On the Defintion, Function and Measurement of Decentralization in Cryptocurrencies

Christopher Aubin, Antonio Knez, Pranay Anchuri (Advisor)

caubin@princeton.edu, aknez@princeton.edu

Department of Computer Science, Princeton University, 20 April 2022

Abstract

Decentralization has been widely acknowledged as a foundational property of blockchain systems. However, there's no clear and effective means of measuring how decentralized these systems truly are. The Nakamoto Coefficient, a metric proposed by Balaji S. Srinivasan, is regarded as one of the better attempts to define a means of measuring the decentralization of a system, in particular blockchain-based systems. In this paper we build upon this metric, arguing that the choice of subsystems that ought to be considered when calculating a system's Nakamoto Coefficient should be motivated by a set of properties that we would like to achieve and preserve through decentralization. We propose what we believe are the most important properties of cryptocurrencies, and then suggest that the decentralization of subsystems is only important insofar as it contributes to those desirable properties of the system. We then calculate the Nakamoto Coefficients of Bitcoin, Ethereum, and Solana using our approach and compare the results to previous attempts to calculate those same coefficients, demonstrating how our approach solves some of the existing issues with the Nakamoto coefficient and leads coefficients that more accurately represent the decentralization of systems. We believe that ideas and methodologies discussed in this paper could help to facilitate future academic discourse regarding decentralization in blockchains.

1. Introduction

Bitcoin has become a cultural phenomenon. While it was by no means the first attempt to create a digital currency, it has experienced wider adoption and demonstrated more staying-power than any of its predecessors[1][2]. One possible explanation for this is the fact that, unlike those predecessors, Bitcoin is entirely removed from any and all existing systems of governance (banks/ financial institutions, regulatory/ compliance agencies, national/ state governments, etc)[2]. This independence from existing systems is often argued to be Bitcoin's *raison d'être*. This is certainly the belief of the authors, and this belief will inform our discussion of the definition, function and measurement of decentralization in cryptocurrencies.

We will tackle these topics by first defining decentralization. We will then explain why we believe that it is important that cryptocurrencies are decentralized. We will proceed to address some prominent attempts to measure decentralization in blockchains, and indicate where we believe that they could be improved. Finally we will hone in on the work of Balaji S. Srinivasan, who introduced the *Nakamoto Coefficient* and the *modified Nakamoto Coefficient* [4], with the goal of making three key contributions to the academic literature surrounding bitcoin and cryptocurrencies.:

I. **We will propose a framework for selecting the essential subsystems, the decentralization of which ultimately determine the *Nakamoto Coefficient* and the *modified Nakamoto Coefficient* of a system**.

II. **We will use that framework to re-measure the Nakamoto Coefficients of Bitcoin and Ethereum using the same data as was used by Srinivasan and compare our results to Srinivasan's.**

III. **We will use that framework to measure the Nakamoto Coefficient of Solana**. Recently the Nakamoto Coefficient was used as the basis for arguing that Solana is in fact more decentralized than Bitcoin and Ethereum. Intuitively, this seems unlikely, so an indicator of the success of our framework could be it producing a result more in line with our intuition that Solana is more centralized than Bitcoin and Ethereum.

2. Related Work

2.1. Defining decentralization

Despite its importance, as a term decentralization has devolved into something of a buzzword. It is often used in a vague and misleading manner. Projects that are far from decentralized get away with using the adjective just because they are blockchain-adjacent (see the abundance of layer two solutions that promise both increased speed and no loss of decentralization). This problem is only exacerbated because decentralization is not a term that is endemic to cryptocurrencies, blockchains or even the field of Computer Science. It is a term used in various academic fields and industries, and because it has different meanings in different contexts, it is a term that many people mistakenly believe that they understand in the context of computer systems, and, by extension, blockchain/ cryptocurrencies. However, its meaning in this context has proved rather challenging to define with any semblance of precision. Perhaps a logical first step in doing so is to look to the dictionary, which defines decentralization as "the dispersion or distribution of functions and powers"[3]. *While this provides a foundation upon which to understand decentralization, moving from this abstract definition to something measurable in computer systems is a non-trivial task.*

The Meaning of Decentralization, Vitalik Buterin

Ethereum co-founder Vitalik Buterin wrote a widely read, well-regarded blog post on the meaning of decentralization in the context of computer systems, in particular in cryptocurrencies[5]. In it he proposed that there are 3 dimensions of decentralization:

- **Architectural decentralization:** A measure of the dispersion or distribution of the underlying physical machines in the system. Architectural decentralization takes into account the number of nodes/ links in the system, the geographical distribution of those nodes, and what proportion of those nodes/ links could fail without compromising the system.
- **Political decentralization:** A measure of the dispersion or distribution of the governance/ decision-making power of a system. Political decentralization primarily concerns the number of individuals, organizations or entities that govern the system.
- **Logical decentralization**: A measure of the dispersion or distribution of the behavior of a system. Logical decentralization is the trickiest/ least intuitive of Vitalik's proposed dimensions of decentralization. It concerns the degree to which the constituent machines in a system behave like a single machine, or, conversely, as individual machines.

Vitalk proceeds to explain that blockchains are *politically* and *architecturally decentralized* but *logically centralized*. They are *politically decentralized* because decision-making power is distributed across nodes and not held by a particular entity(ies), organization(s) or a subset of participants. They are *architecturally decentralized* because blockchains have no set of nodes upon which the system relies in particular or that have elevated privileges, and thus there are no particular infrastructural points of failure. They are *logically centralized* because they ultimately behave like a single database

For the most part this is the definition of decentralization that we will be using in this paper. However, we would be remiss if we didn't point out a minor issue that we have with part of

Vitalk's post. While Vitalik may have implied as much, we feel the need to clarify that not all blockchains are politically and architecturally decentralized (hence the importance of measuring decentralization).

2.2. Why is decentralization important?

The pursuit of decentralization is often deemed virtuous in its own right. Indeed, academic papers and blockchain enthusiasts are often guilty of assuming that decentralization for decentralization's sake is important, and the discussion of why decentralization is important and the advantages of decentralization as opposed to centralization is often neglected. However, some papers make the argument, with which the authors agree, that, as opposed to being a virtue in its own right, decentralization is only important insofar as it provides advantages over its centralized alternative.

The Meaning of Decentralization, Vitalik Buterin

While the primary focus of Vitalik's aforementioned blog post *The Meaning of Decentralization* is defining decentralization, he also makes some insightful contributions to the discussion of why decentralization is important, largely from a technical perspective, in blockchain systems. He points to three particularly useful consequences of decentralization:

- **Fault tolerance**: Decentralized systems are less likely to fail than centralized systems because they rely on a collection of separate components that are unlikely to fail simultaneously.

- **Attack resistance**: Decentralized systems are more difficult and expensive to attack (either with the goal of destroying or manipulating the system) because attacks need to consider multiple targets, as opposed to a single central target.
- **Collusion resistance**: It is much harder for participants/ bad actors in decentralized systems to collude to act in ways that benefit them at the expense of other participants.

We (the authors) do not have much to add with regards to these properties. Although we do not believe that they are all of the advantages of decentralized systems, as will become evident later in this paper, we agree that they are all important and useful consequences of decentralization.

Fault tolerance and *attack resistance* are two well known and somewhat accepted advantages of decentralized systems (not just blockchains), and are often cited to be consequences of the *architectural decentralization* of a system. Vitalik makes the interesting observation that, in order to avoid common mode failure, or the failure of two seemingly separate/ decentralized parts of a system as a result of common dependencies, in the case of blockchains, *political decentralization* is likely also necessary to more rigorously achieve *fault tolerance* and *attack resistance*.

Collusion resistance is a more novel concept, and particularly relevant to blockchains and cryptocurrencies. Vitalik defines collusion as undesired coordination, and notes *Collusion resistance* is a property that is difficult to design for, and can only really be achieved with extreme *political decentralization* and extreme discoordination between users. Vitalik mentions

multiple potential mechanisms to achieve some degree of *collusion resistance*, but acknowledges flaws in all of his suggestions.

The Case for a Small Allocation to Bitcoin, Wences Casares

Wences Casares is an advocate for Bitcoin who's belief in the cryptocurrency is informed by his own experiences, having been repeatedly let down by government mismanagement of the supply of money. In particular, as an Argentinian, over the course of his life he has been subject to the devaluation of the Argentinian Peso in the 1980s, hyperinflation between the 1970s and 1980s (the Argentinian Peso experienced an average annual inflation rate of 300% between 1975 and 1990) [7] and the Argentinian Great Depression between 1998 and 2002, in particular the nation-wide freezing of bank accounts known as the Corralito[8]. Accordingly, in his open letter *The Case for a Small Allocation to Bitcoin* the three particularly useful consequences of decentralization that he notes are largely related to *political decentralization* and independence from the governments, financial institutions and regulatory entities that hold a monopoly over the decision-making power in typical monetary systems[6]. They are:

- **Sovereignty**: Wences makes the argument that decentralization is the means by which Bitcoin achieves sovereignty, which he defines, in line with its dictionary definition, as the fact that Bitcoin is not beholden to any existing systems of governance, and the fact that it is self-goverenaning, obeying only its own rules. He makes the observation that no single individual, government or entity, or, for that matter, no collection thereof, can impose rules on a sovereign system. He makes the observation that even if all of the world's political, economic, religious and ideological leaders came together they could not change the Bitcoin protocol without approval of the majority of Bitcoin users.

- **Inflation resistance/ scarcity**: The total supply of Bitcoin is fixed and unchangeable. Wences argues that there has never been another system or resource for which this has been the case, and notes that typically any assurances with regards to the supply of money or resources have been provided by, and have thus involved trusting, a third party.
- **Censorship resistance:** Wences points out that no-one can change a transaction on the Bitcoin Blockchain, that no one can keep the Bitcoin Blockchain from accepting new transactions, and that there is no way to prevent anyone from transacting on the blockchain. He argues that Bitcoin is the first monetary system that users can be certain will always allow them to transact without trusting a third party.

Wences's focus on the shortcomings of traditional monetary systems as opposed to the technical advantages of decentralization (despite his strong technical background) is unique. His arguments for *censorship resistance* and *sovereignty* are particularly compelling. While we recognise the value of *inflation resistance* and *scarcity,* in particular in the case of Bitcoin, which, as Wences points out, could become a sort of global settlement solution, we are not convinced that this is a feature that all cryptocurrencies absolutely need to have. While scarcity and inflation resistance has its advantages, inflation does serve a purpose, namely, it is used to stimulate spending, investment and borrowing. As alternative cryptocurrencies emerge to serve different purposes and as cryptocurrencies experience wider adoption and are used to transact more regularly, we could well need some inflationary cryptocurrencies.

Denationalization of Money, Friedrich Hayek

The mention of Friedrich Hayek, a 20th century Austrian economist and Nobel laureate who didn't live to see the dot com boom, in a paper discussing an internet technology that is still very much in its infancy, may come as something of a surprise. However, the theoretical foundations of bitcoin have been frequently traced back to the Austrian School of Economics. Hayek in particular was a fervent advocate for the decentralization of money, which was the topic of his 1976 book *The Denationalization of Money* [9]. Bitcoin and cryptocurrency's independence from financial and governmental infrastructures aligns with aspects of Hayek's vision for competitively issued private moneys, and thus it is not surprising that *The Denationalization of Money* provides some valuable insights into why exactly Bitcoin and cryptocurrency's aforementioned independence may prove advantageous.

This independence not only distinguishes Bitcoin and cryptocurrencies from previous digital currencies, but also from most instances of money. It has seemingly been a common belief among economists since as early as the 18th century that one of the core functions of government is to create and manage a monetary mechanism and to issue money, and by extension that the government has to control monetary policy and that each country has to have its own currency. Hayek makes an argument to the contrary that aligns with Casares's sentiments: **Government has failed and will continue to fail to supply a dependable means of payment and has, in practice, been responsible for destabilizing currencies and for inflation.** Hayek proceeds to argue that money, and in particular **the supply of money, not unlike other commodities, is better governed by competition between private entities** and by self-interest than by a government that we assume to be benevolent. Freedom from traditional governance and the

competitive supply of moneys are two properties that cryptocurrencies achieve through decentralization, and are two of the reasons that we, the authors, believe that decentralization in cryptocurrencies is important.

2.3. Measuring Decentralization

In our reading we encountered two particularly interesting attempts to measure decentralization, namely Balaji S. Srinivasan's Quantifying Decentralization and Paul Sztorc's Measuring Decentralization. In Quantifying Decentralization Srinivasan takes a general approach to measuring the decentralization of a cryptocurrency, and because of this he proposes measuring the decentralization of every subsystem of a system, even those for which decentralization provides no benefits over centralisation. On the other end of the spectrum, in Measuring Decentralization Sztorc measures decentralization in a hyper-focused context, namely only insofar it guarantees his "knowing that he has been paid".

Quantifying Decentralization, Balaji S. Srinivasan

Quantifying Decentralization by Balaji S. Srinivasan has become something of a seminal work in the blockchain and cryptocurrency space. In it he introduces the *Nakamoto Coefficient* and the *modified Nakamoto Coefficient*, as the minimum number of actors an attacker would need to compromise a blockchain system, and therein the decentralization of that blockchain system. The metric is motivated by the Gini coefficient and Lorenz curve, and is expressed as a single number. To find that number, Srinivasan poses the following process:

1. Enumerate the essential subsystems of a decentralized system; in the case of blockchain these can include mining, nodes, token ownership, developers, exchanges etc.

2. Derive the number of actors that need to be corrupted in order to compromise each subsystem.
3. Finally, the minimum of these numbers serves as a measure of the effective decentralization of the system.

The difference between the *Nakamoto Coefficient* and the *modified Nakamoto Coefficient* is that, in the case of the *Nakamoto Coefficient,* the percentage of a subsystem that needs to be compromised in order for that subsystem to be considered compromised is >50%, likely because in proof of work chains control of >50% of the hashing power allows an entity to double spend and censor users of the network in what has become known as a *sybil* or *51% attack* (a misnomer, because this attack is possible with control of >50%, as opposed to 51%, of the network's hashing power), while, in the case of the *modified Nakamoto Coefficient,* that threshold must be determined on a subsystem-by-subsystem basis.

We are particularly intrigued by the *modified Nakamoto Coefficient,* which we believe as a metric and method for measuring the decentralization of a blockchain is incredibly promising. Sirnivasan knowingly neglects to delve into two important aspects of calculating the *modified Nakamoto Coefficient* of a blockchain, which he leaves as areas for future work:

I. The challenge of selecting the essential subsystems of a chain

II. The problem of determining the appropriate compromisation threshold for each of those subsystems.

Later in this paper we will focus on the former of the two unsolved issues.

Active members of the cryptocurrency/ blockchain community will no doubt have noticed the frequency with which the Nakamoto Coefficient is used to argue in favor of the decentralization of new blockchain/ cryptocurrency projects[11][12]. It has been abused so as to mischaracterized the decentralization, or lack thereof, of systems because it is widely misunderstood. If we are not careful we may witness it devolve to something of a buzzword (buzzmetric?), not unlike the term decentralization. To this end we would argue that the misleading use of Nakamoto Coefficient is not so much a consequence of flaws in Srinivasan's proposed method as a consequence of incorrect calculations of the Nakamoto Coefficient, especially in the case of smaller/ newer projects for which not much data is publicly available (we will demonstrate this when we discuss the Nakamato Coefficient off Solana and measure it ourselves). To truly measure the Nakamoto coefficient of a blockchain/ cryptocurrency requires an intimate understanding of the process of calculating The Nakamato Coefficient and a significant understanding of that chain/ cryptocurrency, and as such there is a significant amount of chain-specific work and research that needs to be done prior to even attempting to measure any chain's Nakamoto Coefficient. Because of how nuanced this is and because of the number of intricacies that need to be accounted for, quantitative measurements ought to be supplemented with equally extensive qualitative analysis and research.

There is also the issue of the availability of the appropriate data to accurately measure the centralization of each subsystem, or rather the lack thereof. For example, due to the pseudonymous nature of Bitcoin and Ethereum and the fact that anyone can have arbitrarily many seemingly unrelated addresses or run an arbitrary number of seemingly unrelated mining stations, how can we measure accurately measure ownership distribution and hashing power,

essential subsystems of Bitcoin and Ethereum respectively. This means that the Nakamoto Coefficient of a system can be difficult to measure. However, this does not mean that it is conceptually flawed.

Measuring Decentralization, Paul Sztorc

In his paper *Measuring Decentralization* [10], Paul Sztorc takes an entirely different approach to measuring decentralization to Srinivasan. In it, Sztorc defines the process of "money" as "knowing you've been paid". He proceeds to argue that when measuring the decentralization of a currency, the more local this process is, and the fewer third parties that need to be trusted in the process, the more decentralized the currency. Hence, decentralization in the context of cryptocurrencies is the cost of knowing you've been paid without relying on any third parties. According to Sztorc, and aligning with Satoshi's vision for Bitcoin, a truly decentralized currency needs to be "pure P2P', meaning that the network must not have any nodes with any form of superior authority. Each node should be able to verify yall transactions, past and present, by itself, without having to trust the judgment of another node. Thus, each node must possess all the data and all the rules of the network; each node has to be a *full node*. Sztorc calls this the *cost of the option to create a full node*, or CONOP.

In contrast with Srinivasan's approach to measuring decentralization, which measured the decentralization of an entire system in a far more general sense, Sztorc's approach to measuring decentralization is hyper focused, and only considers decentralization insofar as it contributes to knowing that one has been paid. While one may reasonably come to the conclusion that only one of these two methods could be correct, it is our argument that they both have their merits. In

particular, we argue that Sztorc's approach would correspond, to some degree, to an instance of the Nakamoto coefficient in which only subsystems that contribute to knowing that one has been paid are deemed essential. It is this aspect of Sztorc's approach that we find particularly compelling. **We will ultimately use the idea of measuring decentralization only insofar as it contributes to achieving particular, desirable, system properties.**

However, while we are in agreement with particular aspects of Sztorc's work, there are parts of his article with which we disagree. In particular, we disagree with his views on mining. Sztorc claims that decentralization of hashing power is not essential to his notion of decentralization as "miners can't steal anyone's Bitcoin", and because sybil attacks are expensive and hence unlikely. Furthermore, even if the attacker succeeds, Sztorc argues that the attacker is likely to get caught, after which other users will flag the chain as invalid, and censor all further transactions from the attacker's address. We believe that there are several issues with this view of mining. The fact that sybil attacks are so expensive is only true because the Bitcoin network is already heavily decentralized in terms of hashing power distribution. His assertion that the double spending and censorship attacks possible when one entity holds >50% of hashing power will be easily and efficiently detected is far from guaranteed, and doesn't take into account the damage that a single entity holding >50% of hashing power will do the trust that users have inn a system, and the cost and effect of that deterioration of trust. Finally, the idea of the network censoring transactions from certain addresses is incredibly problematic. It is one thing if users elect not to transact with an address because of its behavior, but for the network to prevent an address from transacting would speak volumes to and severely undermine that network's censorship resistance, which is arguably one of the inherent and fundamental properties of truly

decentralized systems. Sztorc only addresses the latter objection, and does so, we argue, unsatisfactorily, by saying that the blacklist-conditions need to be publicly discussed and agreed-to in advance.

3. Approach

In this section we will cover the primary academic contribution and the main idea of this paper. *Quantifying Decentralization* intentionally didn't address how to select the essential subsystems of the system when calculating the *Nakamoto Coefficient* or the *modified Nakamoto Coefficient* of a system. This a problem that Srinivasan knowingly left unsolved, stating:

> *Crucially, a different choice of essential subsystems will change these values. We certainly don't argue that the particular choice of six subsystems here is the perfect one for measuring decentralization…* ***The selection of which essential subsystems best represent a particular decentralized system will be a topic of some debate that we consider outside the scope of this post.***

We would like to suggest a framework for selecting those essential subsystems.

In Vitalik's paper he makes an argument that aligns with Wences's and Sztorc's and with which the authors agree; **as opposed to being a virtue in its own right, decentralization is only important as a means of achieving a particular end. If this is the case, then it stands to reason that we need only measure it in so far as it contributes to achieving that end,** as opposed to measuring general decentralization, or decentralization for the sake of decentralization. However, there remains the problem that is identifying the end that is achieved

by decentralization. **We will propose what we think are the most important properties of a cryptocurrency, and then suggest that the decentralization of subsystems is only important insofar as it contributes to those desirable properties of the system.** As such, we propose the following amendment to Srinivasan's method for calculating the *Nakamoto Coefficient* or the *modified Nakamoto Coefficient* of a system:

1. Enumerate the essential subsystems of the system in question.
2. For each subsystem, **determine whether or not its decentralization contributes to any desirable properties of the system.**
3. For each subsystem **the decentralization of which does contribute desirably to the system**, determine how many entities one would need to compromise to control each subsystem.
4. Use the minimum of these minimums as a measure of the effective decentralization of the system.

3.1. Desirable properties of Cryptocurrencies

Cryptocurrencies are an alternative to typical currencies, and, in accordance with the argument made by Wences Casares, we believe that a significant part of their value is predicated upon their separation from (and lack of) governing/ regulatory third parties that hold decision-making power and that have, historically, demonstrated their inability to supply a dependable means of payment. Thus, to that end, two essential properties of cryptocurrencies are:

1. **Trustless:** It must not be the case that you have to trust a third party at any point during your participation in the system.

2. **Censorship resistance:** It must be the case that anyone can participate in the system, and that the ledger is immutable, so that no-one can remove or edit your participation in the system.

These two properties are of the utmost importance for a cryptocurrency to truly achieve *political decentralization* and *sovereignty,* which are necessary to avoid being beholden to the whims of a governing third party.

Banks/ financial institutions, regulatory/ compliance agencies, national/ state governments, etc do not exist exclusively to mis-manage the supply of money. They also play a significant role in securing our monetary system. By removing them, the onus is on cryptocurrencies to ensure that they can survive accidents, disasters and attacks. Accordingly, cryptocurrencies need to be:

3. **Byzantine fault tolerance:** A Byzantine fault is the term used to describe the situation in which a component in a system (typically a distributed system) fails, but that failure is difficult (or even impossible) to detect and may cause the component to behave unreliably (for example, the component may broadcast incorrect information to the rest of the system) [13]. Formally, a system is said to be tolerant of Byzantine faults if the following is true of the system:

   Consider a system of $\boldsymbol{n}$ *components,* $\boldsymbol{t}$ *of which are dishonest, and assume only point-to-point communication between all of the components in the system. Whenever a component* $\boldsymbol{n_i}$ *tries to broadcast a value* $\boldsymbol{x}$*, all of the other components are allowed to discuss with each other and verify the consistency of the broadcast until they eventually*

settle on a common value ***y****. The system is said to resist Byzantine faults if a component* $\boldsymbol{n_i}$ *can broadcast a value* ***x****, and:*

I. If $\boldsymbol{n_i}$ *is honest, then all of the honest components in the system agree on the broadcast value* ***x****.*

II. In any case, all honest components agree on the same value ***y****. [13][14]*

Fault tolerance and *attack resistance* as discussed and defined by Vitalik in *The Meaning of Decentralization* are encompassed by Byzantine fault tolerance.The importance of these properties is somewhat self-evident; a currency wouldn't be particularly useful if it were the case that you could lose access to your balance in the case of a disaster or your money could be stolen in the case of an attack.

The reasoning behind consensus mechanisms often assumes that participants behave in accordance with the *uncoordinated choice model*, or that they behave as individual actors and make their decisions independently[5]. A consequence of removing third parties and adopting the peer-to-peer governance of cryptocurrencies is that systems become vulnerable to collusion between like-minded parties or bad actors. On top of that, the pseudonymous nature of blockchains and cryptocurrencies removes accountability as a deterrent for bad actors. Hence, cryptocurrencies need to posses the following property:

4. **Collusion resistance:** The system should be resistant to collusion between participants who aim to behave in such a way as to benefit themselves at the expense of other participants in anything but the most dire of circumstances.

Collusion resistance is not encompassed by Byzantine fault tolerance because Byzantine fault tolerance considers only a small number of bad actors (namely, less than a third of the components/ participants). In the case of cryptocurrencies, we would like it to be the case that even collusion of larger groups of bad actors (possibly even the majority of participants) is unlikely. To some extent this is achieved with incentive engineering. In cryptocurrencies all participants are incentivized to behave well because not doing so would jeopardize the value of the cryptocurrency, and because participants in cryptocurrency systems hold/ are paid in that cryptocurrency, this would be detrimental to them.

Collusion resistance can be likened to the fault tolerance of the community that governs the protocol. Without collusion resistance a currency would be unreliable, and would not be a good store of value or medium of exchange. Therefore, safety measures ought to be put in place so as to prevent this kind of behavior. Furthemore, these safety measures necessitate a sufficient level of decentralization in their respective subsystems in order to be effective.

4. Results

Using our framework, we will re-attempt to calculate the minimum modified Nakamoto Coefficient of Bitcoin and Ethereum (using the same data as was used in the Quantifying Decentralization post, so as to be sure that any changes are exclusively a result of our framework), and attempt to calculate the minimum modified Nakamoto Coefficient of Solana. We decided to re-attempt the calculation of the minimum modified Nakamoto Coefficient of Bitcoin and Ethereum to see whether our framework caused those values to change, and we decided to

attempt to calculate the minimum modified Nakamoto Coefficient of Solana because recently the Nakamoto Coefficient was used as the basis for arguing that Solana is in fact more decentralized than Bitcoin and Ethereum[15]. Intuitively, this seems unlikely and thus this result could be construed as incredibly problematic for advocates of the Nakamoto Coefficient. If it is the case that Solana is less decentralized than Bitcoin and Ethereum but has a higher Nakamoto Coefficient, that would surely undermine the Nakamoto Coefficient's usefulness as a metric for measuring decentralization. If, using our framework, we produce a Nakamoto Coefficient for Solana that is in line with our intuition that Solana is more centralized than Bitcoin and Ethereum, this could indicate that our framework has contributed to the refinement of the Nakamoto Coefficient.

Subsystem	Measure	Bitcoin (Gini)	Bitcoin (Nakamoto)	Ethereum (Gini)	Ethereum (Nakamoto)
Mining	Block reward	0.4	5	0.82	3
Client	Unique codebases	0.915	1	0.92	1
Developer	Commits to main client	0.79	5	0.91	2
Exchange	24 hour volume	0.83	5	0.85	5
Node	Distribution across countries	0.84	3	0.85	4
Owner	Distribution across addresses with >$500k [Jul 2017]	0.65	456	0.76	72
Maximum Gini / Min Nakamoto		**0.915**	**1**	**0.92**	**1**

Figure 1: Nakamoto coefficient results by Balaji S. Srinivasan [4]

4.1. Bitcoin

We will consider each of the 6 subsystems suggested by Srinivasan (Mining, Clients, Developers, Exchanges, Nodes, Ownership) and reason about how, or indeed, whether, decentralization of each of those subsystems ensures or jeopardizes Bitcoin's maintaining the aforementioned desirable properties.

Mining Decentralization (by Block Reward)

Mining is the foundation of the Bitcoin ecosystem. It is how blocks are validated and added to the chain and, by extension, it is the system that powers the fundamental functions of the cryptocurrency. Centralization of mining is incredibly problematic. If mining power is distributed such that a single entity has over 25% of the hashing power, selfish mining becomes possible (although this is just a means of winning a larger proportion of the block rewards than your hashing power dictates that you should. It negatively affects other miners, as the distribution of block rewards is a zero-sum game, and it undermines the integrity of the system to some extent, but it does not affect ordinary Bitcoin users beyond that)[16]. In this case it is not clear that any of our desired properties are violated. However, an entity with above 50% of the hasing power can double spend his/ her/ their coins and censor participation of particular addresses in the system, directly undermining multiple of our "desirable properties". **Mining decentralization is definitely an essential subsystem,** the decentralization of which needs to be considered when calculating Bitcoin's Nakamoto Coefficient.

When discussing the decentralization of hashing power in proof of work chains, in particular in the case of Bitcoin, the issue of mining pools and their power is inevitably raised. Bitcoin advocates will often argue that miners can leave pools easily and quickly, and point to the example that in 2014 the Bitcoin mining pool Ghash.io briefly held over half of the total Bitcoin mining hash power. However, miners voluntarily left the pool to maintain Bitcoin's integrity before any significant damage was done[17]. However, in March 2013 a change in Bitcoin clients between version 0.7 and 0.8 meant that block 225430 was accepted by clients/miners running 0.8, but not by clients/miners running 0.7, causing a fork in the chain. The Bitcoin

developers decided that the chain that did not contain block 225430 should be the accepted chain, despite it being the shorter of the two chains. They convinced mining pools, in particular BTC Guild, to support this decision by running 0.7 and it was so. The fact that this was possible undermines the claim that Bitcoin is (or at least was) a decentralized system and that the network is governed by the majority of the hashing power. Fewer than 10 entities made a decision in opposition to the majority of the computing power in the network, demonstrating the power of mining pools[1]. This does, however, mean that calculating the *mining* coefficients for Bitcoin (and Ethereum) requires only considering the hash power of mining pools, which is exactly what Srinivasan did in *Quantifying Decentralization* to produce a ***mining* coefficient of 5 for Bitcoin.**

Ownership Decentralization (by address)

A single entity that controls enough Bitcoin could try to manipulate the network by holding/ dumping the coins to affect the price and liquidity in such a way that benefits them, however it is not clear that this would undermine the desirable properties that we listed. It likely would not. Accordingly, **ownership decentralization is not an essential subsystem** that we would include in the measurement of the Bitcoin's Nakamoto coefficient.

This is not to say that if a malicious entity held enough Bitcoin that entity couldn't do some serious damage. Controlling sufficient Bitcoin could significantly undermine Bitcoin's utility for other users. However, this would involve compromising the users most heavily invested in the success of the network (afterall, the value of that entity's significant Bitcoin balance is predicated upon the success of the network), which would likely prove incredibly challenging.

Node Decentralization (by country)

While node decentralization in general is certainly important, measuring node decentralization is a non-trivial problem. **Node decentralization by country is certainly not an adequate representation of general node decentralization.** While the geographical distribution of hashing power (which is essentially what node decentralization referred to in *Quantifying Decentralization*) is not entirely insignificant, it is something that should really be accounted for when measuring hashing decentralization.

If a nation (for example China, where the majority of Bitcoin mining takes place) were to successfully ban mining the hashing difficulty would just be adjusted so that it would become far easier (and thus more profitable) to mine blocks, which would be sufficient to attract miners from other parts of the world to mine Bitcoin and in turn cause the network to resume normal function. A more troubling consequence of geographical centralization of hashing power is the possibility of a nation taking control of the hashing power therein. In the case of China, in 2019 this would have resulted in the nation controlling 75% of the hashing power (although now no nation contains a majority of the Bitcoin hashing power)[18]. However, merely measuring node decentralization by country does not account for the difficulty/ likelihood of a nation taking control of all of the miners/ hashing power/ hardware within its borders, what percentage of those miners the nation could realistically take control of, or whether, in the case of no nation containing a majority of the hashing power, nations could/ would collude to do this. Such undertaking would be incredibly expensive, incredibly challenging and would involve incredible coordination, and we can reasonably assume that its likelihood is remarkably low. Furthemore, such a large-scale operation would not be inconspicuous, and miners would likely have more

than enough time to respond. Hence, while we agree that node decentralization is an essential subsystem, **node (hashing) decentralization *by country* is not an essential subsystem** that we would include in the measurement of the Bitcoin's Nakamoto coefficient.

Exchange Decentralization (by 24hr volume)

Exchange decentralization is only a problem because of the prevalence of custodial exchanges (custodial exchanges essentially own and control their users' balances, so a malicious exchange could deploy those resources in any way that it chooses, for example by holding/ dumping the coins to affect the price and liquidity in such a way that benefits them), making it essentially a special case of ownership decentralization (although that considering exchanges by 24hr volume does not account for this). Thus, we argue that **exchange decentralization is not an essential subsystem** that we would include in the measurement of the Bitcoin's Nakamoto coefficient.

Developer Decentralization (by commits to core)

Satoshi had this to say about Bitcoin:

> *"The nature of Bitcoin is such that once version 0.1 was released, the core design was set in stone for the rest of its lifetime."* [19]

Developers' power over the Bitcoin ecosystem is limited, and it is not clear that the total number of commits to core is an adequate measure of the power/ influence of a developer. As with changes to clients, developer commits need to be accepted by the majority of participants to take effect, and due to the nature of version control, can be "rolled back" in the case that a commit is

problematic. Thus, the decentralization of this subsystem has very little to do with preserving our desirable properties, and accordingly **developer decentralization is not an essential subsystem** that we would consider when measuring the decentralization of Bitcoin.

Client Decentralization (by codebase)

There aren't very many Bitcoin clients, and many of the popular clients are just forks of Bitcoin core. However, many (including the authors) argue that that (forking) is the responsible way to create separate clients. It allows for clients to differ in minor ways - anyone can fork the codebase and offer an alternative solution - without having unexpected consequences with regards to consensus. Satoshi had this to say about truly different clients:

> *"I don't believe a second, compatible implementation of Bitcoin will ever be a good idea. So much of the design depends on all nodes getting exactly identical results in lockstep that a second implementation would be a menace to the network."* [19]

Whatsmore, it is entirely possible to change Bitcoin clients in the case that you detect or are made aware of an issue with the client that you are running (and the damage that a malicious client could possibly do while maintaining eligibility to participate in the network seems minimal) or to rollback to previous versions of open source clients that are known to work. Thus, **client decentralization is not an essential subsystem** that we would consider when measuring Bitcoin's Nakamoto Coefficient.

4.2. Ethereum

As was the case with Bitcoin, we will consider each of the 6 subsystems suggested by Srinivasan (Mining, Clients, Developers, Exchanges, Nodes, Ownership) and reason about how, or indeed, whether, decentralization of each of those subsystems ensures or jeopardizes Ethereum's maintaining the aforementioned desirable properties. Due to differences from system to system, the same subsystem can play different roles from blockchain to blockchain, and, accordingly, different subsystems will be deemed essential in different blockchains. For example, in the case of Ethereum, we need to connsider the upcoming switch to Ethereum 2.0 and the accompanying adoption of a Proof of Stake consensus mechanism.

Mining Decentralization (by Block Reward)

As observed in the case of Bitcoin, until Ethereum's change to of Proof of Stake, hashing power (what Srinivasan is essentially measuring by considering the distribution of the block reward) is certainly one of the subsystems that need to be decentralized. Similarly to Bitcoin, because of the sheer power that mining pools have, decentralization of this subsystem is really just determined by the hashing power of large mining pools. After The Merge hashing power will be replaced by staking power (the amount of staked ETH). In either case, the subsystem related to a blockchain's consensus mechanism, be it mining or staking, should be considered when calculating the Nakamoto Coefficient. **Ethereum has a *mining* coefficient of 3.**

Ownership Decentralization (by address)

As noted in our related reading section, it is very hard to validate how accurate measures of ownership decentralization are because a single entity might be the owner of an unlimited

number of addresses. In the case of Proof of Work blockchains, while centralized ownership of tokens could have some adverse effects, it does not directly impact the desired properties that we've listed. Accordingly this was not a subsystem that we deemed essential in our analysis of Bitcoin. However, in the case of Ethereum, once the network switches to Proof of Stake, ownership distribution will be closely tied to staking power. Although it will not be the case that all users stake their ETH, either becayse staked tokens become an illiquid assets or because users rather allocate their tokens to DeFi protocols that have better returns, the Nakamoto Coefficient is concerned with measuring the mere possibility of these large holders overtaking the network. In Proof of Stake blockchains, centralizatio of ownership implies centralization of political power and undermines the trustlessness and the censorship resistance of the chain. **Therefore, this subsystem needs to be decentralized.**

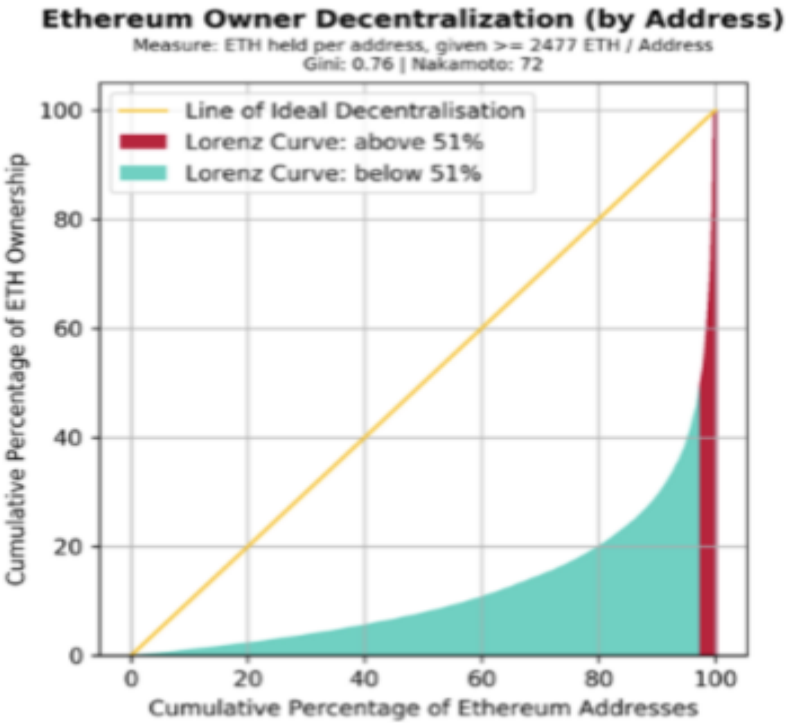

Figure 2: Lorenz Curve graph from Quantifying Decentralization paper

The figure above depicts the percentage of large ETH wallets one needs to compromise in order to own 51% of all ETH tokens [4]. This is what Srinivasan considered when he measured Ethereum's ownership decentralization. However, one needs to own only 33% of tokens to be

able to compromise a Proof of Stake blockchain. Hence, according to the graph that Srinivasan used, the top 1-2% of all large wallets own 33% of all ETH tokens. That corresponds to about 30 wallets. We acknowledge that these are just rough estimates, but more precise calculations are not necessary at the moment as it's reasonable to assume that the results of this subsystem won't affect the final minimum Nakamoto Coefficient of Ethereum. **Hence, the Nakamoto coefficient for ownership distribution of Ethereum is 30.**

Node Decentralization (by country)

As we stated when discussing this subsystem's role in Bitcoin, node decentralization in general is certainly important, but node decentralization by country is not an adequate representation of general node decentralization. It is still true that merely measuring node decentralization by country does not account for the nuances of a nation taking control of all of the hashing power within its borders. After Ethereum switches to Proof of Stake, token ownership distribution by country will be the equivalent to node decentralization before the switch. This would be even more difficult to seize than hashing power, and acquiring enough ETH to compromise the network is unfeasibly expensive, even for nation states. **Accordingly, we found that node (hashing) decentralization by country is not an essential subsystem that we would include in the measurement of the Ethereum's Nakamoto coefficient.**

Exchange Decentralization (by 24hr volume)

Similarly to Bitcoin, while Ethereum is Proof of Work the decentralization of this subsystem does not affect our desired properties. As we noted, custodial exchanges essentially own and control their users' balances, so exchange decentralization is essentially a special case of

ownership decentralization, which means that when Ethereum adopts a Proof of Stake consensus mechanism, exchange decentralization will be an essential subsystem. However, it is stil the case that considering exchanges by 24hr volume does not account for this issue. Thus, as was the case with Bitcion, we argue that **exchange decentralization is not an essential subsystem**.

Developer Decentralization (by commits to core)

In the case of Bitcoin we disregarded this subsystem because developer commits need to be accepted by the majority of the users to take effect, and they are somewhat decentralized in this sense. Whatsmore, because Bitcoin isn't a system that is evolving significanly, developers do not have the capacity to change Bitcoin very much. This is not the case for Ethereum. Not only is Ethereum receiving regular, significant updates, the Ethereum founders Vitalik Buterin (and to some extent Joseph Lubin) hold significant sway in the community. This undermines the trustlessness of Ethereum. **However, as we noted, this subsystem as measured by commits to core, does not affect our desired properties, and should not be used in calculating Nakamoto coefficient.** To truly reflect developer decentralization the analysis of this subsystem needs to be both quantitative and qualitative.

Client Decentralization (by codebase)

As was the case with Bitcoin, the level of decentralization of this subsystem has little effect on the desired properties of the system as a whole. The properties of the network are determined by the client(s) run by the majority of the network participants, so in the case that any issues are detected with the newer versions participants can just roll back to any previous versions that are

known to be stable. **Accordingly, this subsystem should not be considered when calculating the Nakamoto Coefficient for Ethereum.**

4.3. Solana

Solana's touted Nakamoto coefficient of 19 [11] (which has since increased to 21) represents the number of nodes that would need to compromised in order to control 33% of the staked Solana (at which point a 33% attack becomes possible, the proof of stake equivalent of a 51% attack when the latency of the network is accounted for)[20]. There are multiple problems with this measure of Solana's Nakamoto coefficient.

This is a measure of the number of nodes that would need to be compromised in order to control 33% of the staked Solana. However, this does not correspond to a subsystem that the Nakamoto Coefficient considers. It is a hybrid of two subsystems, the *ownership* subsystem and the *nodes* subsystem. While it is true that one can initiate a 33% attack if one controls >33% of the staked coins (i.e. what is being measured above), if one does not control >33% of the total coins then the attack can be effectively resisted if Solana owners who are not staking their Solana elect to stake their Solana (in which case the attacker will no longer control >33% of the staked coins). To mount the true equivalent of a proof of work 51% attack the attacker needs to compromise enough *addresses* such that he or she ends up in control >33% of the market cap of the proof of stake cryptocurrency that he or she is attacking.

Considering that it is estimated that ~48% of the total Solana supply is held by "insiders" and 33% is reserved for "ecosystem development"[21], it is quite possible that Solana's *ownership*

coefficient is <21 (especially given that when calculating the Nakamoto coefficient if proof of work coins, mining pools tend to be grouped together, so it would be reasonable to group together "insiders" when calculating Solana's Nakamoto coefficient).

It is also likely that Solana's *nodes* coefficient is <21, in particular because the nodes being counted in the are *validator nodes*, not *full nodes.* Validator nodes store only the current state of the chain, as opposed to the full history of the chain, which full nodes store. Consequently, while the system can recover from node failure provided that at least one full node remains, this is not the case with validator nodes. In the case of Solana there is not much in the way of full node information available, but it appears that the entirety of Solana's network relies upon 4 trusted nodes, run by the foundation[22]. This is grounds for a *nodes* coefficient of 1.

Even if you do not take into account the issue of validator nodes vs full nodes, the requirements for running a Solana validator node are extremely high (the bold requirements are the most significant requirements)[23]:

CPU

- ***12 cores / 24 threads, or more***
- *2.8GHz, or faster*
- *AVX2 instruction support (to use official release binaries, self-compile otherwise)*
- *Support for AVX512f and/or SHA-NI instructions is helpful*
- *The AMD Zen3 series is popular with the validator community*

RAM

- ***128GB, or more***
- *Motherboard with 256GB capacity suggested*

Disk

- *PCIe Gen3 x4 NVME SSD, or better*

- *Accounts: 500GB, or larger. High TBW (Total Bytes Written)*
- *Ledger: 1TB or larger. High TBW suggested*
- *OS: (Optional) 500GB, or larger. SATA OK*
- *The OS may be installed on the ledger disk, though testing has shown better performance with the ledger on its own disk*
- *Accounts and ledger can be stored on the same disk, however due to high IOPS, this is not recommended*
- *The Samsung 970 and 980 Pro series SSDs are popular with the validator community*

Network

- ***Internet service should be at least 300Mbit/s symmetric, commercial. 1GBit/s preferred***

In order to participate in consensus, a vote account is required which has a rent-exempt reserve of 0.02685864 Solana. Voting also requires sending a vote transaction for each block the validator agrees with, which can cost up to ***1.1 Solana per day ($110 per day, or $40000 per year)*** .

By all accounts the requirements to run a full node are even higher, **including needing a minimum of 17tb of adequately fast storage**. Compare the above requirements to the requirements to run a Bitcoin full node:

CPU

- *Negligible*

RAM

- *2 gigabytes of memory (RAM)*

Disk

- *7 gigabytes of free disk space, accessible at a minimum read/write speed of 100 MB/s.*

Network

- *A broadband Internet connection with upload speeds of at least 400 kilobits (50 kilobytes) per second*
- *An unmetered connection, a connection with high upload limits, or a connection you regularly monitor to ensure it doesn't exceed its upload limits. It's common for full nodes on high-speed connections to use 200 gigabytes of upload or more a month. Download usage is around 20 gigabytes a month, plus around an additional 340 gigabytes the first time you start your node.*

6 hours a day that your full node can be left running. (You can do other things with your computer while running a full node.) More hours would be better, and best of all would be if you can run your node continuously.

One of the major consequences of these requirements is that Solana validators need to run in data centers. Currently >33% of nodes reside in just 3 data centers, which is clearly far from decentralized[24]. Whatsmore, the amount of Solana needed to break even when running a validator node is so significant and therefore prohibitive that the Solana Foundation runs a subsidization program, awarding grants of 25,000 Solana to validators meeting specific criteria (a validator would need $1 million worth of Solana staked without subsidization just to break even) [21] which means that the Solana Foundation indirectly chooses many validators on the network. Clearly even if one considers validator nodes the claims with regards to *nodes* coefficient of Solana quickly fall apart.

Solana's requirements mean that it could prove incredibly challenging for Solana to achieve decentralization of nodes. Many would argue that a cryptocurrency cannot be decentralized if it is not the case that running a full node is feasible for most regular people (this is the fundamental argument of Paul Sztorc's aforementioned paper).

Having discussed the issues with the existing attempts to measure Solana's Nakamoto Coefficient, we will attempt to calculate Solana's Nakamoto Coefficient.

Mining/ Staking Decentralization

Solana uses Proof of Stake as its consensus protocol, as opposed to Ethereum and Bitcoin, which use Proof of Work. As such, Solana doesn't have a "mining" subsystem. As is the case with Bitcoin and Ethereum, the consensus protocol is the foundation of the network, and dictates what data is added to the blockchain. Solana's equivalent of mining is staking, and just like the mining pools that exist for Bitcoin and Ethereum, Solana staking pools have emerged. Because of its importance in determining the state of the blockchain, this subsystem ought to be sufficiently decentralized in any blockchain in order for the system to have desired properties of decentralization. In the case of Solana, the 21 largest validators account for 33% of all staked Solana on the network. This means that these entities could theoretically censor/halt the network if they colluded or got corrupted. **Hence, the Nakamoto Coefficient for this subsystem of Solana is 21.**

Ownership Decentralization

In the case of Proof of Stake networks, and by extension Solana, ownership decentralization is particularly important because all holders could in theory stake their balances. Therefore, if large holders colluded and staked their Solana, they could significantly lower the Nakamoto Coefficient of the mining subsystem. The 101 richest wallets hold 33% of all tokens[25]. However, this data is not reliable as it does not reflect the possibility of a single entity owning multiple addresses. It is also grossly inconsistent with the initial distribution of Solana tokens.

While one could argue that the Solana foundation insiders and other large holders are strongly incentivized to preserve the integrity of the network, as the value of their tokens depends on the network being perceived as reliable, Solana's initial distribution was heavily centralized, with 48% of tokens being given to insiders (the founders, the early team member, and early investors) and 33% of tokens were reserved for "ecosystem development". As mentioned earlier, given that when calculating the Nakamoto coefficient of proof of work coins, mining pools tend to be grouped together, it would be reasonable to group together "insiders" when calculating Solana's Nakamoto coefficient, and **accordingly Solana would have an *ownership* coefficient of 1**.

Node Decentralization

Earlier we noted the centralisation of Solana nodes. The most important observations were:

I. It appears that the entirety of Solana's network relies upon 4 trusted nodes, run by the Solana foundation. At the very minimum is grounds for a *nodes* coefficient of 4. It could also be grounds for a *nodes* coeffficient of 1.

II. Currently >33% of all validator nodes reside in just 3 data centers, which is clearly far from decentralized, and grounds for a *nodes* coeffficient of 3.

III. The Solana Foundation indirectly chooses most of the validators on the network. This could also be considered grounds for a *nodes* coeffficient of 1.

Accordingly, we argue for a *nodes* coeffficient of 1.

Exchange Decentralization

As we noted, custodial exchanges essentially own and control their users' balances, so exchange decentralization is essentially a special case of ownership decentralization, and is thus important

for Proof of Stake chains. However, measuring the number of tokens controlled by an exchange is incredibly difficult. It appears that, as of 22 April 2022, about one quarter of Solana's trading takes place on Binance, and that Coinbase accounts for a twelfth of Solana's trade volume [26]. If ownship distribution across exchanges is in proportion to trade volume, then compromising Binance and Coinbase would mean gaining control over 33% of the Solana in existence, which would mean that Solana has an exchange coefficient of 2. However, as noted earlier, insiders likely hold near to 70% of the Solana in existence, and they almost certainly hold very little of that Solana on custodial exchanges. So, Solana's centralization of ownership may in fact mean that no exchange, or even combination of exchanges, controls sufficient Solana to compromise the network. **Accordingly, while this subsystem is essential, it will not affect Solana's Nakamoto Coefficient, which will be determined by one of the other, less centralized, essential subsystems.**

Developer Decentralization

Solana is heavily centralized in this regard, with most commits pushed by a group of half a dozen developers. Furthermore, the Solana blockchain is maintained by Solana Labs, a (single) company based in San Francisco. The Solana Foundation is the only entity developing core nodes on the blockchain. This means that the future trajectory of the system depends heavily on the decisions made by individuals leading Solana Labs and their investors. Again, the decentralization of decision-making is hard to quantify and needs to be supplemented by qualitative analysis. One can reasonably conclude that the system whose core is maintained by a single legal entity cannot claim to be decentralized, and that **Solana has *developer* coefficient of 1.**

However, it is worth noting that Solana is still in it's infancy (the network is still in its beta stage) so this centralization of developers is perhaps necessary, as designing and implementing a project is likely still best done in a centralized manner because is unfeasible for a very large number of developers will build a blockchain from scratch without a handful of people making important decisions and helping to coordinate the community. However, once the system matures, these developers should take the hands-off approach in order for the system to be truly decentralized..

Client Decentralization

As was the case with Bitcoin and Ethereum, the level of decentralization of this subsystem has little effect on the desired properties of the system as a whole. The properties of the network are determined by the client(s) run by the majority of the network participants, so in the case that any issues are detected with the newer versions participants can just roll back to any previous versions that are known to be stable. **Accordingly, this subsystem should not be considered when calculating the Nakamoto Coefficient for Solana.**

5. Analysis

Using our framework, we calculated the minimum Nakamoto coefficient to be 5 for Bitcoin, 3 for Ethereum, and 1 for Solana. These coefficients differ from Srinivasan's findings of 1 for both Bitcoin and Ethereum, as well as Giacaglia's somewhat perplexing Nakamoto coefficient of 19 for Solana [11]. Our findings are better aligned with our intuitive understandings of the systems in question than Srinivasan and Giacaglia's. In particular in the case of Solana, it made no sense

that a system developed, led, and to great extent owned by a single entity could have a higher Nakamoto Coefficient, and thus be more decentralized, than a system such as Bitcoin. This is promising insofar as it suggests that our framework did in fact result in Nakamoto Coefficients that better represent the decentralization of the underlying systems.

6. Conclusion and Further work

In this paper we made three contributions to the academic literature surrounding bitcoin and cryptocurrencies:

I. **We proposed a framework for selecting the essential subsystems, the decentralization of which ultimately determine the *Nakamoto Coefficient* and the *modified Nakamoto Coefficient* of a system.** We based this framework on the notion that decentralization is only important as a means of achieving a particular end, and therefore we need only measure it in so far as it contributes to achieving that end. We proposed what we think are the most important properties of a cryptocurrency, namely **Trustless**, **Censorship resistance**, **Byzantine fault tolerance** and **Collusion resistance**, and then suggested that the decentralization of subsystems is only important insofar as it contributes to those desirable properties of the system.

II. **We used that framework to re-measure the Nakamoto Coefficients of Bitcoin and Ethereum, and in both cases our changes increased the measured Nakamoto Coefficients.**

III. **We used that framework to measure the Nakamoto Coefficient of Solana**, and the result was a Nakamto Coefficient that was lower than those of Bitcoin and Ethereum, and

therefore in line with our intuition that Solana is more centralized than Bitcoin and Ethereum (as opposed to previous attempts to measure Solana's Nakamoto Coefficient, which came to the conclusion that Solana is more decentralized than Bitcoin and Ethereum).

While the Nakamoto Coefficient is a great tool for measuring of decentralization of blockchains, there is still room for further work. Besides the issue of the choice of subsystems, which we attempted to address in this paper:

I. Work needs to be done with regards to the choice for thresholds of particular subsystems, as the proportion of participants that need to be compromised in order to compromise the network varies from subsystem to subsystem.

II. Work needs to be done with regards to analysing blockchains and interpreting the data stored therein. This could contribute to more accurate measurements of the decentralization of the subsystems of blockchains, and by extension more accurate measurements of the Nakamoto Coefficients of blockchains. An example of such work is the research that is being done with regards to deanonymising cryptocurrency owners (clustering wallet addresses that are flagged as potentially belonging to the same entity with some degree of confidence) which could affect how we measure the decentralization of token ownership and mining power.

III. Finally, as blockchains get increasingly integrated across industries and infrastructure and as cryptocurrencies experience wider adoption, the benefits and shortcomings of decentralized systems will become more clear, which will further inform the desired properties of blockchains and by extension our selection of essential subsystems.

Bibliography

[1] Gervais, A., Karame, G. O., Capkun, V., & Capkun, S. (2014). Is bitcoin a decentralized currency? *IEEE Security & Privacy*, *12*(3), 54–60. https://doi.org/10.1109/msp.2014.49

[2] Narayanan, A., & Clark, J. (2017). Bitcoin's academic pedigree. *Communications of the ACM*, *60*(12), 36–45. https://doi.org/10.1145/3132259

[3]“Decentralization.” *Merriam-Webster.com Dictionary*, Merriam-Webster, https://www.merriam-webster.com/dictionary/decentralization. Accessed 22 Apr. 2022.

[4]B. S. Srinivasan, “Quantifying Decentralization,” *Medium.com*, 27-Jul-2017. .

[5]V. Buterin, “The Meaning of Decentralization,” *Medium.com*, 06-Feb-2017. .

[6]W. Casares, “The Case for Small Allocation to Bitcoin.”

[7]K. F. Veigel, *Governed by emergency: Economic policy-making in Argentina, 1973-1991*. 2005.

[8]“Argentina Since Default: The IMF and the Depression,” *Argentina since default: The IMF and the depression, by Alan B. Cibils, Mark Wesibrot, and Debayani Kar, Sept 2002*. [Online]. Available: https://cepr.net/documents/publications/argentina_2002_09_03.htm. [Accessed: 22-Apr-2022].

[9]H. F. A. von, *Denationalization of money: The argument refined*. London: Inst. of Economic Affairs, 1978.

[10] P. Sztorc, “Measuring Decentralization,” *TruthCoin*, 09-Sep-2015. .

[11]G. Giacaglia, “How Solana scales maintaining decentralization,” *Mirror.xyz*, 15-Dec-2021. .

[12]“R/polkadot - nakamoto coefficient,” *reddit*. [Online]. Available: https://www.reddit.com/r/Polkadot/comments/qmgi89/nakamoto_coefficient/. [Accessed: 22-Apr-2022].

[13] L. Lamport, “The weak Byzantine generals problem,” *Journal of the ACM*, vol. 30, no. 3, pp. 668–676, 1983.

[14]M. Pease, R. Shostak, and L. Lamport, “Reaching agreement in the presence of faults,” *Journal of the ACM*, vol. 27, no. 2, pp. 228–234, 1980.

[15]“R/solana - solana is more decentralized than Ethereum or bitcoin,” *reddit*. [Online]. Available: https://www.reddit.com/r/solana/comments/nj19ep/solana_is_more_decentralized_than_ethereum_or/. [Accessed: 22-Apr-2022].

[16]I. Eyal and E. G. Sirer, “Majority is not enough: Bitcoin Mining is Vulnerable,” *Communications of the ACM*, vol. 61, no. 7, pp. 95–102, 2018.

[17]Y. Pritzker, *Inventing bitcoin: The technology behind the first truly scarce and decentralized money explained*. United States, 2019.

[18]“Bitcoin mining map,” *CCAF.io*. [Online]. Available: https://ccaf.io/cbeci/mining_map. [Accessed: 22-Apr-2022].

[19]*Transactions and scripts: DUP HASH160 ... Equalverify Checksig*. [Online]. Available: https://bitcointalk.org/index.php?topic=195.msg1611#msg1611. [Accessed: 22-Apr-2022].

[20]“Solana Beach,” *Dashboard*. [Online]. Available: https://solanabeach.io/validators. [Accessed: 22-Apr-2022].

[21]T. Craig, “How Decentralized is Solana?,” *cryptobriefing*, 16-Oct-2021. .

[22]J. Lopp, “2021 altcoin node sync tests,” *Cypherpunk Cogitations*, 03-Sep-2021. [Online]. Available: https://blog.lopp.net/2021-altcoin-node-sync-tests/. [Accessed: 22-Apr-2022].

[23]“Validator requirements: Solana docs,” *Solana Docs Blog RSS*. [Online]. Available: https://docs.solana.com/running-validator/validator-reqs. [Accessed: 22-Apr-2022].

[24]“Mainnet Data Centers,” *Solana Validators*. [Online]. Available: https://www.validators.app/data-centers?locale=en&network=mainnet. [Accessed: 22-Apr-2022].

[25]“Solana Beach,” *Dashboard*. [Online]. Available: https://solanabeach.io/supply. [Accessed: 22-Apr-2022].

[26]“Solana Exchanges Sol Markets: Buy & Sell solana,” *Solana Exchanges SOL Markets | Buy & Sell Solana | CryptoRank.io*. [Online]. Available: https://cryptorank.io/price/solana/exchanges. [Accessed: 22-Apr-2022].