Analysing Cryptocurrency Communities on Twitter – A Multilayer Network Study

MSIN0074 Network Analysis by SRN 22086573

Table of Contents

01

INTRODUCTION

BACKGROUND
MOTIVATION
RELATED WORKS
DATA

BACKGROUND

What is Cryptocurrency?

- Cryptocurrencies are digital or virtual currencies that are decentralized and operate independently of traditional financial institutions.
- Cryptocurrencies are not controlled by any government or financial institution.

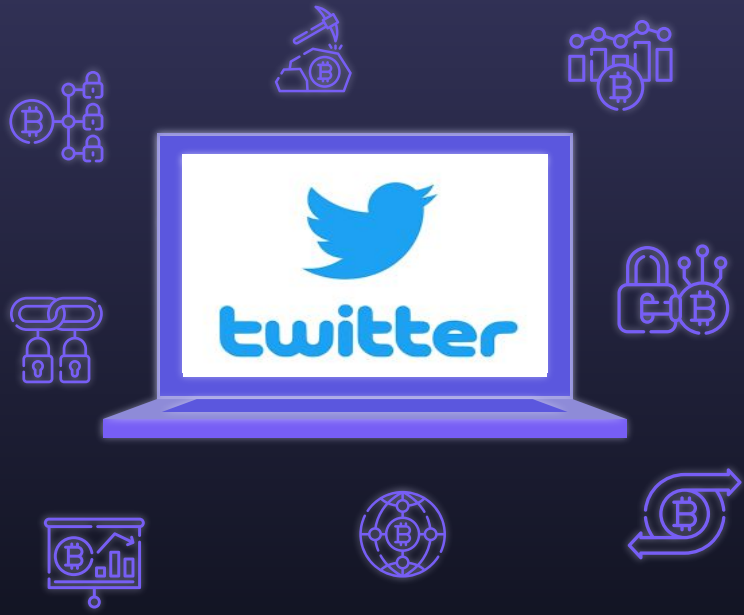
How to make transaction?

- Cryptocurrencies are bought and sold on cryptocurrency online platforms that allow users to trade different cryptocurrencies for other cryptocurrencies or for traditional currencies like US dollars or euros.

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

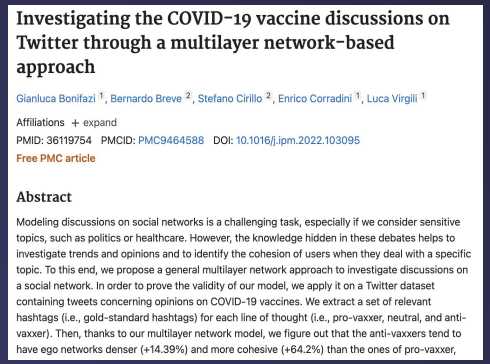
Market Cap Rank

MOTIVATION



Cryptocurrencies have emerged as a new asset class and an alternative to traditional financial systems. As a result, a variety of online communities and social media platforms have formed around them, providing valuable opinions about cryptocurrencies.

RELATED WORK



Investigating the COVID-19 vaccine discussions on Twitter through a multilayer network-based approach

Gianluca Bonifazi [1], Bernardo Breve [2], Stefano Cirillo [2], Enrico Corradini [1], Luca Virgili [1]

Affiliations + expand

PMID: 36119754 PMCID: PMC9464588 DOI: 10.1016/j.ipm.2022.103095

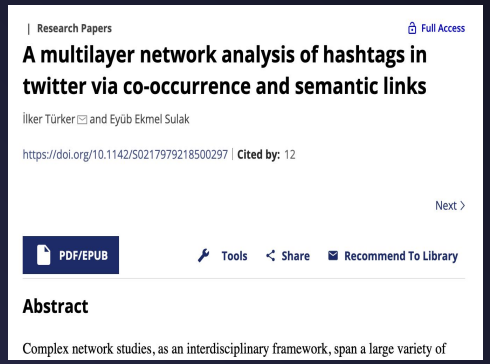
Free PMC article

Abstract

Modeling discussions on social networks is a challenging task, especially if we consider sensitive topics, such as politics or healthcare. However, the knowledge hidden in these debates helps to investigate trends and opinions and to identify the cohesion of users when they deal with a specific topic. To this end, we propose a general multilayer network approach to investigate discussions on a social network. In order to prove the validity of our model, we apply it on a Twitter dataset containing tweets concerning opinions on COVID-19 vaccines. We extract a set of relevant hashtags (i.e., gold-standard hashtags) for each line of thought (i.e., pro-vaxxer, neutral, and anti-vaxxer). Then, thanks to our multilayer network model, we figure out that the anti-vaxxers tend to have ego networks denser (+14.39%) and more cohesive (+64.2%) than the ones of pro-vaxxer,

Using Network Science to explore hashtag culture on Instagram

A practical walk-through of how to scrape hashtags from Instagram and model their relationships with one another using NetworkX



Research Papers Full Access

A multilayer network analysis of hashtags in twitter via co-occurrence and semantic links

İlker Türker and Eyüb Ekmel Sulak

https://doi.org/10.1142/S0217979218500297 | Cited by: 12

Next >

PDF/EPUB Tools Share Recommend To Library

Abstract

Complex network studies, as an interdisciplinary framework, span a large variety of

MULTILAYER NETWORK ANALYSIS ON VARIOUS SOCIAL MEDIA

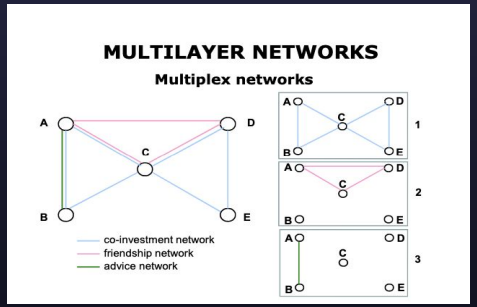
These studies all utilize network analysis to gain insights into the organization and dynamics of conversations on social media and identify opportunities for businesses.

DATA

```
RangeIndex: 79105 entries, 0 to 79104
Data columns (total 11 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   id                   79099 non-null  object
 1   text                 79085 non-null  object
 2   user_id              79065 non-null  object
 3   timestamp            79065 non-null  object
 4   retweet_count        79065 non-null  float64
 5   favorite_count       79055 non-null  float64
 6   in_reply_to_user_id  14908 non-null  object
 7   twt_hashtags         79065 non-null  object
 8   user_name            79063 non-null  object
 9   followers_count      79065 non-null  float64
 10  friends_count        79045 non-null  float64
dtypes: float64(4), object(7)
```

The data includes a total of 79,105 records and covers a period of 11 days, from March 8th, 2023 to March 18th, 2023.

Mainly use text and hashtags with list of cryptocurrencies to build multilayer network.



02

METHOD

DATA SCRAPING
DATA PREPROCESSING
NETWORK ANALYSIS
CENTRALITY CORRELATION ANALYSIS

DATA SCRAPING

- Data was obtained from multiple sources including **CoinMarketCap and Twitter**.
- **CoinMarketCap API** was used to scrape the top 100 coins by market capitalization.
- These coin names were used as keywords to scrape related data from Twitter.
- The **Tweepy** library was used to authenticate with the **Twitter API**, resulting in the acquisition of 79105 data points for analysis.
- For each coin, a list of keywords including the abbreviation was defined to search for relevant tweets. This enabled the analysis of complex language data related to the top cryptocurrencies in the corpus.

DATA PREPROCESSING

Data cleaning is necessary before conducting network analysis. These steps ensure accurate and meaningful results for informed decision-making.

- Steps included removing duplicated or null rows, converting text to lowercase, removing punctuation, and tokenising with **Natural Language Toolkit (NLTK)** . Negation cues were handled, stop words were removed, and contractions were replaced. Words were lemmatised to reduce inflectional forms with **NLTK** and meaningless words were removed.
- Only English sentences were extracted using the **Langdetect** library. After the whole process, 59097 data points were used for analysis.

Natural Language Toolkit (NLTK) is a widely used library in NLP that provides a comprehensive set of tools for text analysis including vast collection of pre-built corpora and models that can be used for this project.

NETWORK ANALYSIS – STRUCTURE

Multilayer Network was used as the main method and it composed of two layers.

1. **Hashtag-Hashtag Layer**
   - Nodes: List of unique hashtags
   - Edges: Co-occurrence of two hashtags in same tweet

2. **Cryptocurrency-Cryptocurrency Layer**
   - Nodes: List of unique cryptocurrencies
   - Edges: Co-occurrence of two cryptocurrencies in same tweet

3. **Hashtag-Cryptocurrency Multilayer Network**
   - Nodes: List of unique entities
   - Edges: Co-occurrence of two entities in same tweet

NETWORK ANALYSIS – MODELLING

For the models and algorithms,

- The **Py3Plex** library was used to create a multilayer network object because of its ability to handle large and complex multilayer networks.
- For each layer in the multilayer network, **NetworkX** was used as a library for working with graphs and networks.

These models will provide efficient algorithms for computing various network measures, such as centrality and community detection, which were used to gain insights into the structure and dynamics of the network.

NETWORK ANALYSIS – MODELLING

For **Community Detection**, a plot was used to visualise the multilayer network and highlight communities using the **Community library and Louvain algorithm**.

The **Louvain algorithm** was used to detect communities in the network by optimising modularity, which measures the strength of division of a network into communities. The nodes belonging to each community were printed to identify groups of nodes that are more densely connected to each other than to other nodes in the network.

NETWORK ANALYSIS – MODELLING

This project set the number of communities to keep at 10, as this strikes a balance between granularity and meaningfulness. It also helps to simplify the network and make it more manageable for further analysis.

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.

NETWORK ANALYSIS – METRICS

Centrality Measures was used to identify important nodes.
Degree, closeness, and betweenness centrality scores were calculated for each node.

- **Degree centrality** measures how many connections a node has.
- **Closeness centrality** measures how easily a node can reach other nodes.
- **Betweenness centrality** measures how often a node appears on the shortest path between other nodes.

CENTRALITY CORRELATION ANALYSIS

A **Correlation Matrix and Heatmap** was created to understand the relationships between centrality measures and other metrics such as sentiment, market capitalization rank, and amount of crypto mentions in the tweet corpus.

- Frequency tables of tweet mentions of each crypto using regular expressions.
- Sentiment analysis performed to gain insights into the overall sentiment of the tweets about each coin with **TextBlob** library.

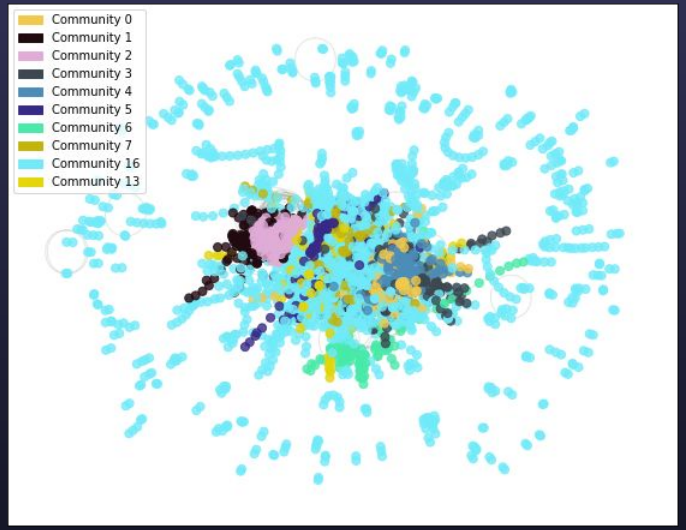
TextBlob is pretrained machine learning model to calculate sentiment scores.
This analysis will provide a comprehensive understanding of how network centrality is related to other factors such as sentiment, market capitalization, and popularity in tweets.

03

RESULTS

HASHTAG-HASHTAG LAYER
CRYPTO-CRYPTO LAYER
HASHTAG-CRYPTO NETWORK
CENTRALITY CORRELATION ANALYSIS

HASHTAG–HASHTAG LAYER



Network Visualisation with Community

This network has 5144 nodes and 11411 edges

- Nodes: List of unique hashtags
- Edges: Co-occurrence of two hashtags in same tweet

HASHTAG-HASHTAG LAYER

- The hashtags 'crypto', ‘blockchain’, 'bitcoin', and 'ethereum' are the most important and influential hashtags in the network, in all three centrality measures.

- **NFTs and Web3**, which are comparatively new to the market are important concepts with significant potential for growth and development.

```
Top 10 nodes with highest degree centrality:
Degree centrality of crypto : 0.08528784648187633
Degree centrality of bitcoin : 0.05989532855204497
Degree centrality of nft : 0.057763132389998056
Degree centrality of btc : 0.04245008722620663
Degree centrality of blockchain : 0.04012405504942818
Degree centrality of ethereum : 0.03934871099050203
Degree centrality of eth : 0.03547199069587129
Degree centrality of cryptocurrency : 0.03314595851909285
Degree centrality of web3 : 0.02946307423919364
Degree centrality of nfts : 0.028881566194999028

Top 10 nodes with highest closeness centrality:
Closeness centrality of crypto : 0.3378667120393272
Closeness centrality of nft : 0.3222406290482675
Closeness centrality of bitcoin : 0.32034974219579676
Closeness centrality of ethereum : 0.31252670863637766
Closeness centrality of btc : 0.31232301971047954
Closeness centrality of blockchain : 0.31083737507618103
Closeness centrality of eth : 0.31072540195835824
Closeness centrality of cryptocurrency : 0.3037659233118758
Closeness centrality of web3 : 0.3036162322549815
Closeness centrality of bnb : 0.3019158963375577

Top 10 nodes with highest betweenness centrality:
Betweenness centrality of crypto : 0.1794872852765143
Betweenness centrality of nft : 0.11111254723212333
Betweenness centrality of bitcoin : 0.11078490661399373
Betweenness centrality of blockchain : 0.06445110222877887
Betweenness centrality of ethereum : 0.05491400518737397
Betweenness centrality of btc : 0.04904816677009997
Betweenness centrality of web3 : 0.040751613297022996
Betweenness centrality of cryptocurrency : 0.04074663831104262
Betweenness centrality of eth : 0.03952523132764694
Betweenness centrality of nfts : 0.03744299660704651
```

Centrality Measures

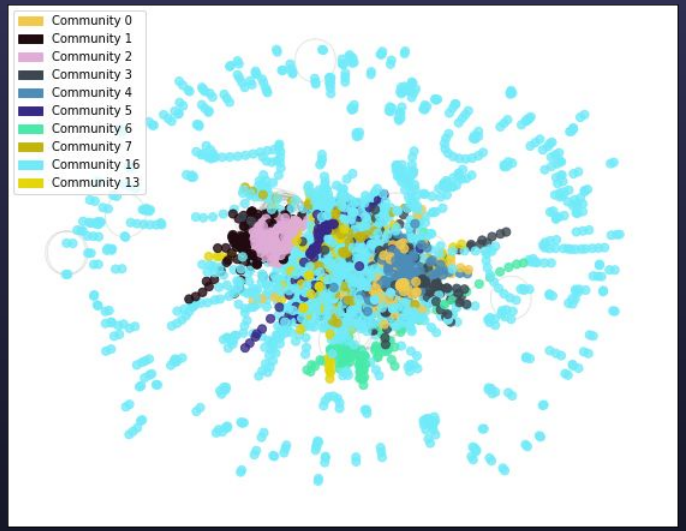
HASHTAG-HASHTAG LAYER

```
Community 2: ['eth', 'btc', 'bnb', 'cryptocurrency', 'xrp', 'matic', 'sol', 'doge', 'eos', 'stx']
Community 4: ['nft', 'nfts', 'nftcommunity', 'giveaway', 'tezos', 'nftart', 'polygon', 'metaverse', 'nftgiveaway', 'nftcollection']
Community 0: ['crypto', 'cryptonews', 'ai', 'news', 'trading', 'chatgpt', 'cryptotwitter', 'terraclassic', 'cryptocommunity', 'cryptocurrencymarket']
Community 3: ['bitcoin', 'creditsuisse', 'borsa', 'dolar', 'forex', 'kripto', 'forextrader', 'smartcontracts', 'tonblockchain', 'fed']
Community 1: ['ethereum', 'binance', 'airdrop', 'kucoin', 'arbitrum', 'hodl', 'bybit', 'mexc', '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', 'immutablex', 'stack', 'thegraph', 'synthetix']
Community 6: ['web3', 'toncoin', 'auction', 'username', 'tonkeeper', 'icp', 'telegram', 'domain', 'sustainability', 'ic']
Community 7: ['blockchain', 'vechain', 'ethereumclassic', 'cryptoinvesting', 'coinmarketcap', 'interoperability', 'etcarmy', 'immutable', 'vechainhasnfts', 'iot']
```

Hashtag Layer Community

The community structure can help businesses identify potential partnership opportunities and inform their product development and marketing strategies.

HASHTAG-HASHTAG LAYER

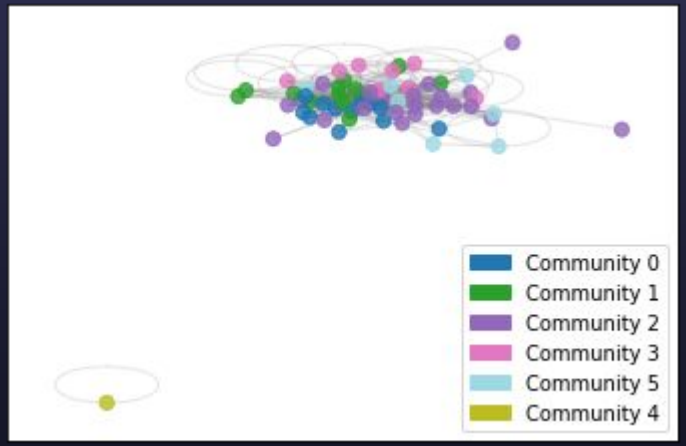


Network Visualisation with Community

For instance, there are a strong presence of AI and crypto news-related topic in community 0, which is marked as yellow dots.

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.

CRYPTO-CRYPTO LAYER



Network Visualisation with Community

This network has 63 nodes and 420 edges

- Nodes: List of unique cryptocurrency
- Edges: Co-occurrence of two cryptocurrencies in same tweet

CRYPTO-CRYPTO LAYER

Comparing to the HASHTAG-HASHTAG LAYER, CRYPTO-CRYPTO LAYER layer can be related to the **direct trading between different cryptocurrencies**. The direct trading enables investors and traders to exchange one cryptocurrency for another without having to go through a traditional currencies.

Hence, this layer may be more **investor-focused**, as mentions of multiple cryptocurrency may represents influential within the trading aspect of the market.

CRYPTO-CRYPTO LAYER

- Bitcoin and Ethereum are top ranked but BNB, XRP and Polygon are also highly ranked in this layer unlike the hashtag-hashtag layer.

- This implies crypto that are particularly influential within this layer, such as **BNB, XRP, Solana and Polygon** may be more relevant for investors or traders who are focused on trading directly.

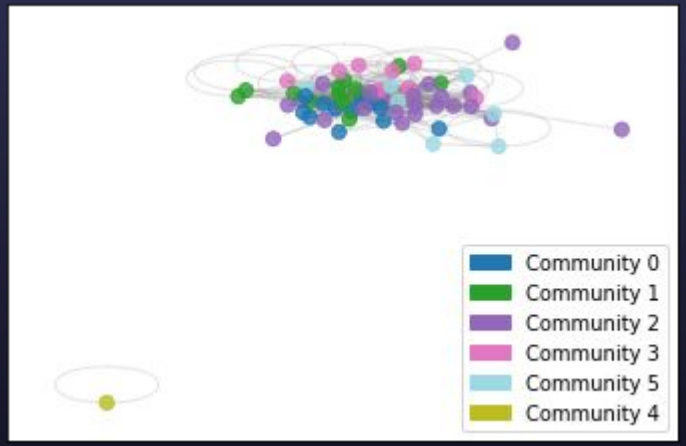
```
Top 10 nodes with highest degree centrality:
Degree centrality of bitcoin : 0.7096774193548387
Degree centrality of bnb : 0.6935483870967741
Degree centrality of ethereum : 0.6612903225806451
Degree centrality of xrp : 0.5161290322580645
Degree centrality of polygon : 0.43548387096774194
Degree centrality of solana : 0.43548387096774194
Degree centrality of dogecoin : 0.3709677419354839
Degree centrality of cardano : 0.3548387096774194
Degree centrality of kava : 0.3225806451612903
Degree centrality of neo : 0.3225806451612903

Top 10 nodes with highest closeness centrality:
Closeness centrality of bitcoin : 0.7502016129032257
Closeness centrality of bnb : 0.7409398645957787
Closeness centrality of ethereum : 0.7144772765745007
Closeness centrality of xrp : 0.6384694577899794
Closeness centrality of polygon : 0.6187229797140007
Closeness centrality of solana : 0.6187229797140007
Closeness centrality of cardano : 0.588393421884883
Closeness centrality of dogecoin : 0.582680864390855
Closeness centrality of dydx : 0.582680864390855
Closeness centrality of kava : 0.5770781637717122

Top 10 nodes with highest betweenness centrality:
Betweenness centrality of bnb : 0.14514172304541273
Betweenness centrality of bitcoin : 0.13910557393384018
Betweenness centrality of ethereum : 0.1098096536920728
Betweenness centrality of polygon : 0.042177389325426 64
Betweenness centrality of conflux : 0.036338998439654525
Betweenness centrality of xrp : 0.03614327610718526
Betweenness centrality of optimism : 0.035518476902411
Betweenness centrality of dydx : 0.03272331213930583
Betweenness centrality of solana : 0.03224517071906181
Betweenness centrality of stellar : 0.023821780233991025
```

Centrality Measures

CRYPTO-CRYPTO LAYER

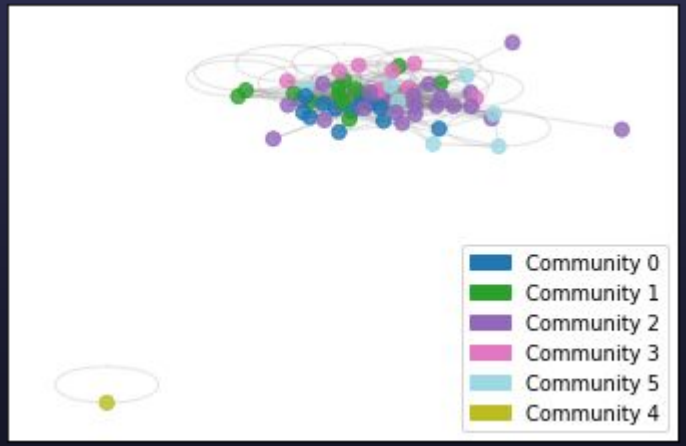


Network Visualisation with Community

This network has 6 communities.

- Community 0: [Ethereum, Xrp, Polkadot, Dogecoin, Cardano, Litecoin, Stellar, Dash, Monero, Zcash, Quant, Gatetoken]
- Community 1: [Bitcoin, Solana, Avalanche, Vechain, Tron, Dai, Usdd, Algorand, Aptos, Ecash, Iota, Chiliz, Thorchain]
- Community 2: [Bnb, Eos, Singularitynet, Cronos, Optimism, Aave, Kava, Toncoin, Hedera, Gmx, Magic, Conflux, Okb, Filecoin, Neo, Flow, Mina, Zilliqa, Immutable, Dydx, Casper, Osmosis]
- Community 3: [Fantom, Polygon, Tezos, Klaytn, Apecoin, Loopring, Decentraland, Pancakeswap]
- Community 4: [Multiversx]
- Community 5: [Chainlink, Cosmos, Uniswap, Bitdao, Maker, Trueusd, Synthetix]

CRYPTO-CRYPTO LAYER

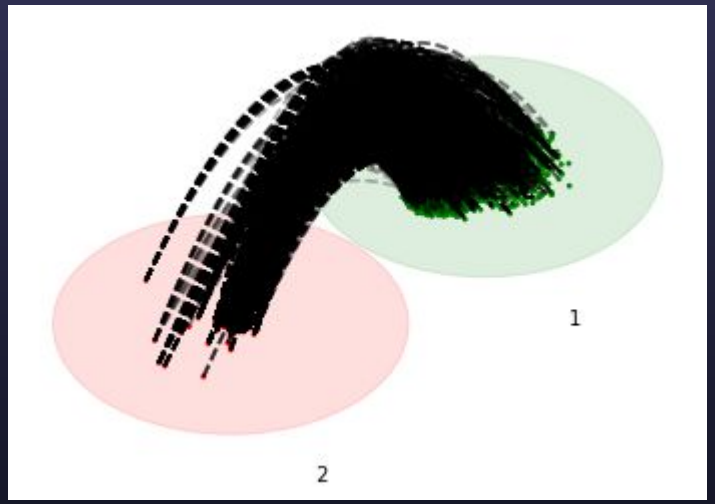


Network Visualisation with Community

Community 4 only includes one cryptocurrency, MultiversX.

Regarding the visualisation, community 4 far away from other communities, indicates that it is not closely connected to them. This means that there are few or no links between MultiversX and other crypto within the broader crypto market.

HASHTAG-CRYPTO NETWORK



Multilayer Network Visualisation

This network has 2 layers, 5221 nodes and 101343 edges. (1: Hashtag, 2: Cryptocurrency)

- Nodes: List of unique entities
- Edges: Co-occurrence of two entities in same tweet

HASHTAG-CRYPTO NETWORK

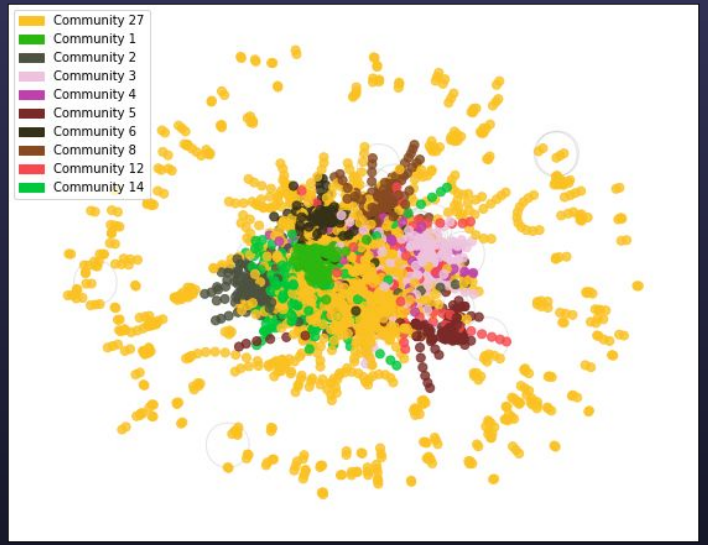
```
Community 3: ['nft', 'nfts', 'nftcommunity', 'tezos', 'giveaway', 'nftart', 'nftgiveaway', 'a
rt', 'nftdrop', 'nftcollection']
Community 1: ['eth', 'btc', 'bnb', 'binance', 'matic', 'xrp', 'doge', 'sol', 'usdt', 'altcoin
s']
Community 6: ['cryptocurrency', 'cryptonews', 'trading', 'news', 'invest', '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']
```

Interplay Layer Community

The top nodes for each community varied in terms of the cryptocurrencies and topics they were associated with.

HASHTAG-CRYPTO NETWORK

For example, the top nodes in Community 3, which are marked as light pink dots were related to NFT topics. This suggests that each community had its own distinct theme or focus.



Multilayer Network Visualisation with Community

HASHTAG-CRYPTO NETWORK

Node	Centrality
(bitcoin, 2)	2.122422
(ethereum, 2)	1.207983
(crypto, 1)	1.109626
(bitcoin, 1)	0.997708
(bnb, 2)	0.994843
(usdd, 2)	0.720015
(btc, 1)	0.536860
(dai, 2)	0.525974
(ethereum, 1)	0.525783
(xrp, 2)	0.508594

Centrality Measures

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.

HASHTAG-CRYPTO NETWORK

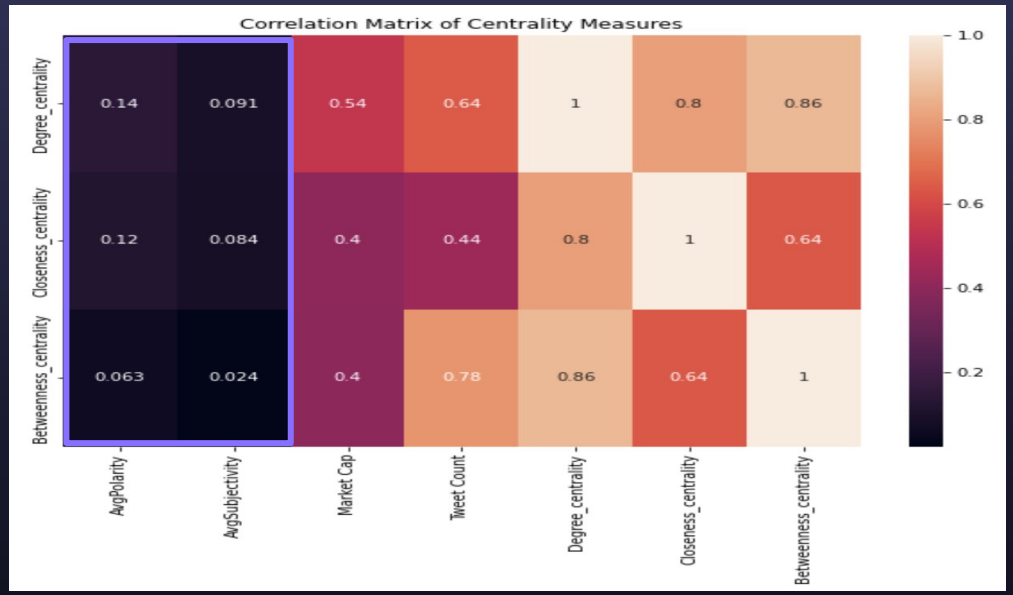
Node	Centrality
(bitcoin, 2)	2.122422
(ethereum, 2)	1.207983
(crypto, 1)	1.109626
(bitcoin, 1)	0.997708
(bnb, 2)	0.994843
(usdd, 2)	0.720015
(btc, 1)	0.536860
(dai, 2)	0.525974
(ethereum, 1)	0.525783
(xrp, 2)	0.508594

Centrality Measures

One possible reason is that both USDD and DAI are stablecoins. Stablecoins are cryptocurrency with stable value relative to other asset or currency, such as the US dollar or gold.

Stablecoins are often used as a bridge currency for trading between different cryptocurrencies, which could explain their strong connections with nodes in both layers of the network.

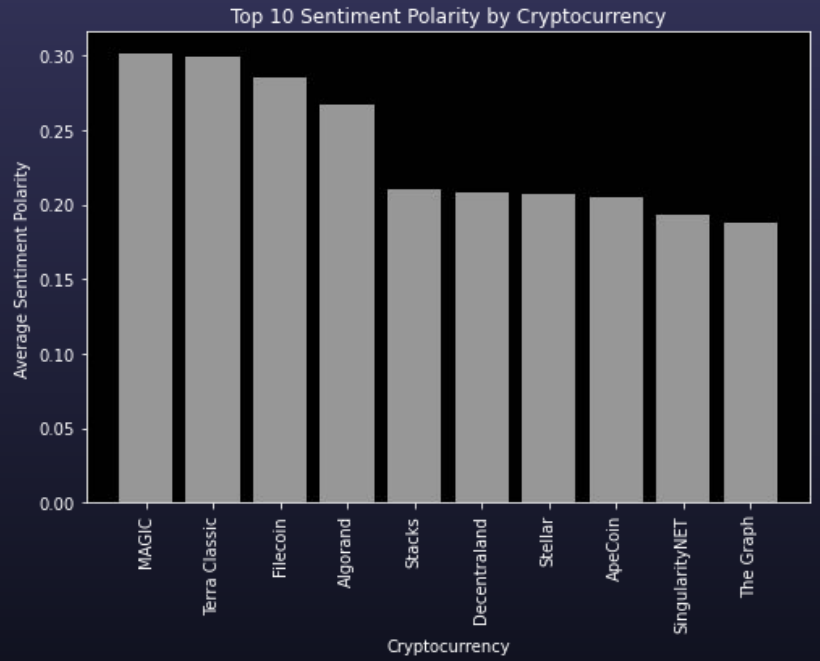
CENTRALITY CORRELATION ANALYSIS

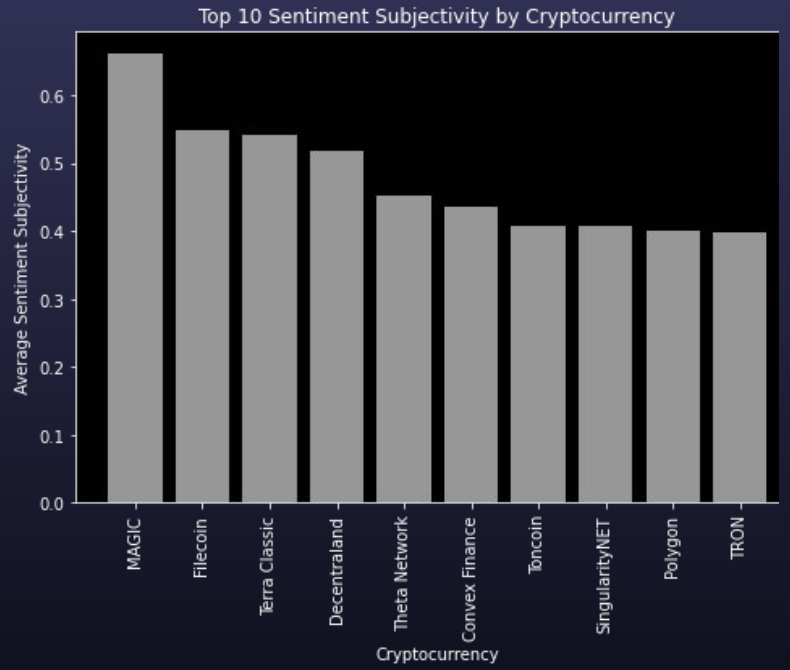


Correlationship Heatmap

- Weak positive relationship between the centrality and the sentiment measures (around 0.1)

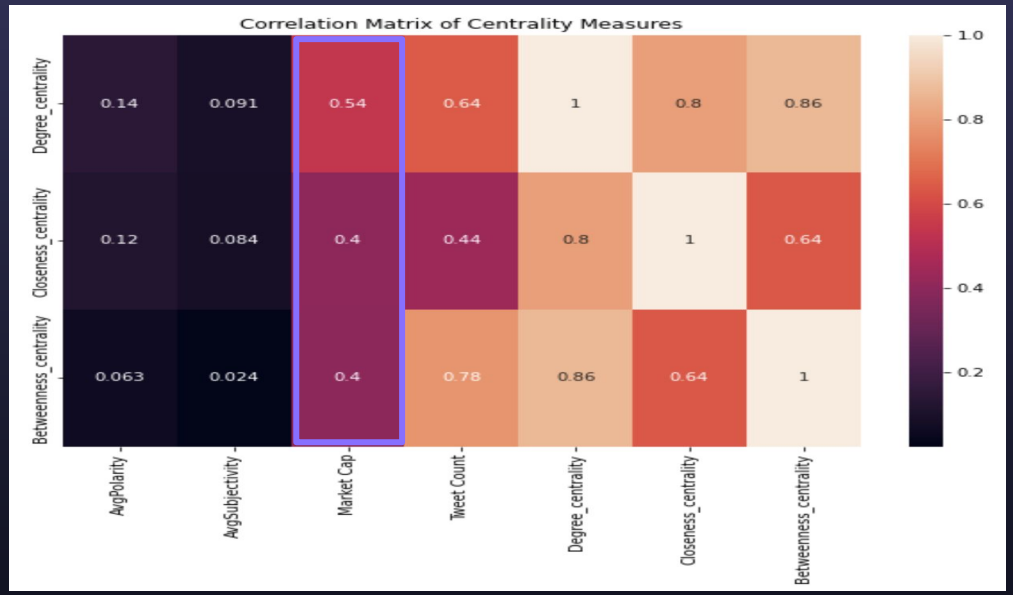
CENTRALITY CORRELATION ANALYSIS





Top 10 Ranked Crypto by Sentiment Score

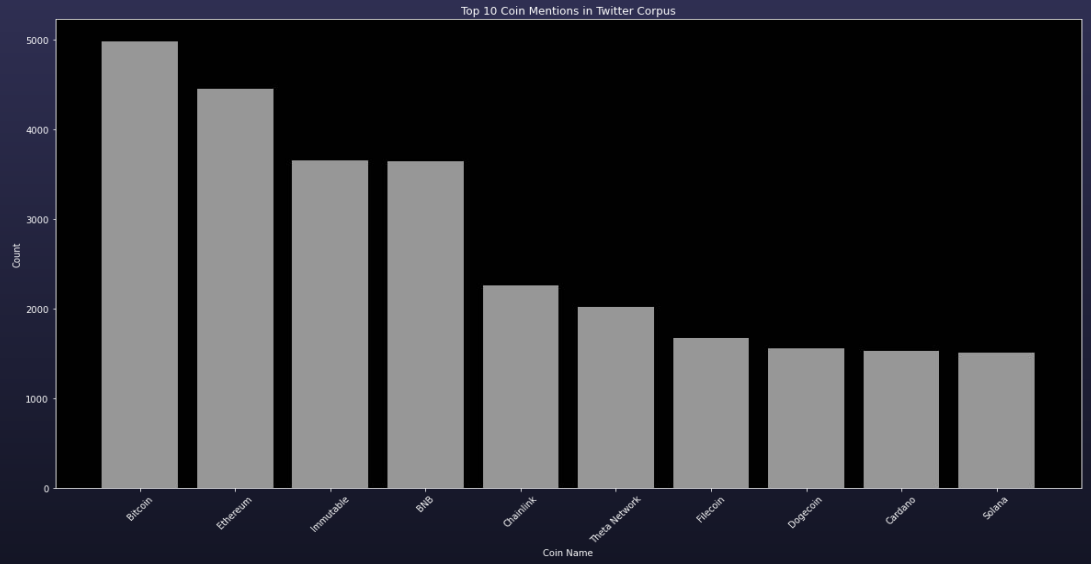
CENTRALITY CORRELATION ANALYSIS



Correlationship Heatmap

- Strong positive relationship between the degree centrality and market capitalization (0.54)

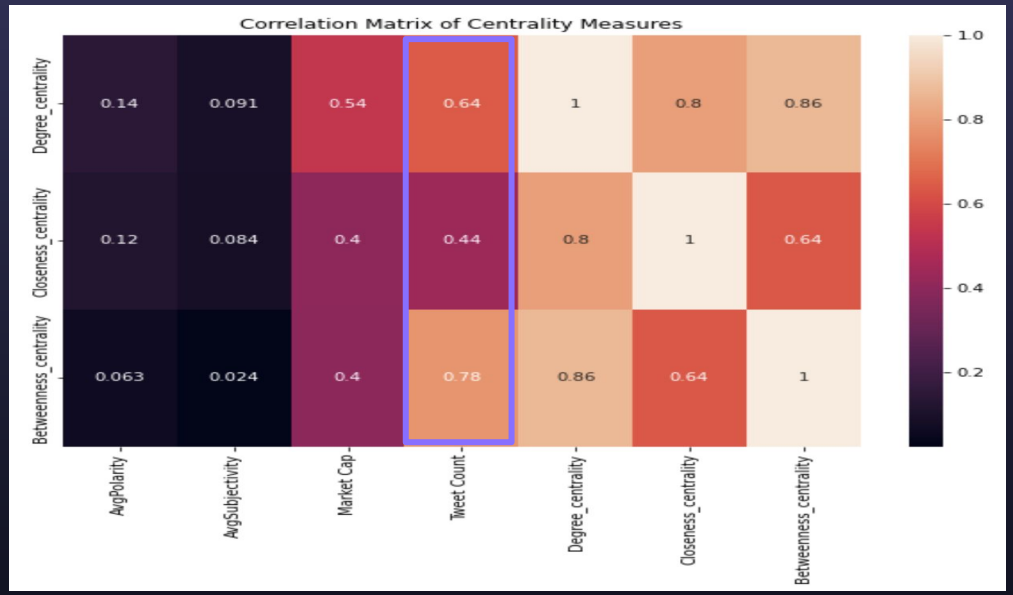
CENTRALITY CORRELATION ANALYSIS



Top 10 Ranked Crypto by Tweet Mentions

- Crypto with higher market cap are typically more established and have a larger user base, leading to more connections (degree).

CENTRALITY CORRELATION ANALYSIS



Correlationship Heatmap

- Strong positive relationship between the betweenness centrality and tweet count(0.78)

CENTRALITY CORRELATION ANALYSIS

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

Market Cap Rank

- Crypto frequently mentioned on Twitter may act as intermediaries in transactions and facilitate connections between different parts of the network.

04

CONCLUSION

LIMITATIONS & FURTHER WORK
DISCUSSION

LIMITATIONS

1. It was restricted to data from Twitter, which may not fully capture the broader cryptocurrency conversation on social media such as Reddit or Discord.

2. The data covers a relatively short period of 11 days, which limits our ability to fully understand the temporal dynamics of the networks and sentiment.

DISCUSSIONS

While there are limitations to the study, 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 businesses identify potential partnership opportunities and inform their product development and marketing strategies in the cryptocurrency market.

THANK YOU

QUESTIONS ARE WELCOMED