

# Analyzing the Musical Characteristics of Viral TikTok Songs

Justin Sech  
jsech1@jhu.edu  
Johns Hopkins University

## ACM Reference Format:

Justin Sech. 2021. Analyzing the Musical Characteristics of Viral TikTok Songs. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 INTRODUCTION

### 1.1 Motivation

Technology has always been a major influence in how we experience music — both in how we listen and how it is created. The invention of the radio brought the sub 3 minute “pop” song format. Long-playing vinyl allowed artists to craft the full length “album” that we know today. Mixtapes came from the accessibility of cassette tapes and CDs, and now instant, on-demand Internet streaming brings an unlimited new wealth of variety to how we consume and create music.

The recent surge in the short-form video social media site, TikTok [9], has made an undeniable impact on the current state of the music industry. In December 2018, the never before heard artists, Lil Nas X, released his song “Old Town Road”. Thanks to TikTok’s ease of virality, “Old Town Road” quickly amassed millions of plays, reached the top of the Billboard charts, and Lil Nas X was subsequently nominated for 6 Grammy Awards, winning 2 [3]. Lil Nas X and “Old Town Road” may be the most prominent example, but a song making its viral debut on TikTok is not an isolated incident.

### 1.2 Research Goals

This paper investigates the dynamic between TikTok and the rest of our modern day music industry. Specifically, this work aims to answer the following research questions:

- (1) Do songs trend on TikTok before they reach the Spotify or Billboard charts
- (2) Is there a measurable distinction in songs that first trend on TikTok compared to songs trending equally across all platforms or that reach TikTok later

This work explores Question 1 by analyzing Trending Sound data scraped from TikTok against archival music charts from various Spotify and Billboard chart categories. Question 2 is subsequently evaluated using meta data and audio analysis data obtained from the Spotify API.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*Conference'17, July 2017, Washington, DC, USA*

© 2021 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 2 BACKGROUND

### 2.1 Related Work

While there is an increasing amount of research focused on TikTok’s influence on youth [1] and the general world of social media, TikTok’s growing significance in the music industry remains relatively unstudied. Hobbist research articles, such as [11], have assessed the musical-characteristics attributed to general TikTok songs, and other research has evaluated the social network structure of Spotify [5]. However, there is a lack in research evaluating cross-platform comparisons between trending songs on TikTok versus other platforms. This work aims to fill that research gap by comparing trending information across TikTok, Spotify, and Billboard.

### 2.2 ERGM

Exponential Random Graph Modeling (ERGM) [4] is a statistical method used to understand the critical factors and structural features in the formation of a network. ERGMs create a probability distribution over all possible network structures to determine the driving effects of the actual network structure. This distribution can be represented as:

$$P(Y = y | \theta) = \frac{\exp(\theta^T s(y))}{Z(\theta)} \quad (1)$$

where  $Y$  is the network,  $s(y)$  is a vector of network parameters,  $\theta$  is a vector of model parameter weights, and  $Z(\theta)$  is a normalizing constant.

## 3 DATASET

### 3.1 Data Collection

All of the data collected in this work comes from 3 independent sources.

**Tokboard.** Tokboard.com [10] is an unofficial web archive of trending TikTok song data. Tokboard constantly fetches data directly from TikTok and tracks the 100 most viewed sounds on TikTok per month. This work scraped Tokboard for all of its monthly data ranging from August 2018 through April 2021 as this was as much information hosted on the site at the time of this research. We will call this data the set of TikTok Trending Charts.

**Billboard.** The Billboard Hot 100 [2] is a weekly ranking based on sales, radio play, and streaming numbers in the United States. For this paper, the Billboard website’s archive of its weekly Hot 100 charts was scraped for all chart information for the week of June 28th, 2018 to April 24th, 2021.

**Spotify.** Similar to Billboard, Spotify publishes its own platform-based charts. The Spotify Viral 50 [7] tracks “trending” songs based on Spotify streaming data per week. The Spotify Top 200 [6] is a

definitive ranking of the 200 most played songs per week on the streaming platform. Both of these charts were scraped for the weeks of June 28th, 2018 to April 24th, 2021.

Spotify also provided all metadata for each artist and song as well as specific audio features that served as the basis of analysis for this work. Both types of information were fetched directly through the Spotify API for every song in the set of all songs scraped from the TikTok, Billboard, and Spotify charts.

### 3.2 Network Creation

The first step in building any network was identifying Viral TikTok Songs—songs that appeared on the TikTok Trending Chart before their debut on either Billboard or Spotify charts. Since the TikTok trending data is in per-month form and the Billboard and Spotify charts are weekly, a Viral song, as defined in this paper, appears on the TikTok Trending Chart at least one month in advance of its debut on either Spotify or Billboard charts.

Two adjacency matrices were created based on the set of Viral TikTok Songs

- (1) *Song  $\times$  Chart*: assigns a link between a Viral Song and the chart(s) it's TikTok debut preceded.
- (2) *Song  $\times$  Genre*: assigns a link between a song and each of the genres designated to that song's Artist

### 3.3 Network Attributes

Every song in each network was attributed a list of audio features obtained from Spotify's API [8].

- *song\_popularity*: a measure of a song's popularity from 0 to 100
- *artist\_popularity*: a measure of an artist's popularity from 0 to 100; calculated from the popularity of all of the Artist's songs
- *followers*: an artist's number of followers
- *danceability*: how suitable a song is for dancing; 0.0 to 1.0
- *energy*: a measure of a song's intensity and activity; 0.0 to 1.0
- *key*: indicates a song's key signature; 0=C, 1=C#, etc.
- *loudness*: overall loudness of a song in decibels
- *mode*: 1 indicates major, 0 indicates minor
- *speechiness*: a measure of the presence of spoken word in a song; 0.0 to 1.0
- *acousticness*: a confidence measure from 0.0 to 1.0 of whether the track is acoustic.
- *instrumentalness*: predicts whether a track contain no vocals
- *liveness*: detects the presence of an audience in the song recording; 0.0 to 1.0
- *valence*: a measure of a songs "musical positiveness"; 0.0 to 1.0
- *tempo*: algorithmically estimated tempo
- *duration\_ms*: the length of the song in milliseconds
- *time\_signature*: the song's estimated time signature

## 4 EXPERIMENTAL SETUP

ERGM methods were used to evaluate the critical network features for both *Song  $\times$  Chart* and *Song  $\times$  Genre* networks. Analyzing the characteristics of *Song  $\times$  Chart* network will help us understand

what aspects of a song contribute to it TikTok Virality and *Song  $\times$  Genre* will identify distinguishing characteristics between genres. A restricted model of just the network structure was first evaluated to test the structural effects of edges in the networks. Both networks are directed and acyclic without any mutual ties or triadic effects. Therefore only the network edges were used in the restricted model.

After accounting for network structure with the restricted model, an unrestricted model was run to assess all Network Attributes. This unrestricted model will allow us to understand which non-structure features about our network are most prominent in creating links. In other words, assessing this model will identify which song characteristics are most significant and distinguishable in Viral TikTok songs.

## 5 ANALYSIS

### 5.1 Experimental Results

Analysis of both *Song  $\times$  Chart* and *Song  $\times$  Genres* networks illustrates that nearly every song characteristic can be considered significant in distinguishing Viral TikTok songs from the rest of the dataset. We will elaborate on these specific differences in the following section.

Attribute	Song $\times$ Chart		Song $\times$ Genre	
	Est.	Sig @ 0.01	Est.	Sig. @ 0.01
edges	1.510e+01	Yes	1.137e+00	No
artist_popularity	3.176e-02	Yes	3.919e-03	Yes
followers	-3.416e-08	Yes	9.910e-09	Yes
song_popularity	-4.228e-02	Yes	2.653e-03	Yes
danceability	-4.794e+00	Yes	2.423e-01	No
energy	-4.378e+00	Yes	3.355e-01	Yes
key	2.373e-01	Yes	1.312e-01	No
loudness	1.155e-01	Yes	1.926e-02	Yes
mode	1.951e-01	Yes	8.004e-02	No
speechiness	7.860e-02	No	2.460e-01	No
acousticness	-4.983e-01	Yes	1.872e-01	Yes
duration_ms	1.784e-06	Yes	7.425e-07	Yes
instrumentalness	-9.476e-02	No	1.125e+00	No
liveness	4.119e+00	Yes	2.636e-01	Yes
valence	3.347e+00	Yes	1.483e-01	Yes
tempo	-2.758e-03	Yes	1.068e-03	Yes

Table 1: ERGM Results

### 5.2 Viral TikTok Song Profile

#### • Song Attributes

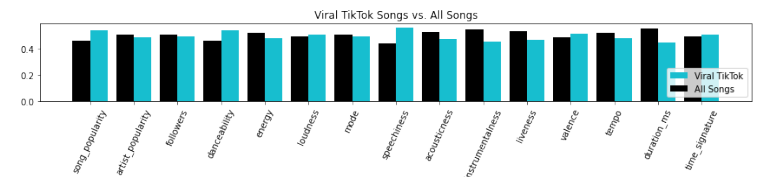


Figure 1: Attributes of Viral TikToks vs All Songs

The musical characteristics of Viral TikTok songs are notably distinguishable from the rest of the Billboard and Spotify Charts.

As seen in Figure 1, TikTok songs are significantly more danceable, more vocal (less instrumental), and are overall shorter songs. Intuitively, this aligns with our general understanding of TikTok as a platform. A significant portion of TikTok is dedicated to replicating choreographed dances over the same "trending" song. Additionally, TikTok videos are a maximum of 60 seconds in length, so it makes sense that shorter songs (with shorter verses/choruses/etc) are better suited for the platform. Other features of Viral TikTok Songs are that they are generally at slower tempos with less "energy". Viral TikTok songs are more frequently in minor keys, however, they are also happier (higher "valence") than the average Chart song. Trivially, Viral TikTok Songs have a higher song popularity—this makes sense as they are "viral" songs—however, Viral TikTok Artists are measurably *less popular* than the artists appearing on the Billboard and Spotify charts. This supports the theory that a song from an unknown artist (such as Lil Nas X's "Old Town Road") can make its debut on TikTok before blowing up (going viral) in the rest of the music industry.

### Genre Breakdown

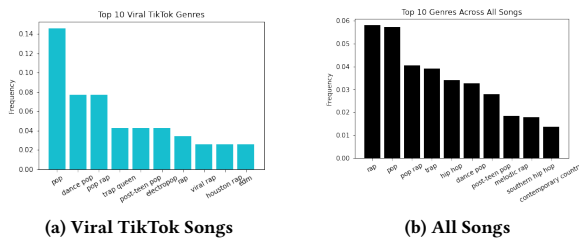


Figure 2: Genres of Viral TikToks vs. All Songs

Comparing the genre distribution of Viral TikTok Songs to that of the average song in the full dataset leads to additional, higher-level insights about what songs become most popular on TikTok compared to the rest of the music industry. As shown in Figure 2, rap is the most common genre that appears across both Billboard and Spotify but it is significantly less common for Viral TikToks. Instead, Viral TikTok Songs most notably favor "dance pop" and "pop" more than twice as frequently than on the Billboard or Spotify Charts. This discovery pairs well with the higher "danceability" score of Viral TikTok Songs in the previous section. Additionally, "country", "hip hop", and "southern hip hop" do not appear in the 10 most frequent TikTok genres at all, and rather "viral rap", "edm", and "houston rap" (designated to A-list rapper Travis Scott) take their place. These genres are likely to be more danceable ("edm" meaning electronic dance music) and more singable (rap) which matches how we have already described the TikTok platform.

## 6 CONCLUSION

This paper presents one of the first cross-platform analyses between TikTok and the modern music industry. We have shown that there is an identifiable set of songs that build momentum on TikTok prior to making a "mainstream" breakthrough on either the Spotify or Billboard record charts. This is a significant finding as we have identified a targeted location for "underground" music.

Additionally, this work distinguishes what makes a Viral TikTok Song noteworthy against the average Charting song. We found that Viral TikTok Songs are generally well fit for the platform itself—they are danceable, singable, and short enough to capture peoples attention. These distinctions seem to fit the specifications of an archetypal pop song and it is unsurprising that songs that leverage these metrics are most likely to be a viral hit.

This work presents a relatively shallow dive into the world of what makes a song go viral and there is still significant analysis and reasons to be done. The dataset collected for this work alone is extremely rich and provided many other points of analysis that are beyond the scope of this paper but may serve as a strong basis for future work.

## REFERENCES

- [1] Arif Mushtaq Badriya Abdul Jaffar Sadia Riaz. "Living in a Moment: Impact of TicTok on Influencing Younger Generation into Micro-Fame". In: *Journal of Content, Community, and Communication* (2019). URL: <https://www.amity.edu/gwalior/jccc/pdf/jccc-12-19-19.pdf>.
- [2] LLC Billboard Media. *Billboard Hot 100*. URL: <https://www.billboard.com/charts/hot-100>.
- [3] Andrew R. Chow. "It Feels Like I'm Chosen to Do This." *Inside the Record-Breaking Rise of Lil Nas X*. URL: <https://time.com/5652803/lil-nas-x/>.
- [4] Johan Koskinen Dean Lusher and Garry Robins. "Exponential random graph models for social networks: Theory, methods, and applications". In: *Cambridge University Press* (2012).
- [5] Tobin South. "Network Analysis of the Spotify Artist Collaboration Graph". In: (2018). URL: [https://vrs.amsi.org.au/wp-content/uploads/sites/78/2018/04/tobin\\_south\\_vrs-report.pdf](https://vrs.amsi.org.au/wp-content/uploads/sites/78/2018/04/tobin_south_vrs-report.pdf).
- [6] Spotify. *Spotify Top 200*. URL: <https://spotifycharts.com/regional>.
- [7] Spotify. *Spotify Viral 50*. URL: <https://spotifycharts.com/viral/>.
- [8] Spotify. *Spotify Web API Documentation*. URL: <https://developer.spotify.com/documentation/web-api/reference/#object-audiofeaturesobject>.
- [9] TikTok. URL: [tiktok.com](https://tiktok.com).
- [10] Tokboard. URL: <https://tokboard.com>.
- [11] Josh Viner. *How To Make A Viral TikTok*. URL: <https://medium.com/the-dopamine-effect/how-to-make-a-viral-tiktok-song-ece56a838a30>.