EE374 Scaling Blockchains **Stanford, Spring 2020**

Homework #1*

Due: 20-April-2020, 11:59pm – Gradescope entry code: `92J6Y3`

Please upload your answers timely to Gradescope. Start a new page for every problem. We strongly suggest LaTeX to type your answers. For the programming/simulation questions you can use any reasonable programming language (please no assembly, brainfuck, etc. ☺). Comment your source code and include the code and a brief overall explanation with your answers. A tentative point distribution (in % of the total) is provided in brackets. For most problems there is more than one valid way of solving them!

1. (20%) In this question we discuss the stochastic modelling of the mining times of Bitcoin.
   a) What is a reasonable model for the distribution of the time between two consecutive mining events? Explain. What is the variance of the inter-mining time under this model?
   b) Using data from `https://btc.com/block`, estimate the variance of the mining time. Is it close to what your model in part (a) predicts? Explain if there is any significant discrepancy. State carefully how you perform the estimation and justify why you estimate this way.
2. (15%) The total hashrate of the Bitcoin network on January 1, 2018 was 14.4 EH/s.
   a) Estimate the threshold in the hash inequality on that day from this fact. Compare this with the true threshold. Why might there be a discrepancy?
   b) Assume all mining was conducted using back then state-of-the-art Antminer S9 hardware, which delivers 14 TH/s at a power consumption of 1372 W. What was the power consumption of the Bitcoin network? Find a comparable country in `https://en.wikipedia.org/wiki/List_of_countries_by_electricity_consumption`.
3. (60%) Let's analyse and simulate the longest chain protocol under two attacks.
   a) Simulate the longest chain protocol under Nakamoto's private attack, and estimate the confirmation error probability for $k = 5, 10, 15, 20$ and for adversarial hash power fraction $\beta = 0.3, 0.45$. Compare your results with Nakamoto's table. Explain if there is any significant discrepancy.

      Hint: Make sure to repeat the runs of the protocol sufficiently often to generate reliable estimates of the error probabilities.

*Version: 3 – Last update: 14-April-2020

b) Consider another attack which is Nakamoto's private attack combined with a *pre-mining phase*. The attack is focused on reverting a transaction TX included in the i-th block of the public chain.

- *Pre-mining phase*: Starting from the genesis block, the attacker starts mining blocks in private to build a private chain. When the first honest block h_1 is mined on the genesis block, the attacker does one of two things: i) If the private chain is longer than the public chain at that moment, then the adversary continues mining on the private chain; ii) if the private chain is equal or shorter then the public chain, the attacker abandons the private chain it has been mining on and starts a new private chain on h_1 instead. The attacker repeats this process with all honest blocks $h_2, h_3, \ldots h_{i-1}$.
- *Private attack phase*: After block h_{i-1} is mined, the attacker will start Nakamoto's private attack from the current private chain it is working on, whether it is off h_{i-1} or the one it has been working on before h_{i-1} depending on which is longer.

Answer the following questions:

i. Suppose $\beta = 0.3$. What is the probability that the attacker will switch to h_1 when it is mined? What is the expected depth at which the attacker is mining when h_1 arrives?

ii. Suppose honest (h) and adversarial blocks (a) are mined in the order:

$$a_1, a_2, h_1, a_3, h_2, h_3, a_4, a_5, h_4.$$

Draw the evolution of the block tree, always including both honest and adversary blocks.

iii. Let D^a_{i-1} be the depth at which the adversary is mining just before the $(i-1)$-th honest block arrives. Let $G_{i-1} := D^a_{i-1} - i + 1$ be the advantage the adversary has over the public chain just after that time. The distribution of G_{i-1} depends on i. What happens when $i \to \infty$?

iv. Argue that this attack is strictly stronger than the pure Nakamoto's private attack. Be precise what this means.

v. Simulate this attack for large i and estimate the confirmation error probability for $k = 5, 10, 15, 20$ and adversarial hash power fraction $\beta = 0.3, 0.45$. Compare these results with those in part (a). Are there significant differences?

vi. (Bonus) In Lecture 3, we bounded the confirmation error probability under Nakamoto's private attack and show that it exponentially decreases with k, with a bound c on the exponent. Show that the confirmation error probability under the attack with pre-mining is also exponentially decreasing, satisfying the same bound on the exponent. (Hint: consider all possibilities by which the adversary switches in the pre-mining phase.)

4. (5%) Please give us feedback.

   a) Roughly how much time did you spend on this homework?

   b) What is your study background (major, degree, years into the degree)?

   c) What was your most favorite problem on this homework and why?

   d) What was your least favorite problem on this homework and why?

   e) Any other feedback (lecture, office hours, homework, ...)?