An Analysis of the Evolution of the Twitter Discussion of Non-Fungible Tokens

Francois Cilliers
Department of Information Science
Stellenbosch University, South Africa
20073445@sun.ac.za

Johanna Engelhard
Department of Information Science
Stellenbosch University, South Africa
22098038@sun.ac.za

Gina N. Lamprecht
Department of Information Science
Stellenbosch University, South Africa
ginalamp@sun.ac.za

ABSTRACT

The non-fungible token (NFT) market has increased exponentially over the last year. We capture the largest community within the NFT discussion on Twitter from 1 February to 31 May 2021 using NetworKit's Louvain algorithm. Its data is explored by investigating tweet frequency over time, and applying BTM topic modelling and VADER sentiment analysis. The discussion is kick-started in the middle of March 2021, and reaches an all-time high on the 5th of May. NFT projects are the largest topic in the conversation. The average sentiments in the largest community and the largest topic are found to be positive throughout the period of analysis, with the largest topic being more positive than the overall sentiment of the community. By addressing the question of how the sentiment within the NFT discussion for topics changes over time, we gain insight into the overall standing and development of the NFT discussion on Twitter during this time.

KEYWORDS

Twitter, cryptocurrencies, NFT, non-fungible token, sentiment analysis, topic modelling, community detection

1 INTRODUCTION

The rise of cryptocurrencies during the Covid-19 pandemic [23] saw investors, hit by the market recession, delve into new areas of e-commerce not affected by the pandemic. Understanding how this new venture was received by the very active cryptocurrency Twitter sphere is key to determining the interest, validity and longevity of non-fungible tokens.

A non-fungible token (NFT) is defined as a "unit of data . . . that certifies a digital asset to be unique and therefore not interchangeable, while offering a unique digital certificate of ownership" [29]. These include digital art, such as images, videos, and music, as well as virtual gaming assets, software licences, and physical assets such as luxury goods and cars [34]. While these digital assets can be duplicated, the original is unique and often can have collectors' value [9]. As with regular art [8], their value cannot be measured objectively, but is rather set individually by the highest bidder. NFTs can also include contracts that allow artists to earn royalties when their original artwork changes ownership or is exhibited in virtual museums. Similar to cryptocurrencies [30], NFTs are recorded by a blockchain, which serves as a digital tamper-resistant public ledger that ensures legitimacy [34].

NFTs have been traded since 2017, but started to experience an increase in popularity in 2021 when NFT sales exploded in February and March [4]. Throughout the previous year, monthly NFT sales averaged around $1 million on OpenSea, a major NFT marketplace. In January 2021, the sales volume was around $8 million, followed by $95 million in February [19]. The total market value of NFTs in March 2021 was $550 million, with $200 million traded in the month of March alone [15].

Little is currently known about the overall structure of the market and it's evolution [29]. Critics say NFTs are the next big craze losing all its worth when the hype dies down. Others believe this is the future of ownership, where the ownership status of all kinds of digital and non-digital assets — from event tickets to houses — will eventually be tokenized [9].

In light of the sudden surge and current activity on the NFT market, this study will explore the question of how the sentiment within the NFT discussion for topics changes over time. To do this, this study will investigate the sentiment and overall topics that are discussed relevant to NFTs to allow further insight into the overall standing and development of the NFT discussion on Twitter from 1 February to 31 May 2021.

2 BACKGROUND LITERATURE

In order to address the research objective, the study aims to utilize community detection, topic modelling and sentiment analysis to gain insight into the discussion on Twitter surrounding NFTs.

Our scope is limited to Twitter, a microblogging service. We have two key reasons for this choice. First, according to a paper by Irsyad and Rakhmawati [22], "Twitter is now considered as one of the fastest and most popular communication media and is often used to track current events or news". Second, Twitter's content is publicly available, making it easier to access than other sites such as Facebook [2], which has more rigid security protocols, and with its user activity being much higher compared to other sites such as Reddit [10].

There exist commonly used and readily available tools that are utilized to perform community detection [28], topic modelling [43] and sentiment analysis [44]. Since we chose Twitter as our source for this study, the data we want to explore has attributes that are typical of social media. For example, short texts and slang usage are characteristic. This means that certain methods of community detection, topic modelling and sentiment analysis are more suitable to our research than others. We carefully evaluated the available options for each of the three analyses to determine which tools are most appropriate to this study.

2.1 Community detection

According to Papadopoulos et al. [33] and Traag et al. [39], community detection is a useful methodology for analysing the structure of complex networks. It has been used in multiple disciplines, including sociology, biology and computer science [16]. In social networks, a relevant problem is to discover user communities that share common characteristics [37], and the development of the techniques used for community detection of Twitter data has resulted in new ways to gain valuable information [3, 26].

Community detection is a graph-based problem that looks for groups of vertices that are more closely connected to each other than the rest of the graph [3, 14, 26]. In social network analysis, explicit relationships, such as structure-based relationships and interaction-based relationships [14, 26], are commonly used for community detection. Structure-based following-follower network relationships stay relatively static on small timescales, while interaction-based network relationships form a more dynamic picture of community structure. There are also other types of communities, such as topic-based communities, which consist of users who talk about the same things, and sentiment-based communities, where users feel the same way about certain topics [26]. Furthermore, there are also multifaceted question-oriented community detection methods which combine the different methods [14].

2.1.1 Twitter Networks. Twitter has multiple interaction-based network relationships, including tag, reply, retweet and quote relationships. Twitter saves all of these interactions in the data obtained from the Twitter API by means of the `@username` mention functionality, along with an extra field in the tweet object to define the type of interaction [36]. Therefore, this paper uses the mention functionality within the tweet object to create the interaction-based network relationships for community detection [13, 42]. Since it is also possible to mention multiple users in a single tweet, these interactions can then be used to increment existing edge weights, or establish new edges [26].

2.1.2 Algorithms. Now that we have established the type of the network on Twitter that we are interested in, we require an algorithm to perform community detection within these networks. There are a couple of available computational approaches to achieve this, including traditional methods such as hierarchical clustering, divisive methods such as Girvan and Newman's algorithm, and modularity-based methods such as greedy and extremal optimization [16].

For the objective of this study, there are three key requirements of a community detection routine. Firstly, it must be fast and efficient as a slow algorithm would be technically infeasible within the scope of this study. Secondly, the method must be suitable for large graphs, as that is the nature of our dataset. Lastly, the procedure must return results that are interpretable within the technical limitations of this study.

According to Sanchez et al. [35], the "Louvain Modularity is one of the most widely used methods to extract communities from networks of any kind". Louvain is a heuristic modularity maximization scheme which has been shown to outperform most community detection methods in terms of computation time [6], and is especially applicable to large networks [35]. Although it has been proposed that the Leidel algorithm can be faster and more efficient compared to Louvain [39], the communities produced could be difficult to interpret [12]. Furthermore, through the use of modularity, which is a measurement of the clusterability of the collection of nodes, the algorithm produces very good quality communities [6]. However, a drawback is the fact that when the modularity of communities are deemed relatively similar they could be clustered together, making it difficult to analyse [39]. Even so, it remains one of the most efficient algorithms in terms of run-time.

2.2 Topic modelling

To gain further insight into the conversation surrounding NFTs, we aim to investigate which overall topics relevant to NFTs are the focus of the discussion on Twitter. Topic modelling is a statistical modelling approach that achieves this through using Natural Language Processing (NLP) to discover abstract topics that occur in a collection of documents [43]. We consider LDA and BTM as topic modelling algorithms, as LDA is a conventional topic modelling method [43, 46] that is often used when modelling topics within text and BTM is designed specifically to work with shorter text [24].

Latent Dirichlet Allocation (LDA) is a modelling technique that views documents as random mixtures of hidden topics which are seen as a probability distribution over words [43]. LDA has mostly been used for topic extraction on documents with a bulk of text (usually consisting of a few hundred words long) and has been shown to be effective for longer text corpora. However, Jonsson et al. [24] state that "LDA may not necessarily perform well when working with documents that are short in length".

Biterm Topic Models (BTM) was created specifically for shorter text analysis [24]. BTM is a word co-occurrence based topic modelling algorithm that learns topics by identifying biterms — the patterns found in word pairs throughout the document [45].

Jonnson et al. [24] and Yan et al. [46] applied different topic modelling techniques on the same datasets, including BTM and LDA, and found that while LDA performed better for longer documents, BTM gives the best result when working with short documents. The research finds that BTM performs better for shorter text, because it uses word co-occurrence patterns to enhance the topic learning and uses the aggregated patterns in the whole corpus to solve sparse word co-occurrence patterns that may occur on a document level [46].

It seems that even though LDA is a conventional method it is not the best-suited for shorter texts or social media, whereas BTM seems like the preferred option [43].

2.3 Sentiment analysis

In addition to determining the overall topics within the discussion surrounding NFTs, this study aims to investigate the sentiment of the discussion to allow further insight into its overall standing and development over time. We aim to achieve this by performing sentiment analysis on the relevant tweets over the selected time period. Sentiment analysis quantifies an expressed opinion or sentiment using NLP techniques [17]. On Twitter, opinions are expressed in short text, and the tweets usually include colloquialisms, slang, mixed language, hashtags, mentions, links, emoji, and punctuation. The latter two can enhance the accuracy of the NLP polarity score, while the others would need to be addressed in preprocessing to make the data machine-readable for analysis [12]. Further challenges that come with doing sentiment analysis on tweets include: text length, topic relevance, incorrect language use or slang, data sparsity, negating words, stop words, tokenization, and multi-modal content. These are challenges that make sentiment analysis on Twitter different from sentiment analysis on other types of mediums such as blogs or articles [17].

A lexicon-based approach is a sentiment analysis method that has one lexicon containing words whose polarities have already been manually assigned in word banks [25], making it an unsupervised and domain independent approach, leading to a more robust performance over different domains and texts [47]. We consider the lexicon-based approaches VADER (Valence Aware Dictionary and sEntiment Reasoner) [21] and TextBlob. In comparison to TextBlob and other Natural Language Toolkit (NLTK) [32] sentiment analysis tools, VADER outperforms these tools in speed, grammatical, and syntactical conventions for expressing and emphasizing sentiment intensity [7, 44]. With VADER having been specifically trained on social media data, it is also found to have the best list of lexical features for finding semantics in micro blog texts [44]. The use of VADER has been shown to be the current most accurate tool for Twitter data sets [44], outperforming even individual human raters [21].

3 RESEARCH METHOD

The objective of this paper is to understand how the discussion regarding NFTs in the Twitter community changes. As we wish to get an understanding of the NFT discussion, we decided to extract the largest community as a sample of the data, as the overall data is too diverse and immense to analyse within our scope. In order to understand the discussion, we require Twitter data relating to NFTs such that we can fulfil four conditions: capturing the largest community within the discussion, and analysing the topics, tweet frequency, and sentiment in the largest community. We further wish to understand the discussion by focusing on the largest topic within this community in order to compare its frequency and sentiment with the overall results. In the following section we will elaborate on the four needs and specify at each step why that method is used.

3.1 Data collection

To obtain the Twitter data relating to NFTs, we used the Twitter API v2 [40] to collect tweets containing the keyword “NFT” over the period of 1 February to 31 May 2021. We finished the data collection in the start of June and chose these dates to represent a snapshot of the NFT discussion due to noticing an increase in the NFT discussion on Twitter in March [29], and we added the additional month prior for comparison. We used the Twitter API’s academic research access [41] to collect the data such that we could get more precise and complete global, real-time and historical data.

A Tweet object includes elements such as the date it was published, the tweet text, the author’s basic information, the interactions with the tweet, and the conversation that the tweet is referencing [42].

3.2 Data preprocessing

As mentioned in Section 2.3, performing analysis on tweets poses challenges associated with characteristics of social media, such as the short text length, slang, links and stop words. To make the data better-suited for topic modelling and sentiment analysis, a few steps of preprocessing are necessary. We converted our data into a graph in order to apply community detection. Since we aim to analyse only the largest community, we first extract the largest community’s tweet data, which is then further cleaned for sentiment analysis and topic modelling.

3.2.1 Community detection. From background research we determined that Louvain would be the most appropriate network analysis method for our dataset [6, 12, 35]. We implemented Louvain using the NetworKit toolkit which provides a parallel implementation of the Louvain method [31] and has been shown to perform well on large networks [38]. A user-to-user graph was constructed along mentions that is weighted according to the frequency of mentions. The algorithm used to create this graph is that for each pair of users A and B, for each tweet by A (that mentions B), increase the weight of the edge between A and B. If the weight is 0 (i.e. no mentions at all), then there is no edge. Further, mentions of deleted accounts were not included, retweets were treated as a tweet by the original author, and a quote tweet was treated as if it was a tweet by both authors. This graph was used as input for the NetworKit library, and the communities were determined using all the defaults in the library with each user being assigned to a community.

The scope of this paper is limited to the largest community, which we have defined to be the community with the most number of users. We find this community in the collected dataset using NetworKit’s Louvain algorithm and extract the users and tweet interactions of this community. For each user in this community, we filter to only the tweets that have mentions of other users within the same community. We did this to reduce the dataset while still capturing the core of the discussion on NFTs.

3.2.2 Data cleaning. For the purpose of this analysis we first removed all retweets and non-English tweets. Since retweets are duplicates of the original tweet, they were removed to avoid redundancy. Non-English tweets were also removed as we consider only one language within the scope of this study, and results are more comprehensive and easier to interpret when they are limited to one language.

The cleaning of the tweets for the BTM topic modelling included the removal of mentions and hashtags, crypto wallet addresses, HTML entities, and stop words, as well as ensuring that the tweets consist of only letters. We used the default English stop words in the Python NLTK library and added two of our own stop words: nft and rt. As all tweets in our dataset contain the word “nft”, it would dwarf our topics, since we collected our tweets based on the criteria that the tweet contains the NFT keyword. For similar reasons we removed “rt”, as it represents the word “retweet”, and occurs quite frequently since all tweets that are posted as a reply to another tweet that is a retweet contains this tag.

The cleaning of the data used for sentiment analysis is the same as for the topic modelling, except that stop words and punctuation are not removed, and the tweet is not limited to consist of only letters. This is because punctuation (such as exclamation marks), stop words such as “very”, and emoji have an impact on sentiment, but are irrelevant for topic modelling.

3.3 Topic modelling

By reviewing comparisons between LDA and BTM [24, 46], we determined that the BTM topic modelling algorithm is the most appropriate for our dataset as we have a short text corpus. The BTM R Cran Package [45] released in July 2021 has the added benefit of enabling us to link document identifiers, such as tweet IDs, to specific topics, which is useful if wanting to determine sentiment on a specific topic, as it allows us to perform sentiment analysis on tweets relating to that topic. We applied this BTM package to our cleaned largest community dataset for multiple topics ranging from 2 to 75. To get the likelihood of how well the biterms fit to the BTM model for a given number of topics, we used the LogLik function within the BTM package. After finding that the more topics we added, the better fitted the LogLik values determined the model was, we decided to use the Elbow Method [20] to find the optimal point such that we limited the diminishing returns of our model. We used the Python kneed package [5] to implement the Elbow Method, and used its built-in basic data normalization to determine the best suited number of topics for our dataset.

As we wish to get an understanding of the NFT discussion, we decided to extract the largest topic as a sample of the data, as the overall data is too diverse and immense to analyse within our scope. We did this by first running the most optimal number of topics through the BTM model, and then further isolating the largest topic from these. We defined the largest topic to be the topic with the most tweets having the highest probability of being associated with it. Once we knew which topic was the largest, we were able to isolate the tweets associated with this topic due to the BTM method recording the tweet IDs along with the probability that they are associated with a topic. This allowed us to run further data analysis on it, such as sentiment analysis and frequency analysis.

3.4 Tweet frequency

To determine the level of activity within the NFT discussion on Twitter and how that engagement changes over time, we consider tweet frequency. Tweet frequency is determined by counting the number of tweets posted per day, which is possible due to a tweet's post date being a feature associated with the tweet data as described in the data collection section. For the overall tweets, the data used is that of the cleaned tweets for topic modelling purposes, such that we can accurately compare the frequency of the overall tweets with an individual topic's tweets. We determine the tweet frequency of each topic by extracting the tweet data per topic similar to the method described in the previous section in extracting the tweet data of the largest topic.

3.5 Sentiment analysis

Through the consultation of relevant literature, we determined that the VADER sentiment analysis algorithm is the most appropriate for our dataset as it is specially attuned to find semantics in micro blog text [7, 21, 44]. We applied VADER to the extracted, cleaned tweets of the largest community to determine whether the sentiment of the tweets is of positive, neutral, or negative polarity on average.

As we are particularly interested in how sentiment changes, we next had a look at how the overall sentiment varies over time. The dataset was divided into 34 segments of equal size of about 12,000 tweets per segment. The segment sizes were chosen such that the number of segments does not exceed computational and analytical capabilities of this study, while still retaining sufficient granularity to be able to analyse the trend of the data. We then performed sentiment analysis on each segment and plotted the average compound value such that we could see how the polarity changes over time.

Similarly, we applied sentiment analysis on the largest topic's tweets both overall and over time in order to compare it with the largest community as a whole. For these tweets, the trade-off between computational and analytical capabilities and granularity lead to a choice of 40 segments of around 2,500 tweets each.

4 RESULTS

The results obtained from the community detection, topic modelling, tweet frequency, and sentiment analysis explore the environment of the NFT discussion on Twitter and lead to insights that will be presented in the following sections.

4.1 Community detection

Our final network consists of 1,290,652 vertices and 4,936,209 edges, where the number of vertices represents the total number of users that posted tweets containing the keyword "NFT", and the edges represent their "mention" interactions. Of this, there are 143,722 users in the largest community such that it comprises 11.13% of the total number of users. From 1 February to 31 May there are 459,479 tweets made relating to NFTs within the largest community, which constitutes 2.90% of the total number of tweets (15,827,186) made in this period containing the keyword "NFT". This leads to an average of 3.2 tweets per user in the largest community, compared to an average of 12.3 tweets per user in the full network. Keeping in mind that the largest community tweets do not take tweets into consideration that were sent by members of the community to users outside of the community, this shows that the NFT conversation contained within the community is less active than the entire general conversation is on average.

4.2 Topic modelling

To determine which overall topics relevant to NFTs are the focus of the discussion on Twitter, we present the results produced by the BTM topic modelling process. Due to the discussion being immense and very diverse, we further take a closer look at the largest topic in order to get an understanding of the conversation.

4.2.1 Overall. First we look at the overall topics during the analysis period. The Elbow Method applied to the BTM's normalized LogLik values determined that the optimal number of topics for the model is 11. Appendix B shows the top terms of the BTM output with the probability that they are associated with the topic. Further analysis shows that topic 11 is the largest topic, with 29.84% of tweets having the highest probability of being associated with it, as per Table 1 in Appendix A.

4.2.2 Largest topic. Topic 11's terms particularly revolve around the term "project", with the strength of the connections between the terms being represented by the thickness of the lines connecting the words in Figure 1. In particular, the terms "good" and "project"

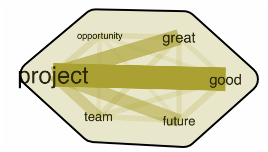

Figure 1: BTM model: Topic 11.

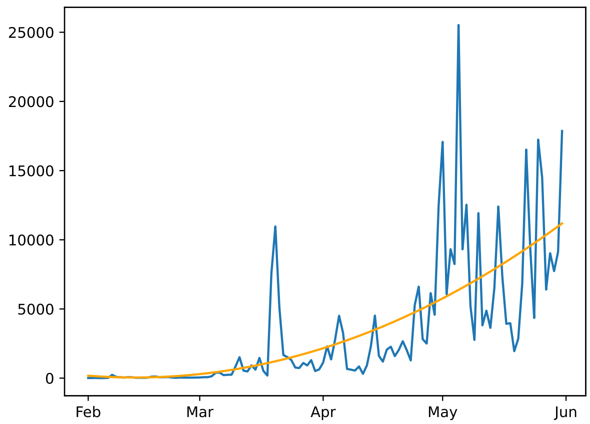

Figure 2: Overall Tweet Frequency (Number of tweets / Month).

occur together the most frequently, with the terms "great" and "project" occurring together the second most frequently. The terms that have the highest probability of being associated with topic 11 are shown in Table 12 (Appendix B), with the probability of them being associated with the topic in the "Probability" column.

4.3 Tweet frequency

To gauge how the level of activity of the NFT discussion on Twitter changes, we report on how the tweet frequency develops over time. To better understand the overall activity of the discussion, we further explore the change in tweet frequency of the largest topic and how it correlates with the overall tweet frequency.

4.3.1 Overall. Despite high fluctuation, the tweet frequency of the largest community has a definite general upward trend from the start of our analysis in February, as visualized in Figure 2. For the first month there was hardly any mention of NFTs, but the conversation increased by 2235.7% from February to March. After the spike in interest in NFTs in mid-March, the discussion reached an all-time high in early May, with the average tweet frequency steadily increasing.

The spike in March may be attributed to the sale of a digital artwork titled "Everyday's — The First 5000 Days" by an artist known as Beeple. It was sold for $69.3 million by Christie's [29] on 11 March 2021 in "the first ever sale by a major auction house of a piece of art that does not exist in physical form" [18]. According to a tweet posted by Christie's later that day, "22 million people tuned

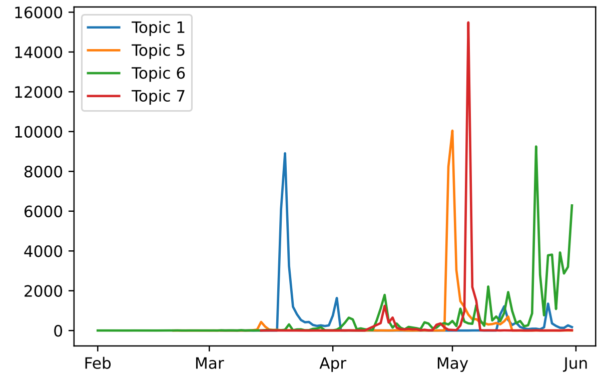

Figure 3: Tweet Frequency of Topics 1, 5, 6, and 7 (Number of tweets / Month).

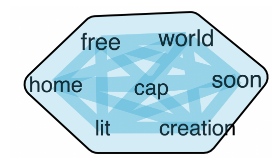

Figure 4: BTM model: Topic 5.

in for the final moments of @Beeple's historic sale" [11], with the auction price being the third-highest achieved for a living artist at the time [29]. Looking at topic 1's top words in Table 2 (Appendix B) and the tweet frequency in Figure 3, we gather that this event is captured by topic 1. From the activity of this topic and the overall tweet frequency, we deduce that this is the topic that kick-starts the NFT discussion on Twitter.

After the initial spike in overall tweet frequency in March, there are another five peaks that draw special attention: on 1, 5, 22, 25 and 31 May the number of NFT-related tweets reached over 15,000.

When inspecting topic 5's word co-occurrences in Figure 4, we notice that the weights between all the words are almost the same. Thus, the probability of a term being associated with any other term shown is very similar, and we can deduce that the terms listed all frequently occur together. This is further supported when examining a sample tweet on the 1st of May:

> "A lit world of free creation that will soon be the home of #blockchain #NFT is here!No cap!#cyberworld #NFT #Airdrop #testnet @cybermiles https://t.co/dH8Ivalf8y"

This tweet is one of many that differ only in its hashtags, mentions and links. When cleaned, all of them look identical: "lit world free creation soon home cap", which matches the terms that occur in topic 5 exactly. Together with the tweet frequency of topic 5 during this time (shown in Figure 3), allows us to attribute the spike in tweet frequency at the start of May to topic 5.

Shortly after topic 5's peak, the overall tweet frequency reached an all-time high on the 5th of May. Similar to the tweets of the previous peak, this spike consists of tweets that are all similar to

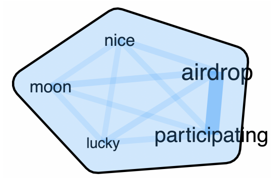

Figure 5: BTM model: Topic 6.

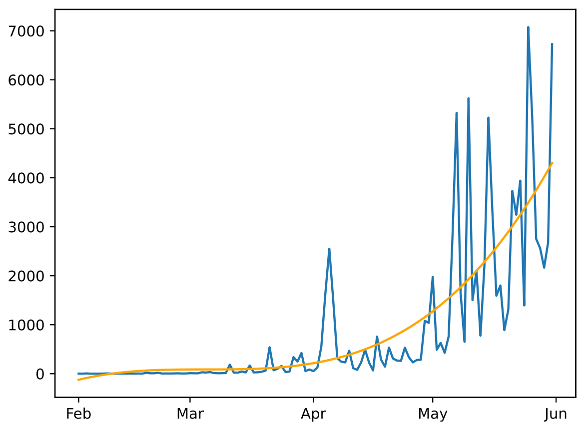

Figure 6: Topic 11 Tweet Frequency (Number of tweets / Month).

each other. The following shows an example of a typical tweet during this time:

> "@renft_protocol is a multi-chain liquidity solution platform that breaks NFT assets into shards. https://t.co/FWsqxkdMkv #nft #bsc #defi #airdrop @cryptochromeorg @BEFOREAROWAPOC1 @DcashOfficial"

Following the same process as before, we can attribute the majority of the NFT discussion on this day to topic 7 based on its top terms seen in Table 8 (Appendix B) and frequency over time (Figure 3).

Towards the end of May, the reason for increased tweet frequency is less straightforward. Throughout this time, but especially on May 22, 25 and 31, there is little variety of tweets that again only differ in their hashtags, mentions and links. After the cleaning process, the reduced tweets "I am participating in" and "great project" occur frequently. We mainly attribute this increase in activity to two topics: topic 6 and topic 11. This is based on their spike in tweet frequency, shown in Figure 3 and 6 for these two topics respectively. In Figure 5 we see that the conversation in topic 6 revolves around the participation in airdrops, an NFT distribution method [27]. Airdrop in this specific context refers to an NFT giveaway. We determine this as the line between the terms "airdrop" and "participate" are thicker than the lines between the rest of the terms, indicating that these terms frequently occur in the same tweet. By linking this with the co-occurring increase in topic 11's activity (Figure 6), we can verify that participation in airdrops of various projects has become more popular. We will further analyse the largest topic in the following section.

Topics 1, 5, 6, 7, and 11 are by far the most discussed, with their peak frequencies all exceeding 7000 tweets whereas other topics' peak frequencies only range from a couple hundreds to low thousands. We also know that topic 11 is the largest topic overall, suggesting that as the discussion on Twitter surrounding NFTs increases, it focuses on this topic.

From the above, we can summarize that the discussion on Twitter surrounding NFTs is kick-started by the sale of Beeple's digital artwork in March and after a few smaller short-lived spikes in tweet frequency, the conversation moves towards the more general topics of NFT projects and participating in airdrops.

4.3.2 Largest topic. Let us have a closer look at its tweet frequency visualized in Figure 6 to understand how the largest topic develops over time in relation to the overall tweet frequency.

We see a general increase in the topic's prevalence as time goes on, which follows the general community trend. During the largest community's initial peak in mid-March, topic 11 did not contribute much to the overall discussion. We further see that at the topic's initial peak in early April, it dominates almost the entire overall conversation. The tweet frequency of both the largest topic and the largest community show an increase over May. At the largest topic's peak at the end of May, it dominates just under half of the largest community's discussion. As discussed above, topic 11 increases in tweet frequency in conjunction with topic 6 in May, as they are closely related. The terms in topic 11 contain words that are found in the majority of tweets throughout the time frame that this study analyses. Therefore, topic 11 encompassed many of the NFT-related topics that were discussed and had the highest overall tweet frequency for the duration.

4.4 Sentiment analysis

Performing sentiment analysis on the relevant tweets allows us to determine the overall sentiment towards the NFT discussion and gain further insight into its overall standing. Since this research also aims at investigating the development of the discussion over time, we performed sentiment analysis on time segments to further explore how the sentiment towards NFTs changes between the start of February and the end of May 2021. The following sections present the results and highlight some key findings.

4.4.1 Overall. The average compound VADER sentiment score for the analysed data is 0.557, showing that the largest community's general sentiment surrounding NFTs is positive. Since VADER classifies 0 as neutral and 1 as the most positive sentiment, we see that the resulting sentiment score is on the slightly higher end of the positivity spectrum. This is in line with cryptocurrency sentiment, which has been shown to be invariably overall positive regardless of whether the price increases or decreases [1].

4.4.2 Largest topic. To explore what may contribute to the overall average sentiment, we analysed the average sentiment of the largest topic. The largest topic's overall Twitter discussion sentiment is positive, with an average compound VADER sentiment of 0.674. This is almost 20% more positive than the largest community's

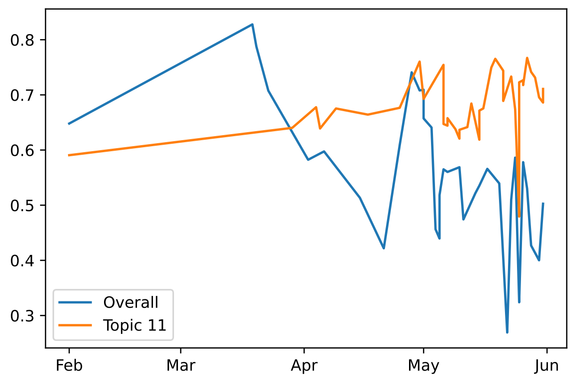

Figure 7: Overall and Topic 11 Sentiment over time (VADER sentiment score / Month).

overall sentiment score of 0.557. Looking at the words that best describe the largest topic, the reason for the positive classification becomes apparent: words such as 'good', 'great', 'best', 'wonderful' and 'excellent' capture positive sentiment and can be found in the top 15 words associated with the largest topic (see Table 8 in Appendix B). Since the largest topic comprises roughly a third of all the largest community's NFT-related tweets, its sentiment proportionally impacts the average overall sentiment. The positive sentiment of the overall discussion can therefore partially be attributed to the positive sentiment surrounding the largest topic.

4.5 Sentiment over time

Now that the average sentiment has been explored, we aim to determine how the sentiment changes over time and how that compares to the average sentiment.

4.5.1 Overall. Although the largest community's overall sentiment is always positive, it shows high fluctuation over 27% of VADER's measurement range of [-1, 1]. In the middle of March there is a peak in positivity, which correlates with the sudden increase in tweet frequency in the NFT discussion. However, after this initial peak it trends downward in positivity until just before the all-time high in tweet frequency in the start of May.

The straight lines surrounding the initial peak in overall sentiment are caused by the binning approach used. Since the segments are based on the number of tweets, each segment stretches over different lengths of time based on the frequency of tweets. The first segments both encompass more than a month, as the tweet frequency is very low in February and early March as shown in Figure 2. As the tweet frequency increases, the period per segment used for analysing the sentiment becomes shorter.

The peak in overall sentiment mid March can be attributed to the same topic that led to the increase in tweet frequency over this time: topic 1, which surrounds the first auction of a digital artwork. The extreme dips in overall sentiment in May are perhaps unexpected, as one might anticipate that the largest topic's increase in prominence and overall positive sentiment would direct the community's sentiment to be positive during this time. However, on May 22 and May 25 the average sentiment dropped to 0.268 and 0.316 respectively. May 22 and 25 are also the days when there is a significant spike in tweet frequency due to topic 6 and 11, with the most frequently occurring words "project", "participating" and "airdrop" being neutral in sentiment and therefore decreasing the positivity of the average sentiment.

4.5.2 Largest topic. The straight line in Figure 7 for topic 11 over the first two months can be explained similarly to that of the overall sentiment. At the largest topic's peak in tweet frequency we also find a significant dip in tweet sentiment. Similar to the overlap of the overall tweet frequency spike and sentiment dip, the same phenomenon for the largest topic can be explained in like manner. Tweet frequency is inflated with words of neutral sentiment, leading to an average sentiment that is notably less positive than the general average sentiment for the largest topic, though it is still more positive than that of the overall tweets for the same time frame.

Although the sentiment is consistently more positive for the largest topic than for the overall tweets as the discussion progresses, we can see no clear correlation other than the dip on the 25th of May between the fluctuations in sentiment for the largest topic and the overall sentiment over time.

5 CONCLUSION

To allow further insight into the overall standing and development of the NFT discussion on Twitter from 1 February to 31 May 2021, this study aimed to address the question of how the sentiment within the NFT discussion for topics changes over time. This was achieved by capturing the largest community within the discussion and analysing the overall topics. The largest topic relates to the word "project" and its sentiment is classified as positive. This study further investigated tweet frequency over time to explore how the level of activity within the discussion developed within the given time frame, and we compared this with the largest topic's activity. The discussion is kick-started in the middle of March 2021, and the tweet frequency reaches an all-time high on the 5th of May. Additionally, the average sentiment in the largest community was considered, as well as the change in sentiment over time, both of which are found to be positive. Furthermore, we found that the largest topic's sentiment is generally more positive than the overall sentiment. These methods have shown to be effective in providing a deeper understanding of the overall standing and development of the discussion surrounding NFTs on Twitter.

5.1 Limitations

In this research, we explore the discussion on Twitter surrounding NFTs. Although we have a better understanding of the discussion, there are some limitations and future expansions to consider. The limitations and future work mostly coincide. In terms of limitations, we have three key factors that are related to scope. Firstly, we consider only English tweets, secondly, we consider only the tweets by the largest community, and lastly, the scope is limited to a defined time frame. Although we have gained insight into the NFT discussion on Twitter through our exploratory research, it is limited to a very specific sample set, which is not guaranteed to represent the NFT discussion as a whole.

5.2 Future work

Future research could explore and compare a different sample of the NFT discussion to gain further insights into the discussion as a whole. While this study leads to a deeper understanding of the shape of the NFT discussion on Twitter, further research could explore what shapes it. Future projects could attempt to answer questions such as what portion of the discussion is automated by bots, who the major influencers and trend-setters are, and how the conversation on Twitter surrounding cryptocurrencies correlates to the conversation surrounding NFTs.

ACKNOWLEDGMENTS

This work is based on the research supported in part by the National Research Foundation (NRF) of South Africa (Grant Numbers MND201019567727 and MND200627537546) and are hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

NRF National Research Foundation

REFERENCES

[1] Jethin Abraham, Daniel Higdon, John Nelson, and Juan Ibarra. 2018. Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Science Review* 1, 3 (2018), 1.

[2] Wasim Ahmed, Peter A Bath, and Gianluca Demartini. 2017. Using Twitter as a data source: An overview of ethical, legal, and methodological challenges. *The ethics of online research* (2017).

[3] Benjamin R C Amor, Sabine I Vuik, Ryan Callahan, Ara Darzi, Sophia N Yaliraki, and Mauricio Barahona. 2016. Community detection and role identification in directed networks: understanding the Twitter network of the care. data debate. In *Dynamic networks and cyber-security*. World Scientific, 111–136.

[4] Lennart Ante. 2021. Non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations. *Available at SSRN 3904683* (2021).

[5] Kevin Arvai. 2020. Welcome to kneed's documentation! — kneed 0.6.0 documentation. https://kneed.readthedocs.io/en/stable/

[6] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.

[7] Venkateswarlu Bonta and Nandhini Kumaresh2and N Janardhan. 2019. A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis. *Asian Journal of Computer Science and Technology* 8, S2 (2019), 1–6.

[8] Holger Bonus and Dieter Ronte. 1997. Credibility and economic value in the visual arts. *Journal of cultural economics* 21, 2 (1997), 103–118.

[9] Usman W Chohan. 2021. Non-Fungible Tokens: Blockchains, Scarcity, and Value. *Critical Blockchain Research Initiative (CBRI) Working Papers* (2021).

[10] Dongho Choi, Ziad Matni, and Chirag Shah. 2016. What social media data should i use in my research?: A comparative analysis of twitter, youtube, reddit, and the new york times comments. *Proceedings of the Association for Information Science and Technology* 53, 1 (2016), 1–6.

[11] ChristiesInc. 2021. "22 million people tuned in for the final moments of @Beeple's historic sale this morning, which totaled $69.3 million. Relive it from the artist's POV in this link! #beeple #digitalart #digitalartist #artist #art #thefirst5000days #nft https://t.co/XaREV5Fdvu" / Twitter. https://twitter.com/ChristiesInc/status/1370132264030498827?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1370132264030498827%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fwww.reuters.com%2Flifestyle%2Fdigital-only-artwork-fetches-nearly-70-million-christies-2021-03-12%2F

[12] Laurenz A Cornelissen, Clarice de Bruyn, Maphiri K Ledingwane, Pieter P A Theron, Petrus Schoonwinkel, and Richard J Barnett. 2019. Cross-Sample Community Detection and Sentiment Analysis: South African Twitter. In *Proceedings of the South African Institute of Computer Scientists and Information Technologists 2019 (SAICSIT '19)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3351108.3351135

[13] Laurenz A Cornelissen, Petrus Schoonwinkel, and Richard J Barnett. 2019. A Socio-Informatic Approach to Automated Account Classification on Social Media. In *Proceedings of the 10th International Conference on Social Media and Society (SMSociety '19)*. Association for Computing Machinery, New York, NY, USA, 24–32. https://doi.org/10.1145/3328529.3328541

[14] David Darmon, Elisa Omodei, and Joshua Garland. 2015. Followers are not enough: A multifaceted approach to community detection in online social networks. *PloS one* 10, 8 (2015), e0134860.

[15] Michael Dowling. 2021. Is non-fungible token pricing driven by cryptocurrencies? *Finance Research Letters* (2021), 102097.

[16] Santo Fortunato. 2010. Community detection in graphs. *Physics reports* 486, 3-5 (2010), 75–174.

[17] Anastasia Giachanou and Fabio Crestani. 2016. Like it or not: A survey of Twitter sentiment analysis methods. https://doi.org/10.1145/2938640

[18] Elizabeth Howcroft. 2021. After first quarter frenzy, NFT market shows signs of stabilising | Reuters. https://www.reuters.com/technology/after-first-quarter-frenzy-nft-market-shows-signs-stabilising-2021-05-04/

[19] Elizabeth Howcroft. 2021. Digital-only artwork fetches nearly $70 million at Christie's | Reuters. https://prod.reuters.com/lifestyle/digital-only-artwork-fetches-nearly-70-million-christies-2021-03-12/

[20] Hestry Humaira and Rasyidah Rasyidah. 2020. Determining The Appropiate Cluster Number Using Elbow Method for K-Means Algorithm. https://doi.org/10.4108/eai.24-1-2018.2292388

[21] C Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (5 2014), 216–225. https://ojs.aaai.org/index.php/ICWSM/article/view/14550

[22] Akhmad Irsyad and Nur Aini Rakhmawati. 2020. Community detection in twitter based on tweets similarities in indonesian using cosine similarity and louvain algorithms. *Register: Jurnal Ilmiah Teknologi Sistem Informasi* 6, 1 (2020), 22–31.

[23] Hadar Yoana Jabotinsky and Roee Sarel. 2020. How crisis affects crypto: Coronavirus as a test case. *Available at SSRN 3557929* (2020).

[24] ElJónsson and Jake Stolee. 2015. An evaluation of topic modelling techniques for twitter. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Short Papers)*. 489–494.

[25] Mohammed Kaity and Vimala Balakrishnan. 2020. Sentiment lexicons and non-English languages: a survey. *Knowledge and Information Systems* (2020), 1–36.

[26] Alron Jan Lam. 2016. Improving Twitter Community Detection through Contextual Sentiment Analysis of Tweets. *ACL 2016* 30 (2016).

[27] Loic Lesavre, Priam Varin, Peter Mell, Michael Davidson, and James Shook. 2019. A taxonomic approach to understanding emerging blockchain identity management systems. *arXiv preprint arXiv:1908.00929* (2019).

[28] Claudio D G Linhares, Jean R Ponciano, Fab´S F Pereira, Luis E C Rocha, Jose Gustavo S Paiva, and Bruno A N Travençolo. 2020. Visual analysis for evaluation of community detection algorithms. *Multimedia Tools & Applications* 79 (2020).

[29] Matthieu Nadini, Laura Alessandretti, Flavio Di Giacinto, Mauro Martino, Luca Maria Aiello, and Andrea Baronchelli. 2021. Mapping the NFT revolution: market trends, trade networks and visual features. *arXiv preprint arXiv:2106.00647* (2021).

[30] Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder. 2016. *Bitcoin and cryptocurrency technologies: a comprehensive introduction.* Princeton University Press.

[31] NetworKit. 2018. Community Detection with NetworKit. https://networkit.github.io/dev-docs/notebooks/Community.html

[32] NLTK. [n. d.]. Natural Language Toolkit. https://www.nltk.org/

[33] Symeon Papadopoulos, Yiannis Kompatsiaris, Athena Vakali, and Ploutarchos Spyridonos. 2012. Community detection in social media. *Data Mining and Knowledge Discovery* 24, 3 (2012), 515–554.

[34] Ferdinand Regner, André Schweizer, and Nils Urbach. 2019. *NFTs in Practice-Non-Fungible Tokens as Core Component of a Blockchain-based Event Ticketing Application Completed Research Paper.* Technical Report. 1 pages.

[35] Daniel López Sánchez, Jorge Revuelta, Fernando la Prieta, Ana B Gil-González, and Cach Dang. 2016. Twitter user clustering based on their preferences and the Louvain algorithm. In *International Conference on Practical Applications of Agents and Multi-Agent Systems.* 349–356.

[36] Petrus Schoonwinkel. 2020. Classification of social media event-discussions using interaction patterns : a social network analysis approach.

[37] Wendel Silva, Ádamo Santana, Fábio Lobato, and Márcia Pinheiro. 2017. A methodology for community detection in Twitter. In *Proceedings of the International Conference on Web Intelligence.* 1006–1009.

[38] Christian L Staudt, Aleksejs Sazonovs, and Henning Meyerhenke. 2016. NetworKit: A tool suite for large-scale complex network analysis. *Network Science* 4, 4 (2016), 508–530.

[39] Vincent A Traag, Ludo Waltman, and Nees Jan Van Eck. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports* 9, 1 (2019), 1–12.

[40] Twitter Developer Platform. [n. d.]. What's new with v2? https://developer.twitter.com/en/docs/twitter-api/early-access
[41] Twitter Developer Platform. [n. d.]. Academic Research product track. https://developer.twitter.com/en/products/twitter-api/academic-research
[42] Twitter Developer Platform. [n. d.]. Data Dictionary. https://developer.twitter.com/en/docs/twitter-api/data-dictionary/object-model/tweet
[43] Ike Vayansky and Sathish A P Kumar. 2020. A review of topic modeling methods. *Information Systems* 94 (2020), 101582.
[44] Wan Nur Syahirah Wan Min and Nur Zareen Zulkarnain. 2020. Comparative Evaluation of Lexicons in Performing Sentiment Analysis. 2 (11 2020), 14–20.
[45] Jan Wijffels. 2021. BTM - Biterm Topic Modelling for Short Text with R. https://github.com/bnosac/BTM
[46] Xiaohui Yan, Jiafeng Guo, Yanyan Lan, and Xueqi Cheng. 2013. A biterm topic model for short texts. In *Proceedings of the 22nd international conference on World Wide Web*. 1445–1456.
[47] Nor Nadiah Yusof, Azlinah Mohamed, and Shuzlina Abdul-Rahman. 2015. Reviewing classification approaches in sentiment analysis. In *International conference on soft computing in data science*. 43–53.

A TWEETS PER TOPIC

Table 1: Tweets per Topic.

Topic	Tweets	Contribution
11	121144	29.84%
6	75836	18.68%
1	38567	9.50%
9	36918	9.09%
5	36494	8.99%
10	32161	7.92%
7	29644	7.30%
4	14986	3.69%
3	8812	2.17%
2	7918	1.95%
8	3541	0.87%

B TOP WORDS PER TOPIC

Table 2: Topic 1 Top Terms.

Term	Probability
help	0.0765932584
create	0.0752196791
need	0.0746937398
share	0.074609809
usd	0.0745396973
bid	0.0745004754
auction	0.0744559698
first	0.0743899225
earnings	0.0743891096
final	0.0743425717
wants	0.0739070664
official	0.0736674673
phemex	0.0735170828

Table 3: Topic 2 Top Terms.

Term	Probability
play	0.11124904
get	0.09395147
tokens	0.08207307
worth	0.06303378
use	0.05751216
signed	0.05623925
code	0.05487475

Table 4: Topic 3 Top Terms.

Term	Probability
doo	0.28291515
go	0.06451099
let	0.06099705
share	0.02071613
buy	0.02045811
shark	0.02012841
moon	0.02006

Table 5: Topic 4 Top Terms.

Term	Probability
project	0.05904157
best	0.04616973
luck	0.04094105
crypto	0.03254974
sir	0.03020066
giveaway	0.0272609
good	0.02666029
success	0.02353585
interested	0.02346208
hope	0.02266318
one	0.02251565

Table 6: Topic 5 Top Terms.

Term	Probability
world	0.136968016
soon	0.136473705
free	0.135751393
creation	0.130843503
cap	0.130755762
home	0.130541355
lit	0.130407273

Table 7: Topic 6 Top Terms.

Term	Probability
airdrop	0.101029844
project	0.098079009
participating	0.069404457
nice	0.043365505
moon	0.04007942

Table 8: Topic 7 Top Terms.

Term	Probability
chain	0.122160913
platform	0.099189889
assets	0.09643413
liquidity	0.096012013
multi	0.09579441
shards	0.095187412
solution	0.095157962
breaks	0.095073975

Table 9: Topic 8 Top Terms.

Term	Probability
attack	0.06286176
enemy	0.03107096
th	0.02991238
increase	0.02437502
damage	0.02335061
lottery	0.02307588
range	0.02243269
hp	0.0220323

Table 10: Topic 9 Top Terms.

Term	Probability
airdrop	0.076195508
ultraman	0.057788948
crypto	0.057316117
aman	0.057118218
referrals	0.05583511
auction	0.055012626
participate	0.054965833
referral	0.054921815
curve	0.054908313
someone	0.054626538
receive	0.054606009
also	0.053574159
get	0.051610889

Table 11: Topic 10 Top Terms.

Term	Probability
first	0.036918017
one	0.036359734
get	0.034316558
gen	0.024123843
living	0.024078903
project	0.012239602
token	0.009188273

Table 12: Topic 11 Top Terms.

Term	Probability
project	0.1502
good	0.0439
great	0.0413
future	0.0367
team	0.0337
best	0.0213
opportunity	0.0159
thanks	0.0154
better	0.0153
hopefully	0.0146
strong	0.0145
think	0.0121
wonderful	0.0117
excellent	0.0108