*Selfish mining pools*

Analysis of AntPool and F2Pool

Isaac Vawter, Alexey Shablygin

College of Information and Computer Science

University of Massachusetts Amherst

Amherst, MA

ivawter@umass.edu, ashablygin@umass.edu

*Abstract*—In this paper we will analyze blockchain data from two of the biggest mining pools, AntPool and F2Pool, to identify mining activity that is indicative of selfish mining. We look for indication of selfish mining by comparing the theoretical probability and observed probability of mining *n* blocks sequentially given their respective mining power. In addition, we conduct a Monte Carlo simulation to determine the likelihood that sequential block discovery by each miner occurred as expected given their mining power.

# Introduction

The objective of this paper is to observe if the frequency of sequential block discovery by individual bitcoin miners is indicative of honest or selfish mining. Selfish mining occurs when a miner uses an algorithm1 that allows them to take advantage of bitcoin’s proof of work system to mine multiple blocks in a row, by mining blocks in secret and only releasing them at the opportune time. This causes competing miners to waste their mining power by not mining from the most recent block, because that block is kept hidden by the selfish miner. Since the selfish mining algorithm1 requires miners to publish sequences of blocks all in a row, we aim to detect it by analyzing the frequency of block discovery sequences of different lengths by individual miners.

# Previous Work

This paper makes use of “***Majority Is Not Enough Bitcoin Mining Is Vulnerable***”1, particularly their definition of selfish mining, the algorithm to perform selfish mining, and findings on probability of selfish mining given computing power of a miner.

# Methods

In order to conduct our analysis do this we collected data about all of the blocks mined between December 1st, 2015 and April 20th, 2016, including information about which of those blocks were mined by AntPool and F2Pool. We collected our data using three main sources – blockchain.info, antpool.com and f2pool.com. We scrapped blockchain.info to get all of the mined blocks for our time period. Antpool.com and f2pool.com both provided us with data about the blocks each pool mined respectively. We assume that both pools are honest and reported every mined block on their respective websites. After we collected all of block data, we developed a Python program to focused on partitioning blocks into different sample sizes (2000, 1000, 500, 100 blocks), in order to account for fluctuations in mining power and the chance that miners may not be mining selfishly all the time. Initially we compared the observed probability **P** against the theoretical probability **P\*** of sequentially discovering **n** blocks in a row based the observed number of blocks mined **x,** and total number of sequences of length n within the sample out of the sample size **y** for each mining pool using the following formulas:

**P = (x/y)n**

**P\* = x / (y - (n-1))**

and performing a Monte Carlo simulation on each sample to determine expected number of sequential block discoveries of varying lengths based mining power of each pool.

# Results

# Conclusion

## Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in Eq. 1, using a right tab stop. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in

 

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “Eq. 1” or “Equation 1”, not “(1)”, especially at the beginning of a sentence: “Equation 1 is . . .”

# References

1. Eyal, I., & Sirer, E. G. (2014). Majority is not enough: Bitcoin mining is vulnerable. In Financial Cryptography and Data Security (pp. 436-454). Springer Berlin Heidelberg.