

On Sybil-Proof Mechanisms

Minghao Pan

California Institute of Technology, Pasadena, CA, USA
Flashbots, Pasadena, CA, USA

Bruno Mazonra

Flashbots, Barcelona, Spain

Christoph Schlegel

Flashbots, Zurich, Switzerland

Akaki Mamagishvili

Offchain Labs, Zurich, Switzerland

Abstract

We show that in the single-parameter mechanism design environment, the only non-wasteful, symmetric, incentive compatible and Sybil-proof direct mechanism is a second price auction with symmetric tie-breaking. Thus, if there is private information, lotteries or other mechanisms that do not always allocate to a highest-value bidder are not Sybil-proof or not incentive compatible. Moreover, we show that our main (im)possibility result extends beyond linear valuations, but not to multi-unit object allocation with unit-demand bidders. We also provide examples of mechanisms (with higher interim payoff for the bidders than a second price auction) that satisfy all of the other axioms and a weaker, Bayesian notion of Sybil-proofness. Thus, our (im)possibility result does not generalise to the Bayesian setting and we have a larger design space: With Sybil constraints, equivalence between dominant strategy and Bayesian implementation (that holds in classical single-parameter mechanism design without Sybils) no longer holds.

2012 ACM Subject Classification Theory of computation → Algorithmic game theory and mechanism design

Keywords and phrases Sybils, Mechanism Design, Blockchain

Digital Object Identifier 10.4230/LIPIcs.AFT.2026.16

Category Extended Abstract

Related Version *Full Version*: <https://arxiv.org/abs/2407.14485>

1 Introduction

Where creating new identities is cheap and verifying them is hard, mechanism design must consider that a participant may enter a mechanism under several identities. Such a manipulation is called a Sybil attack, and a mechanism immune to it *Sybil-proof*. Blockchain systems, where participation is permissionless, are both rich in deployed mechanisms and susceptible to Sybil attacks.

When allocating a private good, a natural Sybil-proof mechanism is an auction that assigns the good to a highest-value bidder. Auctions may, however, be undesirable if fairness, wider participation, or other distributional goals matter. Event tickets, for example, are often sold below the market-clearing price, making the primary sale a target for middlemen who, if quantities per buyer are capped, buy through Sybils [1]. In blockchain systems, randomness in allocating block proposal rights is desirable to avoid power concentration, and proportional selection has been characterised there by Sybil-proofness, symmetry, non-wastefulness and collusion-proofness [2, 4].

It is therefore natural to ask whether there are Sybil-proof mechanisms for private goods other than auctions. Theorem 1 gives a strong negative answer: in the classical Myersonian setting, Sybil-proofness collapses a large design space of monotonic rules to a single one.



© Minghao Pan, Bruno Mazonra, Christoph Schlegel, and Akaki Mamagishvili;
licensed under Creative Commons License CC-BY 4.0

8th Conference on Advances in Financial Technologies (AFT 2026).

Editors: Aggelos Kiayias and Maria Kyropoulou; Article No. 16; pp. 16:1–16:3

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

This can be read as a trilemma: a non-wasteful mechanism in the presence of Sybils is either not incentive compatible, or not Sybil-proof, or centralising. Most related to our work, [6] prove there is no Pareto-efficient, false-name-proof combinatorial auction. We do not require Pareto efficiency, as we ask whether non-trivial distributional goals can be achieved at all; but, as a consequence of our main theorem, our axioms imply Pareto efficiency.

One way out is to require only that no bidder gains from introducing Sybils in equilibrium. Under this Bayesian notion there are mechanisms which satisfy all the remaining axioms yet do not always allocate to a highest-value bidder (Theorem 3). Thus, in contrast to the classical setting [5], Bayesian implementation is strictly weaker than dominant strategy implementation under Sybil constraints.

2 Model and main result

A **(direct) mechanism** specifies, for each finite set $\mathcal{N} \subset \mathbb{N}$ of **bidders** (some real agents, some possibly Sybils), an **allocation rule** $x^\mathcal{N} : \mathbb{R}_+^\mathcal{N} \rightarrow \Delta(\mathcal{N})$, where $\Delta(\mathcal{N}) := \{x \in \mathbb{R}_+^\mathcal{N} : \sum_{i \in \mathcal{N}} x_i \leq 1\}$, and a **payment rule** $p^\mathcal{N} : \mathbb{R}_+^\mathcal{N} \rightarrow \mathbb{R}^\mathcal{N}$. The share $x_i^\mathcal{N}(v)$ for reported **values** $v \in \mathbb{R}_+^\mathcal{N}$ is a probability of obtaining an indivisible good or a share of a divisible good, and bidder i has utility $U_i = v_i x_i - p_i$.

We impose **payment normalisation**, $p_i^\mathcal{N}(0, v_{-i}) = 0$; **non-wastefulness**, $\sum_{j \in \mathcal{N}} x_j^\mathcal{N}(v) = 1$; **symmetry**, $x_j^\mathcal{N}(v) = x_{\pi(j)}^\mathcal{N}(\{v_{\pi(i)}\}_{i \in \mathcal{N}})$ for every permutation π of $\mathcal{N}$; and **incentive compatibility** in dominant strategies, which, given payment normalisation, is equivalent to monotonicity of $x^\mathcal{N}$ with the Myerson payments $p_j(v) = v_j x_j(v) - \int_0^{v_j} x_j(z, v_{-j}) dz$. Our **Sybil-proofness** axiom is weaker than the false-name-proofness of [6], which also allows the deviator to misreport: we ask only for immunity to attacks in which a bidder reports truthfully from his original account and creates a single Sybil with an arbitrary bid; the characterisation therefore holds a fortiori under false-name-proofness. For all $v \in \mathbb{R}_+^\mathbb{N}$, finite $\mathcal{N} \subset \mathbb{N}$, $i \in \mathcal{N}$, $j \in \mathbb{N} \setminus \mathcal{N}$ and $u \geq 0$,

$$v_i x_i^\mathcal{N}(v) - p_i^\mathcal{N}(v) \geq v_i \left(x_i^{\mathcal{N} \cup \{j\}}(v, u) + x_j^{\mathcal{N} \cup \{j\}}(v, u) \right) - p_i^{\mathcal{N} \cup \{j\}}(v, u) - p_j^{\mathcal{N} \cup \{j\}}(v, u).$$

► **Theorem 1.** *A direct mechanism is payment normalised, non-wasteful, symmetric, incentive compatible, and Sybil-proof if and only if it is a second price auction with symmetric tie-breaking.*

The proof turns Sybil-proofness into a lower bound on payoffs: under Myerson payments, $U_j(v) = \int_0^{v_j} x_j(z, v_{-j}) dz$, and a bidder who reports u against $n - 1$ bidders reporting $v < u$ can add a Sybil bidding v , which by symmetry and non-wastefulness wins with probability $\frac{1}{n}[1 - x_1]$ at a price of at most v . Ruling out these deviations for every n gives $\limsup_n x_1(u, v, \dots, v) = 1$ and hence $U_1(u, v) \geq u - v$ for two bidders, while non-wastefulness pins $\frac{1}{u} \int_0^u U_1(u, v) dv$ to $\frac{u}{2}$, so that $U_1(u, v) = u - v$ and $x_1(z, u) = 0$ for every $z < u$. Induction extends this beyond two bidders; symmetry then forces uniform tie-breaking and incentive compatibility the second-price payments.

3 Bayesian Sybil-proofness

Bidders now hold a prior over bid profiles, so that both valuations and the number of real bidders can be unknown. Conditional on the set of real bidders, private values are i.i.d. F , twice differentiable and strictly increasing on its support $(0, \bar{v})$. **Bayesian Sybil-proofness**, similar to the notion in [3], requires that Sybils do not benefit a bidder in expectation if all others report truthfully and use a single identity.

► **Example 2.** Let Z_{n-1} be the maximum of an i.i.d. sample of size $n - 1$ from F , let $y^* = \min\{y : (\bar{v} - x)f(x) \leq 1 - F(x) \ \forall x \in [y, \bar{v}]\}$, and for each integer $n \geq 2$ define the *ticket price* $a_n = \max\{y^*, E[Z_{n-1}], F^{-1}(\frac{n-2}{n-1})\}$. The **lottery mechanism with increasing ticket price** for n bids allocates as follows: if all or all but one bids are below a_n , the item is allocated by a second price auction with uniform tie-breaking; if at least $m \geq 2$ bids are at least a_n , then each of those m bidders wins with probability $1/m$ and the winner pays a_n .

The mechanism uses a monotonic allocation rule with Myerson payments and is hence dominant strategy incentive compatible; it is symmetric and non-wasteful by construction. As long as $y^* < \bar{v}$, that is, as long as the tail of F close to $\bar{v}$ is no heavier than that of the uniform distribution on $[0, \bar{v}]$, it differs from a second price auction with positive probability for every n , and allocating randomly at a fixed price among qualifying bidders yields a weakly higher ex-interim utility. For F uniform on the unit interval, a lottery among at least two bidders occurs with probability $1 - a_n^n - n(1 - a_n)a_n^{n-1}$, which tends to $1 - 2/e \approx 0.26$.

► **Theorem 3.** *The lottery mechanism with increasing ticket prices is Bayesian Sybil-proof. In particular, if $y^* < \bar{v}$, then there exists a mechanism that is payment normalised, non-wasteful, symmetric, incentive compatible, and Bayesian Sybil-proof that is non-equivalent to the second price auction.*

4 Further results and open questions

Indirect mechanisms cannot circumvent the impossibility: the full version establishes a revelation principle for Sybils, in dominant strategies and in Bayes-Nash equilibrium. The characterisation extends beyond linear valuations, to types with convex and superadditive interim utility, but not to multi-unit allocation with unit-demand bidders, where mechanisms other than the uniform-price mechanism satisfy all axioms. Open questions include how well equilibria of Sybil-proof mechanisms approximate distributional objectives, what the general class of Bayesian Sybil-proof mechanisms is and what the admissible class is when Sybils are costly.

References

- 1 Eric Budish and Aditya Bhawe. Primary-market auctions for event tickets: Eliminating the rents of “bob the broker”? *American Economic Journal: Microeconomics*, 15(1):142–170, 2023.
- 2 Xi Chen, Christos Papadimitriou, and Tim Roughgarden. An axiomatic approach to block rewards. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies*, pages 124–131, 2019. doi:10.1145/3318041.3355470.
- 3 Yotam Gafni, Ron Lavi, and Moshe Tennenholtz. Vcg under sybil (false-name) attacks—a bayesian analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 1966–1973, 2020. doi:10.1609/AAAI.V34I02.5567.
- 4 Jacob D Leshno and Philipp Strack. Bitcoin: An axiomatic approach and an impossibility theorem. *American Economic Review: Insights*, 2(3):269–286, 2020.
- 5 Alejandro M. Manelli and Daniel R. Vincent. Bayesian and dominant-strategy implementation in the independent private-values model. *Econometrica*, 78(6):1905–1938, 2010.
- 6 Makoto Yokoo, Yuko Sakurai, and Shigeo Matsubara. The effect of false-name bids in combinatorial auctions: New fraud in internet auctions. *Games and Economic Behavior*, 46(1):174–188, 2004. doi:10.1016/S0899-8256(03)00045-9.