{SM_{public}}(t)$.

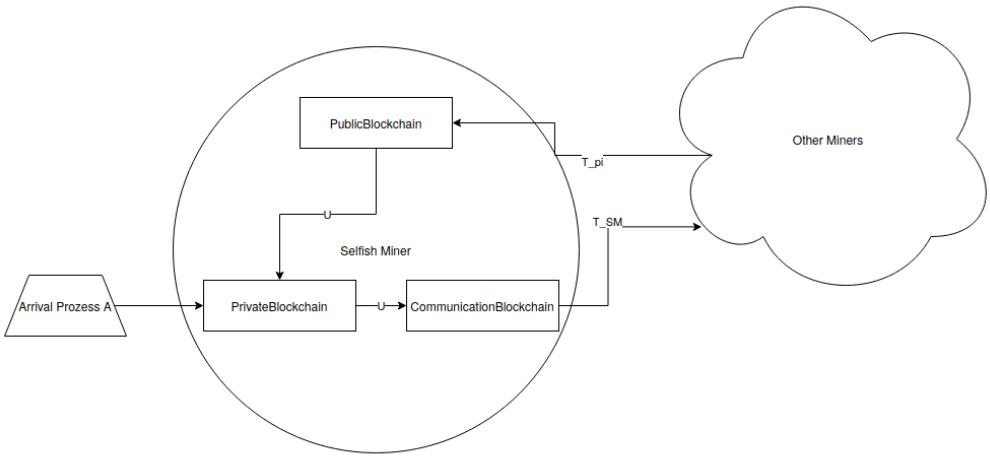

Figure 2.1: keine finale skizze, nochmal überarbeiten etc.

The second task of U is updating $G_{SM_{comm}}(t)$ according to $G_{SM_{private}}(t)$, which is the heart of the selfish mining protocol described by [Eyal and Sirer, 2013].
Let s be the state variable determining selfish mining actions [Eyal and Sirer, 2013].
Then s can be described as a difference between $G_{SM_{private}}(t)$ and $G_{SM_{public}}(t)$.

$$max\_dist(G_{p_i}(t)) := d(j, 0), j \in L_{p_i}(t) \tag{2.7}$$

$$s(t) := max\_dist(G_{SM_{private}}(t)) - max\_dist(G_{SM_{public}}(t)) \tag{2.8}$$

Let t_{inc} refer to the set of times, where s increased and analogous t_{dec} refer to the set of times, where s decreased. Let $f_{-1}(t)$ be a function that outputs the point in time, where s changed the latest before t. U can then be characterized through three kind of update actions. This can be used to model the selfish mining protocol desribed by Eyal and Sirer.

1. Assume $t \in t_{inc}$ and $s(t) \geq 2$, then U updates $G_{SM_{comm}}(t)$, such that $G_{SM_{comm}}(t) = G_{SM_{private}}(t)$.
2. Assume $t \in t_{dec}$ and $s(t) = 0$, then U updates $G_{SM_{comm}}(t)$, such that it includes the subgraph induced by the nodes on the paths between $L'_{SM}(t)$ and 0.
3. Assume $t \in t_{dec}$ and $s(t) = -1$, then U updates $G_{SM_{comm}}(t)$, such that $G_{SM_{public}}(t) \subseteq G_{SM_{comm}}(t)$.
4. Assume $t \in t_{inc}$, $s(t) = 1$, $s(f_{-1}(t)) = 0$, $s(f_{-1}(t)^-) = 1$, then U updates $G_{SM_{comm}}(t)$, such that it includes the subgraph induced by the nodes on the paths between $L'_{SM}(t)$ and 0.

hier fehlt der contest pub transition zu honest mining kann man, darf man T_{SM} modifizieren?

Chapter 3

Contribution

Forschungsfrage etc –> aufschlüsseln
Simulationen eingehen
Model checking, messungen am bitcoin netzwerk mache, darauf eingehen

Chapter 4

Evaluation

Chapter 5

Related Work

andere netzwerk analysen paper
andere selfish mining strategien und deren analyse etc.
andere currencies und selfish mining
pool mining etc. ?

Chapter 6

Conclusion

Bibliography

Qianlan Bai, Xinyan Zhou, Xing Wang, Yuedong Xu, Xin Wang, and Qingsheng Kong. A deep dive into blockchain selfish mining. Cryptology ePrint Archive, Report 2018/1084, 2018. https://eprint.iacr.org/2018/1084.

Simon Barber, Xavier Boyen, Elaine Shi, and Ersin Uzun. Bitter to better — how to make bitcoin a better currency. In Angelos D. Keromytis, editor, *Financial Cryptography and Data Security*, pages 399–414, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. ISBN 978-3-642-32946-3.

Ittay Eyal and Emin Gün Sirer. Majority is not enough: Bitcoin mining is vulnerable. *CoRR*, abs/1311.0243, 2013. URL http://arxiv.org/abs/1311.0243.

Aditya Gopalan, Abishek Sankararaman, Anwar Walid, and Sriram Vishwanath. Stability and scalability of blockchain systems, 2020.

Ayelet Sapirshtein, Yonatan Sompolinsky, and Aviv Zohar. Optimal selfish mining strategies in bitcoin. *CoRR*, abs/1507.06183, 2015. URL http://arxiv.org/abs/1507.06183.

Florian Tschorsch and Bjorn Scheuermann. Bitcoin and beyond: A technical survey on decentralized digital currencies. *IEEE Communications Surveys and Tutorials*, 18:1–1, 03 2016. doi: 10.1109/COMST.2016.2535718.