A Robot Economy for Music Without Any Middleman

Tim Wissel

A Robot Economy for Music
Without Any Middleman

by

Tim Wissel

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on XXX at XXX.

Student number: 4460251
Project duration: March 6, 2020 – XXX
Thesis committee: Dr. ir. J. Pouwelse,   TU Delft, supervisor

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

cover/logo_black.pdf

Preface

Preface...

Tim Wissel
Delft, March 2020

Contents

1

Introduction

Most audio streaming services are run by companies, incentivized to make money. They take large cuts of the subscription money from its users. As a result, the artists receive a low compensation. The distributors Spotify, iTunes and Google Play take on average a 25% cut for signed records and 40% cut for unsigned records. This thesis investigates the feasibility and usability of an audio streaming service without a central distributor.

This thesis proposes a solution in the form of a decentralized system which uses a decentralized autonomous organization (DAO) to operate. Listeners, artists and robots form this DAO. The DAO has a shared responsibility for distributing content. In this system, its users (artists and listeners) share audio files and metadata without any middleman. Additionally, users can give donations to artists while the system does not take a cut of these donations. The user can use this system to discover, search and play audio files, targeted at music and podcasts.

Section X describes the design of the system, section Y its implementation and Z its testing results.
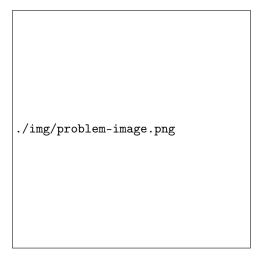

Figure 1.1: Artist compensation inconsistency

2

Problem description

The most widely used music streaming services, with the largest music catalogs, run centralized, proprietary and lcosed-source software. The companies owning these services have an ever increasing amount of power in the music streaming industry due to the amount of personal data they have about listeners and artists

<div style="background:orange">Add other reasons</div>

. Because of their power, they can ask high commission fees or lock artists to one platform. As a result, artists receive low compensation. The processing and storing of user data is nontransparent, as both the database and code are closed for people on the outside. As companies make money from data, their data-gathering methods are expected to become more disruptive for user privacy.

How can we design a music streaming service as alternative to Big Tech that distributes the power from one authority to its users?

2.0.1. Intermediaries take a large share
Artists publishing their content on Big Tech music platforms such as Google Music, Spotify and iTunes receive low compensation, because the intermediaries take a large share. Specifically, these companies take on average a 25 percent cut for signed records, and a 40 percent cut for unsigned records[1].

<div style="background:orange">expand</div>

2.0.2. User privacy
Big Tech companies obtain personal usage data to be able to improve their service, but also to to sell the data to third parties for a profit. In this process, users must heavily trust the company running the service to handle their data exactly as stated in their privacy policy. In the scope of music streaming services, user data such as browsing activity, and friends and sharing activity is obtained. For instance, Spotify saves personal usage data such as "search queries [...], streaming history, playlists you create, your library, your browsing history, and your interactions with the Spotify Service, content, other Spotify users." and shares this data with advertising parties, stated in their privacy policy[2]. Google Music "shares, processes, and maintains information about your usage, access, and playback of Your Music [...], playback activity related to items you preview and buy in the Google Music Store ("Store Usage"); and about the songs you share and listen to in connection with Google Music Social Recommendations [...]" as described in their privacy policy[3].

2.0.3. Data usage
Following the lack of privacy comes issues with what companies do with all the user data they gather. Widely known is the use of this data for targeted advertising[2] and for selling as a profit[5]. A risk in this process is that the third parties may use this private information for malicious purposes[5]. From the perspective of the user, there is no transparency for whether this happens. According to [3], privacy may be breached even when a service is willing to protect a user's privacy, because state-of-the-art de-anonymization methods

---

[1]`https://www.theguardian.com/technology/2015/apr/03/how-much-musicians-make-spotify-itunes-youtube`
[2]`https://www.spotify.com/us/legal/privacy-policy`
[3]`https://music.google.com/about/privacy.html?em_x=22`

do not fully make users anonymous, depending on the features stored in the database. Centralized software services are subject to a single point-of-failure. In this context this involves the risk of security breaches: if a malicious party gains access to its database, all of the records can be leaked at once which can lead to a large-scale privacy breach. Furthermore, [1] shows that such an event can leak personal data of users who are not part of the original database.

2.0.4. Control of data
The GDPR contains the right for individuals to have their data erased from any platform. However, a user taking this action cannot be certain of this happening on request, as access to the companies' database is not available from the outside. Furthermore, the company implements disclosure preferences in the way they see fit, which may not be fine-grained.

2.0.5. Content censoring
The company running the software is free in how and which content to censor. In addition, their content censoring policy may be changed at any time. Recent examples exist such as the disappearance of Li Zhi[4], who published songs about democracy and social issues in China. All of China's main streaming sites removed his songs. In 2019, Apple Music removed content from their platform by singer Jacky Cheung, who referenced the tragedies of Tiananmen Square in his songs.

2.0.6. Recommendation of content
The Big Tech music companies recommend content that best fits their business model, which may be contrary to what fits the user best. The companies can promote or dis-promote content by their choosing. For example, on Spotify, brands are able to sponsor playlists. "A car company might sponsor a popular driving playlist on Spotify"[4]. As the companies run closed-source software, the recommendation engine they use are a black box to the user. They are not obliged to explain the algorithms used for this. Small, independent artists may suffer from labels that invest large amounts of money to have their content promoted.

2.0.7. Security and fault tolerance
As the service and software are proprietary, and the running code is closed-source, there are security risks. Specifically, the cryptography and security mechanisms used internally can not be inspected by people outside of the company.

`expand`

2.0.8. Resiliency
At any point in time, the company running proprietary software can change, add or remove features. Its users do not necessarily have a vote in this. When a software service is sold to a different owner, the new owner can completely change direction for the service, which makes the service prone to large, possibly unwanted changes. Moreover the company can decide to take down the service in its entirety. For example: In 2017, Pandora discontinued running its service in New Zealand and Australia[5]. In this case, users can lose all their data stored on the service.

`Add sources of this happening`

2.0.9. Platform locking
As an example from YouTube, a company can disallow content creators to publish their work on other platforms, resulting creators to be locked to one platform.

`Find sources of this happening`

---

[4]https://www.independent.co.uk/news/world/asia/tiananmen-square-china-li-zhi-singer-disappears-anniversary-protests-a89
html

[5]https://www.businessinsider.nl/pandora-shutting-down-services-australia-new-zealand-2017-7?
international=true&r=US

3

Related work

3.0.1. Decentralized audio streaming services

Multiple decentralized audio streaming services exist. Examples are Audius[1], Resonate[?] and eMusic[2]. All of these systems have in common that they save metadata and identifiers of audio files on a blockchain, and save the audio files in an off-chain database. All these off-chain databases are structured like IPFS[3] with a company-run centralized interface between the user interface and the database. For a system to be fully decentralized, this layer should be removed. These solutions are closed source. Moreover, they use their own cryptocurrency to pay their artists which is an unstable income.

3.0.2. Decentralized content delivery networks

Decentralized content delivery networks are being investigated by multiple systems such as VideoCoin[4] and DCDN[5]. Most of these start-ups use blockchain technology and their own-released cryptocurrency as a token to pay nodes that serve the content. This means that the incentive for running a node depends on the value of those cryptocurrencies, so this is an unstable situation for workers.

A fully decentralized audio streaming service requires sharing and streaming audio files over a network of nodes in which any participant can start and run a node. An example of such network is BitTorrent. The challenge with BitTorrent acting as a streaming service is that the requirement from the user perspective is to have low latency for streaming and buffering media files. For each file, the peer discovery algorithm is run, which is a slow-start algorithm. It also relies on having enough seeders per file available.

Torrent files contain a list of chunks, which represent the different parts of the related file. These chunks are called torrent pieces. Flawless streaming of media files over BitTorrent requires a smart algorithm to predict what file is requested next, and what torrent pieces should be loaded. BitTorrent relies on trackers to perform peer discovery. However, trackers are a central point of failure. To make the system more decentralized, a solution using independent trackers and a gossip protocol[?] can be used.

3.0.3. Incentives for file spreading

In a DAO, the party responsible for hosting and spreading of files is not well-defined. To tackle the tragedy of the commons, entities should be incentivised just enough for the system to be sustainable and usable, but no more. An example incentive system is bandwidth tokens[?].

---

[1] https://audius.co
[2] https://eMusic.com
[3] https://ipfs.io/
[4] www.videocoin.io
[5] https://www.dcdn.com/

Bibliography

[1] C Dwork. Differential privacy, vol. 4052 (jul 2006).

[2] Paul Jessup. Big data and targeted advertising, 2012.

[3] Arvind Narayanan and Vitaly Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125. IEEE, 2008.

[4] Robert Prey. Nothing personal: algorithmic individuation on music streaming platforms. *Media, Culture & Society*, 40(7):1086–1100, 2018.

[5] Jamie Yap. User profiling fears real but paranoia unnecessary, 2011.