Analysing music artists collaboration network

Antarlina Mukherjee

This manuscript was compiled on December 14, 2022

This project researches the music artist collaboration network for lyrical music. The outcome of the study confirms presence of assortativity i.e. artists tend to collaborate in their league of popularity. Additionally, the lyrics of collaborative music is influenced strongly by the Hip hop genre artists. The study found empirical evidence towards the presence of Friendship paradox indicating that some artists are more likely to collaborate than others. Finally, the audio attributes of collaborations borrow speechiness, energy, tempo from the Hip hop artists whereas acousticness is borrowed from the non Hip hop artists.

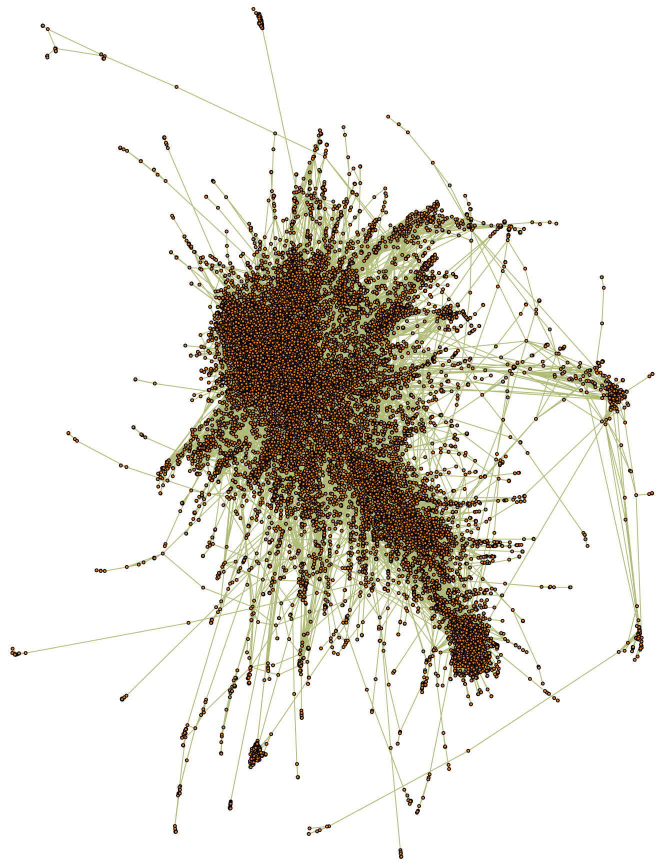
Network Analysis | Social Graphs | Musical artists collaboration |

Music has forever been an expression of human emotions. The lyrics captures impressions about the world we live in and the world we want to be in. Further, the synthesis of music attributes such as the instruments, the tempo etc. with the lyrics are used to complement the mood the song tries to create. While artists tend to their individual music styles, collaboration between artists has become very commonplace as all parties try to combine their art. The key idea behind studying artist collaboration is to understand the dynamics of artist collaborations. This includes investigating any, or the lack of, assortativity amongst the collaborators with regards to their popularity and study how the produced music resembles or departs from their individual music styles. The outcome of the analysis also tries to create a probably roadmap of their collaboration trajectory given their style of music and popularity.

Results

Network Definition. A non-directional weighted network is created with each node representing an artist. An edge connects two artists when they have been credited with the same song. Node attributes are genre tags, follower count on Spotify and popularity index as defined by Spotify. Count of shared songs is defined as the "weight" attribute for the edges. Only artists making lyrical music (i.e. excluding orchestras, instrumental artists) were kept in scope of the analysis to be able to analyse the lyrical sentiments of their music.

Network creation and sampling. After collecting the relevant data, the network is created as seen in Figure 1 with the edge and node definitions as stated above (henceforth referred to as the *Super network*).It can be seen that network is fairly big and almost resembles a dragonfly; it's head being a combination of one or more very big clusters. The ends of its wings or legs have smaller sub-networks which are loosely connected to the torso (probably owing to smaller edge weights).The network was sampled using backbone algorithm (1) to make the analysis computationally economical. The results obtained are based on the backbone-sampled network and the findings are expected to represent the super-network.



(a) Network resembles a dragonfly; it's head being a combination of large clusters. The ends of its wings have smaller sub-networks

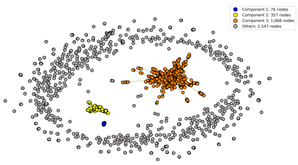
Fig. 1. Super network of music artists

Dominant genres. Figure 2a shows the top three subgraphs that were identified in the sampled graph, of which the largest graph of 1,038 nodes is chosen for analysis. Figure 2b shows the genre name wordcloud of the largest subgraph. It is seen that Hip hop/Rap and its sub-variants(gangster, underground,

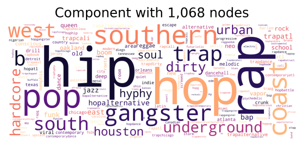
Significance Statement

The outcome of this study sheds light on the nature of artist collaborations in music. The deviations or similarities in the choice of lyrics from the individual styles are studied. Communities of artists identified shows which artists tend to work together. A deep dive into the communities reveals some patterns of association that can predict how an artists given their style of music and popularity is likely to collaborate and what is more likely to succeed amongst a choice of collaborations.

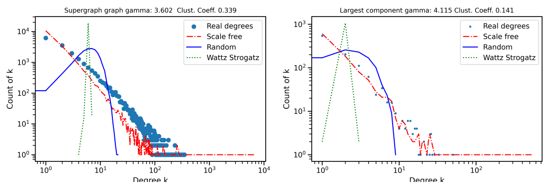
trap etc.) dominate the network followed by Pop. Rap is a very collaborative genre as rappers collaborate either with other fellow rappers or say, Pop singers for their choruses.



(a) Annotating top three subgraphs in the backbone sampled network



(b) Genre term frequency wordcloud of the largest component



(c) Super network vs largest component degree distribution

Network	Size	#neighbours	#Friends neighbours
Super network	19,383	916	84
Largest component	1,068	872	128

(d) Friendship Paradox

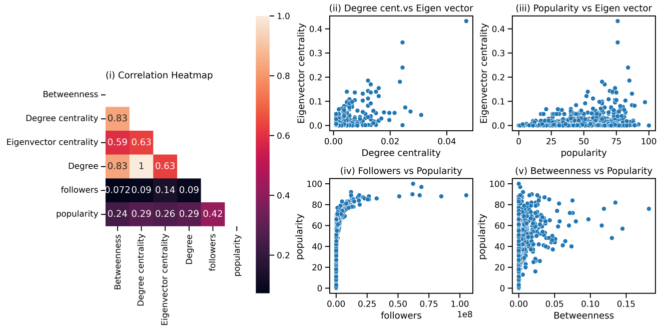
Fig. 2. Structure and degree distribution of sampled network and largest component (a)The biggest component is at the centre, followed by the second component in yellow and the third component in blue. The remaining components are arranged in the outer circle (b)Hip hop and its variants and Pop are the most collaborative genres (c)Degree distribution for both super-network and backbone-sampled largest component are heavy tailed, however the γ > 3 therefore not scale free and a giant component has appeared (supercritical regime) (d) Friendship Paradox test shows 90% of times the neighbours have more neighbours

Degree distribution. Figure 2c shows the degree distributions of the super-network consisting of 19,383 nodes and 66,605 edges vs the largest component of the sampled network. Additionally, scale free variant (with the same number of nodes), Random(Erdos-Renyi) network and the Wattz Strogatz degree distributions for similar network specifications have been plotted alongside the real network. Needless to say, Wattz Strogatz makes the poorest estimation of the network. It also becomes apparent that the degree distribution of the network lies somewhere between that of a random network and a scale free network; in effect it is a heavy-tailed distribution. The γ value >3 indicates that both the super-network and the largest component are outside the scale free regime and more so demonstrate the "small word" phenomenon of the random regime (2)(3) The network appears to be analogous to the scientific collaboration network (4) as in both cases actual interactions between the nodes have been considered towards defining an edge. All real world networks are not scale free (5) Upon running the powerlaw fit diagnostics , it is seen that the distribution fit is (exponential) < lognormal < powerlaw. Therefore the network has some some random as well as scale free characteristics. The inclination towards a random behaviour for artist collaboration seems intuitive as the degree of a node is determined by the actual interaction between two artists that have collaborated on a song. There are finite possibilities to it than say, creation of an edge between two artists who have appeared on the same text or one artist following the other on social media (in-degree distribution) where the possibilities may not be bounded. It is seen that the super-network is in the *supercritical regime* which means that a giant component exists in it that is similar to a network in itself. This is most likely the network of the lyrical genres that was identified.

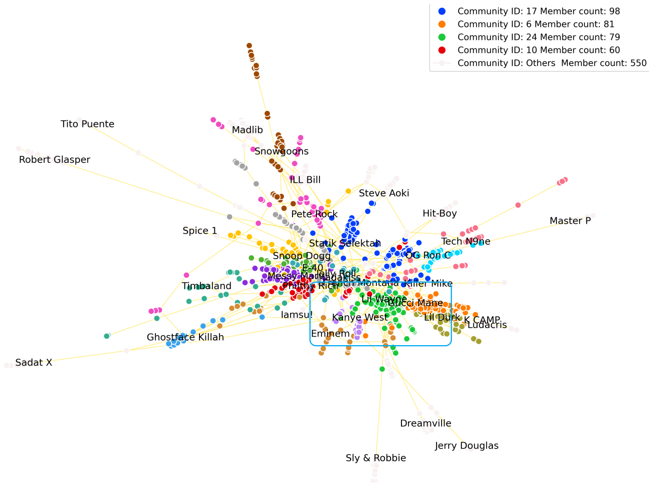
Friendship Paradox. An important characteristic of heavy tailed distribution is the manifestation of Friendship Paradox wherein a new incoming node has a higher chance of connecting with an existing node with very high degree than any other node(as opposed to the hypothesis of equal chance of connecting with any node). Presence of Friendship Paradox indicates that some artists have collaborated a lot more than the others in the network. The test for friendship paradox is conducted for both the super-network and the largest component. It is seen that Friendship Paradox is seen to exist 90% of the times in both cases. This, however, cannot necessarily imply that the high degree artists are more sought-after for collaboration. It may just be their style of music that calls for collaboration and may not have a direct relationship with their popularity.

Centrality, popularity and follower counts. Figure 3a(i) shows a heatmap of the correlation across the measures. Figure 3a(ii)-(v) are scatter plots to look for any possible non-linear variation which may not be detectable by correlation. Degree centrality and Eigen vector centrality(Figure 3a(ii) have moderate correlation (0.63 per Figure 3a(i)). It indicates that not all nodes that have high degree would also be "ranking" higher than others. Subsequent study of communities reveal more. Popularity and Eigen vector centrality don't appear to have any relation. Therefore, a higher collaboration ranking has little to do with popularity. This is understandable for the other genres genre but it takes considerable contemplation for Rap/Hip hop Genre as it thrives on collaborations. It could be explained by the phenomenon of many HipHop artists collaborating on one song; so they may not be hugely famous individually but the song might still do well (be more popular), or the lesser known rappers try to collaborate with as many rappers as they can but the more successful rappers keep their circle small. A similar sentiment is echoed when comparing betweenness and popularity, a very high collaboration has little to do with popularity. Finally, unlike as it may seem, follower count and popularity measure aren't correlated either as "followers" is an actual count of users who follow the artist on Spotify(assuming that synthetic increase of follower count makes up for a small proportion and doesn't affect the overall metric) and popularity is a calculated measure based on how often is an artist listened to. This may point towards a problem with reliability of the follower count (may have been increased synthetically).

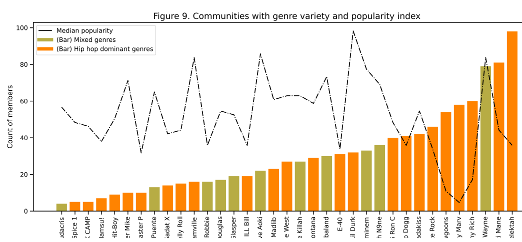
Network communities & popularity. To understand the network better, the Louvain algorithm is deployed to identify communities within the network. Figure 3c shows the communities by size. The largest community is 10% of the total nodes whereas the smallest community has 4 nodes. The three biggest communities are 55% bigger than the next set of communities. Further the plot segregates communities into ones that are dominated by Hip hop artists dominant vs the ones that are shared with genres (also referred to as mixed genres or non Hip hop genres). The median popularity is also



(a) Centrality vs popularity and followers



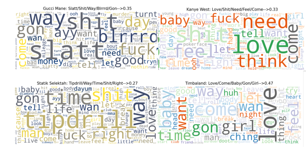
(b) Communities identified by Louvain



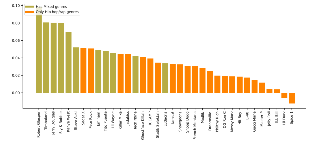
(c) Community & popularity & genres

Fig. 3. Centrality, Popularity and follower counts a(i)indicates popularity has little to do with degree, centrality or even follower counts.a(ii)-a(v) scatter plots indicate that high Eigenvector centrality or degree "ranking" doesn't influence popularity.Very low betweenness can have very high popularity which could be artists who are very successful in their solo work. Follower counts may be synthetic or static as it does not reflect the popularity. (b) Communities of Lil Wayne, Gucci Mane, Eminem, Kanye west (demarcated in the Figure for quick reference) are collocated and lie in the middle indicating central role in the network and relatedness in their communities (d) Small and/or mixed genre communities have higher popularity compared to big Hip hop dominated communities

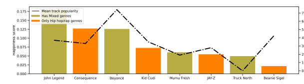
(a) TFIDF wordcloud of Hip hop dominant vs mixed genre communities



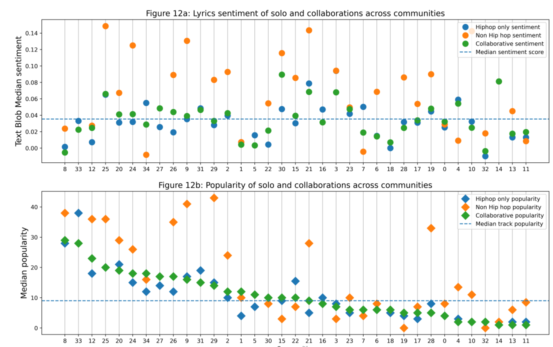
(b) TFIDF wordclouds of 4 communities



(c) Textblob sentiment score across communities



(d) Artist wise lyrics sentiment within a community



(e) Sentiment values solo vs collaborative

Fig. 4. Lyrics sentiment of collaborative songs (a)Hip hop dominant genre wordcloud(left) have sadder tokens compared to the mixed genre wordclouds(right) (b)Communities on the right have happier words whereas the ones on the left have swear words and references to crimes.(c) Sentiment scores sorted in descending order show mixed genre communities have happier lyrics than their counterparts (d)Non hip hop genre artists have higher average sentiment score than their Hip Hop counterparts in their respective solo songs. (e) [top] Non Hip hop genre sentiments are generally above the median sentiment reference line whereas the sentiment of the Hip hop artists closely follows the sentiment of collaborations.Solo trends varying directly with the collaboration sentiment indicate that artists with similar lyrical styles prefer collaborating with each other. [bottom] In the communities in the left half, the Hip hop artists benefited more from the collaboration that their counterparts. In some communities in the middle, the non Hip hop artists may have gained more popularity by collaborating with their Hip hop counterparts

plotted on another axis. It is clearly seen that the biggest communities have low popularity whereas popularity peaks for mixed-genre and small communities.

Lyrics sentiment across communities. Evaluating lyrics based on TFIDF and wordclouds alone is not significant as the highest TFIDF values seen are very low (0.28 and 0.44 respectively for Figure 4a) therefore NLP methods which better represent the sentiment of a document via a lexicon based sentiment analyser or word embedding based contextual sentiment analyser was called for. Three methods were deployed to calculate the sentiment scores of the song lyrics (higher the sentiment score, higher is the happiness quotient). The figure 4c bar plot shows a clear trend that the collaborations across mixed genres have a more positive sentiment than the communities with Hip hop-dominant artist collaborations. Comparing methods, the happiness score based on LabMT file offers the least reliable insight although it scores the artist in almost the same way and therefore the ranking seems unreliable. NLTK Vader and Text Blob sentiment values have a lot more variance and the order of artists is almost the same (detailed analysis available in the explainer notebook). The Textblob sentiment is built on top of NLTK corpora and is faster than NLTK Vader and is therefore chosen as the preferred method for sentiment analysis for this study.Another method to explore was Flair (6) which is a Vector space model and therefore utilises word-embeddings to capture the context of the sentences. The signature embedding "Flair embedding" is supported by contextual string embedding i.e. it interprets the meaning of words given the context it is used in (7). The method was kept out of scope for this study due to computational limitations.

The word clouds of some select communities are shown in Figure 4b with the mixed genre communities color-coded with a brighter color scheme. Words such as *love*,*baby*,*feel*, *heart*,*girl* are seen in the wordclouds of the mixed genre communities which point towards a romantic inclination. On the other hand, curses and material wealth such as *money*,*diamond*, or crime such as *murder* are mentioned in the Hip hop dominant genres. The TFIDF word clouds are however not conclusive evidence as the TFIDF values are itself very low (0.27 to 0.47) which indicates lower exclusivity across communities.

In order to be more confident about the conclusions, the communities were bifurcated into Hip hop dominant and mixed genre communities and their average sentiment scores were passed through a permutation or randomisation test. This is to check the null hypothesis that the sentiment scores of both these categories come from the same distribution. The test shows that the difference between the distributions is significant at α value < 0.05 and therefore the null hypothesis is rejected thus substantiating the difference between sentiment scores of the two classes.

Contrast solo music styles with collaboration styles. The above results indicate that Hip hop genres are , generally, more sad, atleast lyrically and the behaviour changes when they are collaborating with other genres. This further encourages the investigation of lyrics sentiment of solo artists from hip hop and other respective genres and how that compares against the sentiment of the collaborations. To be able to compare how a Pop artist,say, writes their lyrics for their solo tracks and how do the lyrics change when they collaborate with a Hip hop artist is one of the factors that would define the character of the communities. For this purpose, one sample community of sizeable member count is chosen. Figure 4d shows the average sentiment value of randomly chosen artists (four non Hip hop and four Hip hop artists) from the community. More specifically, the sentiment for each artist is calculated based on their solo tracks. It can be clearly seen that the non Hip hop artists have a higher sentiment score and popularity as well. The wordclouds for Hip hop and non-Hip hop artists' lyrics in Figure 4a supports the observation about the sentiment to an extent. The non Hip hop word cloud have happier words such as *dance*, *love*, *time*, *baby*, *girl* ie about the lighter aspects of life. The latter, appear to be curse and have words *bad*, *hard* etc. There are many words common between the wordclouds; however the curses and the mention of material elements and sadder adjectives render the Hip hop genre wordcloud a grim or realistic tone (than a dreamy tone). This shows a distinction between the lyric styles of Hip hop artists and non Hip hop artists.

Audio attributes. Given the success international music has received in recent years, the relevance of the music composition in determining the mood of the song has increased manifold. It is therefore rudimentary to study audio features to understand the nature of music different artists make in their solos and how that changes when they collaborate. Figure 5a shows the histograms of the various audio attributes of the songs. It can be seen that some of the attributes are multimodal indicating that it could be a mix of multiple distributions. At the onset, valence is ignored as it is Spotify's method of measuring the happiness of the track which could be a combination of both lyrics and audio features and no information is available about it.Instrumentalness and liveness are not relevant for this study and mode has very low variance. From Figure 5b it can be seen that loudness and energy are highly correlated therefore just one can be analysed for any patterns. The remaining attributes are analysed community-wise for any difference in patterns between the collaborations, the solo tracks of the Hip hop artists and those of the non Hip hop artists.

Figure 5b shows that the collaborations have lower tempo compared to the non Hip hop solo songs and therefore follow the beats frequency of the Hip hop artists. This falls in line with the trend observed with speechiness (lower tempo allows for rapping) where the collaborations are as verbose as the Hip hop artists and are way higher than the non hip hop counterparts. The danceability of collaborations are higher than non Hip hop artist solos and follow that of hip hop solos as the genre is also immensely popular as a dance form of the same name. Energy shows a similar pattern to danceability although no linear relationship(correlation) was observed.Finally, acousticness seems to be higher for the non hip hop solos and that for the collaborations and the hip hop solos follow each other. This makes sense as raps are mostly driven by beats and other genres are more about melody although some new age rappers are changing that (8).

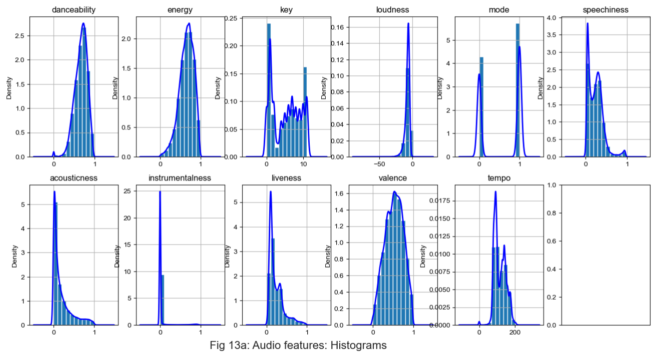
Methods

Tools & Packages. Network analysis package *Networkx* is used to build the graphs. Louvain algorithm is chosen over Infomap as it is comparatively faster (9).Also, the Infomap algorithm loses scalability as the number of communities generated rise exponentially with network size. Louvain algorithm also creates internally disconnected communities (10) it may merge two communities that are connected from outside the community. Similarly a connecting node between two dense communities might be moved to a different community altogether. Sentiment analysis was carried out using three methodologies *viz.,* LabMT, NLTK Vader Sentiment Analyser & TextBlob Sentiment Analyser, all of which are Lexicon based methods. Another package, Flair, that is based on word-embeddings is also known to give very good results as word embeddings are better at capturing the context. However, the method couldn't be used as it is very time and resource intensive.

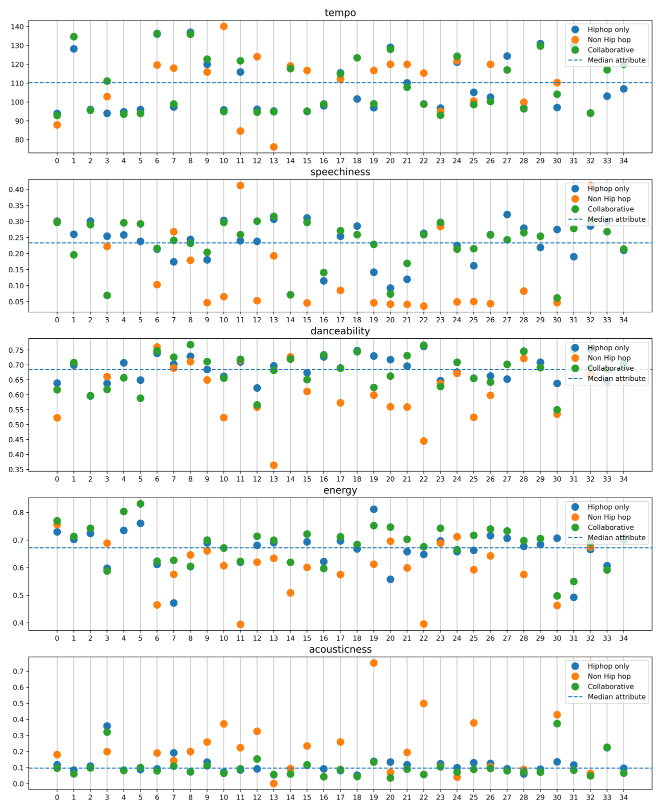
Data sources. Data was sourced from three places. The list of artists was compiled from Wikipedia. The artists were searched in the Spotify database via the Spotipy API and the tracks performed by those artists were queried into the Genius.com data via the Lyrics Genius API. Github link : https://github.com/antarlinam/SG

Discussion

Conclusion. The music artists collaboration network follows a heavy tailed degree distribution but not necessarily a scale free network as music collaborations have physical limitations and therefore cannot be scale free. Friendship Paradox that is found to exist does not imply popularity (as may be native to intuition) but suggests the working style of the Hip hop artists. The study has also found evidence of assortativity amongst the artists. The popular artists are more likely to collaborate



(a) Histograms of audio attributes



(b) Solo vs collaboration across communities

Fig. 5. Centrality Popularity and follower counts (a) Some of the attributes are multi-modal indicating that it could be a mix of multiple distributions (b) The collaborations have lower tempo compared to the non hip hop songs and therefore follow the beats frequency of the hip hop artists. This falls in line with the trend observed with speechi-ness (lower tempo allows for rapping) where the collaborations are as verbose as the hip hop artists and are way higher than the non hip hop counterparts. The danceability of collaborations are higher than non hip hop artist solos and follow that of hip hop so-los as the genre is also immense popular as a dance form of the same name. Energy shows a similar pattern to danceability although no linear relationship(correlation) was observed.Finally, acousticness seems to be higher for the non hip hop solos and that for the collaborations and the hip hop solos are in line

with other popular artists and that could be motivated by an incentive of success on the music charts. Further, two major styles of lyrics have emerged, one that is from the house of Hip hop and is characterised by darker elements manifested through curses and negative connotations. The other house is that of Pop music which is the second biggest genre collaborating in music. The sentiment conveyed through lyrics is more about love, relationships and at times have a sad undertone but continue to have a far happier sentiment compared to the hip hop counterparts. The collaborations between the hip hop artists and their counterparts from other genres borrow the lyrical style from the hip hop artists. As rap artists are bound to be verbose, they command the lion's share of the words spoken in the song and therefore determine the sentiment lyrically. However, it is interesting that lyrics are losing relevance and the audio attributes determine the mood of the song. The acousticness of the collaborations are borrowed from the non Hip hop styles whereas the other attributes such as speechiness, tempo, danceability are borrowed from the Hip hop genre.

Assumptions. It is assumed that the data sourced from Spotify via the API is comprehensive and is representative of the music collaboration network. All artists credited in the song have participated in similar capacity and therefore are equal stakeholders in the lyrics and composition. The popularity index of Spotify is representative of the artist's actual popularity amongst the masses. The use of curses and reference to crimes (or reference to acts that are often related to it) imply a negative tonality. The sub variants of Hip hop are mostly region based (11) and are largely similar in lyrics and composition. The backbone sampling method is able to produce a downsized version of the super-network that does not contrast with any patterns that would have otherwise been observed while studying the super-network itself.

Future directions. An important aspect of music is how the song is split between the participant artists during a collaboration. The role of each type of artists would be better understood if the chorus and raps were analysed separately. As Hip hop and Pop were the pre dominant genres in the collaboration network, the dynamics of Hip hop and other genres may have been neglected which could be different from the results obtained in this study. Additionally, the speed of rap delivery could also play a role in defining the sentiment of the song, which nowadays is far beyond the lyrics. Additionally, word embeddings- based NLP models could be used to identify the topics in order to tag the songs. The song tags, along with technical details of the music composition could be useful to predict the popularity. Additionally, it would be interesting to study the halo effect collaborations have on the popularity of an artist by compared the artist's popularity before, during and after the release of the collaborated music and the duration of the halo effect.

ACKNOWLEDGMENTS. The author of the report would like to thank Prof. Sune Lehmann Jorgensen for providing the opportunity to prepare this project. They would additionally like to thank the TA Mia, Silvia, Louis, Nojus for guiding through the course.

References

1. M Ángeles Serranoa, backbone of complexweighted networks. *https://www.pnas.org/* (2009).
2. Barabasi, *Network Science Book.* (2016).
3. Barabasi, The γ dependent properties of scale-free networks. (2016).
4. MEJ Newman, The structure of scientific collaboration networks. *pnas* (2001).
5. ADBA Clauset, Scale-free networks are rare. *https://www.nature.com/* (2001).
6. A Akbik, An easy-to-use framework for state-of-the-art nlp. *ACL Anthol.* (2020).
7. A Vidhya, Introduction to flair for nlp:. (2019).
8. Fuelrocks.com, Interesting hip hop sounds on acoustic guitar. (2022).
9. VL Dao, Community structure. *HAL Arch.* (2008).
10. VA Traag, From louvain to leiden. *nature.com* (2019).
11. Wikipedia, Types of rap. (2022).