Analyzing Cryptocurrency Communities on Twitter: A Multilayer Network Study

*Submitted to Prof. Soong Moon Kang for MSIN0074 Network Analysis by SRN 22086573

***Abstract*—This report analyzes the cryptocurrency networks on Twitter to gain a deeper understanding of their structures, centrality, and community detection. The Multilayer Network is composed of the Hashtag-Hashtag layer and the Crypto-Crypto layer. This project investigates the relationships between co-occurrence patterns of hashtags and cryptocurrencies in these networks and explores the interplay between them. The analysis focuses on centrality measures and community detection techniques. The goal is to provide insights into the cryptocurrency conversation on social media and identify opportunities for businesses operating in this space.**

I. Introduction

A. Cryptocurrency

Cryptocurrencies are a type of digital or virtual currency that utilizes cryptographic techniques to secure and validate transactions and to control the creation of new units. Unlike traditional fiat currencies, cryptocurrencies are decentralized and not controlled by any government or financial institution, which means they operate independently of traditional financial institutions. To make transactions with cryptocurrencies, users can buy and sell them on cryptocurrency online platforms that allow users to trade different cryptocurrencies for other cryptocurrencies or for traditional currencies like US dollars or euros. Transactions are recorded on a public digital ledger called a blockchain, which ensures that transactions are secure, transparent, and irreversible.

In recent years, cryptocurrencies have emerged as a significant and rapidly growing asset class, leading to the formation of numerous online communities and social media platforms around them. These communities have played a critical role in shaping the cryptocurrency market and influencing its direction. With the increasing use of social media, analyzing the structure and dynamics of these communities becomes vital in understanding how information and opinions spread and how they impact the market.

B. Multilayer Network

A multilayer network is a network composed of multiple layers, where each layer represents a different type of relationship between nodes. In the case of analyzing the crypto market on social media, each layer can represent a different type of interaction or relationship between entities, such as hashtags, users, or cryptocurrencies. In a multilayer network, nodes can exist in multiple layers simultaneously, and edges can connect nodes across different layers. This allows for a more comprehensive analysis of the network, as it takes into account multiple types of relationships and their interactions, which can is the advantage of using a multilayer network compare to a single-layer network. Analyzing a multilayer network involves examining each layer separately, as well as exploring the interactions between the layers. This can involve identifying common nodes across layers, analyzing how nodes in one layer influence nodes in another layer, and exploring the underlying structure of the network. In the context of cryptocurrencies and social media, a multilayer network can be used to explore the relationships between different types of entities, such as hashtags and cryptocurrencies, and identify how they relate to each other. It can also help identify key influencers, subgroups within the network, and trends within the conversation.

Overall, the use of a multilayer network in this context can add significant value by providing a more comprehensive analysis of the cryptocurrency conversation on social media. By identifying trends, influential nodes, and communities within the network, businesses and investors can make informed decisions and develop effective strategies in this rapidly evolving industry.

II. Related Work/Literature Review

Several studies have explored the structure and characteristics of networks in social media, including their centrality and community detection. Bonifazi et al. (2022) analyzed the structure and dynamics of the Twitter conversation on COVID-19 using a multilayer network composed of the user-user layer and the hashtag-hashtag layer, which is similar to this project. They identified the most influential users and hashtags in the conversation and examined the communities that emerged in the network. Türker and Sulak (2018) focused on tag networks on Twitter and presented a two-layer analysis, similar to this project. They used the co-occurrences of hashtags in a single entry (tweet) into links, which is similar to this study. Ho (2020) proposed that Network Science and graph theory can be used to effectively analyze hashtag culture on Instagram, which can allow for a deeper understanding of conversations and themes found across the platform, similar to this study.

Furthermore, studies on cryptocurrencies have also become popular topics for social media analysis. Shen et al. (2019) used Twitter data to predict Bitcoin prices and found that Twitter data can be a useful predictor of cryptocurrency prices. Decker and Wattenhofer (2013) analyzed information propagation in the Bitcoin network using network science and found that the network structure plays a crucial role in information dissemination. These studies all utilize network analysis to gain insights into the organization and dynamics of conversations on social media and identify opportunities for businesses operating in the crypto space.

However, specifically, the use of a multilayer network in analyzing cryptocurrency on social media can add significant value. By examining each layer separately and exploring the interactions between layers, a multilayer network analysis can provide a more comprehensive understanding of the cryptocurrency conversation on social media.

III. DATA

The data includes tweets and hashtags information related to cryptocurrencies, covering a period of 11 days, from March 8th, 2023 to March 18th, 2023, and includes a total of 79,099 records.

A. Data Scrapping

In this study, data were obtained from multiple sources including CoinMarketCap and Twitter. The CoinMarketCap API was used to scrape the top 100 coins by market capitalization, and these coin names were used as keywords to scrape related social media data from Twitter. The Tweepy library was used to authenticate with the Twitter API, resulting in the acquisition of 79099 data points for analysis. Then for each coin, a list of keywords including the abbreviation was defined to search for relevant tweets. For example, 'btc' is also recognized as Bitcoin in further analysis. This enabled the analysis of a complex language data related to the top cryptocurrencies in the corpus.

B. Network from the Data

The network being built from this data is a multilayer network with two layers: a crypto-crypto layer and a hashtag-hashtag layer. In the hashtag-hashtag layer, the nodes represent hashtags extracted from the tweets within the dataset. The links between the nodes are determined based on their co-occurrence in one tweet. This means that if two hashtags appear in the same tweet, a link will be created between the two nodes representing them in the network.

In the crypto-crypto layer, the nodes are the cryptocurrency listings obtained from the CoinMarketCap API. The edges between the nodes are based on their co-occurrence in a tweet. For instance, if two cryptocurrency listings are mentioned together in a tweet, an edge will be created between the two nodes representing them in the network. Essentially, the multilayer network connects hashtags with cryptocurrencies, allowing for the exploration of the relationship between the two. Overall, this dataset and network provide valuable insights into the use and popularity of cryptocurrencies on social media.

IV. MODEL/ALGORITHM/METHOD

Python is the primary programming language used in this project. This project involves several steps of data processing, network analysis, and network centrality correlation analysis. The data manipulation and analysis tasks are mainly performed using the popular Python libraries NumPy and Pandas. Data visualization is an essential aspect of the project, which is mostly done using the powerful plotting libraries, Matplotlib and Seaborn. Other specific libraries and algorithms used in each step of the process are discussed in detail in each part.

A. Data Preprocessing

After data collection, it is essential to clean and preprocess the data before conducting network analysis. In this study, the data cleaning process involved several steps to ensure that the data were reliable and valid for analysis. Firstly, duplicated or any null text rows were removed to ensure that only complete data were used. As hashtags and tweets will be main source for analysis, the Natural Language Toolkit (NLTK) was used as it is a widely used library in natural language processing that provides a comprehensive set of tools for text analysis that can be easily integrated into Python code. Additionally, it has a vast collection of pre-built corpora and models that can be used for various natural language processing tasks, such as The WordNetLemmatizer class, which is used for lemmatizing the words in this project.

To make the text more uniform, the text were converted to lowercase and punctuation was removed. Then text was tokenized using the NLTK library to split it into individual words, and negation cues were handled to ensure that the sentiment of the text was accurately captured for further sentiment analysis. Stop words such as "for", "a/an", "and", "the" were removed, and contractions were replaced with their expanded form to make the text more understandable. And the words were lemmatized to reduce inflectional forms to their base or root form. For example, the original sentence of "I didn't like the movie. It was too boring and long" converted to "do not like movie boring long". Furthermore, meaningless words were removed for data cleaning to ensure that the data used for analysis were relevant and meaningful. This includes the words that are frequently shown in sentences but not meaningful in this project such as "be","u","de","do". Lastly, only English sentences were extracted using the Langdetect library to ensure that the data were consistent and could be analyzed without the influence of other languages. After the whole process, 59097 data points were finally used for analysis. By applying these steps, the study can obtain accurate and meaningful results, which can be used to make informed decisions.

B. Network Analysis

1) Modelling: In this project, network analysis was used as the main algorithm to investigate the relationships between cryptocurrencies and the hashtag corpus. To create a multilayer network model composed of two layers, two lists of unique hashtags and cryptocurrencies were defined, and edges were established between hashtags and cryptocurrencies, hashtags and other hashtags in the same tweet, and cryptocurrencies and other cryptocurrencies in the same tweet. After combining them into one dataframe, the Py3Plex library was used to create a multilayer network object. Py3Plex is the ability to handle large and complex multilayer networks. For each layer in the multilayer network, NetworkX which is a library for working with graphs and networks is used and it provides tools for analyzing their properties are used. These libraries provide efficient algorithms for computing various network measures, such as centrality, clustering, and community detection, which can be used to gain insights into the structure and dynamics of the network.

2) Centrality: Centrality measures play a crucial role in analyzing multilayer networks, and they are particularly useful in gaining insights into the dynamics of complex systems like the cryptocurrency market. Degree centrality, closeness centrality, and betweenness centrality provide different perspectives on the importance of nodes in the network. Degree centrality measures the number of edges that are incident to a node, while closeness centrality measures the inverse of the sum of the shortest path distances between a node and all other nodes in the network. Betweenness centrality measures the number of shortest paths that pass through a node. By calculating these centrality measures for each node in the network, the project was able to identify the most influential and important cryptocurrencies within the market, as well as the interconnections between them. This information can be valuable for investors, traders, and other stakeholders who are interested in gaining a deeper understanding of the cryptocurrency market and the factors that impact it.

3) Community Detection: Community detection is a powerful tool for identifying subgroups of nodes in a network that are more densely connected to each other than to other nodes in the network. In this project, the Louvain algorithm was used to detect communities in the multilayer network. The Louvain algorithm is a widely used community detection algorithm that optimizes modularity, a measure of the strength of the division of a network into communities.

To identify the most significant communities in the network, this project set the number of communities to keep at 10. This was done to focus the analysis on the most prominent and meaningful communities in the network. The number was set to 10 because it strikes a balance between granularity and meaningfulness. It also helps to simplify the network and make it more manageable for further analysis. To determine communities to keep, this study identified the communities with the largest number of nodes using the partition generated by the Louvain algorithm. Nodes belonging to smaller communities were merged into larger ones to create a new partition. By this, the new partition preserves the most significant and well-structured communities in the network while reducing the complexity of the overall network. Finally, the modularity of the new graph was computed to assess the strength of the division of the network into the identified communities. The modularity value indicated the effectiveness of the Louvain algorithm in identifying significant communities in the network. After identifying the top 10 communities by size, this project went on to analyze the nodes within each community to identify the most important and influential nodes using degree centrality. By examining the most important nodes within each community, businesses and investors can identify key influencers and opinion leaders and use this information to inform their strategies and decision-making processes.

Overall, the use of several libraries, including pandas, Py3Plex, NetworkX, and community, allows for the efficient construction and analysis of the multilayer network. These libraries provided a comprehensive set of tools for handling data, building complex network models, and performing advanced network analysis. By combining these libraries, the study was able to gain insights into the complex relationships between cryptocurrencies and social media, which would have been difficult to obtain through traditional data analysis methods.

C. Network Centrality Correlation Analysis

Analyzing the relationships between network centrality measures and other metrics in the data is meaningful for several reasons. Firstly, it allows gaining insights into the dynamics and relationships specifically between cryptocurrencies, which can help us better understand how they interact within the larger cryptocurrency network. Additionally, by comparing the centrality measures to other variables such as sentiment scores, market capitalization rank, and amount of crypto mentions in the tweet corpus, the study can better understand how these factors are related and potentially impact the overall network structure. This information can be valuable for investors, traders, and other stakeholders who are interested in gaining a deeper understanding of the cryptocurrency market and the factors that affect it.

To gain insights into the relationships between the network centrality measures of the crypto-crypto layer and other metrics in the data, a correlation matrix was created. These metrics include sentiment score, market capitalization rank and amount of crypto mentions in the whole tweet corpus. For this analysis, additional data analytics were performed. A frequency table of tweet mentions of each crypto was created using regular expressions to match coin names and their associated abbreviations in the text corpus. Sentiment analysis was also performed using the TextBlob, which is pretrained machine learning model to calculate sentiment scores. Sentiment scores of polarity and subjectivity for each coin and the overall average are calculated. While the polarity score measures the positivity or negativity of a text, ranging from -1 (most negative) to 1 (most positive), subjectivity score ranges from 0 (most objective) to 1 (most subjective). The sentiment analysis was used to gain insights into the overall sentiment of the tweets about each coin and to compare the sentiment scores with the network centrality measures in the data. The resulting insights were then used to create a correlation matrix. Correlation coefficients were calculated between the crypto-crypto centrality measures (degree centrality, closeness centrality, and betweenness centrality) and other variables in the data (tweet count, market capitalization rank, and sentiment score). The resulting correlation matrix was visualized as a heatmap using the Seaborn library, allowing for a better understanding of the relationships between the different metrics and centrality measures. This analysis provided a comprehensive understanding of how network centrality is related to other factors such as their sentiment, market capitalization and popularity in tweets.

V. Results and findings

The results section of this report will primarily focus on the result of the multilayer network analysis, including the hashtag-hashtag network, crypto-crypto network, and hashtag-crypto network, as well as the centrality correlation analysis. The discussion will cover the structure of each network, the main network centralities, and community detection. The implications of the results for businesses will also be discussed. Figure 1 shows the top 10 currencies by market capitalization, which is based on CoinMarketCap when the data scrapping was conducted. Bitcoin and Ethereum are the top 2 ranked cryptos according to it, which will be used for further analysis in this study.

Coin	Market Cap Rank
Bitcoin	1
Ethereum	2
BNB	3
XRP	4
Cardano	5
Dogecoin	6
Polygon	7
Solana	8
Polkadot	9
Shiba Inu	10

Fig. 1. Top 10 Ranked Coin by Market Cap

A. Hashtag-Hashtag Layer

The Hashtag-Hashtag layer is a representation of the relationships between hashtags used in the cryptocurrency space on social media platforms. The network is visualized as a graph in Figure 2. It has 5144 nodes and 11411 edges, with each node representing a hashtag and the edges indicating a co-occurrence of two hashtags in a social media post. The graph uses different colors for the nodes to show different communities of hashtags within the network.

1) Centrality: Figure 3 reveals that the hashtags 'crypto', 'blockchain', 'bitcoin', and 'ethereum' are the most important and influential hashtags in the network, in terms of all three centrality measures. This indicates that these concepts are widely used and highly connected to other nodes in the

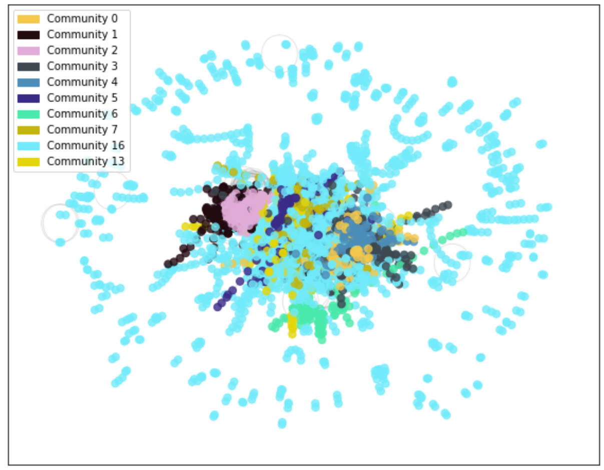

Fig. 2. Hashtag Network Visualisation

```
Top 10 nodes with highest degree centrality:
Degree centrality of crypto : 0.08496986194827921
Degree centrality of bitcoin : 0.06085942057165079
Degree centrality of nft : 0.056581761617732845
Degree centrality of btc : 0.042387711452459655
Degree centrality of blockchain : 0.04024888197550068
Degree centrality of ethereum : 0.03849893058526152
Degree centrality of eth : 0.03655454015166246
Degree centrality of cryptocurrency : 0.03266575928446432
Degree centrality of web3 : 0.029943612677425627
Degree centrality of nfts : 0.029943612677425627

Top 10 nodes with highest closeness centrality:
Closeness centrality of crypto : 0.3358516762789542
Closeness centrality of nft : 0.3210295618014303
Closeness centrality of bitcoin : 0.32062527240587324
Closeness centrality of ethereum : 0.3119590999140034
Closeness centrality of btc : 0.31133078349927223
Closeness centrality of eth : 0.3109728806011556
Closeness centrality of blockchain : 0.31083887907260005
Closeness centrality of web3 : 0.3034509256262275
Closeness centrality of cryptocurrency : 0.30283519117036684
Closeness centrality of bnb : 0.3016532305692001

Top 10 nodes with highest betweenness centrality:
Betweenness centrality of crypto : 0.1756776302861538
Betweenness centrality of nft : 0.11337869330180281
Betweenness centrality of bitcoin : 0.11278214494717319
Betweenness centrality of blockchain : 0.06471463474630786
Betweenness centrality of ethereum : 0.05685011676931576
Betweenness centrality of btc : 0.04792468113973199
Betweenness centrality of eth : 0.04259920374532121
Betweenness centrality of web3 : 0.04207879219698968
Betweenness centrality of cryptocurrency : 0.039515404037893186
Betweenness centrality of nfts : 0.03827883701708601
```

Fig. 3. Hashtag Centrality

network, and are central to the cryptocurrency and blockchain market.

This is not surprising as Bitcoin and Ethereum are most highly ranked by market capitalisation. However, the hashtags related to NFT and Web3 are also found in the top 10 ranked in all three centralities. NFTs, represented by the hashtag 'nft', are digital assets that are unique and non-interchangeable, and are often used to represent digital art, music, or other collectables. NFTs have gained significant attention in recent years, with high-profile sales of NFT artwork and music. Web3 refers to the third generation of the internet, which aims to create a decentralized web using blockchain technology. The centrality values of these hashtags suggest that they are relevant and influential within the cryptocurrency and blockchain market, and have significant potential for growth and development.

```
Community 2: ['eth', 'btc', 'bnb', 'cryptocurrency', 'xrp', 'matic', 'sol', 'doge', 'eos', 's
tx']
Community 4: ['nft', 'nfts', 'nftcommunity', 'giveaway', 'tezos', 'nftart', 'polygon', 'metav
erse', 'nftgiveaway', 'nftcollection']
Community 0: ['crypto', 'cryptonews', 'ai', 'news', 'trading', 'chatgpt', 'cryptotwitter', 't
erraclassic', 'cryptocommunity', 'cryptocurrencymarket']
Community 3: ['bitcoin', 'creditsuisse', 'borsa', 'dolar', 'forex', 'kripto', 'forextrader',
'smartcontracts', 'tonblockchain', 'fed']
Community 1: ['ethereum', 'binance', 'airdrop', 'kucoin', 'arbitrum', 'hodl', 'bybit', 'mex
c', 'gpt', 'bnbchain']
Community 16: ['dogecoin', 'defi', 'zilliqa', 'chainlink', 'bsc', 'ecash', 'cardano', 'xec',
'hedera', 'cosmos']
Community 13: ['computer', 'hack', 'internet', 'python', 'cybersecurity', 'security', 'meme',
'javascript', 'coding', 'tech']
Community 5: ['singularitynet', 'fantom', 'optimism', 'alt', 'conflux', 'altcoin', 'immutable
x', 'stack', 'thegraph', 'synthetix']
Community 6: ['web3', 'toncoin', 'auction', 'username', 'tonkeeper', 'icp', 'telegram', 'doma
in', 'sustainability', 'ic']
Community 7: ['blockchain', 'vechain', 'ethereumclassic', 'cryptoinvesting', 'coinmarketcap',
'interoperability', 'etcarmy', 'immutable', 'vechainhasnfts', 'iot']
```

Fig. 4. Communities in the Hashtag Layer

2) Community Detection: The modularity with 10 communities was over 0.80, indicating that the communities were well-defined and that the nodes within each community were more densely connected than those between communities.

Each community has a unique interest in specific aspects of the cryptocurrency space. Figure 4 shows the communities in the network. This community structure can help businesses identify potential partnership opportunities and inform their product development and marketing strategies. For instance, there are a strong presence of AI and crypto news-related topic in community 0, which is marked as yellow dots in Figure 2. The nodes with high centrality include 'chatgpt', 'ai', 'crypto news' and 'news', suggesting that members of this community may be particularly interested in exploring the intersection of AI and cryptocurrency news.

In summary, businesses operating in the cryptocurrency space can benefit from the insights gained from analyzing the hashtag-hashtag layer. By leveraging the centrality measures and community structure of the network, businesses can identify potential collaboration opportunities, inform their product development and marketing strategies, and target their marketing efforts more effectively.

B. Crypto-Crypto Layer

The Crypto-Crypto layer is a representation of the relationships between cryptocurrencies mentioned in the same tweets. The network is visualized as a graph with 63 nodes and 420 edges, with each node representing a cryptocurrency and the edges indicating a co-occurrence of two cryptocurrencies. It is worth noting that there are some cryptocurrencies in the dataset that are not connected to any other coins in the network. This is why the network has fewer nodes than expected which is 100. The graph uses different colors for the nodes to show different communities of hashtags within the network.

1) Centrality: Figure 5 shows centrality measures in crypto-crypto layer. It seems Bitcoin and Ethereum are top-ranked as hashtag-hashtag layer but BNB, XRP, Solana and Polygon are also highly ranked in all three of centrality, unlike the hashtag-hashtag layer. This implies that those cryptocurrencies appear to be particularly influential within the crypto-crypto layer, suggesting that they may be particularly relevant for investors or traders who are focused on trading cryptocurrencies directly.

```
Top 10 nodes with highest degree centrality:
Degree centrality of bitcoin : 0.7258064516129032
Degree centrality of bnb : 0.7096774193548387
Degree centrality of ethereum : 0.6451612903225806
Degree centrality of xrp : 0.5161290322580645
Degree centrality of solana : 0.45161290322580644
Degree centrality of polygon : 0.43548387096774194
Degree centrality of dogecoin : 0.3709677419354839
Degree centrality of cardano : 0.3387096774193548
Degree centrality of neo : 0.3387096774193548
Degree centrality of singularitynet : 0.3225806451612903

Top 10 nodes with highest closeness centrality:
Closeness centrality of bitcoin : 0.759697835851368
Closeness centrality of bnb : 0.7502016129032257
Closeness centrality of ethereum : 0.7060721062618596
Closeness centrality of xrp : 0.6384694577899794
Closeness centrality of solana : 0.6251680107526881
Closeness centrality of polygon : 0.6187229797140007
Closeness centrality of dogecoin : 0.582680864390855
Closeness centrality of cardano : 0.582680864390855
Closeness centrality of neo : 0.582680864390855
Closeness centrality of dydx : 0.582680864390855

Top 10 nodes with highest betweenness centrality:
Betweenness centrality of bnb : 0.14994573540346837
Betweenness centrality of bitcoin : 0.14362204585847296
Betweenness centrality of ethereum : 0.10243197914063604
Betweenness centrality of optimism : 0.040905778231629754
Betweenness centrality of polygon : 0.03892337085512912
Betweenness centrality of conflux : 0.03539446303583812
Betweenness centrality of xrp : 0.03475739339507685
Betweenness centrality of solana : 0.032289059170157786
Betweenness centrality of dydx : 0.030744984065886282
Betweenness centrality of stellar : 0.023009075925831034
```

Fig. 5. Cryptocurrency Centrality

This distinction between the hashtag-hashtag layer and the crypto-crypto layer highlights the complexity of the cryptocurrency market. The hashtag-hashtag layer represents the co-occurrence of hashtags in social media posts related to cryptocurrency, which is more public-facing, as it represents the broader conversation and discourse around cryptocurrency. On the other hand, the crypto-to-crypto layer in this analysis can be related to the direct trading between different cryptocurrencies. Direct trading between different cryptocurrencies is a key aspect of the cryptocurrency market because it enables investors and traders to exchange one cryptocurrency for another without having to go through a traditional currency. Hence, this layer may be more investor-focused, as mentions of multiple cryptocurrencies may represent influential within the trading aspect of the market. Coins with high centrality measures in this layer are likely to be highly traded and influential within the market, making them important for investors or traders who are focused on trading cryptocurrencies directly. As a result, understanding the coins that are most influential in the crypto-to-crypto layer can be valuable for investors and traders who are looking to make informed decisions about their trades and investments.

Overall, centrality measures in both layers are important and provide different insights into the cryptocurrency market. The hashtag-hashtag layer can help identify the different topics and sub-communities within the broader market so that the businesses operating in the cryptocurrency space can benefit from the insights gained from analyzing this layer, while the cryptocurrency-to-cryptocurrency layer can provide insights into the most influential coins for direct trading between cryptocurrencies.

2) Community Detection: The crypto-to-crypto network was composed of 7 communities. As it can be checked in Figure 6 and 7, community 6 only includes one cryptocurrency, MultiversX and regarding the visualization of the network, this community is far away from others. This means that there are few or no links between MultiversX and other cryptocurrencies or topics within the broader cryptocurrency market, which could further explain its isolation in a separate community.

```
Community 0: ['ethereum', 'xrp', 'polkadot', 'dogecoin', 'eos', 'cardano', 'litecoin', 'stell
ar', 'dash', 'monero', 'zcash', 'quant', 'gatetoken']
Community 1: ['bitcoin', 'tron', 'dai', 'usdd', 'kava', 'zilliqa']
Community 2: ['bnb', 'fantom', 'singularitynet', 'cronos', 'klaytn', 'aptos', 'optimism', 'to
ncoin', 'gmx', 'magic', 'conflux', 'okb', 'filecoin', 'neo', 'mina', 'loopring', 'immutable',
'dydx', 'casper', 'osmosis']
Community 3: ['polygon', 'solana', 'tezos', 'aave', 'flow', 'apecoin', 'decentraland', 'chili
z', 'pancakeswap', 'thorchain']
Community 5: ['avalanche', 'vechain', 'algorand', 'ecash', 'hedera', 'iota']
Community 4: ['chainlink', 'cosmos', 'uniswap', 'bitdao', 'maker', 'trueusd', 'synthetix']
Community 6: ['multiversx']
```

Fig. 6. Communities in the Crypto Layer

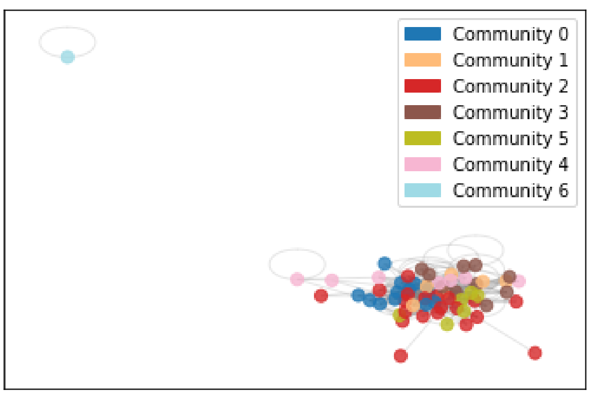

Fig. 7. Cryptocurrency Network Visualisation

C. Interplay Between the Two Layers

Hashtag - Crypto Layer indicates the interplay between two layers: Crypto-Crypto layer and Hashtag-Hashtag layer.

While this Multilayer Network consists of 2 layers, it has 5,221 nodes and 101343 edges. In Figure 8, 1 refers to the hashtag-hashtag layer and 2 refers to the crypto-crypto layer. The degree of the network is 38.82, indicating that each node has an average of 38.82 edges.

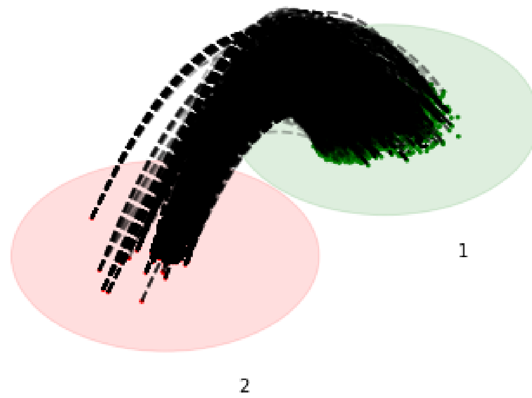

Fig. 8. Multilayer Network Visualisation

1) Centrality: For exploring the interplay between these two networks in the Multilayer Network, degree centrality is used to identify the nodes that are connecting the two layers of the network, and how they may be influencing each other. In Figure 9, the top 10 nodes identified by degree centrality in the core network of the multilayer network are likely to be important or influential hubs in the larger network.

	Centrality
Node	
(bitcoin, 2)	2.139464
(ethereum, 2)	1.208429
(crypto, 1)	1.101341
(bitcoin, 1)	1.009195
(bnb, 2)	0.994444
(usdd, 2)	0.717625
(btc, 1)	0.538697
(dai, 2)	0.528544
(ethereum, 1)	0.525862
(xrp, 2)	0.506130

Fig. 9. Top10 Ranked Multilayer Network Centrality

Interestingly, while Bitcoin, Ethereum and Binance are highly central nodes in the core network, other nodes such as USDD and DAI in the crypto layer also hold significant relevance and popularity within the entire multilayer network. While they may not have ranked among the top 10 nodes in the individual hashtag or crypto layers, their centrality across multiple layers makes them highly influential in the overall analysis.

One possible reason for their high centrality is that both USDD and DAI are stablecoins. Stablecoins are a type of cryptocurrency designed to maintain a stable value relative to another asset or currency, such as the US dollar or gold. Stablecoins offer the benefits of cryptocurrencies, such as fast and cheap transactions, while avoiding the volatility commonly associated with other cryptocurrencies like Bitcoin and Ethereum. They are often used as a bridge currency for trading between different cryptocurrencies, which could explain their strong connections with nodes in both the crypto and hashtag layers of the network. Overall, the Multilayer Network analysis provides valuable insights into the complex relationships between the hashtag and crypto networks, highlighting the crucial role of highly central nodes in influencing the network's structure and dynamics.

2) Community Detection: The resulting multilayer modularity with 10 communities was 0.799, indicating that the communities were well-defined and that the nodes within each community were more densely connected than those between communities. The top 10 nodes with high centrality were identified for each community in Figure 10. The top nodes for each community varied in terms of the cryptocurrencies and topics they were associated with. For example, the top nodes in Community 3, which are marked as light pink dots in Figure 11 were related to NFT topics. This suggests that each community had its own distinct theme or focus.

```
Community 3: ['nft', 'nfts', 'nftcommunity', 'tezos', 'giveaway', 'nftart', 'nftgiveaway', 'a
rt', 'nftdrop', 'nftcollection']
Community 1: ['eth', 'btc', 'bnb', 'binance', 'matic', 'xrp', 'doge', 'sol', 'usdt', 'altcoin
s']
Community 6: ['cryptocurrency', 'cryptonews', 'trading', 'news', 'invest', 'cryptocurrencie
s', 'money', 'lunc', 'cryptotwitter', 'cryptocurrencymarket']
Community 14: ['blockchain', 'future', 'technology', 'hack', 'tech', 'internet', 'computer',
'developer', 'meme', 'javascript']
Community 2: ['bitcoin', 'ecash', 'xec', 'privacy', 'creditsuisse', 'bankingcrisis', 'fed',
'ordinal', 'mining', 'banking']
Community 8: ['singularitynet', 'toncoin', 'fantom', 'alt', 'conflux', 'altcoin', 'immutable
x', 'optimism', 'stack', 'thegraph']
Community 5: ['crypto', 'stpatricksday', 'price', 'exchange', 'svb', 'volt', 'stake', 'crc',
'oscar', 'axie']
Community 12: ['tron', 'polygon', 'arbitrum', 'aptos', 'cosmos', 'magickingdom', 'tronlightcy
clerun', 'waltdisneyworld', 'layer2', 'wdw']
Community 4: ['dogecoin', 'solana', 'shiba', 'shibarmy', 'shibarmystrong', 'shibainu', 'shiba
rium', 'uniswap', 'floki', 'babydoge']
Community 27: ['web3', 'ethereum', 'defi', 'vechain', 'ai', 'metaverse', 'zilliqa', 'cardan
o', 'airdrop', 'chainlink']
```

Fig. 10. Communities in the Whole Network

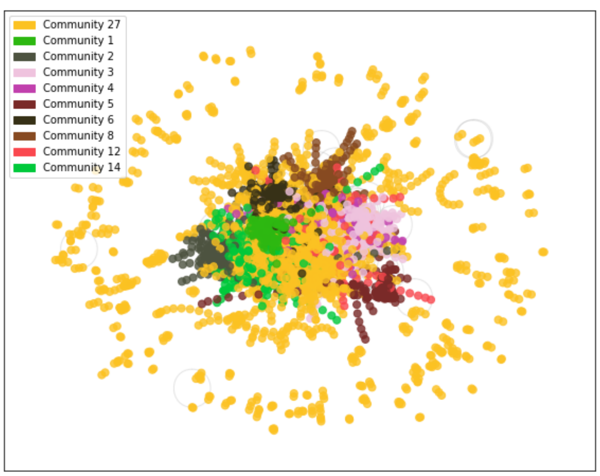

Fig. 11. Multilayer Network Visualisation with Communities

Overall, the analysis of the multilayer network using the Louvain algorithm and degree centrality helped to identify distinct communities within the cryptocurrency conversation on social media. These communities were characterized by their own unique topics and cryptocurrencies and had their own set of influential nodes. This information can be useful for businesses and investors operating in the cryptocurrency industry, as it can help them to better understand the dynamics of the community and make informed decisions.

D. Metrics Correlation Analysis

In Figure 12, a correlation matrix is presented to show the relationship between the centrality measures and various other metrics, including sentiment scores, market capitalization, and tweet mentions. The values in the matrix range from 0 to 1, as there was no negative correlation between variables, with a value of 1 indicating the strongest positive correlation between the two measures. The closer the value is to 1, the stronger the positive correlation between the two measures.

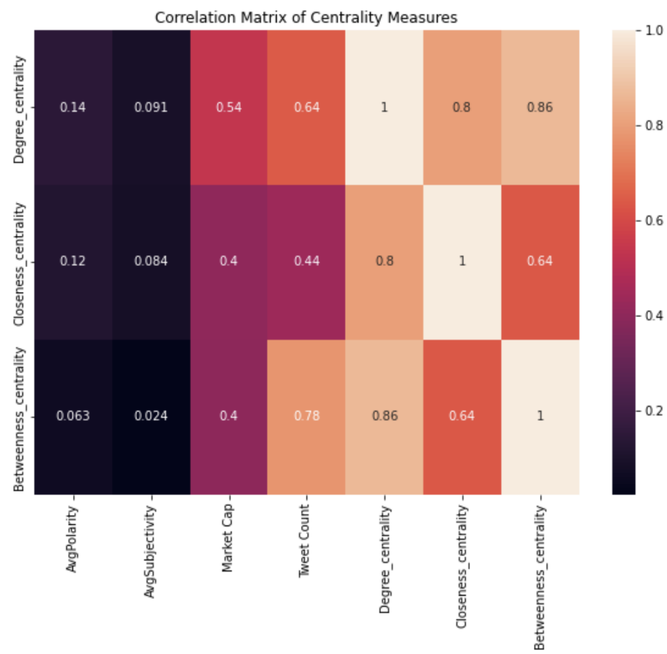

Fig. 12. Metrics Correlation Heatmap

1) Sentiment Analysis: The correlation matrix shows a weak positive relationship between the centrality measures of the cryptocurrency networks and the sentiment analysis measures of average polarity and average subjectivity. Also, as it is shown in Figure 13, highly ranked cryptos by sentiment scores are not likely to align with those by centrality measures. This could be due to the specific nature of the conversation around cryptocurrencies on social media during the 11-day period when the data was collected. It's possible that sentiment and network centrality are more strongly correlated in certain contexts, such as during periods of high volatility or news events.

2) Market Capitalization and Amount of Tweet Mentions: On the other hand, the analysis did reveal meaningful relationships between the centralities and market capitalization, as well as tweet count. Specifically, market capitalization was found to be more strongly related to degree centrality (0.54) than other centralities. As can be checked in Figure 1, more than half of the highly ranked cryptos by market

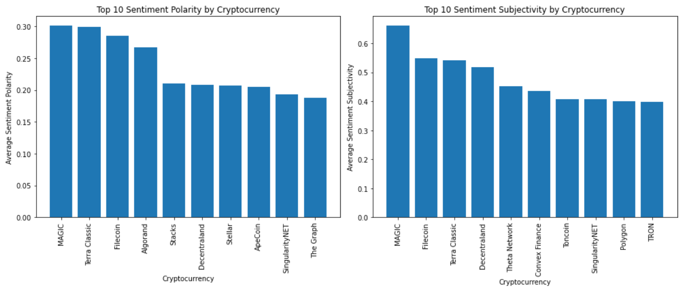

Fig. 13. Top10 Ranked Cryptocurrencies by Sentiment Score

capitalization overlapped with those by centralities. This may be because cryptocurrencies with higher market capitalization are typically more established and have a larger user base, leading to more connections (degree) in the network.

In contrast, tweet count was found to be more strongly related to betweenness centrality (0.78). As can be checked in Figure 14, most highly ranked cryptocurrencies by tweet count are mostly overlapped with those by centralities. This may be because cryptocurrencies that are frequently mentioned on Twitter may act as intermediaries in transactions and facilitate connections between different parts of the network, leading to a higher betweenness centrality.

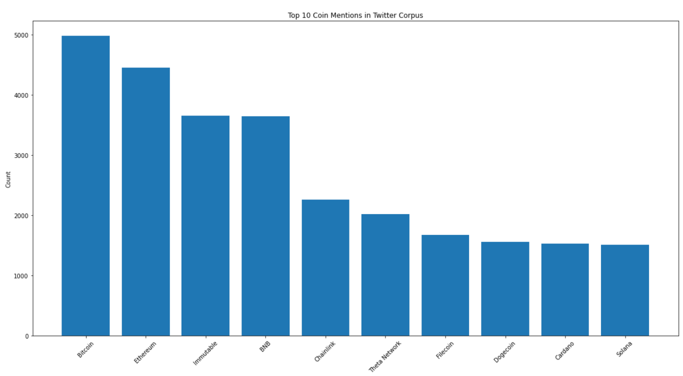

Fig. 14. Top10 Ranked Frequent Cryptocurrencies

Overall, these findings suggest that market capitalization and tweet count are both related to different aspects of a cryptocurrency's centrality in the network. While market capitalization is more strongly associated with degree centrality, tweet count is more strongly associated with betweenness centrality.

VI. Limitations and Future Work

It is important to recognize that there are several limitations to the study. Firstly, the analysis was restricted to data from Twitter, which may not fully capture the broader cryptocurrency conversation on social media. Moreover, the study only focused on three networks and did not include other popular social media platforms such as Reddit or Discord. Secondly, the data covers a relatively short period of 11 days, which limits the ability to fully understand the temporal dynamics of the networks and how the sentiment and conversation around cryptocurrency changes over time.

To address these limitations, future work could expand on the findings in several ways. Firstly, researchers could incorporate data from a variety of social media platforms to provide a more comprehensive understanding of the cryptocurrency conversation. This could also involve exploring how sentiment and network structure vary across different social media platforms. Secondly, incorporating temporal dynamics into the analysis could provide valuable insights into how the cryptocurrency conversation and sentiment evolves over time. This could be achieved by analyzing longer time periods or conducting longitudinal studies. Lastly, exploring the relationship between the structure of the cryptocurrency conversation and cryptocurrency prices or market trends could provide a deeper understanding of the factors driving market dynamics.

Overall, while the study provides valuable insights into the relationship between sentiment and network structure in the cryptocurrency conversation on Twitter, future research can build upon these findings and address the limitations of the study to provide a more nuanced understanding of the cryptocurrency landscape on social media.

VII. Conclusion

In conclusion, the analysis of the cryptocurrency community on Twitter highlights the importance of specific hashtags and their centrality within the network. While Bitcoin and Ethereum are currently the most important and influential concepts within the market, NFT and Web3 are also important concepts with significant potential for growth and development by analysing hashtag-hashtag layer. For the communities detected in this layer, it can be particularly valuable information for business in the crypto environment as various topics are clustered into 10 communities. Understanding the coins that are most influential in the crypto-to-crypto layer, such as BNB, XRP, Solana and Polygon, can be valuable for investors and traders who are looking to make informed decisions about their trades and investments. The analysis also revealed that USDD and DAI show their relevance and popularity within the whole multilayer network. One of the possible contributing factors to their centrality is that USDD and DAI are both stablecoins.

The analysis also reveals meaningful relationships between the centralities and market capitalization, as well as tweet count. Market capitalization was found to be more strongly related to degree centrality, while tweet count was more strongly related to betweenness centrality. These findings suggest that market capitalization and tweet count are both related to different aspects of a cryptocurrency's centrality in the network, providing valuable insights for investors and traders who are looking to make informed decisions.

While there are limitations to the study, such as the restriction to data from Twitter and the limited timeframe of 11 days, future research can build upon these findings and address the limitations to provide a more nuanced understanding of the cryptocurrency landscape on social media. Overall, these insights can help investors, traders and businesses identify potential partnership opportunities and inform their product development and marketing strategies in the cryptocurrency market.

REFERENCES

[1] Bonifazi, G., Breve, B., Cirillo, S., Corradini, E. and Virgili, L. (2022). Investigating the COVID-19 vaccine discussions on Twitter through a multilayer network-based approach. Information Processing Management, [online] 59(6), p.103095. doi:https://doi.org/10.1016/j.ipm.2022.103095.

[2] Türker, İ. and Sulak, E.E. (2018). A multilayer network analysis of hashtags in twitter via co-occurrence and semantic links. International Journal of Modern Physics B, 32(04), p.1850029. doi:https://doi.org/10.1142/s0217979218500297.

[3] Ho, S. (2020). Using Network Science to explore hashtag culture on Instagram. [online] Medium. Available at: https://towardsdatascience.com/using-network-science-to-explore-hashtag-culture-on-instagram-1f7917078e0 [Accessed 22 Apr. 2023].

[4] Shen, D., Urquhart, A. and Wang, P. (2019). Does twitter predict Bitcoin? Economics Letters, [online] 174, pp.118–122. doi:https://doi.org/10.1016/j.econlet.2018.11.007.

[5] Wang, R., Liu, W. and Gao, S. (2016). Hashtags and information virality in networked social movement. Online Information Review, 40(7), pp.850–866. doi:https://doi.org/10.1108/oir-12-2015-0378.

[6] Papadopoulos, S., Kompatsiaris, Y., Vakali, A. and Spyridonos, P. (2011). Community detection in Social Media. Data Mining and Knowledge Discovery, 24(3), pp.515–554. doi:https://doi.org/10.1007/s10618-011-0224-z.

[7] Becker, H., Naaman, M. and Gravano, L. (2011). Beyond Trending Topics: Real-World Event Identification on Twitter. Proceedings of the International AAAI Conference on Web and Social Media, [online] 5(1), pp.438–441. Available at: https://ojs.aaai.org/index.php/ICWSM/article/view/14146.

[8] Zhang, L., Wang, S. and Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4). doi:https://doi.org/10.1002/widm.1253.

[9] Decker, C. and Wattenhofer, R. (2013). Information propagation in the Bitcoin network. IEEE P2P 2013 Proceedings. doi:https://doi.org/10.1109/p2p.2013.6688704.

[10] McNally, S., Roche, J. and Caton, S. (2018). Predicting the Price of Bitcoin Using Machine Learning. 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP). doi:https://doi.org/10.1109/pdp2018.2018.00060.

[11] Guidi, B. (2020). When Blockchain meets Online Social Networks. Pervasive and Mobile Computing, 62(101131), p.101131. doi:https://doi.org/10.1016/j.pmcj.2020.101131.

[12] Swan, M. (2015). BLOCKCHAIN : blueprint for a new economy. O'Reilly Media.

[13] Wu, J., Liu, J., Zhao, Y. and Zheng, Z. (2021). Analysis of cryptocurrency transactions from a network perspective: An overview. Journal of Network and Computer Applications, p.103139. doi:https://doi.org/10.1016/j.jnca.2021.103139.

[14] NLTK (2009). Natural Language Toolkit — NLTK 3.4.4 documentation. [online] Nltk.org. Available at: https://www.nltk.org/.

[15] TextBlob (2018). TextBlob: Simplified Text Processing — TextBlob 0.15.2 documentation. [online] Readthedocs.io. Available at: https://textblob.readthedocs.io/en/dev/.

[16] networkx.org. (n.d.). NetworkX — NetworkX documentation. [online] Available at: https://networkx.org/.