# Proof of Activity: Extending Bitcoin's Proof of Work via Proof of Stake

[Extended Abstract][†]

Iddo Bentov
Computer Science Dept., Technion
idddo@cs.technion.ac.il

Charles Lee
Litecoin Project
coblee@litecoin.org

Alex Mizrahi
chromawallet.com
alex.mizrahi@gmail.com

Meni Rosenfeld
Israeli Bitcoin Association
meni@bitcoin.org.il

## ABSTRACT

We propose a new protocol for a cryptocurrency, that builds upon the Bitcoin protocol by combining its *Proof of Work* component with a *Proof of Stake* type of system. Our *Proof of Activity* protocol offers good security against possibly practical attacks on Bitcoin, and has a relatively low penalty in terms of network communication and storage space.

## 1. INTRODUCTION

The Bitcoin [9] cryptocurrency continues to gather success since its launch in 2009. As a means of exchange, Bitcoin facilitates fast worldwide transactions with trivial fees and without identity theft risks. As a store of value, Bitcoin entails no counterparty risk and no exposure to manipulation of the money supply by central banks, due to its decentralized nature. Many other features can be derived from Bitcoin's cryptographic foundations, including: resilient forms of secret sharing via multi-signature control over an address, various types of contracts that are enforced by the distributed network, trust-free gambling and multi-party computation [1, 4], decentralized stock exchange and prediction market via colored coins [5], and so on.

Thus it is natural to ask which kinds of attacks on the Bitcoin network are likely to be practical, and which ideas would be the most effective in mitigating plausible attacks on a successful cryptocurrency. Since Bitcoin has proven to be quite capable of resisting attacks up until now, our focus is on long term sustainability. More specifically, we are especially concerned with the attack environment after *Proof of Work* (PoW) mining is no longer subsidized via the block reward, and the network needs to be secured via transaction fees acquired from the commerce taking place.

A robust cryptocurrency protocol should strive to provide an incentives structure under which it is in the self-interest of the different participants in the system to sustain its health. In this work, we offer an elaborate extension to the Bitcoin protocol as a remedy to what we argue to be probable security threats, in the sense that an attack on the extended protocol would be much more expensive. Essentially, our analysis and proposed remedy stem from the proposition that PoW miners do not have the right economic incentives to be solely in charge of securing the network, and that stakeholders are fitted to assist in this task. The *Proof of Activity* (PoA) protocol that we propose is also likely to have other features, such as promoting an enhanced network topology and less overall energy consumption. In exchange for the desirable properties that we argue that the PoA protocol achieves, nodes in the PoA network are required to do more work and communication relative to nodes in the Bitcoin network.

The purpose of the PoA protocol is to have a decentralized cyrptocurrency network whose security is based on a combination of *Proof of Work* and *Proof of Stake*. In general terms, *Proof of Work* based protocols give the decision-making power to entities who perform computational tasks, while *Proof of Stake* based protocols give the decision-making power to entities who hold stake in the system. While we contend that *Proof of Stake* based protocols offer consequential advantages, *Proof of Stake* is neither trouble-free nor effective at mitigating all the major risks that a successful cryptocurrency faces. One major risk is centralization, as data centers that are dedicated to PoW computations and transactions verification may outcompete hobbyist miners, due to economies of scale. With *Proof of Stake* systems, the particular risk of PoW data centers is indeed reduced, though other risks remain intact (see Section 2.1), and new kinds of centralization risks are introduced (large stakeholders may try to exhibit control over the system, as well as the more subtle risks that are presented in †). Another major risk is an erosion of fungibility due to blacklisting of "tainted" coins. This risk is orthogonal to *Proof of Stake*, and could be mitigated by mixing [8] or SNARKs [6].

## 2. MOTIVATION

There are different ways in which direct attacks on Bitcoin's pure PoW protocol can be attempted. One kind of an attack is the infamous >50% hashpower attack, where the attacker invests in hardware equipment (ASIC) to obtain more PoW hashpower than all the other Bitcoin miners combined [2,3,9–11]. This attacker could then double-spend by reversing the recent ledger history to defraud merchants,

[†]A full version of this paper is available at http://eprint.iacr.org/2014/452

or carry out a PoW-denial-of-service (PoW-DoS) attack by refusing to include transactions in the blocks that she generates, unless perhaps the transactions conform with the policy that the attacker imposes. In case the attacker is malicious and wishes to destroy or harm the Bitcoin network, she may achieve her objective because either double-spending or PoW-DoS can cause a loss of confidence in the Bitcoin protocol. In the case of a greedy self-interested attacker, she can make direct financial gains via double-spending, or extort others and demand higher fees by carrying out PoW-DoS.

If an attacker has enough resources to obtain >50% of the total hashpower, then it is not unreasonable to assume that she could also obtain e.g. 90% of the total hashpower, which would increase the effectiveness of the PoW-DoS attack. While it is true that the attacker depletes her resources as she carries out PoW-DoS, and therefore the Bitcoin network can survive this attack by simply waiting until the attacker gives up, in practice there could be a snowball effect where honest miners quit as confidence in the network is being lost, thus making it easier for the attacker to obtain the vast majority of the total hashpower.

## 2.1 Tragedy of the Commons related attacks

It is likely that PoW mining will become significantly less lucrative when the block reward subsidy becomes negligible, and the reward consists almost entirely of transaction fees. The rationale behind this stems from an economic phenomenon known as the "Tragedy of the Commons" [7], which is a prevalent problem in the study of economics and game theory. In its most general form, it is a system in which participants have an opportunity to act selfishly, taking action to benefit themselves at the cost of harming their peers. A selfish rational agent will always take such action because she is interested only in her own well-being; but if everyone acts selfishly, everyone will be worse off than if everyone cooperates. Any agreement to mutually cooperated for the benefit of all will not be stable, because every rational participant will prefer to enjoy the fruit of both everyone else's cooperation and her own defection.

There are several instantiations of this phenomenon in relation to funding the operation of the Bitcoin network. It is in the interest of every Bitcoin user that fees will be paid for transactions, to encourage miners to provide the network with a sufficient level of security; however, each user will prefer that others pay fees, while he pays no fee and still enjoys the network security. Without a way to force users to pay fees, the vast majority of the users will avoid paying, and end up with an insecure system that is of use to no one.

The solution lies within the power miners have to reject transactions if the fee paid is not high enough. However, here we have another tragedy of the commons problem, between the different miners. Without protocol-enforced limitations on what can go into a block, a rational miner will prefer to include every fee-paying transaction, even if the fee is very low, because the marginal cost of including a transaction is trivial. If miners accept low-fee transactions, users will have no reason to pay significant fees, and the total fees that can be collected by miners will not be sufficient to cover the cost of PoW mining.

The miners could try to form an agreement to accept only high-fee transactions; in this scenario, users will be forced to pay high fees if they want their transaction included (meaning, fees that are high enough for an adequate network security, while low enough for the market to bear - which is desirable). However, the agreement is not stable - it is in the interest of each miner to defect and accept low-fee transactions, as this action boosts the revenue of this individual miner. If all the miners do so, the bargaining power of miners will be eroded and they will no longer be able to force high fees, reducing the total revenue.

Since the total mining revenue is what funds the overall network security, in this case the network security will be weak. For this reason, maintaining a healthy network requires some protocol-enforced rules protecting miners, as a group, from themselves - such as, a cap on the total value transferred in transactions in each block. If the cap is properly chosen, miners will actually earn more with this kind of a cap in place - by making block space a scarce resource, its price goes up; transactions will have to compete with others for admission, and pay high fees for the privilege. An individual miner cannot break the market by accepting low-fee transactions, as she can only put so many in the block.

For example, let us say there exist in the market 1000 users wanting to send a transaction of 1 BTC each, and willing to pay 1% of the transaction value in fees (the amount per transaction is irrelevant as long as the total BTC and percentage fee stay the same). In addition, there are 1000 users wishing to send 1 BTC, for whom the transaction is not as important and so they are only willing to pay 0.1% in fees. If there is no cap, every miner will want to include all transactions, even if their fee is only 0.1%. There is no way to effectively segment the market, so with 0.1% fee transactions being accepted, this is what will be paid also by those willing to pay more. The total transaction value will be 2000 BTC and the fee paid is 0.1%, for a total of 2 BTC. If, however, a cap is placed at 999 BTC, the miners are no longer free to include all transactions. They must choose which 999 transactions to include, and the equilibrium is that those users who can bear it will pay 1% (if they pay less, a user whose transaction is excluded will offer a higher fee to get ahead of the others, and so on). Thus, 999 BTC will be transacted but with a fee of 1%, and the total revenue is 9.99 BTC.

Still, if the total cost of PoW mining at an adequate security level is more than what the market can bear, just having a protocol-enforced rule such as a block value cap will not be enough. When the transactions volume (and market cap) of a cryptocurrency increases, the total amount of transaction fees will increase too (as derived from the fees that the users can bear), and therefore it may seem that this total fees amount will be sufficient to fund the security of the network. However, as the market cap of the cryptocurrrency grows, the incentives to attack it also increase, hence the cost of maintaining the security of the network is correlated with an increased market cap and transactions volume. Thus, we need a way to have a high ratio of security to transaction fees. The growth of a *Proof of Work* based cryptocurrency should not be expected to make this ratio better, unlike *Proof of Stake* based protocols such as PoA, because the entities that secure the PoA network have fewer expenses and may therefore collect lower fees (due to competition among them). Stated differently, since the overall network security is proportional to the total transaction fees paid, users may disagree on which protocol rules offer the best tradeoff between security and low fees, and *Proof of Stake* relieves some of this tension as network security requires less computational effort and therefore has a lower

cost attached to it.

There is also a third tragedy of the commons problem, as the transaction fees are paid only to the miner who created the block, while the cost of propagating, verifying, and storing the transactions is shared by all the nodes in the network. Miners will prefer keeping every transaction to themselves and collecting a fee for it, while avoiding as much as possible the work of propagating it. Having a value cap for each block will not help in this regard, because the users as a whole may still wish to send a very large amounts of low-value transactions. Here the solution is limits on data size and CPU cycles (currently dominated by ECDSA signature verifications) for each block, which is controversial since many users believe that the block size should accommodate the Bitcoin economy. *Proof of Stake* based protocols offer little help here, as they do not reduce these particular costs.

In the overall scheme of things, a negligible block reward subsidy is likely to imply an environment in which it is easier to carry out PoW-based attacks, because there would be fewer honest miners to compete against. In particular, for a governmental entity who wishes to destroy the competition against the fiat currency that it issues, a clandestine ASIC-based attack could be quite easy. An entity who carries out PoW-based attacks under these circumstances is not likely to possess any significant amount of bitcoins, since Bitcoin stakeholders would be scrambling to keep the network secure in order to protect their fortune. Moreover, the PoW hardware may retain some resale value to the attacker, and also to other miners who prefer to exchange the bitcoins that they earn for fiat currencies and hence have no paramount interest in the soundness of the cryptocurrency. Note that ASIC mining hardware may have resale value too, as it can be repurposed to mine alternative cryptocurrenies. Nowadays, self-interested miners can even delegate their (ASIC) hashpower to a service that automatically switches among the most profitable cryptocurrencies to mine, implying that such miners are oblivious as to whether their hashpower participates in attacks. Therefore, it makes sense to vest part of the power that synchronizes the transactions in the hands of the stakeholders, rather than vesting all of this power in the hands of the PoW miners.

## 2.2 Other types of hazards

Another category of direct attacks is network attacks, in particular network denial-of-service and network isolation. Here the topology of the online nodes in the Bitcoin network is attacked. When the connectivity between the nodes is low, it becomes easier to deny service by flooding miner nodes, or carry out a Sybil attack by isolating and transacting with some specific node. Due to pooled PoW mining, the current topology of the Bitcoin network is barely at the level of a single popular torrent. In contrast, as we will see in Section 3, the PoA protocol incentivizes non-miner nodes to maintain a continual online presence, and this should be quite helpful to the overall network topology. Even if all the Bitcoin miners were using p2pool, Bitcoin's network topology would probably still be worse than the PoA network topology, since many Bitcoin users are not miners (c.f. †).

# 3. PROTOCOL

The primary subroutine that PoA incorporates is called *follow-the-satoshi*, whereby we transform some pseudorandom value into a satoshi (smallest unit of the cryptocurrency) that is picked uniformly among all the satoshis that have been minted thus far. This is done by selecting a pseudorandom index between zero and the total number of satoshis in existence up to the last block, inspecting the block in which this satoshi was minted, and following each transaction that transferred this satoshi to a subsequent address until reaching the address that currently controls this satoshi. Note that this process can be regarded as picking a pseudorandom stakeholder in a uniform fashion, e.g. if Alice has 2 coins and Bob has 6 coins then Alice is 3 times less likely to be picked compared to Bob.

1. Each miner uses her hashpower to try to generate an empty block header, i.e. header data that consists of the hash of the previous block, the miner's public address, and a nonce. This header does not reference any transactions.
2. When a miner succeeds in generating an empty block header, meaning that the hash of her block header data is smaller than the current difficulty target, she broadcasts her block header to the network.
3. All the network nodes regard the hash of this block header as data that deterministically derives $N$ pseudorandom stakeholders. The derivation is done by concatenating this hash with the hash of the previous block and with $N$ fixed suffix values, then hashing each combination, and then invoking *follow-the-satoshi* with each of the $N$ hashes as input.
4. Every stakeholder who is online checks whether the empty block header that the miner broadcasted is valid, meaning that it contains the hash of the previous block and meets the current difficulty. Upon validation, the stakeholder checks whether she is one of the $N$ lucky stakeholders of this block. When the first $N-1$ lucky stakeholders discover that the block derives them, they sign the hash of this empty block header with the private key that controls their derived satoshi, and broadcast their signature to the network. When the $N^{\text{th}}$ stakeholder sees that the block derives her, she creates a wrapped block that extends the empty block header by including as many transactions as she wishes to include, the $N-1$ signatures of the other derived stakeholders, and her own signature for the hash of this entire block.
5. The $N^{\text{th}}$ stakeholder broadcasts the wrapped block to the network, and when the other nodes see that this wrapped block is valid according to the above, they consider it a legitimate extension of the blockchain. The nodes try to extend the longest branch of the blockchain that they are aware of, where "longest" is measured in PoW difficulty as in Bitcoin.

- The fees from the transactions that the $N^{\text{th}}$ stakeholder collected are shared between the miner and the $N$ lucky stakeholders (see † for details).

**Figure 1: The PoA protocol (think $N = 3$)**

# 4. ANALYSIS

If some of the $N$ lucky stakeholders were offline, then other miners will also solve the block and thereby derive $N$ other pseudorandom stakeholders, so the overall difficulty will readjust both according to the total hashpower and according to what fraction of all the stakeholders is online.

Let $E_1 = \{$the $N$ lucky stakeholders that the block derives are under the attacker's control$\}$ and $E_2 = \{$the $N$

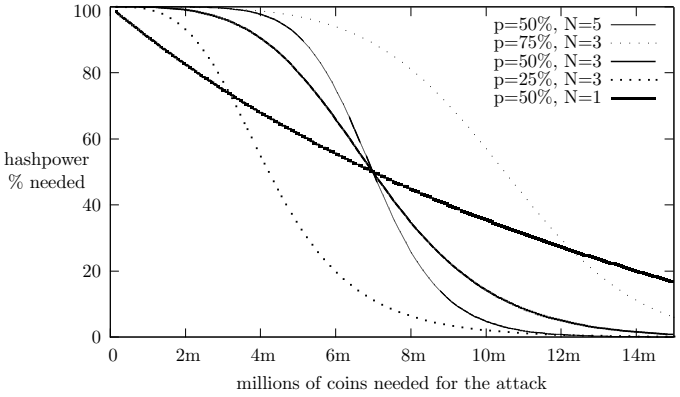
lucky stakeholders that the block derives are honest}. We condition on the event $E_3 = \{$the $N$ lucky stakeholders that the block derives are online$\}$, and note that $\Pr[E_1|E_3] = x^N$ is the probability that $N$ online stakeholders that a mined block derives are under the attacker's control, and that $\Pr[E_2|E_3] = (1-x)^N$ is the probability that $N$ online stakeholders that a mined block derives are honest. This means that on average the attacker will generate a block after $(\frac{1}{x})^N$ nonce attempts that meet the current difficulty target and derive $N$ online stakeholders, while the honest network needs $(\frac{1}{1-x})^N$ such attempts on average. Therefore, if the attacker is fast enough so that she could compute $(\frac{1}{x})^N/(\frac{1}{1-x})^N = (\frac{1}{x}-1)^N$ nonce attempts per one nonce attempt of the honest network, she can generate the blocks at the same average speed as the rest of the network. It follows that if $p$ fraction of the honest stake is online, an attacker with $y$ fraction of the *total* stake needs more than $((\frac{1}{y}-1)\cdot p)^N$ times the hashpower of the honest miners in order to gain advantage over the network, because the speedup factor that the attacker needs is $((1-y)\cdot p)^N/y^N = ((\frac{1}{y}-1)\cdot p)^N$, where $((1-y)\cdot p)^N$ is the probability that $N$ derived stakeholders are both online and honest, and $y^N$ is the probability that $N$ derived stakeholders are controlled by the attacker.

For example, if $N = 3$ and we assume that transaction fees (and PoA add-ons †) incentivize $p = 50\%$ participation level of the honest stakeholders, an attacker who has 88.8% of the total hashpower and $y = 20\%$ of the total stake would still not have an advantage over the honest network: $((1/\frac{20}{100} - 1)\cdot\frac{50}{100})^3 = 2^3$, meaning that the attacker needs to be more than 8 times faster than the rest of the network.

With a PoW-based cryptocurrency, the security is sustained under the assumption that the majority of the mining power that participates is honest. Similarly, the PoA network derives its soundness from the assumption that the majority of the online stake is honest. In Section 2.1 we argue that for a cryptocurrecy to be attack-resistant over the long term, relying on the assumption that the majority of the stake is honest is more conservative than to rely on the assumption that the majority of the hashpower is honest.

To get a rough idea of the cost difference between an attack on Bitcoin and an attack on PoA, let us take for example an `AntMiner S2` ASIC unit that runs at 1 terahash/s and costs about 8 coins. Currently the total hashrate of the Bitcoin network is around 50,000 terahash/s, therefore an attacker needs to have under her control $\approx 50,000$ `AntMiner` units that cost 400,000 coins, in order to have $\frac{1}{2}$ of the total hashrate. Contrast that to e.g. 4.2 million coins that an attacker needs to control in order to have 20% of a total stake of 21 million coins, for gaining just $\frac{1}{3}$ of the online stake in a network in which 50% of the honest stakeholders participate. If we take $N = 3$ and assume that the hashrate of the PoA network is for example $\frac{1}{10}$ of Bitcoin's pure PoW network, i.e. around 5000 terahash/s, then this attacker also needs to control about 40,000 `AntMiner` units with a price tag of 320,000 coins, in order to be 8 times faster than the honest miners in the PoA network. Keep in mind that if the total hashrate of the PoA network is indeed $\frac{1}{10}$ of Bitcoin's pure PoW network, then PoA is much more efficient in terms of energy consumption.

Let us summarize the costs of an attack on PoA by giving exemplary figures in the chart below. The attacker's expenses depend on the honest stakeholders' participation level $p$ and the amplification parameter $N$. We assume that the coins needed for the attack are out of a total of 21 million minted coins.




# 5. CONCLUSION

The PoA protocol seeks to decentralize the power that synchronizes the transactions in a quite pronounced fashion. To monopolize the block creation process, an attacker needs to control a substantial fraction of the total amount of coins that have been generated thus far. We argue that in likely scenarios the cost of an attack would be much higher with the PoA protocol than with Bitcoin's pure PoW protocol. Furthermore, the PoA protocol is likely to accomplish other beneficial properties, namely an improved network topology, low transaction fees, and a more efficient energy usage.

# 6. REFERENCES


[1] Andrychowicz M., Dziembowski, S., Malinowski D., and Mazurek, L. 2014. Secure multiparty computations on bitcoin. In *IEEE Security and Privacy*.

[2] Babaioff, M., Dobzinski, S., Oren, S., and Zohar, A. 2012. On bitcoin and red balloons. In *ACM Conference on Electronic Commerce*, pp. 56–73.

[3] Barber, S., Boyen, X., Shi, E., and Uzun, E. 2012. Bitter to better - how to make bitcoin a better currency. In *Financial Cryptography*, LNCS vol. 7397, pp. 399–414.

[4] Bentov, I. and Kumaresan, R. 2014. How to Use Bitcoin to Design Fair Protocols. In *Crypto 2014*.

[5] Bonneau, J., Clark, J., Felten, E., Kroll, J., Miller, A. and Narayanan, A. 2014. On Decentralizing Prediction Markets and Order Books. In *Workshop on the Economics of Information Security*.

[6] Chiesa, A., Garman, C., Miers, I., Virza, M., Ben-Sasson, E., Green, M., and Tromer, E. 2014. Zerocash: Practical decentralized anonymous e-cash from bitcoin. In *IEEE Security and Privacy (S&P)*.

[7] Hardin, G. 1968. The tragedy of the commons. *Science 162*, 1243–1248.

[8] Maxwell, G. 2013. CoinJoin. *Bitcoin forum thread*. https://bitcointalk.org/index.php?topic=279249.0.

[9] Nakamoto, S. 2008. Bitcoin: A peer-to-peer electronic cash system. *Bitcoin.org*.

[10] Rosenfeld, M. 2012a. Analysis of hashrate-based double-spending. http://arxiv.org/abs/1402.2009.

[11] Sompolinsky, Y. and Zohar, A. 2013. Accelerating bitcoin's transaction processing. In *ePrint* 2013/881.