BITCOIN SELFISH MINING AND DYCK WORDS

CYRIL GRUNSPAN AND RICARDO PÉREZ-MARCO

Abstract. We give a straightforward proof for the formula giving the long-term apparent hashrate of the Selfish Mining strategy in Bitcoin using only elementary probabilities and combinatorics, and more precisely, Dyck words. There is no need to compute stationary probabilities on Markov chain nor stopping times for Poisson processes as it was previously done. We consider also several other block withholding strategies.

1. Introduction

Selfish mining (in short SM) is a particular non-stop strategy of block withholding strategy described in [1] which exploits a flaw in the Bitcoin protocol in the difficulty adjustment formula [4]. The strategy is made of attack cycles. During each attack cycle, the attacker adds blocks to a secret fork and then, broadcasts them to peers with an appropriate timing. This is a deviant strategy from the Bitcoin protocol since an honest miner never withholds blocks and always mines on top of the last block of the official blockchain [3].

As explained in [4] the good objective function based on sound economics principles in order to compare profitabilities of mining strategies with repetition is the revenue ratio $\frac{\mathbb{E}[R]}{\mathbb{E}[T]}$ where R is the revenue of the miner per attack cyle and T is the duration time per cycle. After a difficulty adjustment, this mean duration time becomes equal to $\mathbb{E}[L] \cdot \tau_B$ where L is the number of blocks added to the official blockchain by the network per attack cycle and $\tau_B = 600$ sec. [6]. Thus, the objective function becomes proportional to the long-term apparent hashrate of the strategy $\tilde{q} = \frac{\mathbb{E}[Z]}{\mathbb{E}[L]}$ where Z is the number of blocks added by the attacker to the official blockchain per attack cycle. Several methods have been conceived to compute $\tilde{q}$. In [1], first a stationary probability is computed for a Markov chain. In [4] we use martingale techniques and consider Poisson processes and associated stopping times. The revenue ratio is then computed at once using Doob's stopping time theorem. This last method has the advantage to fit the correct profitability analysis, and to identify the source of

Date: February 4, 2019.

2010 *Mathematics Subject Classification.* 68M01, 60G40, 91A60.

Key words and phrases. Bitcoin; selfish mining; Catalan distribution; Dyck words.

the weakness of the protocol. It allows a Bitcoin Improvement Proposal (BIP) to prevent the attack. It also yields the mean duration time before the attack becomes profitable. This last fact is out of reach with pure Markov chain models.

As usual, the relative hashrate of the honest miners (resp. attacker) is p (resp. q) and γ denotes its "connectivity". We have $p+q=1$, $q<\frac{1}{2}$ and $0\leq\gamma\leq 1$. We consider that whenever a competition occurs between two blocks or two forks, there is a fraction γ of the honest miners who mines on top of a block validated by the attacker.

2. Attack cycle and Dyck word

An attack cycle for the SM strategy can be described as a sequence $X_0\dots X_n$ with $X_i\in\{S,H\}$. The index i indicates the i-th block validated since the beginning of the cycle and letters S,H determine the miner who has discovered this block between the selfish miner (S) and the honest miners (H).

Example 2.1. *The sequence SSSHSHH means that the selfish miner has first validated three blocks in a row, then the honest miners have mined one, then the selfish miner has validated a new one and finally the honest miners have mined two blocks. At this point, the advantage of the selfish miner is only of one block. So according to the SM strategy, he decides to publish his whole fork and ends his attack cycle. In that case, we have $L=Z=4$.*

We are interested in the distribution of L.

Theorem 2.2. *We have $\mathbb{P}[L=1]=p,\mathbb{P}[L=2]=pq+pq^2$ and for $n\geqslant 3$, $\mathbb{P}[L=n]=pq^2(pq)^{n-2}C_{n-2}$ where $C_n=\frac{(2n)!}{n!(n+1)!}$ is the n-th Catalan number.*

Proof. For $n\geqslant 3$, we note that $\{L=n\}$ is a collection of sequences of the form $w=\mathrm{SSX}_1\cdots X_{2(n-2)}H$ with $X_i\in\{S,H\}$ for all i, such that if S and H are respectively replaced by the brackets "(" and ")" then, $X_1\cdots X_{2(n-2)}$ is a Dyck word (i.e., balanced parentheses) with length $2(n-2)$ (see [2]). The number of letters "S" (resp. "H") in w is n (resp. $n-1$). So, we get $\mathbb{P}[L=n]=p^{n-1}q^nC_{n-2}$ (see [2]). Finally, we note that $\{L=1\}=\{H\},\{L=2\}=\{\mathrm{SSH},\mathrm{SHS},\mathrm{SHH}\}$. Hence we get the result. □

Corollary 2.3. *We have $\mathbb{E}[L]=1+\frac{p^2q}{p-q}$*

Proof. It comes from the well know relations

$$\sum_{n\geqslant 0} p(pq)^n C_n = 1 \tag{1}$$

$$\sum_{n\geqslant 0} np(pq)^n C_n = \frac{q}{p-q} \tag{2}$$

that have been already used and proved in [5]. □

We can now compute the apparent hashrate.

Theorem 2.4. *The long-term apparent hashrate of the selfish miner in Bitcoin is*

$$\tilde{q}_B = \frac{[(p-q)(1+pq)+pq]q-(p-q)p^2q(1-\gamma)}{pq^2+p-q}$$

Proof. If $L \geqslant 3$, then we are in the cases where all blocks validated by the selfish miner will end in the official blockchain. So, $Z = L$. If $L = 1$, then $Z = 0$. Moreover, $Z(\mathrm{SSH}) = Z(\mathrm{SHS}) = 2$ and $Z(\mathrm{SHH}) = 0$ (resp. 1) with probability $1-\gamma$ (resp. γ). So,

$$\begin{aligned}\mathbb{E}[Z] &= \mathbb{E}[L] - p - p^2q\gamma - 2p^2q(1-\gamma)\\ &= \mathbb{E}[L] - (p + p^2q + p^2q(1-\gamma))\end{aligned}$$

Using Corollary 2.3 we get:

$$\begin{aligned}\frac{\mathbb{E}[Z]}{\mathbb{E}[L]} &= \frac{p^2q+p-q-(p-q)(p+p^2q+p^2q(1-\gamma))}{pq^2+p-q}\\ &= \frac{[(p-q)(1+pq)+pq]q-(p-q)p^2q(1-\gamma)}{pq^2+p-q}\end{aligned}$$

This is nothing but Proposition 4.9 from [4] which is itself another form of Formula (8) from [1]. □

3. Stubborn Mining

We consider now two other block witholding strategies described in [7]. In the sequel, $C(x) = \frac{1-\sqrt{1-4x}}{2x}$ denotes the generating series for the Catalan numbers $(C_n)_{n\geqslant 0}$.

3.1. **Equal Fork Stubborn Mining.** In this strategy, the attacker never tries to override the official blockchain but when it is possible, he broadcasts the part of his secret fork sharing the same height as the official blockchain as soon as the honest miners publish a new block. The attack cycle ends when the attacker has been caught up and overtaken by the honest miners by one block [5, 7]. We show that the

distribution of $L-1$ is what we have called a (p,q)-Catalan distribution of first type in [5].

Theorem 3.1. *We have* $\forall n \in \mathbb{N}, \mathbb{P}[L=n+1] = p(pq)^n C_n$.

Proof. Indeed, for $n \in \mathbb{N}$, $\{L = n+1\}$ is a collection of sequences of the form $w = X_1 \cdots X_{2n} H$ with $X_i \in \{S, H\}$ for all i, such that if S and H are respectively replaced by the brackets "(" and ")" then, $X_1 \cdots X_{2n}$ is a Dyck word with length $2n$. □

Corollary 3.2. *We have* $\mathbb{E}[L] = \frac{p}{p-q}$

Proof. Obvious by (1) and (2). □

Theorem 3.3. *The long-term apparent hashrate of a miner following the Equal-Fork Stubborn Mining strategy is given by* $\tilde{q} = \frac{q}{p} - \frac{(1-\gamma)(p-q)}{\gamma p}(1 - pC((1-\gamma)pq))$.

Proof. In an attack cycle, all the honest blocks except the last one have a probability γ to be replaced by the attacker. So, $\mathbb{E}[Z|L=n+1] = n+1-\frac{1-(1-\gamma)^{n+1}}{\gamma}$. See Lemma B.1 [5]. Conditionning by $\{L=n+1\}$ for $n \in \mathbb{N}$ and using Theorem 3.1, we then get

$$\mathbb{E}[Z] = \frac{q}{p-q} - \frac{1-\gamma}{\gamma}(1 - pC((1-\gamma)pq))$$

Hence we get the result. □

3.2. **Lead Stubborn Mining.** The strategy looks like the selfish mining strategy but here, the attacker takes the risk of being caught up by the honest miners. When this happens, there is a final competition between two forks sharing the same height. Once the competition is resolved, a new attack cycles starts. In this case, the distribution of $L-1$ is what we have called a (p,q)-Catalan distribution of second type [5].

Theorem 3.4. *We have* $\mathbb{P}[L=1] = p$ *and for* $n \geqslant 1$, $\mathbb{P}[L=n+1] = (pq)^n C_{n-1}$.

Proof. Indeed, we have $\{L=1\} = \{H\}$ and for $n \in \mathbb{N}$, $\{L=n+1\}$ is a collection of sequences of the form $w = \mathrm{S}X_1 \cdots X_{2(n-1)}\mathrm{HY}$ with $X_1, \ldots, X_{2(n-1)}, Y \in \{S, H\}$ and such that if S and H are respectively replaced by the brackets "(" and ")" then, $X_1 \cdots X_{2(n-1)}$ is a Dyck word with length $2(n-1)$. □

Corollary 3.5. *We have* $\mathbb{E}[L] = \frac{p-q+pq}{p-q}$

Proof. Obvious by (1) and (2). □

By repeating the same argument as in the proof of Theorem 3.3 for the computation of $\mathbb{E}[Z]$, we obtain the following theorem [5].

Theorem 3.6. *The long-term apparent hashrate of a miner following the Lead Stubborn Mining strategy is given by* $\tilde{q} = \frac{q(p+pq-q^2)}{p+pq-q} - \frac{pq(p-q)(1-\gamma)}{\gamma} \cdot \frac{1-p(1-\gamma)C((1-\gamma)pq)}{p+pq-q}$

We color the region $(q, \gamma) \in [0, 0.5] \times [0, 1]$ according to which strategy is more profitable, and we obtain Figure 1 [5] (HM is the honest mining strategy).

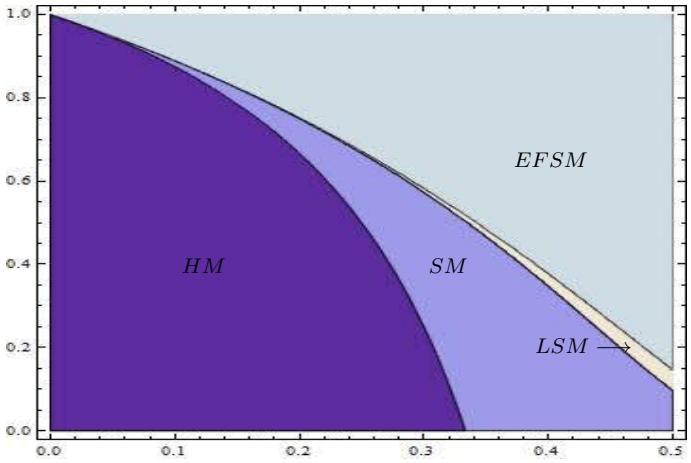

FIGURE 1. Dominance regions in parameter space (q, γ).

REFERENCES

[1] I. Eyal, E. Sirer. *Majority is not enough: bitcoin mining is vulnerable.* International Conference on Financial Cryptography and Data Security, pages 436–454, 2014.
[2] T. Koshy. *Catalan Numbers with Applications.* Oxford University Press, 2008.
[3] S. Nakamoto. *Bitcoin: a peer-to-peer electronic cash system.* Bitcoin.org/bitcoin.pdf, 2008.
[4] C. Grunspan, R. Pérez-Marco. *On profitability of selfish mining.* arXiv:1805.08281v2, 2018.
[5] C. Grunspan, R. Pérez-Marco. *On profitability of stubborn mining.*arXiv:1808.01041, 2018.
[6] C. Grunspan, R. Pérez-Marco. *On profitability of trailing mining.* arXiv:1811.09322, 2018.
[7] K. Nayak, E. Shi, S. Kumar, A. Miller. *Stubborn mining: generalizing selfish mining and combining with an eclipse attack.* IEEE European Symp. Security and Privacy, pages 305–320, 2016.

CYRIL GRUNSPAN
LÉONARD DE VINCI PÔLE UNIV, FINANCE LAB
PARIS, FRANCE,

E-mail address: cyril.grunspan@devinci.fr

RICARDO PÉREZ-MARCO
CNRS, IMJ-PRG, UNIV. PARIS-DIDEROT
PARIS, FRANCE

E-mail address: ricardo.perez.marco@gmail.com