COMS 4995-001 (Science of Blockchains): Homework #9

Due by 11:59 PM on Wednesday, April 23rd, 2025

Instructions:

(1) Solutions are to be completed and submitted in pairs.

(2) We are using Gradescope for homework submissions. See the course home page for instructions, the late day policy, and the School of Engineering honor code.

(3) Please type your solutions if possible and we encourage you to use the LaTeX template provided on the Courseworks page.

(4) Write convincingly but not excessively. (We reserve the right to deduct points for egregiously bad or excessive writing.)

(5) Except where otherwise noted, you may refer to your lecture notes and the specific supplementary readings listed on the course Web page *only*.

(6) You are not permitted to look up solutions to these problems on the Web. You should cite any outside sources that you used. All words should be your own. Submissions that violate these guidelines will (at best) be given zero credit, and may be treated as honor code violations.

(7) You can discuss the problems verbally at a high level with other pairs. And of course, you are encouraged to contact the course staff (via the discussion forum or office hours) for additional help.

(8) If you discuss solution approaches with anyone outside of your pair, you must list their names on the front page of your write-up.

Problem 1

(15 points) This problem concerns the sampling of a leader in a quasi-permissionless proof-of-stake setting. Suppose there are n validators, with respective stakes $s_1, \ldots, s_n$. Denote the total amount of stake by $S = \sum_{i=1}^{n} s_i$. Finally, let $S_0 = 0$. Suppose further that we are given a perfectly uniform random number $r \in [0, 1]$ that falls from the sky. Given r, we would like to sample a validator with probability proportional to stake — i.e., we'd like to select validator i as the leader with probability s_i/S. Consider the following three attempts at doing so.

1. Denote by $S_i = \sum_{j=1}^{i} s_j$ the partial sums up to i of the stake. Suppose we perform the following: We choose validator i as the leader iff $r \in [\frac{S_{i-1}}{S}, \frac{S_i}{S}]$. Does this work (i.e., sample according to the desired distribution)? Why or why not?

2. Suppose we perform the following. For $i = 1, 2, \ldots, n$, do:

   - Flip a biased coin with probability $\frac{s_i}{\sum_{j=i}^{n} s_j}$ of "heads."
   - If it lands "heads," select validator i as the leader and halt.
   - Otherwise, continue.

Does this work? Why or why not?

3. The following approach is known as the "follow-the-satoshi" algorithm. Suppose that one can map a random point $r \in [0, 1]$ to a uniformly random coin in a blockchain protocol, e.g. a uniformly random wei (the smallest unit of currency in Ethereum) or a uniformly random satoshi (the smallest unit of currency in Bitcoin). Denote this coin by c. Now do the following.

   - Let pk be the public key that owns c at this time.
   - If pk is the public key of a validator i amongst the current set of validators, select i as the leader.
   - Else, sample a new r and corresponding coin c and repeat.

   Would this approach work? Why or why not? Separately, is the idea of "the current owner of coin c" well defined in a UTXO-based blockchain like Bitcoin (see Lecture 8)? If it isn't, can you propose some canonical way of making the idea well defined?

Problem 2

(10 points) This problem explores the idea of selecting a leader via a hash-based lottery.

1. Specifically, suppose we select validator i to be a leader if and only if $\mathsf{Sha}_{256}(sign_i(pk_i)||t||r_t) \leq \tau_i$, where the τ_i's are proportional to validators' stakes (in the notation of Problem 1, to the s_i's). Here $sign_i(\cdot)$ indicates the signing algorithm with i's secret key, pk_i is i's public key, t is a time step, and r_t is a 256-bit uniformly random seed (specific to the timestep and known to all validators). If our goal is to select in expectation one leader per timestep, how should we set the τ_i's as a function of $s_1, \ldots, s_n$?

2. Is the distribution over outcomes of the above approach identical in all aspects to the approaches described in Problem 1? If not, discuss some of the advantages and disadvantages of the above leader sampling method compared to those described in Problem 1. Does your answer depend on whether the sampling procedure is being used in the context of selecting leaders in longest-chain consensus vs. in a Tendermint-style protocol?