UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Augusto Gai Dal'Asta

DETECÇÃO E ANÁLISE DE PADRÕES DE CENTRALIZAÇÃO EM REDES DE CRIPTOMOEDAS

Santa Maria, RS
2022

Augusto Gai Dal'Asta

DETECÇÃO E ANÁLISE DE PADRÕES DE CENTRALIZAÇÃO EM REDES DE CRIPTOMOEDAS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciência da Computação da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Bacharel em Ciência da Computação**. Defesa realizada por videoconferência.

ORIENTADOR: Prof. Joaquim Vinicius Carvalho Assunção

Número do TG: 502
Santa Maria, RS
2022

End. Eletr.: agasta@inf.ufsm.br

Augusto Gai Dal'Asta

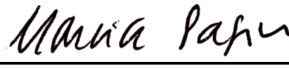
DETECÇÃO E ANÁLISE DE PADRÕES DE CENTRALIZAÇÃO EM REDES DE CRIPTOMOEDAS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciência da Computação da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Bacharel em Ciência da Computação**.

Aprovado em 11 de fevereiro de 2022:

Joaquim Vinicius Carvalho Assunção, Dr. (UFSM)
(Presidente/Orientador)

Carlos Raniery Paula dos Santos, Prof Dr. (UFSM) (videoconferência)



Marcia Pasin, Profª Drª . (UFSM) (videoconferência)

Santa Maria, RS
2022

RESUMO

DETECÇÃO E ANÁLISE DE PADRÕES DE CENTRALIZAÇÃO EM REDES DE CRIPTOMOEDAS

AUTOR: Augusto Gai Dal'Asta
ORIENTADOR: Joaquim Vinicius Carvalho Assunção

Durante a última década, houve o surgimento de um enorme fenômeno no mundo da tecnologia financeira: as criptomoedas. Desde sua primeira concepção em 2008, na forma da *Bitcoin* (NAKAMOTO, 2008), esta inovação propõe um sistema monetário decentralizado através de uma rede *peer-to-peer* de computadores (denominados "mineradores"), que mantém um registro permanente de todas operações realizadas na sua *blockchain*. Apesar de a ausência de uma autoridade central ser um fator chave para qualquer rede de criptomoedas, recentemente pôde-se perceber uma forte centralização em torno de alguns grandes conglomerados de mineração (SAI et al., 2021) em muitas das principais moedas eletrônicas atuais, o que coloca o controle da maior parte deste ecossistema nas mãos de poucos mineradores (CACCIOLI; LIVAN; ASTE, 2016), podendo originar cenários que possam quebrar a confiança do público geral em relação à estas redes. Por exemplo, caso um grupo de mineradores detenha 50% ou mais do poder de processamento de uma *blockchain*, pode-se ocorrer uma falha de segurança nesta rede, denominada como "ataque de 51%". Este cenário permitiria que os membros deste grupo impedissem a confirmação de novas transações e a criação novos blocos na rede alvo, uma vez que eles estariam no controle do mecanismo de consenso desta rede. Tendo em vista esta crescente centralização originada por grandes conglomerados de mineradores, e sabendo dos riscos à integridade destes sistemas que dado evento pode originar, este trabalho visa analisar padrões de centralização de poder de processamento em diversas redes de criptomoedas.

Palavras-chave: Blockchain. Criptomoedas. Log de dados.

ABSTRACT

DETECTION AND ANALYSIS OF CENTRALIZATION PATTERNS IN CRYPTOCURRENCY NETWORKS

AUTHOR: Augusto Gai Dal'Asta
ADVISOR: Joaquim Vinicius Carvalho Assunção

Over the past decade, a huge phenomenon has emerged in the world of financial technology: cryptocurrencies. Since its first conception in 2008, in the form of *Bitcoin* (NAKAMOTO, 2008), this innovation proposes a decentralized monetary system through a peer-to-peer network of computers and maintains a permanent record of all operations carried out in this way by means of blocks. Recently, however, a strong centralization around some large mining conglomerates has been noticed, which puts control of the majority of the blockchain in the hands of a few miners (SAI et al., 2021), which puts control of the majority of this ecosystem in the hands of a few miners (CACCIOLI; LIVAN; ASTE, 2016), which may give rise to scenarios that can break the trust of the general public on these networks. For instance, if a group of miners holds 50% or more of the mining processing power of a blockchain, there may be a security breach in this network, called "51% attack", which would allow members of this group to prevent the confirmation of new transactions and the creation of new blocks in the target network. Furthermore, this conglomerate would still be able to roll back completed transactions as long as they have most of the processing power of this network. In view of this growing centralization caused by large miners, and the risk to the integrity of these systems, this work aims to analyse centralization patterns of processing power in cryptocurrency networks.

Keywords: Blockchain. Cryptocurrency. Data logging.

LISTA DE FIGURAS

Figura 2.1 – Diagramação de um bloco. 16
Figura 2.2 – Diagramação de um *blockchain*. 17
Figura 2.3 – Diagramação de funcionamento de um algorítmo de consenso *Proof-of-Work* em uma *blockchain*. 18
Figura 3.1 – Diagramação da identificação da *pool* de mineração responsável pela mineração de um bloco a partir de sua transação *coinbase*. 25
Figura 3.2 – Curva de Lorenz em relação ao coeficiente de Gini (G). 26
Figura 4.1 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Bitcoin. 29
Figura 4.2 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Bitcoin Cash. 30
Figura 4.3 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Dash. 31
Figura 4.4 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Ethereum. 32
Figura 4.5 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Litecoin. 33
Figura 4.6 – Evolução mensal do Coeficiente de Gini das criptomoedas analisadas. .. 34
Figura 4.7 – Evolução mensal do Coeficiente de Theil das criptomoedas analisadas. . 35
Figura 4.8 – Coeficiente de Nakamoto das criptomoedas analisadas durante o período estipulado. 36

LISTA DE TABELAS

Tabela 3.1 – Informações gerais sobre as criptomoedas analisadas. 23
Tabela 3.2 – Informações técnicas sobre as criptomoedas analisadas. 24

LISTA DE ABREVIATURAS E SIGLAS

API	*Application Programming Interface*
BTC	Bitcoin
BCH	Bitcoin Cash
$DASH$	Dash
ETH	Ethereum
LTC	Litecoin
PoW	*Proof-of-Work*

LISTA DE SÍMBOLOS

G Coeficiente de Gini

T Coeficiente de Theil

N Coeficiente de Nakamoto

SUMÁRIO

1 INTRODUÇÃO 12
1.1 MOTIVAÇÃO 12
1.2 OBJETIVOS GERAIS 12
1.3 OBJETIVOS ESPECÍFICOS 13
2 REFERENCIAL TEÓRICO 14
2.1 BLOCKCHAINS 14
2.1.1 Funções Hash Criptográficas 14
2.1.2 Assinaturas Digitais Criptográficas 15
2.1.3 Conceituando Blockchain 16
2.1.4 Proof-of-Work 18
2.2 CRIPTOMOEDAS 19
2.2.1 Centralização em Redes de Criptomoedas 19
2.2.2 O Ataque de 51% 20
2.2.3 Double-spending 20
2.3 TRABALHOS RELACIONADOS 21
3 METODOLOGIA 23
3.1 CRIPTOMOEDAS ANALISADAS 23
3.2 FERRAMENTAS UTILIZADAS 23
3.3 OBTENÇÃO DOS DADOS NECESSÁRIOS 24
3.4 PROCESSAMENTO DOS DADOS OBTIDOS 25
3.5 ANÁLISE DE PADRÕES DE CENTRALIZAÇÃO 25
3.5.1 Coeficiente de Gini e a Curva de Lorenz 26
3.5.2 Coeficiente de Theil 27
3.5.3 Coeficiente de Nakamoto 27
3.5.4 Considerações Adicionais 28
4 EXPERIMENTOS REALIZADOS 29
4.1 BITCOIN 29
4.2 BITCOIN CASH 30
4.3 DASH 31
4.4 ETHEREUM 32
4.5 LITECOIN 33
4.6 INDEXAÇÃO ATRAVÉS DOS DADOS OBTIDOS 34
5 CONSIDERAÇÕES FINAIS 37
5.1 TRABALHOS FUTUROS 37
REFERÊNCIAS BIBLIOGRÁFICAS 39

1 INTRODUÇÃO

1.1 MOTIVAÇÃO

Durante a última década, houve o surgimento de um enorme fenômeno no mundo da tecnologia financeira: as criptomoedas. Desde sua primeira concepção em 2008, na forma da *Bitcoin* (NAKAMOTO, 2008), esta inovação propõe um sistema monetário decentralizado através de uma rede *peer-to-peer* de computadores (denominados "mineradores"), que mantém um registro permanente de todas operações realizadas na sua *blockchain*.

Figurativamente, uma *blockchain* pode ser definida como uma sequência de blocos, onde cada um deles armazena uma lista de transações. Uma vez que todos os computadores na rede do *Bitcoin* operam sobre os mesmos blocos, e as transações contidas nestes blocos são visíveis para todos membros dessa rede, torna-se muito improvável que uma fraude aconteça nesse sistema.

Apesar da ausência de uma autoridade central ser um fator chave para qualquer rede de criptomoedas, recentemente pôde-se perceber uma forte centralização em torno de alguns grandes conglomerados de mineração (SAI et al., 2021) em muitas das principais moedas eletrônicas atuais, o que coloca o controle da maior parte deste ecossistema nas mãos de poucos mineradores (CACCIOLI; LIVAN; ASTE, 2016), podendo originar cenários que possam quebrar a confiança do público geral em relação à estas redes.

Por exemplo, caso um grupo de mineradores detenha 50% ou mais do poder processamento de uma *blockchain*, pode-se ocorrer uma falha de segurança nesta rede, denominada como "ataque de 51%". Este cenário permitiria que os membros deste grupo impedissem a confirmação de novas transações e a criação novos blocos na rede alvo, uma vez que eles estariam no controle do mecanismo de consenso desta rede. Além disso, este conglomerado ainda poderia reverter transações concluídas enquanto eles possuírem a maioria do poder de processamento da *blockchain* em questão.

1.2 OBJETIVOS GERAIS

Tendo em vista esta crescente centralização originada por grandes conglomerados de mineradores, e sabendo dos riscos à integridade destes sistemas que dado evento pode originar, este trabalho visa analisar de padrões de centralização de poder de processamento em diversas redes de criptomoedas. Além disso, busca-se quantificar o nível desta centralização através de métricas medidoras de desigualdade de distribuição.

1.3 OBJETIVOS ESPECÍFICOS

Para a realização da análise destes padrões de centralização em redes de criptomoedas, este trabalho tem como objetivos específicos:

- Obter dados sobre os blocos das criptomoedas selecionadas para esta pesquisa.
- Processar as informações obtidas nestes blocos, de modo a obter a porcentagem de poder de processamento todos os *peers* com alto poder computacional e conglomerados de mineração identificados.
- Desenvolver um algoritmo para auxiliar na quantificação do nivel de centralização em uma determinada rede de criptomoedas a partir das informações processadas.

2 REFERENCIAL TEÓRICO

Neste capítulo são definidos os conceitos utilizados ao longo deste trabalho. A seção 2.1 define o conceito de *blockchain*, dividindo-se em três subseções: na subseção 2.1.1, é explicado o conceito de funções *hash* criptográficas, e na subseção 2.1.2, esclarece-se a maneira como estes métodos são utilizados em assinaturas digitais. Na subseção 2.1.3, explica-se como estes dois conceitos se unem para formar a ideia de *blockchain* e como esta tecnologia funciona na prática. Na seção 2.1.4, por fim, discorre-se sobre como algumas *blockchains* utilizam de mecanismos de consenso *Proof-of-Work* para garantir a consistência de uma *blockchain* com todos os seus *peers*.

Na seção 2.2, é comentado sobre como criptomoedas fazem o uso de *blockchains* para gerar um livro-razão distribuído reforçado por uma rede heterogênea de computadores. Na subseção 2.2.1, são abordados alguns conceitos de consenso *Proof-of-Work* e como ele pode ser comprometido em um cenário de maior centralização numa rede de criptomoedas. A seguir, nas seções 2.2.2 e 2.2.3, é abordada a ideia de como este fenômeno pode resultar em uma falha de segurança conhecida como "ataque de 51%", e como essa falha pode desencadear um problema de *double-spending* em redes de criptomoedas.

2.1 BLOCKCHAINS

Blockchains podem ser facilmente compreendidas através de seu próprio nome: uma cadeia de blocos. De forma simples, elas funcionam como uma espécie de base de dados que organiza suas informações em blocos individuais com um tamanho limitado, que, quando preenchidos, são adicionados ao último bloco já ligado nesta cadeia. Enquanto um banco de dados tradicional estrutura seus dados em tabelas, as blockchains montam suas informações de forma similar a uma lista encadeada em ordem cronológica de criação.

Esta tecnologia estabeleceu um protocolo seguro e confiável para transações financeiras peer-to-peer através do uso inteligente de funções *hash* e assinaturas digitais criptográficas. Nas seções seguintes, é explicado como estes conceitos unem-se para formar um sistema de livro-caixa descentralizado confiável (NAKAMOTO, 2008).

2.1.1 Funções Hash Criptográficas

Uma função hash é, por conceito, uma função determinística que recebe como entrada um dado de tamanho arbitrário que, após a aplicação desse método, é transformado

em uma *string* de comprimento fixo, denomidada hash. Para que seja possível categorizar este tipo de função como criptográfica, ela deve possuir um sentido único, de forma que a extração de um hash a partir de uma determinada entrada seja computacionalmente simples, porém a única maneira de obter a entrada original a partir de um hash seja por um método de força bruta, testando todos os valores de entrada possíveis nesta função até encontrar uma correspondência com o hash desejado. É essencial para uma função *hash* criptográfica que qualquer alteração no seu valor de entrada resulte em um *hash* completamente diferente, evitando que o novo valor tenha qualquer correlação com o valor original.

2.1.2 Assinaturas Digitais Criptográficas

Pode-se definir uma assinatura digital criptográfica como um protocolo matemático que verifica a autenticidade de informações disponíveis digitalmente. Uma assinatura digital autêntica, onde todos os pré-requisitos são satisfeitos, oferece a um destinatário qualquer razão muito forte para acreditar que a mensagem foi criada por um remetente legítimo e não foi alterada durante o seu envio.

Para gerar uma assinatura digital, é necessário, primeiramente, obter um par de chaves criptográficas assimétricas: uma chave pública, que estará disponível para todos os membros desta rede, e uma chave privada, que somente o usuário terá acesso. Com estas duas informações disponíveis, o algoritmo criptografa a mensagem a ser enviada utilizando a chave privada do remetente. Uma vez que o conteúdo da mensagem enviada depende da chave privada deste usuário, é impossível que essa mesma assinatura seja utilizada para validar uma mensagem diferente.

O destinatário, então, ao receber a mensagem, averigua a legitimidade da mesma através de uma função secundária. Este método recebe como parâmetros a mensagem, a assinatura digital gerada a partir dela e a chave pública do remetente. Por fim, retorna-se um valor booleano, que indica se essa assinatura foi gerada pela chave privada correspondente à chave pública recebida.

É necessário, porém, que este algoritmo garanta uma determinada impossibilidade de que uma assinatura digital válida seja forjada por qualquer usuário que não possua a chave privada de outro, e também que uma assinatura gerada pelo conjunto da mensagem com a chave privada de um remetente possa ser verificada utilizando sua chave pública correspondente.

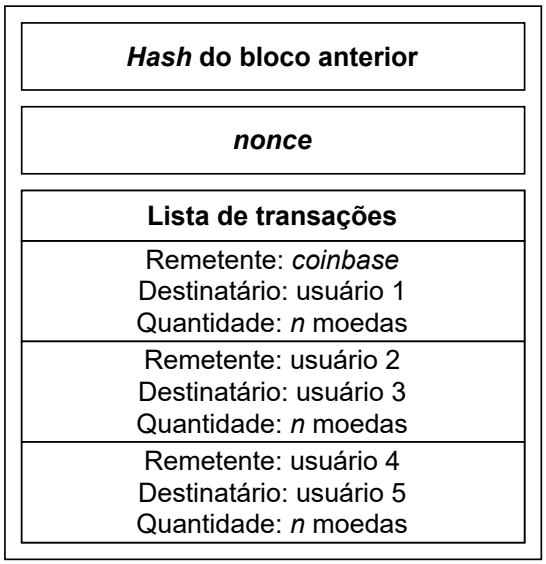
Deste modo, as assinaturas digitais verificam, com razoável confiabilidade, que determinada mensagem foi enviada por um remetente específico. É possível inferir, também, a partir de uma assinatura digital que o remetente da mensagem seja incapaz de negar sua assinatura em uma operação, uma vez que somente um usuário que possua sua chave privada poderia gerá-la com precisão.

2.1.3 Conceituando Blockchain

A *blockchain* foi introduzida como plataforma para a criptomoeda Bitcoin, propondo uma rede descentralizada *peer-to-peer* capaz de manter um livro-razão permanente e imutável sobre todas transações realizadas entre seus usuários (NAKAMOTO, 2008). Neste tipo de rede, seus usuários podem realizar transações assinadas digitalmente a partir de suas respectivas chaves privadas, garantindo aos recipientes destas transações a sua procedência. Estas transações, então, são processadas pelos seus destinatários e, em seguida, armazenadas em um bloco de informação contendo todas as trocas realizadas durante um determinado período de tempo.

Este bloco, por fim, para ser validado, é inserido na última posição da cadeia, contendo uma referência do bloco ao qual ele foi anexado e a lista de transações validadas. É importante ressaltar que toda a informação contida na *blockchain* é de acesso público, o que permite que um usuário verifique o balanço atual de qualquer emissor que esteja transmitindo determinada quantia a ele. Portanto, é impossível que esta transação ocorra se seu emissor não possuir ativos suficientes para completá-la.

Figura 2.1 – Diagramação de um bloco.



Fonte: Produção autoral

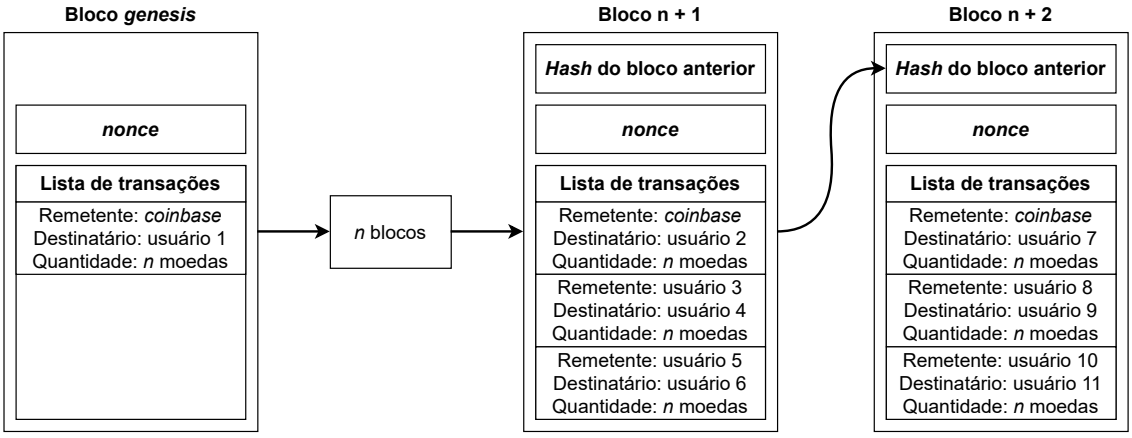
Na Figura 2.1 é mostrado um diagrama apresentando a organização de um bloco de informação. Este bloco possui uma referência ao hash do bloco anterior a ele na *blockchain*, assim como um valor *nonce* (*number only used once*), o qual será abordado mais aprofundadamente na subseção 2.1.4. Este bloco contém, também, uma lista de todas as transações registradas em um determinado período de tempo pelo *peer* que o validou, que, então, recebe um valor predeterminado por esta validação, conhecido como *coinbase*.

No entanto, tratando-se de um sistema distribuído, esta rede está vulnerável a ataques por um ou mais usuários maliciosos. Partindo do princípio que os diferentes nós conectados à *blockchain* terão acesso à diferentes transações, um usuário, com o intuito de gastar múltiplas vezes a mesma moeda, pode divulgar transações diferentes para *peers* distintos. Desse modo, com o intuito de garantir que todos os *peers* da rede contenham os mesmos registros de transação em ordem correta e contornar o problema apresentado, toda *blockchain* deve implementar um mecanismo de consenso entre seus nós para determinar quais transações serão registradas no próximo bloco inserido na cadeia, eliminando qualquer transação inválida, como a do exemplo.

É interessante adicionar que devido à implementação deste mecanismo de consenso, a tecnologia de *blockchain* propõe uma solução ao Problema do General Bizantino (LAMPORT, 1983). Uma vez que as transações são agrupadas em blocos e os nós participantes devem concordar com a autenticidade de qualquer bloco adicionado à *blockchain*, é impossível alterar ou excluir entradas já confirmadas nesta cadeia. O livro-caixa distribuído armazenado através da *blockchain* é, portanto, imutável (CACHIN; VUKOLI, 2017).

Para que uma *blockchain* seja iniciada, é preciso instanciar estaticamente um bloco inicial, conhecido como *genesis*, assim que um *peer* conectar-se a esta rede, de forma que qualquer *peer* que se conecte a partir deste momento contenha, em sua cadeia, a informação deste bloco inicial em comum. A partir disso, um bloco somente poderá ser anexado ao final desta *blockchain* uma vez que todos os *peers* da rede alcançarem um consenso sobre sua adição.

Figura 2.2 – Diagramação de um *blockchain*.



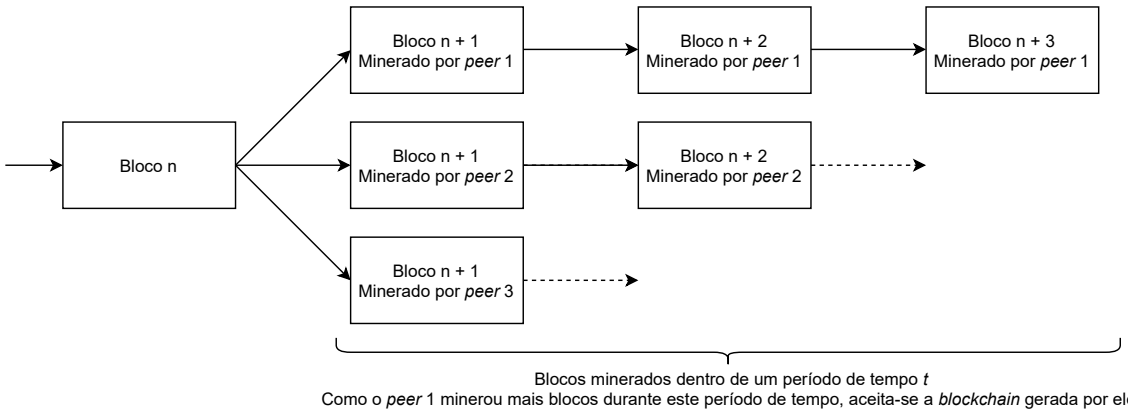
Fonte: Produção autoral

Na Figura 2.2 é mostrado um diagrama apresentando a organização de uma *blockchain*. Nesta cadeia, todos os blocos adicionados, com excessão do bloco *genesis*, devem conter uma referência ao *hash* do bloco anterior. Cada bloco contém, também, uma lista das transações realizadas no período de tempo em que ele esteve ativo, assim como um registro do *peer* que o validou através de sua transação *coinbase*.

2.1.4 Proof-of-Work

Para se obter o consenso sobre o próximo bloco a ser anexado na *blockchain*, no caso do *Bitcoin*, utiliza-se o algoritmo de *Proof-of-Work* (PoW). Neste algoritmo, os *peers* conectados à rede devem resolver um problema altamente custoso computacionalmente, denominado *nonce* (*number only used once*), onde o *peer* que alcançou uma solução primeiro determina o próximo bloco a ser adicionado. Caso dois ou mais nós consigam solucionar este problema simultaneamente, cada um deles divulga a sua versão da cadeia atual. Eles, então, competem entre si resolvendo os blocos seguintes. O *peer* que obtiver a cadeia mais longa entre os concorrentes tem a sua versão aceita e anexada ao final da *blockchain*. Este algoritmo é chamado de *Proof-of-Work* pois a versão da cadeia a ser adicionada é sempre aquela com a maior quantidade de trabalho computacional aplicada (NAKAMOTO, 2008).

Figura 2.3 – Diagramação de funcionamento de um algorítmo de consenso *Proof-of-Work* em uma *blockchain*.



Fonte: Produção autoral

Na Figura 2.3 é apresentado um diagrama mostrando o funcionamento de um algoritmo *Proof-of-Work* em uma *blockchain*. É proposto um cenário de competição, onde três *peers* disputam entre si para decidir qual versão da cadeia será anexada. Uma vez que o *peer* 1 validou mais blocos que os outros *peers*, ou seja, adicionou a maior quantidade de trabalho na cadeia, a versão da *blockchain* que ele gerou é aceita.

Um algoritmo de consenso *Proof-of-Work* também implica que, para que seja feita qualquer alteração em determinada *blockchain*, um *peer* ou um conglomerado de mineração mal intencionado qualquer deve poder resolver este problema altamente custoso computacionalmente de maneira mais rápida que o restante dos nós. Portanto, conclue-se que, a menos que algum agente fraudulento controle 50% ou mais do poder de proces-

samento uma *blockchain Proof-of-Work*, este não conseguirá fazer nenhuma alteração na mesma.

2.2 CRIPTOMOEDAS

Brevemente, criptomoedas são redes descentralizadas baseadas em *blockchain* que originam um livro-razão distribuído reforçado por uma rede heterogênea de computadores, devido à maneira a qual as informações sobre distribuição e criação de blocos são armazenadas. Uma das principais características que definem uma rede de criptomoedas é a ausência de uma autoridade central, como bancos e governos. Isso às torna, teoricamente, imunes à interferências e manipulações de poderes externos.

Utiliza-se o exemplo da *Bitcoin* (NAKAMOTO, 2008) para conceituar *blockchains*, PoW, e outros conhecimentos básicos para a fundamentação teórica deste texto. Esta criptomoeda encontra-se em circulação desde meados de 2008 e atualmente acumula um valor total de mercado de mais de 1 trilhão de dólares (TASKINSOY, 2021). Porém, além da Bitcoin, ainda existem diversas outras criptomoedas, apelidadas de *altcoins*. Devido ao escopo deste trabalho, apenas com o Bitcoin e as principais *altcoins* que implementam algoritmos de consenso *Proof-of-Work*, como a Bitcoin Cash[1], o Dash (DUFFIELD; DIAZ, 2014) o Ethereum (BUTERIN, 2013) e a Litecoin[2], que serão abordadas adiante.

2.2.1 Centralização em Redes de Criptomoedas

Na subseção 2.1.4, é comentado que para qualquer alteração ser realizada em determinada *blockchain*, um *peer* ou um conglomerado de mineração mal intencionado qualquer deve ser capaz de resolver um problema altamente custoso computacionalmente de maneira mais rápida que o restante dos nós conectados a essa rede. Disso, assume-se que a menos que algum agente fraudulento controle 50% ou mais do poder de processamento uma *blockchain Proof-of-Work*, este não conseguirá fazer nenhuma alteração na mesma (NAKAMOTO, 2008).

Com a crescente tendência de centralização de poder de processamento em grandes *pools* de mineração nas redes de Bitcoin e Ethereum (SAI et al., 2021), deve-se nos atentar aos possíveis riscos de segurança que esta progressiva centralização nessas e em outras redes de criptomoedas possa causar, como por exemplo, ataques de 51%, que sucessivamente levariam ao problema do *double-spending*. Estes dois pontos serão abordados nas subseções a seguir.

[1] Bitcoin Cash - https://bitcoincash.org/
[2] Litecoin - https://litecoin.org/pt/

2.2.2 O Ataque de 51%

O ataque de 51% é, em termos simples, uma vulnerabilidade em blockchains, a qual ocorre uma vez que um conglomerado de mineradores obtém 50% ou mais do poder de processamento de uma determinada rede. Esta falha de segurança permite que o agente desse ataque, estatisticamente, sempre obtenha a maior prova de trabalho computacional entre os *peers* da *blockchain*, e, por consequência, daria ao agente deste ataque o poder de evitar que novas transações sejam confirmadas (APONTE-NOVOA et al., 2021), permitindo-lhes interromper o sistema de pagamentos entre todos os usuários. Eles também poderiam de reverter transações próprias que foram concluídas enquanto eles estavam no controle da rede, o que poderia resultar no problema de *double-spending* (SAYEED; MARCO-GISBERT, 2019).

Em um ataque de 51%, embora o invasor possa desencadear o problema de *double-spending*, ele não pode reverter transações efetuadas por outros usuários ou impedí-los de transmitir suas transações para a rede. Além disso, um ataque de 51% é incapaz de criar novos ativos, roubar ativos de outros usuários ou organizações ou alterar a funcionalidade de recompensas por bloco.

2.2.3 Double-spending

O double-spending é, basicamente, um ataque de segurança em uma rede de criptomoedas que permite que uma moeda digital seja gasta duas vezes. Este problema é exclusivo deste meio, uma vez que as informações digitais podem ser reproduzidas facilmente por indivíduos com um conhecimento significativo em uma determinada rede de criptomoedas e o poder de computação necessário para manipulá-la.

Este ataque pode ocorrer quando um usuário 'A' realizar a transferência de uma determinada moeda para outro usuário 'B', mas não informar nenhum outro usuário da rede sobre essa transação. O usuário 'A', portanto, pode criar outra transação com a mesma moeda para outro usuário 'C' e esta ainda será válida sob as regras do sistema. Isso faz com que ambos os usuários recipientes da moeda de 'A' possuam uma mesma moeda, que foi gasta duas vezes pela mesma pessoa (ROSENFELD, 2014).

A maior probabilidade do problema do double-spending acontecer em um cenário real seria na forma de um ataque de 51% (SAYEED; MARCO-GISBERT, 2019). Se um agente controlar mais de 50% de uma *blockchain*, ele poderá realizar transações diversas vezes, revertendo as mesmas no livro-razão descentralizado mantido pela *blockchain* como se as transações iniciais nunca houvessem acontecido.

2.3 TRABALHOS RELACIONADOS

Beikverdi e Song (BEIKVERDI; SONG, 2015) propuseram uma análise quantitativa sobre a tendência de centralização em conglomerados de mineração na Bitcoin. Neste trabalho, os pesquisadores analisaram todos os blocos criados nesta *blockchain* entre 2009 e 2014, e, a partir do cálculo da proporção de uniformidade de uma determinada rede, quantificaram uma medida capaz de determinar o se índice de centralização. Quando este trabalho foi publicado, os autores encontraram um nível de centralização de mais de 30% na rede de Bitcoins, ou seja, aproximadamente um em cada três blocos era validado por um conglomerado de mineração.

Cong, He e Li (CONG; HE; LI, 2020) também realizaram uma análise quantitativa sobre a centralização em conglomerados de mineração na Bitcoin. Este trabalho, porém, propõe uma visão mais economicista sobre o assunto. É proposto, nesta pesquisa, que a centralização da rede em *pools* de mineração geram uma espécie de corrida armamentista tanto entre mineradores individuais quanto entre os conglomerados em questão, consequentemente aumentando o trabalho total aplicado à *blockchain* de redes de criptomoedas *Proof-of-Work*. Os autores desta pesquisa estimaram que nos dias atuais mais de 90% do poder de processamento da Bitcoin é controlada por conglomerados de mineração, ou seja, aproximadamente nove em cada dez blocos era validado por um conglomerado de mineração.

Li, Yang e Tessone (LI; YANG; TESSONE, 2020) também exploraram a questão da centralização em conglomerados de mineração nas redes de Bitcoin, Bitcoin Cash, Ethereum e Litecoin, que também serão abordadas no nosso trabalho. Esta pesquisa, porém, tem como foco a detecção de anomalias entre os *block times* esperados e os *block times* reais em blocos consecutivos destas redes de criptomoedas. É proposto, também, um método estatístico capaz de identificar comportamentos de mineração egoísta (EYAL; SIRER, 2014) a partir da detecção destas anomalias.

Lin, Li, Zhao e Chen (LIN et al., 2021), em seu trabalho, buscaram estimar o grau de descentralização das *blockchains* de Bitcoin e Ethereum durante o ano de 2019, calculando a distribuição do poder de mineração com três métricas diferentes: o coeficiente de Gini, a entropia de Shannon e o índice de Nakamoto. Esta pesquisa demonstrou, também, que observar *blockchains* entre "janelas irregulares" de tempo (por exemplo, entre duas semanas) pode revelar informações adicionais, que acabam sendo ignoradas pelas medições diárias, semanais, mensais e anuais, aumentando assim a eficácia de medições de descentralização em termos de tendências contínuas e situações anormais.

Caccioli, Livan e Aste (CACCIOLI; LIVAN; ASTE, 2016) investigaram, em seu trabalho, os limites físicos dos mecanismos de consenso em *blockchains*. Os autores também examinaram se existem razões de eficiência e escalabilidade que justifiquem a tendência à centralização nestas redes. Para isto, eles estimaram o tempo necessário para o alcance

do consenso da maioria entre todos os *peers*, comparando redes igualitárias com redes centralizadas, para assim, quantificar o efeito destas topologias de rede na propagação de informações.

Sai, Buckley, Fitzgerald e Le Gear (SAI et al., 2021) trazem, eu seu trabalho, uma revisão sistemática da literatura de 89 artigos científicos publicados entre 2009 e 2019 acerca da centralização em sistemas públicos de *blockchain*, inspirados por estudos recentes que demonstraram uma tendência crescente deste fenômeno nas redes de Bitcoin e Ethereum. Nesta pesquisa, os autores destacaram as múltiplas definições e medidas de centralização em *blockchains* presentes na literatura. Eles identificaram, também, diferentes aspectos sobre estas definições de centralização, propondo uma taxonomia abrangente de todos os possíveis problemas causados por este fenômeno.

Aponte-Novoa, Orozco, Villanueva-Polanco, e Wightman (APONTE-NOVOA et al., 2021), propuseram, em sua pesquisa, uma análise quantitativa sobre a centralização de poder de processamento em poucos conglomerados de mineração nas redes de Bitcoin e Ethereum. Este estudo concluiu que, durante o período de tempo estudado, apenas dezoito conglomerados de mineradores controlavam mais de 50% do poder computacional da rede de Bitcoin. Também foi observado que apenas treze *pools* de mineração em Ethereum detinham aproximadamente 75% de todos os blocos validados durante este período.

Através deste trabalho, será realizada uma análise quantitativa das redes de criptomoedas de Bitcoin, Bitcoin Cash, Ethereum, Litecoin e Dash, em busca de padrões de centralização de poder de processamento em conglomerados de mineração através de informações contidas nos blocos de cada uma das *blockchains* das criptomoedas escolhidas. Também foram aplicados índices medidores de desigualdade de distribuição nestes dados, para que seja possível quantificar e comparar o nível de centralização de poder computacional nestas redes.

3 METODOLOGIA

Para a execução desta monografia, a metodologia deste trabalho foi dividida em três etapas. Primeiro, foram definidas as criptomoedas a serem analisadas e as ferramentas que serão utilizadas para a obtenção dos dados sobre todos os blocos já adicionados às *blockchains* selecionadas dentro de um determinado período de tempo. Depois, foi necessário processar estes dados a fim de armazenar somente as informações relevantes para esta pesquisa. Por fim, foi feita a análise destes dados já processados, de forma a quantificar, através das métricas propostas nas próximas seções, a centralização das criptomoedas selecionadas.

3.1 CRIPTOMOEDAS ANALISADAS

Neste trabalho, foram analisadas quatro criptomoedas com a maior capitalização de mercado, e que implementam um mecanismo de consenso *Proof-of-Work*[1]. Também foi analisada uma criptomoeda menos financeiramente expressiva, a Dash, uma vez que ela apresentou resultados interessantes e relevantes ao objetivo desta monografia. A partir disso, buscou-se formar um comparativo entre as principais redes de criptomoedas com uma de menor adoção. Com estas restrições, então, pode-se definir as nossas redes de criptomoedas de estudo a partir da Tabela 3.1:

Tabela 3.1 – Informações gerais sobre as criptomoedas analisadas.

Criptomoeda	**Data de criação**	**Capitalização de mercado**
Bitcoin	03/01/2009	$1 trilhão
Bitcoin Cash	01/08/2017	$10 bilhões
Dash	18/01/2014	$200 milhões
Ethereum	30/07/2015	$500 bilhões
Litecoin	07/10/2011	$14 bilhões

3.2 FERRAMENTAS UTILIZADAS

Para investigar os novos blocos criados em cada *blockchain* das criptomoedas selecionadas para análise, foi utilizado uma *block explorer*, que nada mais é que uma API

[1] De acordo com o website CryptoSlate - https://cryptoslate.com/cryptos/proof-of-work/ (acesso em 30/11/2021).

especializada em prover informações completas sobre cada bloco presente em uma determinada cadeia. Para este trabalho, foi feito o uso da *block explorer* disponibilizada pela Blockchair[2]. Não somente ela disponibiliza informações sobre todas as criptomoedas selecionadas, como contém *endpoints* para o *download* do histórico de blocos já adicionados às *blockchains* das redes de criptomoedas analisadas[3].

3.3 OBTENÇÃO DOS DADOS NECESSÁRIOS

O primeiro passo deste trabalho foi a obtenção dos dados necessários para a análise das criptomoedas selecionadas. Para isso, utilizou-se o *endpoint* do *block explorer* da Blockchair comentado na seção anterior. Tendo o acesso a um histórico de todos os blocos das criptomoedas a serem analisadas, foi possível montar uma base de dados de tamanho flexível. Para esta análise, houve um empenho para gerar um *dataset* com um balanço entre a precisão dos dados obtidos e o desempenho da *database* gerada.

Vale ressaltar que cada uma destas redes de criptomoedas tem um *block time* (tempo de geração estimado por bloco) distinto devido às diferenças entre as implementações e os protocolos utilizados em cada uma delas. Os *block times* das redes de criptomoedas analisadas e os algorítmos de *hash* utilizados por cada uma delas[4], são apresentados na Tabela 3.2:

Tabela 3.2 – Informações técnicas sobre as criptomoedas analisadas.

Criptomoeda	**Algoritmo de *hash***	***Block time***
Bitcoin	SHA256	10 minutos
Bitcoin Cash	SHA256	10 minutos
Dash	X11	2 minutos e 30 segundos
Ethereum	Ethash	15 segundos
Litecoin	Scrypt	2 minutos e 30 segundos

Essa diferença entre os *block times* resulta em uma quantidade diferente de blocos na nossa base de dados, caso fosse feita a análise destas redes em um intervalo de tempo igual. Para garantir o balanço comentado no parágrafo anterior, deve-se analisar individualmente cada uma destas redes de modo que, para cada uma delas, se garanta um desempenho desejável e uma precisão de dados aceitável.

[2]Blockchair - https://blockchair.com/ (acesso em 30/11/2021).

[3]Blockchair *database dumps* - https://blockchair.com/dumps

[4]De acordo com o website CryptoSlate - https://cryptoslate.com/cryptos/proof-of-work/ (acesso em 30/11/2021).

3.4 PROCESSAMENTO DOS DADOS OBTIDOS

Após a obtenção dos dados desejados para este trabalho, foi necessário processá-los para facilitar a extração de informações sobre a centralização destas redes. Dada a natureza da *blockchain*, é impossível identificar o *peer* ou a *pool* que minerou um determinado bloco a menos que o minerador tenha escolhido se identificar. Porém, é comum, por exemplo, que conglomerados de mineração, ou até mesmo nós de alto poder computacional, identifiquem-se através de uma mensagem ou assinatura compartilhada na transação base (*coinbase*) de um bloco.

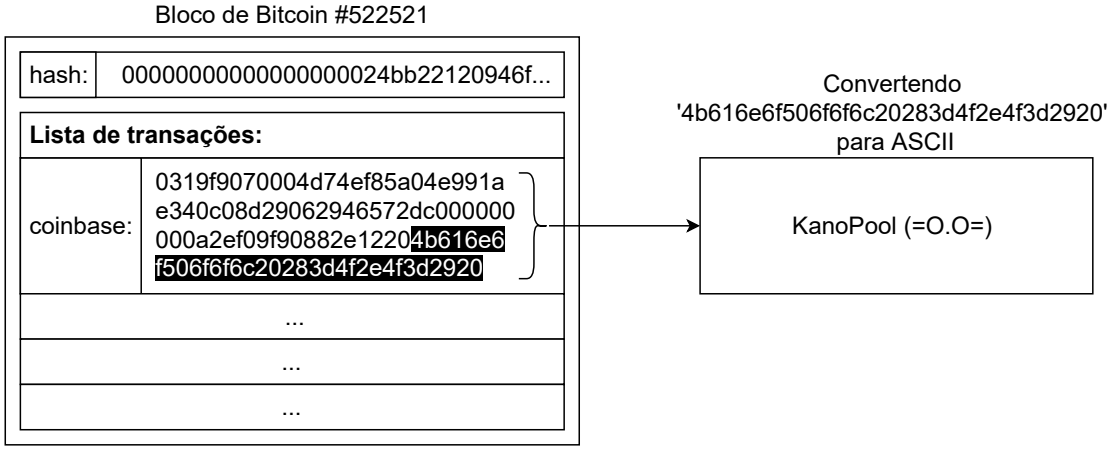


Figura 3.1 – Diagramação da identificação da *pool* de mineração responsável pela mineração de um bloco a partir de sua transação *coinbase*.

A partir da *coinbase* de um determinado bloco, portanto, foi possível determinar, com relativa precisão, quem foi responsável por minerá-lo. Aplicando isso em todos os blocos adicionados a nossa base de dados, é possível assumir qual a porcentagem de blocos minerados em um determinado período de tempo por um *peer* ou uma *pool* conectada à *blockchain* de uma das redes de criptomoedas analisadas.

3.5 ANÁLISE DE PADRÕES DE CENTRALIZAÇÃO

Uma vez processados estes dados, foi feita a análise dos resultados obtidos. É possível determinar o nível de centralização em uma rede de criptomoedas uma vez que na seção anterior foi obtida a porcentagem dos blocos gerados por diversos *peers* com alto poder de processamento e conglomerados de mineração em um determinado período de tempo. A partir destes dados, foram utilizadas três equações para determinar o nível de centralização de uma determinada rede: os índices de Gini, Nakamoto (LIN et al., 2021) e Theil (LI; YANG; TESSONE, 2020).

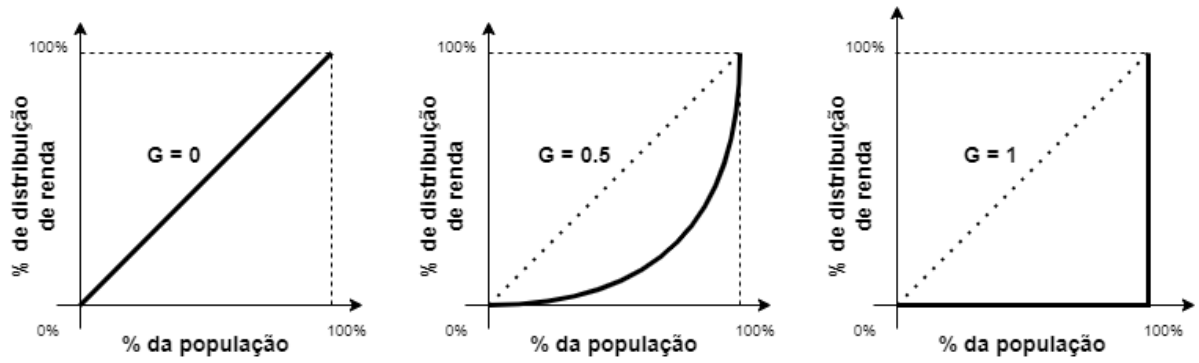
3.5.1 Coeficiente de Gini e a Curva de Lorenz

O coeficiente de Gini é frequentemente usado como um medidor de desigualdade econômica, medindo a distribuição da riqueza entre uma população. Em um cenário de medição da descentralização de uma *blockchain*, pode-se utilizar este índice para indicar a centralização do poder de mineração entre os *peers* e conglomerados de mineração conectados à rede (KWON et al., 2019b).

Para facilitar o entendimento desta métrica, é preciso compreender, primeiro, como o coeficiente de Gini se relaciona com o conceito de curva de Lorenz. Esta curva, no nosso contexto, é dada pela relação entre a porcentagem de *peers* com um poder de processamento menor ou igual a um determinado valor e o poder de processamento total ibservado em uma *blockchain*.

Na Figura 3.2, demonstra-se como a curva de Lorenz relaciona-se com o coeficiente de Gini. Quanto mais a curva diverge de uma linha reta, o coeficiente de Gini aumenta de 0 para 1. A curva de Lorenz, neste contexto, representa a distribuição de renda entre uma determinada população. Intuitivamente, quanto mais próxima de uma linha reta é essa distribuição, mais o coeficiente de Gini tende para 0. No entanto, a medida que a curvatura aumenta, mais este coeficiente tende para 1 (SRINIVASAN; LEE, 2017).

Figura 3.2 – Curva de Lorenz em relação ao coeficiente de Gini (G).



Fonte: Produção autoral

O coeficiente de Gini, representado por G na equação 3.1, pode ser calculado da seguinte forma:

$$G = \frac{\sum_{i,j \in A} |NB_i - NB_j|}{2|A| \sum_{NB_j \in NB} NB_j} \tag{3.1}$$

Onde NB_i refere-se ao número de blocos validados por um *peer* i, $NB = \{NB_i | i \in A\}$ determina a porcentagem de *peers* com um poder de processamento de até um determinado valor e A o poder de processamento total em determinado período de tempo (LIN et al., 2021).

3.5.2 Coeficiente de Theil

O coeficiente de Theil é uma métrica baseada em entropia que expressa, também, uma medida de desigualdade em um determinado sistema (LI; YANG; TESSONE, 2020). Os índices de Gini e Theil se comportam de maneira muito semelhante, diferenciando-se na faixa de valores que eles podem apresentar (enquanto o coeficiente de Gini é representado com um valor entre 0 e 1, o coeficiente de Theil é representado com um valor entre 0 e $ln(N)$, sendo N o número total de *peers* em uma *blockchain*). O valor factual do coeficiente de Theil, porém, não é tão importante, uma vez que o nível de entropia em que esta métrica é baseada é relativo por definição (IVANITSKIY, 2019).

O coeficiente de Theil, representado por T na equação 3.2, pode ser calculado da seguinte forma:

$$T = \frac{1}{N} \sum_{i=1}^{N} ln(\frac{\mu}{x_i}) \tag{3.2}$$

Onde x_i representa o poder de processamento total de N *peers*, e μ representa o poder total de processamento de uma *blockchain*. Nesta equação, se todos os nós de uma rede possuirem o mesmo poder de processamento, o coeficiente de Theil resulta em 0. Agora, se apenas um *peer* desta rede detiver todo este poder, esta equação resulta em $ln(N)$.

3.5.3 Coeficiente de Nakamoto

O coeficiente de Nakamoto estabelece a quantidade mínima de *peers* ou conglomerados de mineração necessários em uma *blockchain* para reunir mais de 51% do poder de processamento deste sistema (SRINIVASAN; LEE, 2017). Intuitivamente, quanto maior o coeficiente de Nakamoto, mais *peers* precisam unir o seu poder de processamento de modo a causar um ataque de 51%. (LIN et al., 2021).

O coeficiente de Nakamoto, representado por N na equação 3.3, pode ser calculado através da seguinte equação:

$$N = min\{k \in [1, ..., k] : \sum_{i=1}^{k} p_i \geq 0.51\}] \tag{3.3}$$

Onde k define a quantidade de *peers* em uma *blockchain*, e $p_1 > ... > p_K$ sejam a proporção controlada por cada um destes K nós em uma determinada *blockchain*, de modo que a soma de todas as proporções seja 1 (SRINIVASAN; LEE, 2017).

3.5.4 Considerações Adicionais

A partir do processamento dos dados obtidos, e da aplicação das métricas discutidas para quantificar a centralização em uma determinada rede de criptomoedas, foi possível atigir o objetivo deste trabalho. Assim, não só foi possível quantificar a porcentagem do poder de processamento que determinados *peers* ou conglomerados de mineração possuem na rede, como também quantificar esta centralização com métricas mais sofisticadas capazes de acessar melhor o risco de um ataque de 51% nestas redes de criptomoedas.

4 EXPERIMENTOS REALIZADOS

Neste capítulo, serão apresentados os dados resultantes da implementação proposta no capítulo anterior. Todos as informações foram obtidas como descrito na seção 4.3, e apresentam, neste conjunto, dados referentes à totalidade dos blocos minerados entre os dias 01/01/2020 e 31/12/2021 para todas as criptomoedas escolhidas, as quais serão analisadas por ordem decrescente de capitalização de mercado. A implementação do restante deste projeto seguiu o padrão proposto na seção 4.4 e 4.5.

4.1 BITCOIN

Bitcoin, por ser a primeira criptomoeda criada e o maior ativo digital atualmente[1], causaria, tanto financeiramente quanto culturalmente, o maior impacto entre todas as criptomoedas analisadas caso fosse vítima de um ataque de 51%. Sendo assim, utilizaremos essa moeda como base de comparação, uma vez que, entre todas as criptomoedas selecionadas, ela é, atualmente, a mais consolidada.

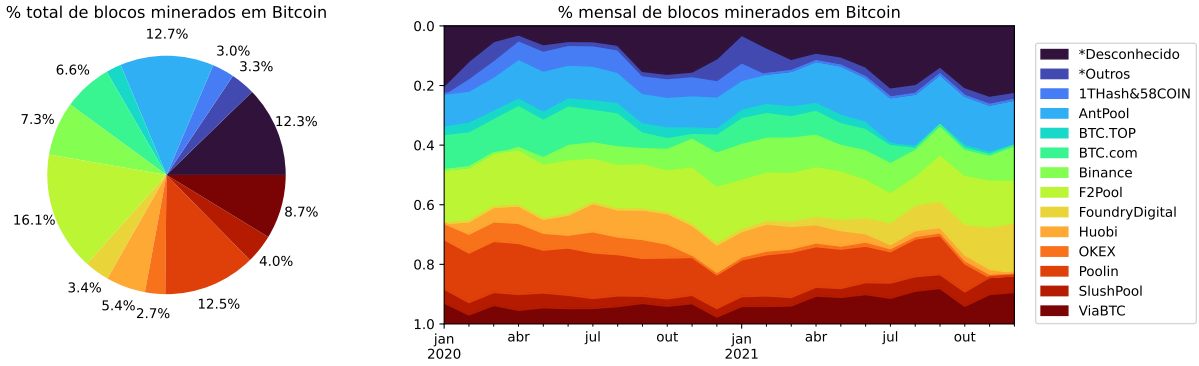


Figura 4.1 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Bitcoin.

Na Figura 4.1, apresentamos a distribuição mensal e total dos blocos minerados em Bitcoin. Os valores em "Desconhecido" tratam-se de endereços que não podem ser traçados à uma entidade de mineração conhecida, enquanto os valores em "Outros" representam conglomerados menos significantes, com menos blocos minerados do que aqueles que foram declarados. Através dessas imagens, podemos perceber que o Bitcoin possui uma concentração de poder de mineração muito dispersa. Ao final de 2021, mais de 20% de todos os blocos minerados nessa rede provinham de mineradores possivelmente não

[1] De acordo com o website CryptoSlate - https://cryptoslate.com/cryptos/proof-of-work/ (acesso em 29/01/2022).

atrelados à nenhuma entidade, enquanto grande parte dos conglomerados de mineração haviam processado entre 10% e 15% dos blocos gerados durante este período.

Podemos ver, também, durante o período analisado, que a validação de blocos na *blockchain* de Bitcoin concentrou-se, principalmente, entre três conglomerados: AntPool, F2Pool e Poolin. Estas *pools* foram responsáveis pela mineração de 41,3% de todos os blocos inseridos durante este período. Para comparação, o resto dos conglomerados desta criptomoeda validaram 44,4% de todos estes blocos.

4.2 BITCOIN CASH

Bitcoin Cash é uma implementação de criptomoeda baseada no código fonte original do Bitcoin. De forma a aumentar a escalabilidade e permitir que mais transações sejam realizadas em comparação à versão original, essa implementação aumentou o tamanho máximo dos seus blocos, de 1MB para 32MB, e definiu um tamanho mínimo de 8MB por bloco (JAVARONE; WRIGHT, 2018). Apesar destas diferenças, a implementação do sistema de mineração do Bitcoin e do Bitcoin Cash continuou a mesma, o que permitiu que os mineradores dessas duas criptomoedas pudessem alternar o uso do seu poder de processamento entre elas (KWON et al., 2019a).

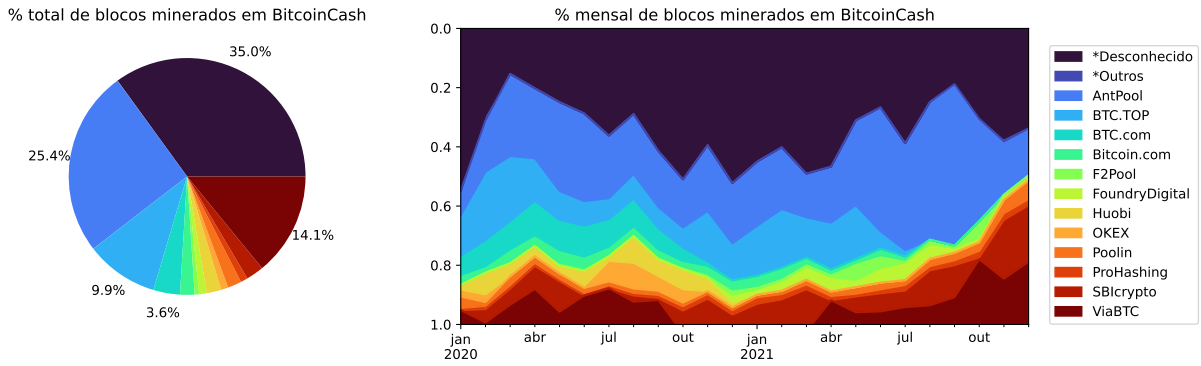


Figura 4.2 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Bitcoin Cash.

Na Figura 4.2, podemos visualizar a distribuição mensal e total dos blocos minerados em Bitcoin Cash. Os valores em "Desconhecido" e os valores em "Outros" seguem o mesmo esquema de representação usado na Figura 4.1. Pode-se perceber que os conglomerados de mineração do Bitcoin Cash passam por vários altos e baixos com suas porcentagens de blocos minerados durante cada mês. Isso ocorre devido à facilidade do redirecionamento do poder computacional dos mineradores de Bitcoin Cash para a rede de Bitcoin, uma vez que ambas implementam a mesma função *hash* criptográfica para validação de seus blocos (KWON et al., 2019a). É notável, também, que apesar de existir

uma grande concentração de poder de processamento em alguns destes conglomerados durante quase todo o período analisado (Antpool, nomeadamente, gerou mais de 40% do total de blocos inseridos na *blockchain* de Bitcoin Cash no mês de setembro de 2021), grande parte dos blocos foram gerados por mineradores desconhecidos, o que é um bom indicador de descentralização.

Mesmo que 35% de seus blocos ainda sejam validados por usuários desconhecidos, o Bitcoin Cash ainda apresenta uma significativa centralização em dois conglomerados de mineração: a AntPool e a ViaBTC, que juntos validaram 39,5% de todos os blocos inseridos na *blockchain* dessa criptomoeda durante o período de tempo analisado. Juntando seu poder de processamento com uma *pool* de menor expressão, como a BTC.TOP, faria com que estes três conglomerados atingissem 49,4% de todos os blocos processados nessa rede.

4.3 DASH

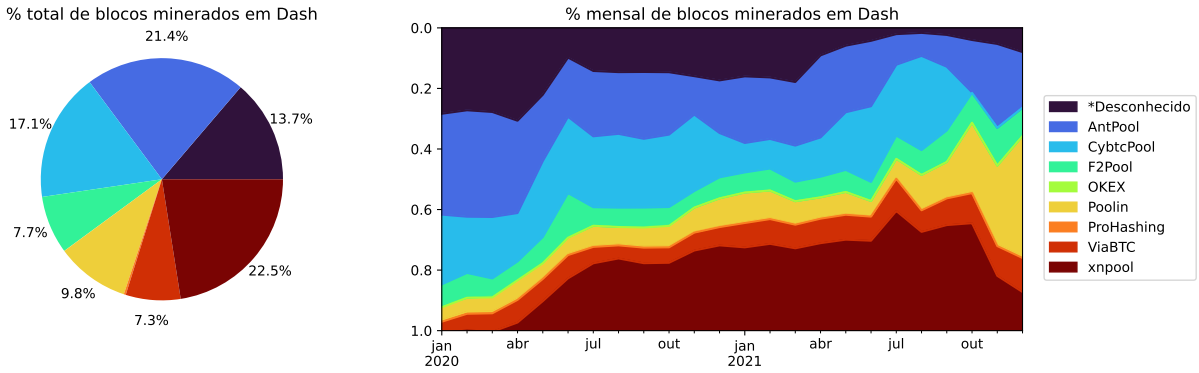


Figura 4.3 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Dash.

Na Figura 4.3, podemos visualizar a distribuição mensal e total dos blocos minerados em Dash. Os valores em "Desconhecido" seguem o mesmo esquema de representação usado na Figura 4.1, porém, como esta criptomoeda não possui um número significativo de conglomerados de mineração ativos, não foi necessário incluir valores em "Outros", com *pools* de menor expressão. A partir destes dados, podemos perceber que, apesar de não existirem muitos conglomerados nesta rede de criptomoedas, entre o grupo selecionado, destacam-se a AntPool, a CybtcPool, e, ao final de 2021, Poolin, que processou aproximadamente 40% dos blocos adicionados à *blockchain* da Dash nesse período.

Dash demonstra uma grande concentração do seu poder total de mineração na mão de apenas três conglomerados de mineração. Durante o período analisado, 61% de todos os blocos inseridos em sua *blockchain* foram processados por uma destas três *pools*:

AntPool, CybtcPool e xnpool. O restante destes conglomerados, juntos, mineraram apenas 24,8% destes blocos.

4.4 ETHEREUM

Ethereum não é somente uma rede de criptomoedas, mas também uma plataforma baseada em *blockchain* para criação de aplicativos decentralizados. Esta rede armazena e valida contratos digitais, os quais permitem que transações e acordos confiáveis sejam realizados entre partes anônimas e díspares sem a necessidade de uma autoridade central, sistema legal ou mecanismo externo de fiscalização (BUTERIN, 2013). Mesmo assim, ainda é possível realizar operações financeiras normalmente na rede de Ethereum, utilizando a criptomoeda Ether. Nesta seção, será proposta uma análise a partir dos dados obtidos para a rede de Ethereum.

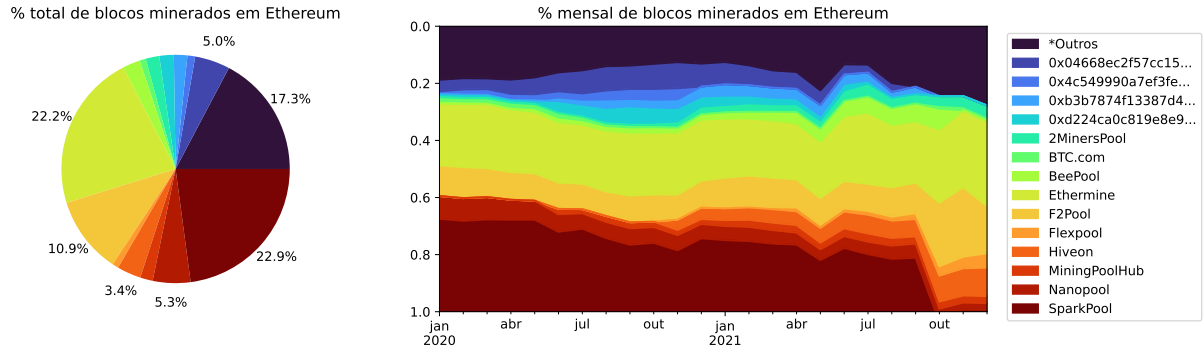


Figura 4.4 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Ethereum.

Na Figura 4.4, podemos visualizar a distribuição mensal e total dos blocos minerados em Ethereum. Nessa rede, ao contrário das outras que iremos analisar, não há nenhum bloco processado por um minerador desconhecido, uma vez que uma referência ao usuário que o minerou é armazenada, na forma de um *hash*, em todo bloco inserido na *blockchain* do Ethereum.

Há uma concentração de poder de processamento, principalmente, em três conglomerados: SparkPool, Ethermine e F2Pool. Pode-se perceber, também, em meados de setembro de 2021, uma queda enorme de poder de processamento da SparkPool, a qual minerou mais blocos durante o período registrado. Isso coincide, temporalmente, com o banimento da compra e venda de critpomoedas na China[2], onde essa *pool* era sediada.

O Ethereum, em comparação ao Bitcoin, apresenta uma concentração muito maior

[2]Banco Central da China declara ilegais todas as transações com criptomoedas - https://g1.globo.com/economia/noticia/2021/09/24/banco-central-da-china-declara-ilegais-todas-as-transacoes-com-criptomoedas.ghtml (acesso em 30/01/2022.

de poder de mineração em poucos conglomerados. Durante o período analisado, somente duas *pools*, a SparkPool e a Ethermine, processaram 45,1% de todos os blocos inseridos na *blockchain*. Isso é maior que a soma de todos os blocos minerados pelo resto dos conglomerados de mineração identificados, que validaram um 37,6% deles.

4.5 LITECOIN

Litecoin é uma criptomoeda *open-source*, considerada uma das primeiras moedas alternativas ao Bitcoin (*altcoins*), uma vez que deriva-se do código fonte desta rede. O Litecoin difere do Bitcoin em dois aspectos: sua taxa de geração de blocos mais rápida e o uso de *Scrypt* ao invés de SHA-256 como função *hash* criptográfica para o seu algoritmo de *Proof-of-Work*. Nesta seção, será proposta uma análise a partir dos dados obtidos para a rede de Litecoin.

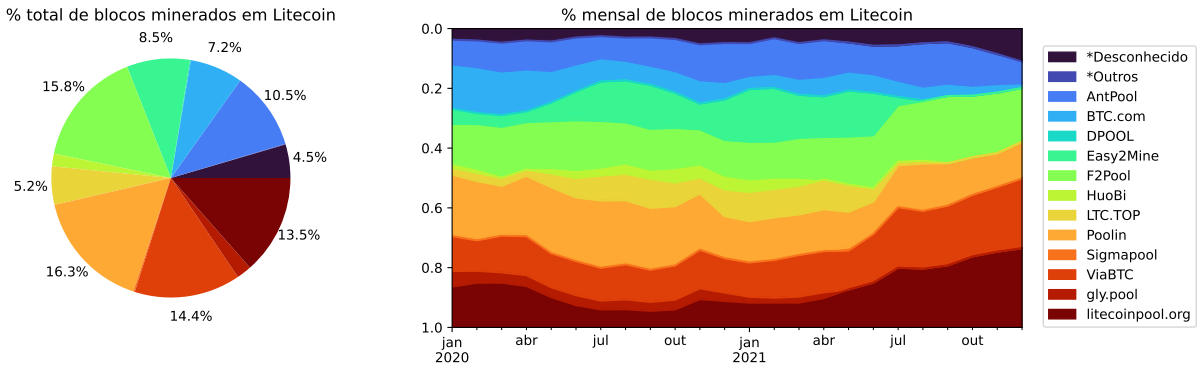


Figura 4.5 – Distribuição percentual mensal e total de blocos minerados por *pools* de mineração em Litecoin.

Na Figura 4.5, podemos visualizar a distribuição mensal e total dos blocos minerados em Litecoin. Os valores em "Desconhecido" e os valores em "Outros" seguem o mesmo esquema de representação usado na Figura 4.1. É visível, nesta rede, que a maioria do seu poder de processamento provém de conglomerados de mineração. Durante quase todo o período registrado, menos de 5% dos blocos foram minerados por usuários desconhecidos, subindo para aproximadamente 10% ao final deste mesmo período. Apesar da quantidade de blocos minerada pelas *pools* apresentadas ser bem distribuída, há, nos últimos 6 meses desses dados, um crescimento significativo de três conglomerados: F2Pool, gly.pool e litecoinpool.org. Outros conglomerados, como Easy2Mine e LTC.TOP apresentam uma diminuição significativa durante este mesmo período. Apesar do Litecoin ter seu poder de mineração mais uniformemente distribuído, ainda pode-se perceber uma concentração entre quatro *pools* de mineração: F2Pool, Poolin, Sigmapool, e litecoinpool.org.

4.6 INDEXAÇÃO ATRAVÉS DOS DADOS OBTIDOS

Para completarmos a análise proposta neste trabalho, devemos quantificar mais precisamente a concentração de poder computacional presente em cada uma das criptomoedas escolhidas. Para isso, nesta seção, aplicaremos as equações discutidas na seção 4.5 às informações já apresentadas neste capítulo, de forma a quantificar o nível de centralização de uma rede de criptomoedas.

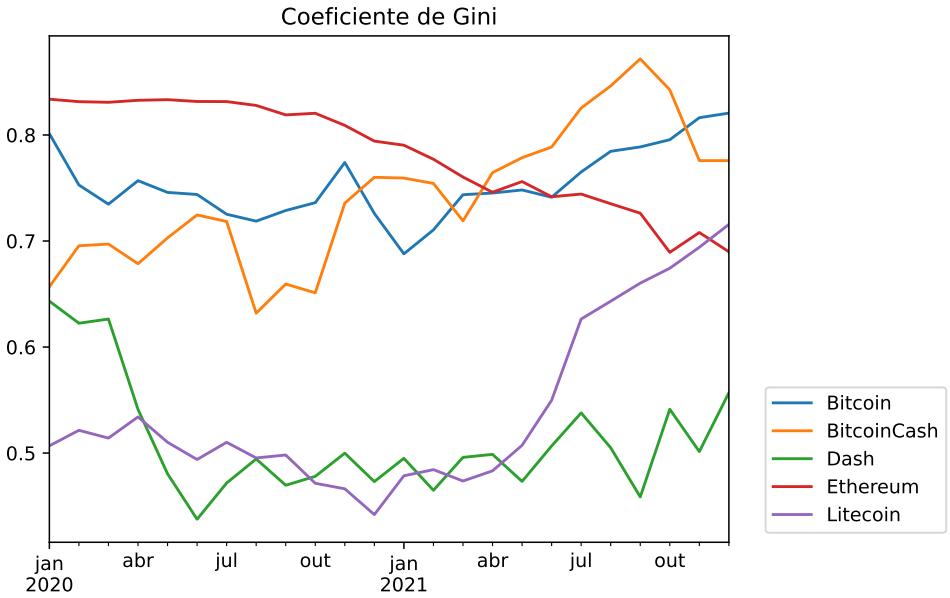


Figura 4.6 – Evolução mensal do Coeficiente de Gini das criptomoedas analisadas.

Na Figura 4.6, observamos a evolução do índice de Gini mensalmente para as cinco criptomoedas analisadas. Podemos notar, através deste coeficiente, que redes com grande disparidade de poder de processamento entre seus conglomerados de mineração acabam obtendo um índice de Gini maior. A grande subida destes valores no final de 2021 para Litecoin comprova isso, uma vez que, durante esse período, esta criptomoeda viu um aumento da concentração do seu poder computacional em menos *pools* de mineração.

Essa crescente também pode ser vista nas redes de Bitcoin e Bitcoin Cash. Mesmo que a maior parte da capacidade de processamento do Bitcoin seja espalhada de forma uniforme entre os seus maiores conglomerados de mineração, uma porcentagem muito ínfima de blocos é processada por mineradores individuais, fazendo com que a disparidade de poder de processamento calculada pelo índice de Gini seja alta. Já o Bitcoin Cash, mesmo possuindo a maior porcentagem de mineradores desconhecidos entre todas moedas analisadas, acaba concentrando grande parte do seu poder de processamento em duas *pools* de mineração (Antpool e ViaBTC), o que também contribui para o aumento desse índice.

Ethereum e Dash, porém, obtiveram uma queda neste índice. Para Ethereum, isso aconteceu devido a brusca diminuição do poder de processamento de conglomerados de

mineração como o SparkPool, que, por vários meses, validou de 25% à 30% de todos os blocos inseridos na *blockchain* de Ethereum. Essa descida permitiu que outras *pools* de menor expressão conseguissem minerar mais blocos, distribuindo dessa criptomoeda. Para Dash, a ascensão de conglomerados de mineração como xnpool e a Poolin distribuíram de forma mais uniforme o seu poder de processamento, o que explica o expressivo declínio em seu índice de Gini em abril de 2020.

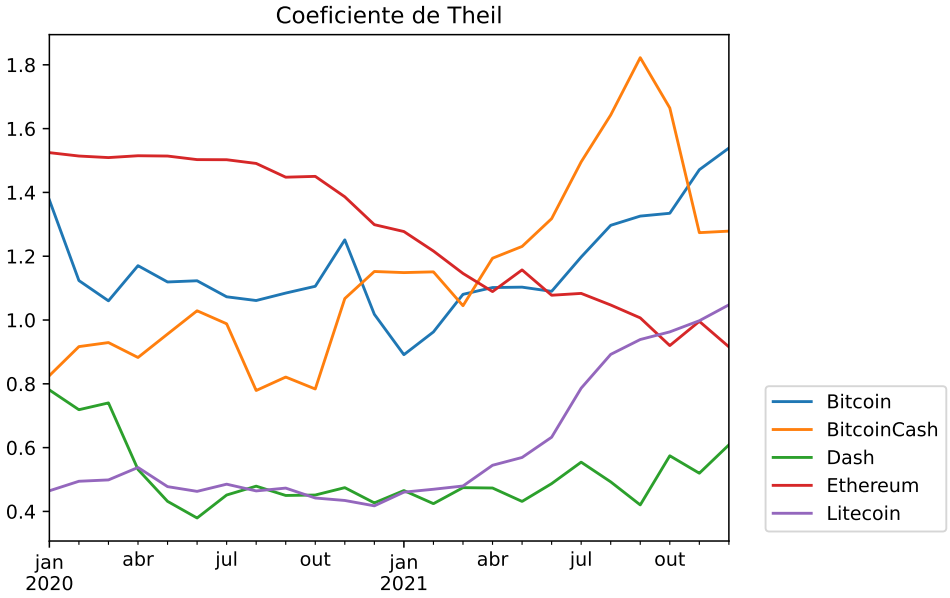


Figura 4.7 – Evolução mensal do Coeficiente de Theil das criptomoedas analisadas.

Na Figura 4.7, observamos a evolução do índice de Theil mensalmente para as cinco criptomoedas analisadas. Este índice atua de forma semelhante ao índice de Gini, porém, comporta-se de maneira mais sensível à distribuição irregular de poder de processamento entre conglomerados (LI; YANG; TESSONE, 2020). Podemos perceber essa diferença ao analisar o Bitcoin Cash, uma vez que, a medida que as suas parcela de blocos minerados vão se concentrando em menos mineradores, seu coeficiente de Theil tende a crescer de forma mais drástica.

O mesmo pode ser percebido ao observar o comportamento do Ethereum, que apresenta uma queda maior que a vista no gráfico do índice de Gini. Uma vez que o poder de processamento nessa criptomoeda começou a se apresentar de uma maneira mais uniformemente espalhada, este índice desceu drasticamente. Criptomoedas como o Bitcoin, o Dash e o Litecoin não apresentaram uma mudança tão brusca em relação à sua representação na Figura 3.1. Isso ocorre porque estas redes de criptomoedas apresentam uma distribuição relativamente uniforme de poder de mineração entre alguns conglomerados durante todo o período analisado. Portanto, a partir da comparação entre estas métricas, pode-se dizer que os coeficientes de Theil de uma criptomoeda tendem a ter uma variação maior do que os seus índices de Gini dependendo do nível de distribuição do poder de processamento em suas redes.

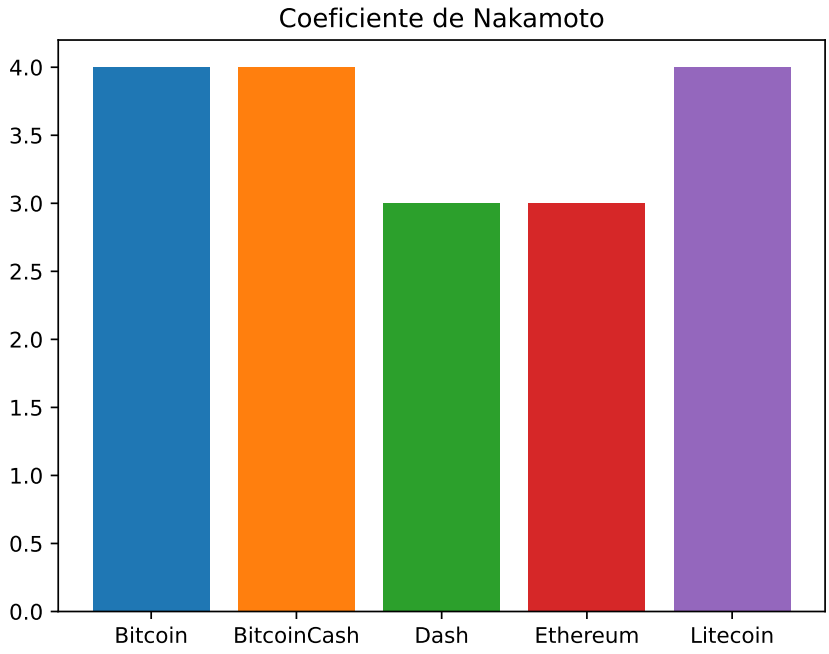


Figura 4.8 – Coeficiente de Nakamoto das criptomoedas analisadas durante o período estipulado.

Na Figura 4.8, observamos o índice de Nakamoto para as cinco criptomoedas analisadas durante o período completo do nosso conjunto de dados. Este coeficiente, de forma simples, indica quantos dos maiores conglomerados de mineração de uma rede devem unir seu poder de processamento em uma só *pool* para que um seja possível realizar um ataque de 51% bem sucedido. A partir deste índice, podemos inferir que, no caso do Bitcoin, Bitcoin Cash e Litecoin, seria necessário que quatro de suas maiores *pools* unam seu poder de processamento para que um ataque de 51% aconteça. No caso do Dash e do Ethereum, só seriam necessário unir a capacidade computacional de três destes conglomerados.

5 CONSIDERAÇÕES FINAIS

A manutenção da descentralização em redes de criptomoedas garante que qualquer pessoa possa desfrutar facilmente de um sistema monetário anônimo, seguro, e sem interferência de nenhuma autoridade central. Em corroboração com este ideal, esta monografia propôe uma análise quantitativa de padrões de centralização de poder de processamento nas redes de Bitcoin, Bitcoin Cash, Dash, Ethereum e Litecoin.

Por meio dos dados obtidos, obteve-se um panorama atualizado sobre a distribuição de poder computacional entre conglomerados de mineração nestas criptomoedas. Nota-se, principalmente, a enorme parcela de capacidade computacional agregada somente por *pools* de mineração nestas redes. É possível perceber, ainda, que muitos destes conglomerados não limitam sua atuação à somente uma criptomoeda. Algumas entidades como Antpool, Poolin e ViaBTC mostram-se entre as maiores validadoras de blocos em quatro das cinco redes estudadas.

Este trabalho também propôs uma análise da tendência de centralização entre estas criptomoedas por meio de métricas medidoras de desigualdade. Utilizando os índices de Gini, Theil e Nakamoto, pôde-se quantificar mais precisamente a distribuição de poder de processamento entre as *pools* de mineração identificadas, e mostrar, de forma mais precisa, a disparidade nestes sistemas.

Por fim, pôde-se perceber através dos dados obtidos neste trabalho que um ataque de 51% partindo de um único conglomerado de mineração, ao contrário do que foi inicialmente proposto, tem uma probabilidade muito remota de acontecer. Em todas as redes estudadas, houve poucos casos onde uma *pool* deteu mais de 20% do poder de processamento de uma criptomoeda durante o periodo de tempo observado.

Porém, assim como foi demonstrado a partir dos coeficientes de Nakamoto de cada uma das criptomoedas analisadas, há, ainda, uma possibilidade deste ataque ocorrer caso grandes conglomerados de mineração decidam agir em conluio para executá-lo. Esta ação em conjunto, ao contrário de um ataque proveniente de uma única *pool*, ainda demonstra um perigo real para as redes estudadas neste trabalho.

5.1 TRABALHOS FUTUROS

Há algumas limitações no conjunto de dados provido pela API utilizada, destacada na seção 4.3. Especificamente, algumas *pools* de mineração mais novas, que constam em outros *block explorers*, não são possíveis de identificar através da transação *coinbase* dos blocos destas redes de criptomoedas. Por isso, para termos um escopo completo sobre dados de centralização, seria necessário utilizar uma base de dados mais precisa,

que providenciasse a informação de todas as transações *coinbase*, ou contasse com a informação de todos os possíveis conglomerado de mineração responsáveis por minerar um bloco.

Também é possível expandir este trabalho investigando outras criptomoedas que utilizam algoritmos de consenso *Proof-of-Work*, uma vez que a API utilizada também não apresentava dados para várias novas moedas digitais em ascensão. Um estudo comparativo entre redes mais consolidadas, como Bitcoin e Ethereum, e novas redes de criptomoedas emergentes, principalmente durante a pandemia de COVID-19 (CORBET et al., 2020), poderia trazer uma visão interessante sobre o mercado atual destes ativos digitais.

Este trabalho aborda centralização de uma rede de criptomoedas somente no âmbito da concentração de poder de processamento e geração de blocos em conglomerados de mineração. Há possibilidade de expandí-lo tratando de fatores como concentração de renda ou volume de transações, como já abordado por Srinivasan e Lee (SRINIVASAN; LEE, 2017) para as redes de Bitcoin e Ethereum.

Além disso, pode-se, ainda, expandir sobre este trabalho ao analisar se os níveis de centralização destas redes dependem de outros fatores envolvendo estas criptomoedas. Por exemplo, é possível propor uma análise comparativa sobre a tendência de centralização entre redes criadas recentemente com outras ativas há mais tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

APONTE-NOVOA, F. et al. The 51% attack on blockchains: A mining behavior study. **IEEE Access**, v. 9, p. 140549–140564, 2021.

BEIKVERDI, A.; SONG, J. Trend of centralization in bitcoin's distributed network. In: **2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)**. [S.l.: s.n.], 2015. p. 1–6.

BUTERIN, V. Ethereum white paper: A next generation smart contract & decentralized application platform. 2013. Disponível em: <https://ethereum.org/en/whitepaper/>.

CACCIOLI, F.; LIVAN, G.; ASTE, T. Scalability and egalitarianism in peer-to-peer networks. In: **Banking Beyond Banks and Money: A Guide to Banking Services in the Twenty-First Century**. [S.l.]: Springer International Publishing, 2016. p. 197–212. ISBN 978-3-319-42448-4.

CACHIN, C.; VUKOLI, M. **Blockchain Consensus Protocols in the Wild**. 2017.

CONG, L.; HE, Z.; LI, J. Decentralized mining in centralized pools. **The Review of Financial Studies**, v. 34, n. 3, p. 1191–1235, 04 2020. ISSN 0893-9454. Disponível em: <https://doi.org/10.1093/rfs/hhaa040>.

CORBET, S. et al. Any port in a storm: Cryptocurrency safe-havens during the covid-19 pandemic. **Economics Letters**, v. 194, p. 109377, 2020. ISSN 0165-1765. Disponível em: <https://www.sciencedirect.com/science/article/pii/S016517652030238X>.

DUFFIELD, E.; DIAZ, D. Dash: A payments-focused cryptocurrency. 2014. Disponível em: <https://github.com/dashpay/dash/wiki/Whitepaper>.

EYAL, I.; SIRER, E. Majority is not enough: Bitcoin mining is vulnerable. In: CHRISTIN, N.; SAFAVI-NAINI, R. (Ed.). **Financial Cryptography and Data Security**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. p. 436–454. ISBN 978-3-662-45472-5.

IVANITSKIY, I. **Distribution of wealth in Cryptocurrencies**. 2019. Disponível em: <https://blog.parsiq.net/distribution-of-wealth-in-cryptocurrencies/>.

JAVARONE, M.; WRIGHT, C. From bitcoin to bitcoin cash: A network analysis. In: **Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems**. New York, NY, USA: Association for Computing Machinery, 2018. (CryBlock'18), p. 7781. ISBN 9781450358385. Disponível em: <https://doi.org/10.1145/3211933.3211947>.

KWON, Y. et al. Bitcoin vs. bitcoin cash: Coexistence or downfall of bitcoin cash? In: **2019 IEEE Symposium on Security and Privacy (SP)**. [S.l.: s.n.], 2019. p. 935–951.

____. Impossibility of full decentralization in permissionless blockchains. In: **Proceedings of the 1st ACM Conference on Advances in Financial Technologies**. New York, NY, USA: Association for Computing Machinery, 2019. (AFT '19), p. 110123. ISBN 9781450367325. Disponível em: <https://doi.org/10.1145/3318041.3355463>.

LAMPORT, L. The weak byzantine generals problem. **J. ACM**, Association for Computing Machinery, New York, NY, USA, v. 30, n. 3, p. 668676, jul 1983. ISSN 0004-5411. Disponível em: <https://doi.org/10.1145/2402.322398>.

LI, S.; YANG, Z.; TESSONE, C. Proof-of-work cryptocurrency mining: a statistical approach to fairness. In: **2020 IEEE/CIC International Conference on Communications in China (ICCC Workshops)**. [S.l.: s.n.], 2020. p. 156–161.

LIN, Q. et al. Measuring decentralization in bitcoin and ethereum using multiple metrics and granularities. In: **2021 IEEE 37th International Conference on Data Engineering Workshops (ICDEW)**. [S.l.: s.n.], 2021. p. 80–87.

NAKAMOTO, S. Bitcoin: A peer-to-peer electronic cash system. 2008. Disponível em: <https://bitcoin.org/bitcoin.pdf>.

ROSENFELD, M. Analysis of hashrate-based double spending. **CoRR**, abs/1402.2009, 2014. Disponível em: <http://arxiv.org/abs/1402.2009>.

SAI, A. R. et al. Taxonomy of centralization in public blockchain systems: A systematic literature review. **Information Processing & Management**, v. 58, n. 4, 2021. ISSN 0306-4573.

SAYEED, S.; MARCO-GISBERT, H. Assessing blockchain consensus and security mechanisms against the 51% attack. **Applied Sciences**, v. 9, n. 9, 2019. ISSN 2076-3417. Disponível em: <https://www.mdpi.com/2076-3417/9/9/1788>.

SRINIVASAN, B.; LEE, L. **Quantifying Decentralization**. 2017. Disponível em: <https://news.earn.com/quantifying-decentralization-e39db233c28e>.

TASKINSOY, J. Bitcoin nation: The worlds new 17th largest economy. 2021. Disponível em: <https://ssrn.com/abstract=3794634>.