# Centralization in Proof-of-Stake Blockchains: A Game-Theoretic Analysis of Bootstrapping Protocols

**Varul Srivastava**, Sankarshan Damle and Sujit Gujar

Machine Learning Lab, International Institute of Information Technology, Hyderabad

May 6, 2024

$6^{th}$ Games, Agents, and Incentives Workshop **(GAIW@AAMAS 2024)**

# Overview



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

# Introduction

- **PoS Blockchain** Public blockchain with $n$ players each holding some stake in the PoS system. Selection as *proposer* or *validator* for any player is proportional to their stake.
  - Incentive for high stake players to follow the protocol
  - Stake proportional to their valuation of the PoS System

# Introduction: Centralization in PoS Blockchains

- PoS blockchains are prone to centralization during two phases:
  - During bootstrapping — allocaiton of stake disproportional to valuation
  - During protocol execution — due to entities like staking pools

| **Blockchain** | Top 5% | Top 10% |
|---|---|---|
| ICON Network | 44.2% | 59.8% |
| Tezos | 24.2% | 40.8% |
| Cosmos | 30.0% | 46.3% |
| Irisnet | 20.9% | 33.4% |
| Kava | 26.4% | 46.8% |

Figure: Centralization in PoS blockchains [1]

# Our Model

- $n$ players — $P = \{p_1, p_2, \dots p_n\}$
- $p_i$ has valuation $\theta_i$ for the PoS system with $\sum_i \theta_i = 1$
- existence of *sybil-identities* is captured through partitions of the set $P$
- $p_i$'s disclosed (elicited) value is $\hat{\theta}_i$ with $\sum_i \hat{\theta}_i = 1$
- Utility is given by the equation

$$U_i((\hat{\theta}, A_i), (\hat{\boldsymbol{\theta}}_{-\boldsymbol{i}}, \boldsymbol{A}_{-\boldsymbol{i}}); \theta_i) = b \cdot \hat{\theta}_i - \underbrace{\Omega(\cdot)}_{\text{Centralization Metric}} \cdot \underbrace{g(\theta_i)}_{\text{Cost of centralization}} \tag{1}$$

# Ideal Bootstrapping Protocol — Properties

"Ideal" Bootstrapping protocol should satisfy – **IR, IC and Decentralized**

- *Individual Rationality (IR).* The protocol is IR if $\forall p_i, \forall \boldsymbol{\theta} \in {\Delta_n}^1, \hat{\boldsymbol{\theta}} \in \Delta_{n-1}, \forall \boldsymbol{A}_{-i}$

$$U((\theta, \{i\}), (\theta_{-i}, \boldsymbol{A'_{-i}})) \geq U((\hat{\theta}_i, A_i), (\boldsymbol{\theta_{-i}}, \boldsymbol{A_{-i}}))$$

- *Incentive Compatibility (IC).* The protocol is IC[2] if $\forall p_i, \forall \boldsymbol{\theta} \in \Delta_n, \hat{\boldsymbol{\theta}}_{n-1} \in \Delta_{n-1}, \forall \boldsymbol{A}_{-i}$

$$\mathbb{E}[U((\theta, \{i\}), (\theta_{-i}, \boldsymbol{A'_{-i}}))] \geq \mathbb{E}[U((0, \emptyset), (\boldsymbol{\theta_{-i}}, \boldsymbol{A_{-i}}))]$$

[1] $\Delta_n$ is n-simplex

[2] Bayesian Incentive Compatibility

# Ideal Bootstrapping Protocol — Properties

- *$(\tau, \delta, \epsilon)$-Decentralization.* [2] A protocol is $(\tau, \delta, \epsilon)$ decentralized for $\tau \in [0, 1], \delta \in [0, 100]$ and $\epsilon \in \mathbb{R}_{\geq}0$ if it follows:
  - *Minimum Participation* — $\frac{|P_t|}{|P|} \geq \tau$
  - *Proportionality* — $\frac{\beta_{max}}{\beta_\delta} \leq 1 + \epsilon$, where $\beta_{max}$ is the maximum scaled stake and $\beta_\delta$ is $\delta^{th}$ percentile scaled stake.
  - *Sybil-proofness* — The ratio $\frac{\beta_{max}}{\beta_\delta}$ cannot be reduced by increasing number of identities for a player.

# CNorm: Quantifying Centralization in presence of strategic players

- Should capture history of transactions across multiple identities and should be *Sybil-resistance*
- CNorm uses Directed Acyclic Graph (DAG) representation of PoS system to capture current and historical states.
- Effective stake for $p_i$ is $\omega_i = c_i + \underbrace{w_{ij} - w_{ji}}_{\text{Net currency influx}}$

$$\Omega = \max_{\boldsymbol{\theta} | \theta_i > 0 \forall i} \sum_{j=1}^{n} |\beta_j(\theta) - \frac{1}{n}|$$

where, $\beta_i(\boldsymbol{\theta}) = \frac{\omega_i / \theta_i}{\sum_{j \in [n]} \omega_j / \theta_j}$

# Centralization Game

- We measure properties of CNorm through Centralization Game $\Gamma_{cent}\langle\{M_C, M_D\}, S_{SA}, e_r, Q(\cdot), \kappa\rangle$
- *Resitance to Sybil Attacks* For CNorm, $\Gamma_{cent}$ returns correct bit with probability $1 - negl(\kappa)$
- *Decentralization* For any IC protocol, low value of CNorm $\Rightarrow$ system is $(\tau, \delta, \epsilon)-$Decentralized.

# CNorm Description

$\Gamma_{\textbf{cent}} = \langle \{M_C, M_D\}, S_{SA}, e_r, Q(\cdot), \kappa \rangle$

**Metric Descriptor** ($M_C$):

❶ Samples $s_0 \in S_{SA}$ and sets $s_1 := e_r(s_0)$ such that $(s_0, s_1) \in S_{SA} \times S_{NSA}$.

❷ Chooses $(a, b) \in_R \{(0,1), (1,0)\}$ and communicates $(s_a, s_b)$ to $M_D$.

**Metric Challenger** ($M_D$):

① Evaluates $v_a := Q(s_a)$ and $v_b := Q(s_b)$.

② Distinguishes between $v_a, v_b$ to get $(a', b')$ using any program $\mathcal{D}$ by $(a', b') \leftarrow \mathcal{D}(v_a, v_b)$.

**Success Probability:** Consider a random variable $D(Q)$, that takes value 1 if $M_D$ successfully guesses $(a, b)$ i.e. $a' = a,\ b' = b$ and 0 otherwise. For $\kappa \in \mathbb{Z}_{\geq 1}$ trials of $\Gamma_{\text{cent}}$, we have $D_\kappa(Q) = 1$ if all $\kappa$ trials are successful. $M_D$'s success depends on the centralization metric it employs. We have,

Ineffective Metric: If $Q(\cdot)$ is an ineffective metric for quantifying centralization, success probability will be as good as a random guess across $\kappa$ trials. Formally, for $\kappa$ trials

$$\Pr(D_\kappa(Q) = 1) \leq \mathsf{negl}(\kappa)$$

Where $\mathsf{negl}(\kappa)$ is a negligible function in $\kappa$.

Effective Metric: If $Q(\cdot)$ is an effective metric, then $M_D$ can distinguish between $(s_0, s_1)$ and $(s_1, s_0)$ with very high probability. Formally, $\forall \mathsf{negl}(\kappa)$

$$Pr(D_\kappa(Q) = 1) > 1 - \mathsf{negl}(\kappa)$$

# CNorm Description

$\Gamma_{\text{cent}} = \langle \{M_C, M_D\}, S_{SA}, e_r, Q(\cdot), \kappa \rangle$

**Metric Descriptor** ($M_C$):

1. Samples $s_0 \in S_{SA}$ and sets $s_1 := e_r(s_0)$ such that $(s_0, s_1) \in S_{SA} \times S_{NSA}$.
2. Chooses $(a, b) \in_R \{(0,1), (1,0)\}$ and communicates $(s_a, s_b)$ to $M_D$.

**Metric Challenger** ($M_D$):

① Evaluates $v_a := Q(s_a)$ and $v_b := Q(s_b)$.

② Distinguishes between $v_a, v_b$ to get $(a', b')$ using any program $\mathcal{D}$ by $(a', b') \leftarrow \mathcal{D}(v_a, v_b)$.

**Success Probability:** Consider a random variable $D(Q)$, that takes value 1 if $M_D$ successfully guesses $(a, b)$ i.e. $a' = a,\ b' = b$ and 0 otherwise. For $\kappa \in \mathbb{Z}_{\geq 1}$ trials of $\Gamma_{\text{cent}}$, we have $D_\kappa(Q) = 1$ if all $\kappa$ trials are successful. $M_D$'s success depends on the centralization metric it employs. We have,

Ineffective Metric: If $Q(\cdot)$ is an ineffective metric for quantifying centralization, success probability will be as good as a random guess across $\kappa$ trials. Formally, for $\kappa$ trials

$$\Pr(D_\kappa(Q) = 1) \leq \mathrm{negl}(\kappa)$$

Where $\mathrm{negl}(\kappa)$ is a negligible function in $\kappa$.

Effective Metric: If $Q(\cdot)$ is an effective metric, then $M_D$ can distinguish between $(s_0, s_1)$ and $(s_1, s_0)$ with very high probability. Formally, $\forall \mathrm{negl}(\kappa)$

$$Pr(D_\kappa(Q) = 1) > 1 - \mathrm{negl}(\kappa)$$

| Centralization Metric | PoS Systems $s_0$ | PoS Systems $s_1$ |
|---|---|---|
| Nakamoto Coefficient ($N$) [40] | 3 | 3 |
| Entropy ($H$) [63] | 0.1405 | 0.1405 |
| Gini Coefficient ($G$) [21] | 0.0804 | 0.0804 |
| C-NORM ($\Omega^\star$) | 0.6 | 0 |

Figure: Distinguishability between Sybil $s_0$ and Non-Sybil $s_1$ system

# Results on Bootstrapping Protocols

- Airdrop is not IC due to which it has very high potential value of CNorm.
- Proof-of-Burm is not IR (due to a phenomenon called *pegging*).
- W2SB is (1) IC, (2) IR and (3) $(\tau, \delta, \epsilon)-$decentralized.

# References

Centralization of stake in pos.
`medium.com/stakin/centralization-of-stake-in-pos-f7ccb8f8254`, 2020.

KWON, Y., LIU, J., KIM, M., SONG, D., AND KIM, Y.
Impossibility of full decentralization in permissionless blockchains.
In *ACM Conference on Advances in Financial Technologies* (2019), p. 110–123.

# Thank You

# Questions?

Machine Learning Lab, IIIT

Paper Link (ArXiV)

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD