What Makes a Music Track Popular in Online Social Networks?

Jing Ren
Singapore Management University
jing.ren.2012@smu.edu.sg

Jialie Shen
Singapore Management University
jlshen@smu.edu.sg

Kauffman, R.J.
Singapore Management University
rkauffman@smu.edu.sg

ABSTRACT

Tens of thousands of music tracks are uploaded to the Internet every day through social networks that focus on music and videos, as well as portal websites. While some of the content has been popular for decades, some tracks that have just been released have been completely ignored. So what makes a music track popular? Can we predict the popularity of a music track before it is released? In this research, we will focus on an online music social network, Last.fm, and investigate three key factors of a music track that may have impact on its popularity. They include: the music content, the artist reputation and the social context of the music. The results suggest that we can predict the future popularity of music with around 80% accuracy using just these three factors. We also found out that in the social networks scenario, the content of the music seems to be an surprisingly important factor that determines the popularity of a track online.

Keywords

Music Popularity Online, Acoustic Content, Topic, UGC

1. INTRODUCTION

With contemporary modes of digital entertainment, people can easily access very large music collections and stream their contents via social networks such as Last.fm, Spotify and YouTube. In comparison to the timeliness and relevance of other kinds of social media, music is a *durable information good.* A music track can bring unexpected value and utility to the listener. And even one strong and widely appreciated song can lead to the rise of a new music superstar, such as "Rolling in the Deep" for Adele, or "Poker Face" for Lady Gaga. Moreover, a classic track can make people remember the singer and encourage them to keep listening or buying the artist's albums, even may years after they were released. Examples include "Hey Jude" by The Beatles, which Billboard named the "10th biggest song of all time" in 2013, although it was first released in 1968.

What are the secret ingredients for music to achieve high popularity? By discovering them, we can predict whether a song will become popular or even predict who may become the next music superstar. Prior research [2, 3] has focused on predicting the popularity of text, image and video, or even a person [1] such as a fashion model or a music artist. However, to our knowledge, there has been no previous work that studied the inherent reasons of why a song became popular. In this work, we will analyze the key factors that determine the popularity of online music. We will leverage music track adoption data from the online music-focused social network, Last.fm, to investigate three key attributes that seem to affect a music track's popularity. These determinants are: the music content, the artist's reputation, and the contemporary social context that was present when the track was released. Social influences act in a pervasive way to advertise the nature of the music that people listen to. So we ask:

- Through the analysis of a track's content, can we predict its popularity to a reasonable degree?
- Does an artist's reputation affect the popularity level of a music track?
- Does the addition of social context in a model improve the prediction of a music track's popularity?

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2889402.

2. WHAT IS MUSIC POPULARITY ONLINE?

There are various ways to define the popularity of a music track. They include: how many times of a song has been listened to; the sales volume of the related album; the amount of time of its radio play; and the music industry awards that it received. We focus on the popularity of a specific track in online social networks, a narrower and more specialized environment. The popularity of music in social networks seems to endure over time. This is a useful indicator of how long a period of time that the track can maintain relevance and appeal to its audience's tastes.

To measure and define the duration of the popularity of a music track, we used data from the *weekly listening logs* that are available via Last.fm. It lists the top-150 music tracks each week in terms of how much they have been listened to. We collected these music rankings from the period May 2005 to December 2013. This led us to download 512 weeks of such data. This resulted in over 12 million streamable tracks from Last.fm. But in the years that we targeted, far fewer made it into the top-150 ranking: just 4,500 tracks or about 0.04% of the total. Figure 1 shows the distribution of appearances of music tracks in the top-150 ranking.

A long tail distribution is evident, with around 80% of the tracks appearing in the ranking for less than 20 weeks. A large proportion of the tracks tend to diminish in the attention they are able to draw from listeners shortly after they are released. Based on descriptive statistics, we can define the popularity of music online in terms of its duration in the top-150 ranking. A music track reaches "*popularity online*" if it achieves a duration of no less than 20 weeks

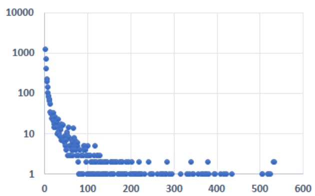

Figure 1: Log-distribution of top-150 tracks (duration in weeks)

in the ranking; and if it appears less than 20 weeks in the top-150 ranking, then we view it as "*not achieving popularity online*".

3. EXPERIMENTS & ANALYSIS

For our empirical research, we sampled 1,961 tracks from 305 different artists. Among them, we labeled 628 tracks as having reached popularity online, according to our application of the operational definition of this construct. For each track, we crawled specific websites to identify three kinds of attributes of music tracks. The related websites included Last.fm, Wikipedia, 7digital, GoogleLyrics and MusicSongLyrics. *Table* 1 shows the three attributes that came out of the collected data.

Attributes	Description
Music Acoustic Content	Tempo (fast, moderate, slow); Melody (pitch, rhythm);
Music Topics	Topic distribution on 5 topics learned from Lyrics using LDA
Artist Reputation	# of 3 big music awards; Associated to 3 big recordlabel
Social Context	Weighted UGC tags; # of comments in the first 5 weeks since released

Table 1: Three attributes of music popularity online.

The first two attributes of the music tracks are content related. These have some degree of semantic meaning related to each of the latent topics learned by an *Latent Direchlet Allocation* (LDA) topic model. We found that more than 51% of the highly ranked music tracks had contents pertaining to "love" and other positive and optimistic emotions, and this percentage increased to about 61% for music tracks that were operationally defined in this research as having achieved popularity online.

To answer our three research questions, we employed several analytical approaches to detect the impact of the three different attributes for the popularity of music online, including: *decision tree analysis* for a set of baseline results; and then *support vector machine* (SVM); *random forest* (RF); and *bagging*. Figure 2 and *Table* 2 present our results for the values of prediction accuracy and kappa.

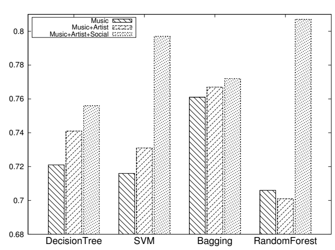

Figure 2: Popularity prediction accuracy.

We observe that predictive power appears to increase from the left to the right in the figure: in other words, from the case of a single attribute to multiple attributes. Although not all of the algorithms that we used appear to have been able to demonstrate noticeable improvements compared to the importance of the pure content attribute of the music, it still is possible to comment on the findings relative to our research questions (RQs).

Kappa	Music	Music+Artist	All
Decision Tree	0	0.233**	0.408**
SVM	0.078	0.121	0.378**
Bagging	0.320	0.355	0.332
Random Forest	0.225	0.217	0.474**

Table 2: Algorithm performance for Kappa. **: $p<0.05$

Music content (RQ1). No matter which algorithm was used, the music content attribute alone achieved more than 70% accuracy for the prediction of music popularity online. The music attributes (or topics) indicate that the popular ones were largely positive, cheerful, and about love and life. We conclude, as a result, that music content is an important determinant for the prediction of music popularity online.

Artist reputation (RQ2). In accuracy and kappa terms, there was only a small improvement compared to pure music content, especially for the Bagging and RF learning methods. Consider the album "Red" by Taylor Swift for example. It was released in 2012, and one track, "We Are Never Ever Getting Back Together", was popular online on its own, although the artist won four music-related Grammy Awards in 2010. In addition, she was associated with a large and powerful record label, the Universal Music Group, an advertising leader. We conclude from this kind of analysis that it appears an artist's reputation is an important determinant for their popularity in general terms, although it may not be so important for their popularity online in the terms we have used to operationalize this construct. For example, the artist's reputation may be an important determinant of the person's ability to win music awards and drive album sales, similar to the star power of leading actors' impacts on movie ticket revenues.

Social context (RQ3). Social context appears to have had a positive impact for the data sample that we analyzed. Social information such as user-generated content (UGC) seems to have been useful in supplementing music content as a predictor for a music track's popularity online. User tagging shows that pop and rock music tracks are more likely to have longer durations of popularity online. And the number of user comments in the first five weeks since a tracks was released were useful for predicting its popularity online.

4. CONCLUSION

We studied the importance of three different determinants of the popularity of music online based on an LDA topic model, and the discovery of three core attributes that seem relevant for the content. We conducted this work using the music social network, Last.fm, and found that the content of the music was an important determinant of a music track's time duration in terms of weeks of popularity online. In addition, the social attention of music listeners appears to be another important determinant of future popularity online based on what happens during the early stage of a music track's diffusion. We plan to study this more deeply through additional analysis of the temporal and dynamic popularity patterns of music tracks in the future.

5. REFERENCES

[1] Park, J., Ciampaglia, J.C., and Ferrara, E. Style in the age of Instagram: predicting success within the fashion industry using social media. *arXiv preprint arXiv:1508.04185*, 2015.

[2] Petrovic, S., Osborne, M., and Lavrenko, V. Rt to win! predicting message propagation in twitter. In *Proc. Intl. Conf. Web Social Learning*, Barcelona, Spain, AAAI Press, Palo Alto, CA, 2011.

[3] Pinto, H., Almeida, J.M., and Gonçalves, M.A. Using early view patterns to predict the popularity of youtube videos. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, ACM Press, New York, NY, 2013, 365-374.