EE374 Internet-scale Consensus **Stanford, Winter 2021**

Homework #1*

Due: Fri, 22-January-2021, 11:59pm – Gradescope entry code: `N8XV23`

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. (30%) In this question we discuss the stochastic modelling of the mining times of Bitcoin.
   a) A reasonable model for the distribution of the time between two consecutive mining events is exponential. Write down this probability density function with specific choice of all parameters to model ideal Bitcoin.
   b) What is the standard deviation of the inter-mining time under this model? What is the ratio of the standard deviation over the mean?
   c) What is the mean of the time it takes to mine 10 blocks? What is the standard deviation, and the ratio of the standard deviation over the mean? Be explicit about any assumptions you made to get this conclusion, and justify your modeling assumptions.
   d) Using data from `https://btc.com/block`, estimate the standard deviation of the inter-block 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. (20%) 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`.

*Version: 1 – Last update: 15-Jan-2021

3. (35%) According to Figure 7 of lecture note #2, the mining pool *F2Pool* has about 20% of the total hash rate of the current Bitcoin network.

   a) What is the distribution of the time we need to wait until a block is mined by *F2Pool*? Justify your answer.

   b) What is the probability that the next block mined is from *F2Pool*? Justify your answer.

   c) At the end of Lecture 2, we consider an attack on the confirmation rule where a transaction enters the ledger as soon as its block enters the longest chain. Alice's transaction is in a block B and she is trying to mine two blocks in private before the honest miners are able to mine a new block on B, so as to remove the transaction from the ledger. Suppose *F2Pool* is trying to perform this attack. What is the probability it will succeed? You can assume the attacker starts mining the private blocks as soon as the transaction enters the mempool, and that the transaction will be in the next block mined by the network. Make explicit any other assumptions you are making in deriving the answer.

4. (15 %) Using the data in `https://www.blockchain.com/charts/` or elsewhere, estimate the following quantities, in the past month on the Bitcoin network:

   a) The average size of a transaction, in bytes;

   b) the average size of a block in Mbytes;

   c) the throughput in transactions per seconds (Tps);

   d) compare this to the throughput of a system like Visa.

   e) Roughly how many transactions are waiting to be confirmed on average?

   f) Any changes in the charts that could be explained by the recent price fluctuations?

   Provide a short description of how you derived the estimates and state any assumptions you have made.