

## UNDERSTANDING BLOCKCHAIN THROUGH SOCIAL MEDIA

*Suhong Li, Bryant University, [sli@bryant.edu](mailto:sli@bryant.edu)  
Richard Glass, Bryant University, [rglass@bryant.edu](mailto:rglass@bryant.edu)*

### ABSTRACT

*Numerous studies have been written on how blockchain works and its potential applications. This study aims to further the understanding of blockchain through social media platforms. This study is based on 6,854,132 tweets from 942,123 Twitter users who have tweeted about this topic between May 2019 and December 2019. The results show that blockchain topic has been widely discussed in various countries and languages. The major topic discussed is cryptocurrencies, the original application of blockchain technology. A sentiment analysis shows that hashtags such as #airdrop, #giveaway, #bigdata and #tron are more favorable in the eyes of the public while the hashtags #cybersecurity, #data breach and #security received the lowest sentiment score. These results also reveal there may be an existing correlation between the market value of a certain cryptocurrency and sentiment on social media. In addition, a retweet network was created using Gephi to understand how Twitter users interact with each other and how messages travel on social media. Six communities emerged with a few key users in each community who were retweeted heavily by the rest of the users in the same group.*

**Keywords:** Blockchain, Social Media, Sentiment Analysis

### INTRODUCTION

The idea of blockchain originated in 2008 as a type of technology that supports Bitcoin protocol – a digital asset designed in a peer-to-peer network using cryptography to verify transactions without the involvement of intermediaries (Swan, 2015). While traditional payment processing is centralized and requires a third party – usually banks or credit card providers to verify the transactions, blockchain uses a trustless network of users who work on solving computational puzzles to verify the transactions. By providing a decentralized and tamper-proof environment, Blockchain is believed to enhance the security, data integrity, and transparency of transfer of information or assets (Nowiński and Kozma, 2017). Although the idea of blockchain appeared in 2008, research on blockchain and its applications did not evolve until 2014 because before then, people only thought of blockchain as an infrastructure that supported Bitcoin protocol (Casino et al., 2018).

Blockchain can be applied to different sectors including financial services, government, supply chain, Internet of Things (IoT), data management, and authentication verification (Casino et al., 2018). Financial institutions such as JP Morgan Chase started to implement Blockchain into their payment processing and eliminate their international transaction cost (Extance, 2015). Walmart implements IBM’s blockchain-based Food Trust to keep track of the origins of the food and make the production and supply chain of the food more transparent, easier, and quicker to access (IBM, 2019; Zuckerman, 2019). Identity verifications such as birth certificates, ownership of intellectual and physical properties, and land registries also utilize blockchain’s tamper-proof and transparent features to improve integrity and reduce fraud (Lemieux, 2016). IoT services also expect to improve their information security, reliability, privacy, and fraud traceability by integrating itself with blockchain to provide a more trusted sharing platform (Reyna et al., 2018). Blockchain is expected to improve people’s lives and reduce economic costs for companies, sectors, and countries as a whole. Despite the fraud-protection that blockchain provides, some researches have focused on potential security issues of blockchain technology. Phan et al.(2019) conducted an analysis based on blockchain fraud discussion on Twitter between November 6, 2018 and December 31, 2018 and found that the most frequently mentioned words in tweets regarding blockchain fraud include scam, tax, combat, fight, bitcoin, crypto, payment, Asia, Japan, Germany, Thailand, and banks. In addition, a sentiment analysis shows that the majority of the tweets (69%) on the discussion on blockchain fraud are negative.

This study aims to further the understanding of blockchain through Twitter, a social media platform. This study is based on 6,854,132 tweets from 942,123 Twitter users between May 2019 and December 2019 who have tweeted

about this topic. This study tries to understand blockchain technology from two perspectives. One perspective is Twitter users. For example, who is tweeting about blockchain? Where do they come from? What language do they speak? How long they have had a Twitter account? Who are the most active users in this topic? Who are the most influencing users in this topic? Another perspective is the content of Blockchain discussion. For example, what are people tweeting about this topic? What are the sentiments about this topic? What is sentiment on major cryptocurrencies? Is there a change of sentiment over time? Are there any communities/clusters of users emerging in this topic?

## LITERATURE REVIEW ON CRYPTOCURRENCIES AND BLOCKCHAIN APPLICATIONS

### Cryptocurrencies

Blockchain is the technology behind bitcoin. Today, the major blockchain based applications are cryptocurrencies. According to Yahoo Finance, there are approximately 5,392 cryptocurrencies being traded with a total market capitalization of \$201billion (as of April 22, 2020). Table 1 shows the top 10 cryptocurrencies in April of 2019 and April of 2020 respectively based on market capitalization. It can be seen that Bitcoin, Ethereum and Ripple/XRP has maintained the top 3 positions in both years. Tether’s market capitalization has tripled and represent the highest increase in market value followed by bitcoin, bitcoin cash and Ethereum. In contrast, ripple and EOS lost about 40% of market capitalization. In addition, Stellar Lumens and Cardano dropped out from the top 10 and were replaced by Bitcoin SV and Tezos. The detailed description of each cryptocurrency can be found in [www.coinbase.com](http://www.coinbase.com) website.

**Table 1.** Top 10 Cryptocurrencies in April 2019 and April 2020

| Top 10 Cryptocurrencies in April 2019 |                      |                       | Top 10 Cryptocurrencies in April 2020 |                    |                       | Ranking Change | % of Change |
|---------------------------------------|----------------------|-----------------------|---------------------------------------|--------------------|-----------------------|----------------|-------------|
| Ranking                               | Crypto.              | Market Capitalization | Ranking                               | cryptocurrency     | Market Capitalization |                |             |
| 1                                     | Bitcoin (BTC)        | 83.8bn                | 1                                     | Bitcoin (BTC)      | 128bn                 | No             | 52.7%       |
| 2                                     | Ethereum (ETH)       | 16.19bn               | 2                                     | Ethereum (ETH)     | 19.4bn                | No             | 19.8%       |
| 3                                     | Ripple (XRP)         | 13.87bn               | 3                                     | XRP (XRP)          | 8.22bn                | No             | -40.7%      |
| 4                                     | EOS (EOS)            | 4.18bn                | 8                                     | EOS (EOS)          | 2.4bn                 | No             | -42.6%      |
| 5                                     | Litecoin (LTC)       | 4.17bn                | 7                                     | Litecoin (LTC)     | 2.6bn                 | -2             | -37.6%      |
| 6                                     | Bitcoin Cash (BCH)   | 3.26bn                | 5                                     | Bitcoin Cash (BCH) | 4.1bn                 | -1             | 25.8%       |
| 7                                     | Binance Coin (BNB)   | 2.70bn                | 9                                     | Binance Coin (BNB) | 2.4bn                 | -2             | -11.1%      |
| 8                                     | Stellar Lumens (XLM) | 2.26bn                |                                       | Bitcoin SV (BSV)   | 3.4bn                 |                |             |
| 9                                     | Tether (USDT)        | 2.066bn               | 4                                     | Tether (USDT)      | 6.4bn                 | +5             | 209.8%      |
| 10                                    | Cardano (ADA)        | 2.04bn                | 10                                    | Tezos (XTZ)        | 1.5bn                 |                |             |

### Blockchain Applications

Blockchain originated as an infrastructure for Bitcoin – cryptocurrency in 2008. However, there are more applications across diverse sectors that are believed to be improved by blockchain technology such as financial services, supply chain, IoT, authentication verification, and data management (Casino et al., 2018).

**Financial Services:** A wide range of business sectors have researched and implemented blockchain into their system such as settlement of financial assets, cross-border payments, securities and derivative transactions (Casino et al., 2018). Ripple, with support from large banks such as Santander, ReiseBank, CIBC, UniCredit and 90+ other banks, utilized blockchain to provide a real-time interbank payment platform which will replace the current SWIFT system (Nowiński and Kozma, 2017). R3 (R3, 2015) also leads a consortium of the world's biggest banks including Barclays and Goldman Sachs and another 200+ firms to establish a distributed ledger in financial systems as well as other areas of commerce (Casino et al., 2018). Deloitte also worked on improving customer benefits by developing solutions such as Smart Identity which improves Know Your Customer (KYC) processes (Extance, 2015; Genkin et al., 2018). Blockchain could also be used to record and complete private securities transactions such as Linq, implemented by Nasdaq (Extance, 2015; Nowiński and Kozma, 2017).

**Authentication Verification:** With the tamper-proof feature that blockchain supports, the technology has been applied to verify the integrity or authentication of information (Casino et al., 2018). In the government sector, blockchain is used for identity verification such as passports, e-identity, birth certificates, or land registration (Reyna et al, 2018). Blockchain also helps prove and protect intellectual properties such as text-based, painting, music, and architecture (Zeilinger, 2016). For example, Ascribe was founded to create a permanent connection between the creator and his/her work, thus making it impossible to change or steal and preventing unauthorized access to the work (Shrier et al., 2016). Block Verify provides services to identify counterfeit goods or fraudulent activities and verify the provenance of luxury goods, pharmaceuticals, diamonds and electronics (Shrier et al., 2016). Especially in countries where the management of data is poor, blockchain helps provide authentication verification because once the data is recorded and added to the blockchain, it is immutable and tamper-proof. This prevents corrupting and fraudulent/stealing activities from occurring.

**Internet of Things (IoT):** Reyna et al (2018) states that blockchain could enrich the IoT with its transparent feature because it makes it easier to trace back activities, thus enhancing security. Moreover, a decentralized peer-to-peer IoT system is expected to allow a higher control of IoT services to keep track of the flow of information, solve the problem related to high maintenance cost caused by the centralized systems, and enable the automated processing of goods and services (Casino, 2018). Blockchain can also improve some sectors of the IoT like how a new IoT E-business model proposed by Zhang et al. (2015) in which Distributed Autonomous Organization (DAO) allows users to obtain coins and exchange sensor data through a blockchain environment (Zhang et al., 2015; Zheng et al., 2018). IBM also uses its proof of concept for Autonomous Decentralized Peer-to-Peer Telemetry, which allows smart-home owners to identify operational issues and update the software by themselves (Zheng et al., 2018).

**Supply Chain:** Blockchain structure improves the transparency and accountability in supply chain, which will increase the productivity and add values to businesses (Ahram et al., 2017). Walmart is expected to implement IBM's Food Trust, which utilizes blockchain to connect the supplies (including growers, processors, distributors) to the customers through a "permissioned, permanent and shared record of food system data" (IBM, 2018). Now with blockchain, the retailers and customers can keep track of where the products come from. If anything happens, it is easy to trace back to the supply chain. Blockchain technology in supply chains eliminate the intermediaries between sellers and buyers (Casino et al., 2018). Blockchain helps to improve robust contract management for fighting information asymmetry among more than two-party logistics and improves communication and transparency across the entire supply chain, food safety, and thus provides better customer services (Casino et al., 2018).

**Data Management:** Data is usually stored and controlled at a trusted party, which is very susceptible to attacks. Many incidents have occurred over the past years, for example, Facebook and its data breach of 540 million users, or Google with its exposure of data of 52.5 million users, and other companies such as Cathay Pacific (9.4 million) and Exactis (340 million) (Leskin, 2018). Zyskind et al. (2015) proposed a blockchain-based data management system that help protect data from privacy issues such as data ownership, transparency and fine-grained access control (Zheng et al., 2018).

## **RESEARCH METHODOLOGY**

The focus of this study is to understand how blockchain is discussed in social media. A Twitter listener using python was created to collect tweets using blockchain as keywords and the listener was run continuously between May 2019

and December 2019. During this period, about 40GB data was collected and includes 6,854,132 tweets from 942, 123 Twitter users. Because of size of the data, it is impossible to process it on a single machine using traditional application. The tweets were uploaded to Amazon S3 storage and was analyzed using the Databrick Spark platform. Using pyspark, the tweets (in JSON format) was read into a spark data frame, cleaned and analyzed using various functions.

## **DATA ANALYSIS AND DISCUSSION**

This section will first discuss the type of users who tweeted blockchain from joining year, language, and number of followers/following, followed by a discussion of major blockchain topics discussed in the Twitter platform. For example, what are the top hashtags in blockchain discussions? What are the sentiments about this topic? What are the sentiments on major cryptocurrencies? Is there a change of sentiment over time? Are there any sub-communities of users in this topic?

### **Twitter User Analysis**

Table 2 shows 400 users joining Twitter in 2006, the year that Twitter was launched, that tweeted about blockchain topic. The number jumped dramatically in 2009, then stayed the same with some minor floatation between 2010 and 2017. It began to increase greatly in 2018 and reached its peak in 2019. There are about a quarter million users who joined Twitter in 2019 and tweeted about blockchain topics. This result may indicate an increased interest in blockchain topic discussion starting in 2018.

Table 3 shows a user' self-reported language. Over 90% of users do not report language in their profile. For people reported, the majority of them are English, followed by Russian, Spanish, French, Indonesian, German, Japanese, Turkish, Italian and Vietnamese. This indicates that the majority of blockchain topic is discussed among these languages.

Table 4 shows that about half of the users have less than 10 followers in Twitter, about one third of the users have between 100 and 1000 followers, and about 15% have between 1000 and 10,000 followers. It can be seen that majority of them have less than 10,000 followers. Table 5 shows that one third of all the users are following less than 100 users, about 46% of all the users are following between 100 and 1000 users, and about 17% of all the users follow between 1000 and 10,000 Twitter accounts. Less than 1% of all the users have more than 100,000 followers/friends.

A pyspark script was created to find the top 10 users who tweeted the most. The results show that five accounts (CodingUK, coolLaTechJobs, Salesukwork, salesjobsusa3 and StartupsBot) were suspended by Twitter due to the violation of Twitter rules. The rest of the five accounts (Theautomatski, Airdropnotecom, XelotX, cybersec\_feeds and Coinhasyou) are platform that post/repost blockchain related news/research. To fight against malicious activity targeting the crucial conversations taking place on Twitter — including elections in the United States and around the world the new rules of tweets disallows simultaneously posting identical or substantially similar content to multiple accounts or simultaneously performing actions such as likes, retweets, or follows from multiple accounts. In addition, the new rules prohibit the use of any form of automation (including scheduling) to post identical or substantially similar content or to perform actions such as likes or retweets across many accounts that have authorized your app.

The similar analysis was used to find the top 10 users who received the most retweets. Those users can be considered as the most influential since retweets reflect a user's ability to generate original high value content. The results show there are 7 top users in the USA and other three users are in India (reach2ratan), Hong Kong (RefineMedium) and Thailand (Tanawatofficial) respectively. For those users in USA, they either offer a cryptocurrency (IOSToken, Justinsuntron, Murthaburke, and jimmyzhong\_iost), act as a cryptocurrency trading exchange (p2pb2b), or provide news related to cryptocurrency (airdropnotecom and ICOAnnouncement).

**Tweets Analysis**

There were 6,854, 132 tweets collected. Among them, 45% (3,081,245) are original tweets and the rest 55% (3,771,887) are retweets. The results show there are more retweets than tweets.

**Table 2.** Number of Users by Joining Year

| Joining Year | # of Twitter Users |
|--------------|--------------------|
| 2006         | 400                |
| 2007         | 6,759              |
| 2008         | 17,475             |
| 2009         | 68,758             |
| 2010         | 56,851             |
| 2011         | 63,716             |
| 2012         | 54,802             |
| 2013         | 47,799             |
| 2014         | 47,339             |
| 2015         | 45,438             |
| 2016         | 50,139             |
| 2017         | 75,478             |
| 2018         | 116,725            |
| 2019         | 290,836            |

**Table 3.** Number of Twitter User by Language

| Language   | # of Twitter Users |
|------------|--------------------|
| N/A        | 904453             |
| English    | 70642              |
| Russian    | 4980               |
| Spanish    | 4892               |
| French     | 3284               |
| Indonesian | 3000               |
| German     | 1920               |
| Japanese   | 1411               |
| Turkish    | 1405               |
| Italian    | 1181               |
| Vietnamese | 1063               |
| Other      | 4470               |

**Table 4.** Number of Users by Followers

| Number of Followers of Each User | Number of Users | % of User |
|----------------------------------|-----------------|-----------|
| Less than 100                    | 484535          | 51.43%    |
| Between 100 and 1000             | 292394          | 31.04%    |
| Between 1000 and 10,000          | 139534          | 14.81%    |
| Between 10,000 and 100,000       | 23122           | 2.45%     |
| Between 100,000 and 1,000,000    | 2281            | 0.24%     |
| Greater than 1,000,000           | 268             | 0.03%     |

**Table 5.** Number of Users by Friends

| Number of Friends of Each User following | Number of Users | % of User |
|--|-----------------|-----------|
| Less than 100                            | 334352          | 35.49%    |
| Between 100 and 1000                     | 433296          | 45.99%    |
| Between 1000 and 10,000                  | 165874          | 17.61%    |
| Between 10,000 and 100,000               | 8305            | 0.88%     |
| Between 100,000 and 1,000,000            | 303             | 0.03%     |
| Greater than 1,000,000                   | 4               | 0.00%     |

***Tweets by Week Day and by Hour***

Figure 1 shows that most people tweeted on Tuesday and Thursday. The chart shows that people tweeted less on the weekend (Saturday and Sunday). In addition, there are more tweets generated around the middle of day (hours 12-16) and less tweets during the night (hours 20-23, hours 0-4).

### Sentiment Analysis on Blockchain Discussion

The AFINN lexicon is a list of English terms manually rated for valence with an integer between -5 (negative) and +5 (positive) by Finn Årup Nielsen between 2009 and 2011. It currently includes 2,477 words. For example, words such as superb, outstanding, and excellent receive a 5 rating, while words such as bastard and prick receive a -5 score.

A sentiment analysis with AFINN Lexicon was conducted based on the original tweets. First, each tweet was split into words, then all sentiment words were extracted from each tweet, and then the average score of all sentiment words were calculated for each tweet. A pyspark script was created to conduct the analysis and the results show that the majority of tweets (84%, 1,564,704) are positive (with a score greater than zero), about 14% (253,648) are negative and the rest are 2% (45,704) and are neutral.

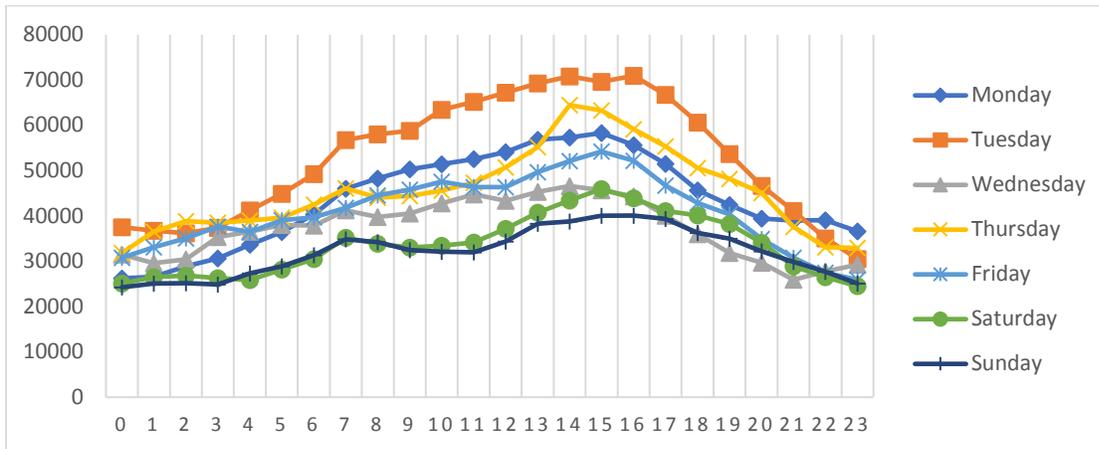


Figure 1. Number of Tweets by Week day and Hour

A further analysis was conducted to identify top hashtags based on frequency and sentiment. The first column of Table 6 shows the top 20 hashtags in blockchain topics. It can be seen that the top 3 hash tags are #crypto, #bitcoin and #eth and they are all related to cryptocurrencies. The hashtags can be classified into four categories. The first category is hashtags related to cryptocurrencies which include #crypto, #bitcoin, #eth, #binance, #tron, and #xrp; the second category represents blockchain based platform/services/opportunities such as #isot, ico, #airdrop, #giveaway, # jobs, and #startup; the third category is for hashtags relating to other emerging technologies including #ai, #fintech, #iot, and #bigdata; and the last category is hashtags related to security which include hashtags such as #cybersecurity, #security, and #databreach. In addition, #satoshi was mentioned more than 45,000 times. Satoshi can be used to represent Bitcoin's smallest, indivisible part. It can also be used to refer to Satoshi Nakamoto, the name used by the presumed pseudonymous person or persons who developed bitcoin. It can be concluded that the dominant discussions on blockchain are still associated with cryptocurrencies, especially bitcoin and Ethereum. In addition, blockchain security has also received a lot of attention on Twitter.

Regarding the overall sentiment score of each hashtag (see Table 6), #airdrop received the highest score, followed by #giveaway, #satoshi, #bigdata and #tron, showing people have more favorable view toward those topics. In contrast, #cybersecurity, #databreach, #fintech, #binance, and #security received lowest sentiment score, showing a more negative view of those topics.

In addition, a sentiment analysis was conducted to evaluate the frequency and sentiment of 12 leading cryptocurrencies (discussed in pervious section of this paper) from May to December of 2019 and the results are shown in Figure 2 and Figure 3. From Figure 2, it can be seen that Bitcoin and Ethereum were tweeted the most followed by Ripple and EOS.

It is interesting to note that there are more tweets in June, September and November, and less tweets in the summer (July and August) and Christmas period (December).

Figure 3 shows the overall sentiment of the top 10 cryptocurrencies. It can be seen that tether has the highest sentiment score in August of 2019 and December of 2019. It should be noted Tether also has the highest increase in market capitalization from \$2.6 billion in April, 2019 to \$6.4 billion in April, 2020. Cardano has the highest sentiment score in June 2019 and its sentiment score has dropped since July of 2019. It is no longer in the top 10 cryptocurrencies list as of April of 2020. In addition, the sentiment score of Tezos rose and reached the highest sentiment score among all cryptocurrencies in October of 2019 and it climbed to the 10<sup>th</sup> position in April, 2020. The results suggest there may be an existing correlation between the market value of crypto-currencies and the amount of sentiment on social media.

**Table 6.** Top 20 hashtags and Overall Sentiment Rating of Each Hashtag

| <b>Hashtag</b> | <b>Number Of Tweets</b> | <b>Overall Rating</b> | <b>% of Negative</b> | <b>% of Neutral</b> | <b>% of Positive</b> |
|----------------|-------------------------|-----------------------|----------------------|---------------------|----------------------|
| #crypto        | 686009                  | 1.55                  | 12.4%                | 1.9%                | 85.7%                |
| #bitcoin       | 442443                  | 1.33                  | 18.9%                | 2.8%                | 78.4%                |
| #eth           | 161098                  | 1.38                  | 16.5%                | 1.7%                | 81.8%                |
| #airdrop       | 130562                  | 2.39                  | 3.7%                 | 0.5%                | 95.8%                |
| #ai            | 120075                  | 1.47                  | 11.2%                | 2.8%                | 86.0%                |
| #tron          | 112968                  | 1.68                  | 6.7%                 | 2.8%                | 90.5%                |
| #fintech       | 102074                  | 1.20                  | 14.3%                | 5.5%                | 80.2%                |
| #giveaway      | 95653                   | 2.25                  | 1.6%                 | 5.5%                | 92.9%                |
| #ico           | 91122                   | 1.41                  | 13.0%                | 3.6%                | 83.4%                |
| #iot           | 85231                   | 1.41                  | 13.5%                | 1.9%                | 84.6%                |
| #cybersecurity | 59741                   | -0.43                 | 62.3%                | 3.7%                | 34.0%                |
| #satoshi       | 45675                   | 1.93                  | 5.2%                 | 1.7%                | 93.0%                |
| #bigdata       | 40937                   | 1.69                  | 10.2%                | 1.2%                | 88.6%                |
| #xrp           | 39037                   | 1.41                  | 17.7%                | 5.2%                | 77.1%                |
| #binance       | 30126                   | 1.24                  | 14.2%                | 1.0%                | 84.8%                |
| #jobs          | 30017                   | 1.44                  | 11.7%                | 1.9%                | 86.4%                |
| #security      | 26541                   | 1.27                  | 15.2%                | 4.0%                | 80.9%                |
| #databreach    | 25859                   | 0.32                  | 34.0%                | 2.1%                | 63.8%                |
| #iost          | 25598                   | 1.46                  | 14.2%                | 0.0%                | 85.7%                |
| #startup       | 24046                   | 1.42                  | 15.5%                | 3.5%                | 81.0%                |

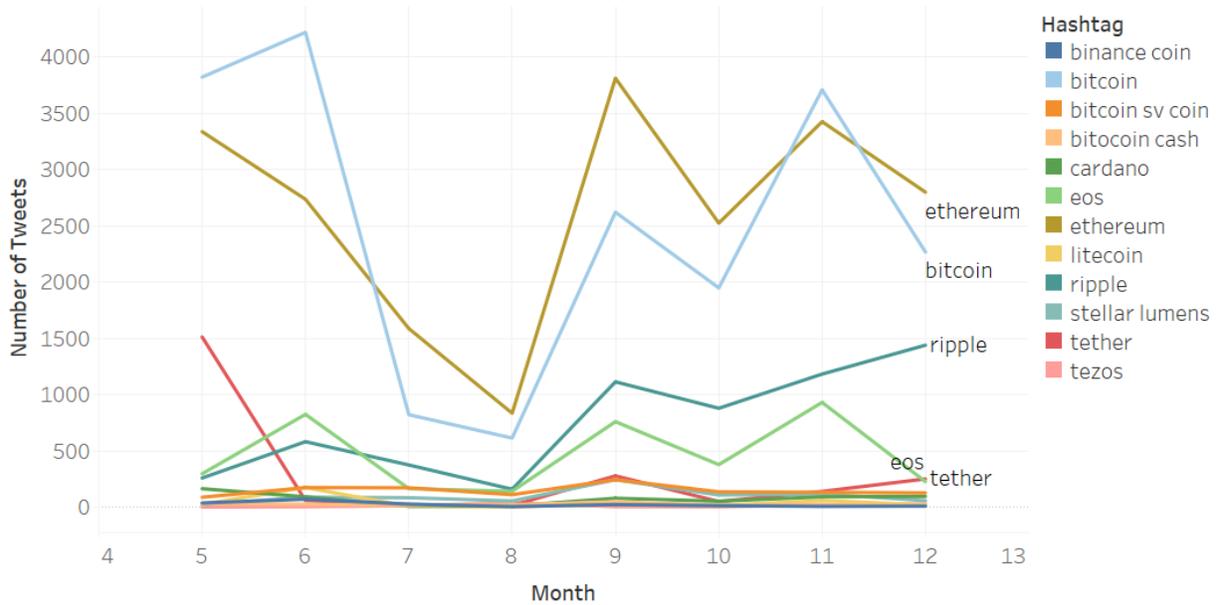


Figure 2. Total Number of Tweets for Top 10 Cryptocurrencies between May and December 2019

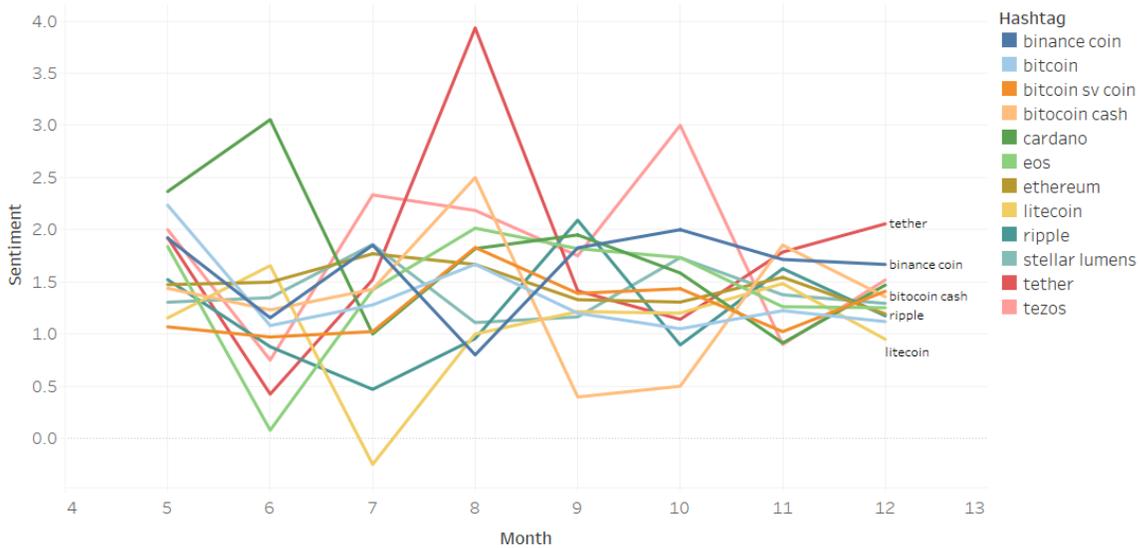


Figure 3. Overall sentiment of selected crypto-currency between May and December of 2019

### Blockchain Retweet Network analysis

From tweets, we extract the number of retweets received by each user from another user using pypark. A retweet network was created using Gephi, an open-source network analysis and visualization software. In this network, each node represents one Twitter user and the edge (the link between two nodes) represents the number of retweets from one user to another. For simplicity, we only included the top 10,000 edges based on the number of retweets and it ended up representing 7,812 tweet users. Using Fruchterman Reingold as layout, Figure 4a shows the retweet network generated from the dataset. The size of the node represents in degree (the number of retweets received by that node).

The nodes in the same color belong to the same community/group. It can be seen that six communities exist in blockchain topic discussions. The biggest community is in brown, followed by blue, green, orange, pink and dark green. It is interesting to see all the users in the same community all connect to one or two key users in the center and however there is a very limited connection from one community to another.

The original network was filtered to show the top 6 communities and the result is presented in Figure 4b. The chart shows that there are two central users in the brown community. They are ISOToken and jimmyzhong\_iost. Jimmy Zhong is the co-founder of IOSToken. This group represents the follower of IOST, a blockchain technology based platform that provides many online services.

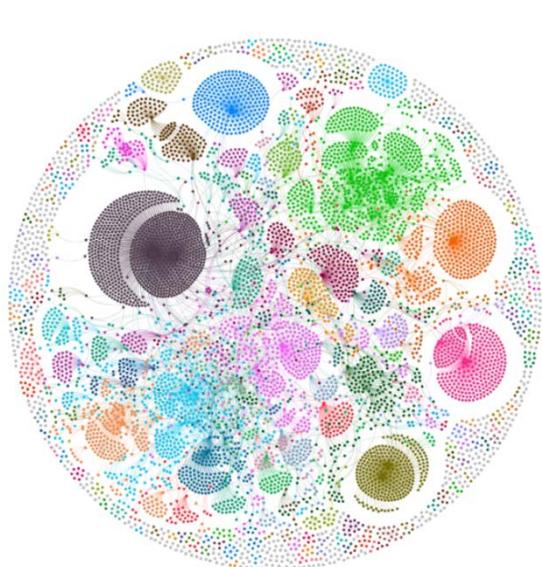
The green cluster has a few dominant users including Soraix, HiveNetCloud, Xeratech, MVSDNA and Xeratech. They represent emerging cryptocurrency trading platforms. For example, Soraix is the world's first trading platform where investors can acquire digital shares of companies entering the stock exchange via the Equity Token Offering. HiveNetCloud aims to building the next Generation of Cloud Computing through utilizing computers during idle times. Xeratech is a community-driven exchange where they listen to feedback & introduce fresh new ideas to make the trading of cryptocurrency a simple straight forward experience. MVSDNA builds the next-generation of interoperable, scalable blockchain technology.

The center node in the orange cluster is Authpaper Coin. According to its website, Authpaper Delivery combines blockchain, BT, cryptography and other current technologies to build a peer-to-peer digital data delivery platform. All data delivered is confidential except to the specified recipients. Data is verified to be unchanged down to every single bit.

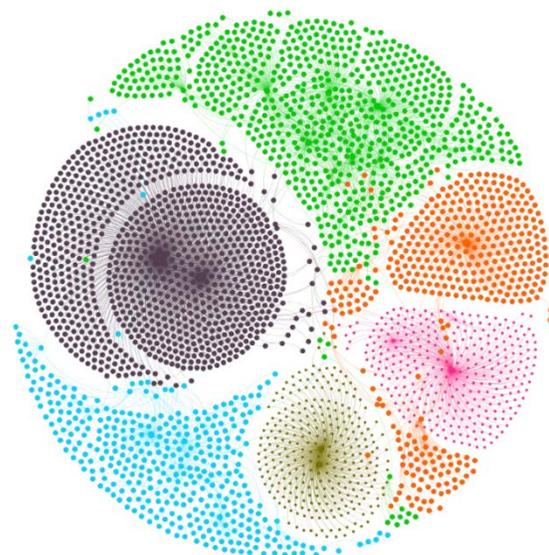
There are two central nodes in the pink cluster: p2pb2b and EDCBlockchain. P2pb2b is a cryptocurrency exchange and EDCBlockchain is a multifunctional blockchain platform that provides tokenization capabilities for small and medium businesses, functionality for fast and anonymous transactions, innovative solutions for loyalty programs and project scaling.

In the dark green community, the two center nodes are STPIINDIA and Omkar\_Raii. STPIINDIA is the account for Software Technology Parks of India (STPI) and Omkar Rail is the Director of STPI. This group represents a specific blockchain group in India. In the blue cluster, there are two central nodes: cointelegraph and coindesk. They are all news platforms for blockchain. Cointelegraph provides high-tech finance, Bitcoin and Blockchain news, analysis and review. CoinDesk is the media platform for the next generation of investors. It explores how cryptocurrencies and digital assets are contributing to the evolution of the global financial system, events, and education.

Therefore, it can be concluded that the major groups in blockchain topic discussions are companies that provide blockchain-based technology platforms, blockchain trading platforms, or news and analysis related to cryptocurrencies and Blockchain.



**Figure 4a** Retweet Network of Blockchain Discussion



**Figure 4b.** Top 6 Clusters in Retweet Network

## CONCLUSION

This study offers valuable insight into Blockchain and cryptocurrencies from the discussions on social media platforms. The results show that blockchain topics have been discussed not only by the users in the United States but also in other countries and other languages. The major topics in blockchain still focus on cryptocurrencies the original application of blockchain. Blockchain can be classified into four categories: cryptocurrencies (bitcoin, eth, binance, and tron), blocked-chain based platform/services/applications (isot, ico, and fintech), emerging technologies (artificial intelligence, Internet of things, and big data) and cybersecurity. In addition, a sentiment analysis shows that people are more favored towards hashtags including airdrop, giveaway, big data and tron; while cybersecurity, data breach and security received the lowest sentiment score. The results also suggest there may be an existing correlation between market capitalization of crypto-currencies and sentiment on social media. A retweet network was created using Gephi to understand how Twitter users interact with each other and how messages travel on social media. Six communities emerged with a few key users in each community who were retweeted heavily by the rest of the users in the same group. It is interesting to note that 5 out of 10 of the accounts that tweeted the most in 2019 were suspended by Twitter due to violation of the Twitter rules.

Future study can include tweets from multiple years to see how the discussion of Blockchain evolves over the years. Future study can also explore how the discussion of blockchain topics on social media impact price change of cryptocurrencies. In addition, sentiment analysis can be improved by analyzing combinations of words instead of just a single word.

## REFERENCES

- Ahram, T., Sargolzaei, A., Sargolzaei, S., Daniels, J., Amaba, B., 2017. Blockchain technology innovations. *IEEE Technology and Engineering Management Society Conference, TEMSCON 2017*pp. 137–141.
- Casino, F., Dasaklis, T. K, Patsakis, C., 2018. A systematic literature review of Blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, Volume 36, 2019, Pages 55-81, ISSN 0736-5853. Retrieved from <https://doi.org/10.1016/j.tele.2018.11.006>.

- Extance, A., 2015. The future of cryptocurrencies: Bitcoin and beyond. *Nature*, 526(7571), 21–23. <https://doi.org/10.1038/526021a>
- Genkin, D., Papadopoul, D., & Papamanthou, C., 2018. Privacy in Decentralized Cryptocurrencies. *Communications of the ACM*, 61(6), 78–88. <https://doi.org/10.1145/3132696>
- IBM, 2019. Retrieved. from: <https://www.ibm.com/Blockchain/solutions/food-trust>
- Lemieux, L. V., 2016. "Trusting records: is Blockchain technology the answer?", *Records Management Journal*, Vol. 26 Issue: 2, pp.110-139, <https://doi.org/10.1108/RMJ-12-2015-0042>
- Leskin, P. 2018. The 21 scariest data breaches of 2018. *Business Insider*. Retrieved from <https://www.businessinsider.com/data-hacks-breaches-biggest-of-2018-2018-12>
- Mavridou, A., Laszka, A., 2018. Tool demonstration: FSolidM for designing secure ethereum smart contracts. CoRR abs/1802.09949.
- Nowiński, W. & Kozma, M., 2017. How Can Blockchain Technology Disrupt the Existing Business Models? *Entrepreneurial Business and Economics Review*. 5. 10.15678/EBER.2017.050309.
- Phan, Linh, Li., S., Mentzer, K. (2019), "Blockchain Technology and the Current Discussion on Fraud", *Issues in Information Systems*, 20(4), pp.8-20.
- Reyna A., Martín C., Chen J., Soler E., Díaz M., (2016). On Blockchain and its integration with IoT. Challenges and opportunities, *Future Generation Computer Systems*, Volume 88, 2018, Pages 173-190, ISSN 0167-739X. Retrieved from <https://doi.org/10.1016/j.future.2018.05.046>.
- Shrier D., Wu W., Pentland A., 2016. Blockchain & Infrastructure (Identity, Data Security). Connection Science & Engineering. *Massachusetts Institute of Technology*. Retrieved on April 04, 2019 from [https://www.getsmarter.com/career-advice/wp-content/uploads/2017/07/mit\\_Blockchain\\_and\\_infrastructure\\_report.pdf](https://www.getsmarter.com/career-advice/wp-content/uploads/2017/07/mit_Blockchain_and_infrastructure_report.pdf)
- Swan, M. and de Filippi, P. (2017), Toward a Philosophy of Blockchain: A Symposium: Introduction. *Metaphilosophy*, 48: 603-619. doi:[10.1111/meta.12270](https://doi.org/10.1111/meta.12270)
- Yahoo Finance, <https://finance.yahoo.com/news/top-10-cryptocurrencies-market-capitalisation-160046487.html>, retrieved on June 7<sup>th</sup>, 2020.
- Zeilinger M., 2016, "Digital art as 'monetised graphics': Enforcing intellectual property on the Blockchain", *Philosophy Technol.*, vol. 31, no. 1, pp. 15-41, 2016.
- Zhang, Y., Wen, J., (2015). An iot electric business model based on the protocol of bitcoin. Proceedings of 18th International Conference on Intelligence in Next Generation Networks (ICIN). pp. 184–191. Paris, France (2015).
- Zheng, Zhibin & Xie, Shaoan & Dai, Hong-Ning & Chen, Xiangping & Wang, Huaimin. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*. 14. 352. 10.1504/IJWGS.2018.095647.
- Zuckerman, M. J. 2019. Walmart and IBM Blockchain Initiative Aims to Track Global Food Supply Chain. Retrieved from <https://cointelegraph.com/news/walmart-ibm-Blockchain-initiative-aims-to-track-global-food-supply-chain>

Zyskind, G., Nathan, O., et al., 2015. Decentralizing privacy: Using Blockchain to protect Personal Data. *Security and Privacy Workshops IEEE*, 180-184, 2015.