Discussion Section #3

April 22, 2025

1. We have seen in discussion section 2 that given the k-deep confirmation rule on Bitcoin, private attack against an honest block B succeeds with probability

$$p \leq \frac{1}{1-4\beta(1-\beta)} \left(4\beta(1-\beta)\right)^k,$$

which can be made arbitrarily small by selecting a large k. Given this information, can you think of an attack on Bitcoin's safety that reverts a block or a transaction with probability 1?

Answer:

Consider an adversary that constantly attempts to create a long private chain of empty blocks. If the public (honest) chain's height passes the height of the adversary's chain, the adversary abandons its current attack and starts building a new private chain off the latest public block (thus, a new iteration of the attack starts). If the private chain becomes k-blocks long before the public chain catches up to it in height, the adversary has succeeded: the private chain is published once the public chain reaches the same height as the private chain. Once this happens, the adversary reverts the confirmed transactions within the replaced part of the public chain. Note that at each iteration of the attack, the adversary succeeds with probability p, which is at least $\beta^k > 0$, and the iterations are independent. Therefore, the adversary eventually succeeds with probability 1, in fact, within expected constant time, which is $\Theta(1/p)$!

The adversary can even revert transactions of its choosing (let's denote this transaction by tx) by doing a modified version of the attack. In the modified version, the adversary again attempts to create a long private chain of empty blocks. However, it restarts the attack if the private chain cannot get k blocks ahead of the public (honest) chain. In other words, it restarts the attack if the adversary's private chain cannot receive k blocks before the public chain receives any. Once the private chain becomes k-blocks longer than the public chain, the adversary sends tx to the mempool, waits for it to be confirmed, and then publishes the private chain, thus reverting tx. Even though this is a much harder attack to pull off, the adversary again succeeds with probability at least $\beta^k > 0$, and new iterations of the attack are independent. Therefore, the adversary again succeeds with probability 1 (eventually), in fact, within expected constant time, which is $\Theta(1/\beta^k)$!

Is Bitcoin not secure?? The catch is that if the protocol is run sufficiently long, there is no way to guarantee that there will not be any attack on Bitcoin, even under an honest majority assumption for the mining power. However, for reasonable time horizons, we can guarantee that there will not be any attack with high probability.

Let k in the k-deep rule be our *security parameter*. After all, we select k to be larger for more confidence that the confirmed blocks will not be reverted. Suppose the protocol is run for a duration of T that is selected as a polynomial function of the security parameter k, i.e., $T = \mathsf{poly}(k)$ for some polynomial $\mathsf{poly}(\cdot)$. Now, no confirmed block B targeted by the adversary can be reverted except with probability $Ce^{-\tilde{c}k}$ for some constants C and $\tilde{c}$, as we have seen in the discussion section 2. Then, by a union bound, we can infer that no confirmed block (or confirmed transaction) can be reverted throughout the duration T except with probability

$$O(T \cdot C \cdot e^{-\tilde{c}k}) = \mathsf{poly}'(k) \cdot e^{-\tilde{c}k},$$

where $\mathsf{poly}'(\cdot)$ is some polynomial in k. Note that again, by selecting k to be sufficiently large, we can make this probability arbitrarily small, as $e^{-\tilde{c}k}$ decays faster than any polynomial's growth. Formally, we observe that no confirmed block (or confirmed transaction) can be reverted throughout the duration the protocol is run, except with *negligible* probability in the security parameter, i.e., except with probability $f(k)$, which satisfies $f(k) = o(1/g(k))$ for any polynomial $g(\cdot)$ in k (here, $o(\cdot)$ denotes the little-o notation).
