

Giovanni Ferreira Martinelli

**Quantificando a disrupção em redes de similaridade musicais: Como os usuários navegam por estas redes**

Belo Horizonte, Minas Gerais

2021

Giovanni Ferreira Martinelli

# **Quantificando a disrupção em redes de similaridade musicais: Como os usuários navegam por estas redes**

Relatório final apresentado à disciplina de Monografia em Sistemas de Informação do curso de graduação em Sistemas de Informação da Universidade Federal de Minas Gerais

Universidade Federal de Minas Gerais  
Instituto de Ciências Exatas  
Departamento de Ciência da Computação

Orientador: Flavio Vinicius Diniz de Figueiredo

Belo Horizonte, Minas Gerais  
2021

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>4</b>
1.0.1	Objetivos	4
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>6</b>
<b>2.1</b>	<b>Estudo da influência e da inovação</b>	<b>6</b>
<b>2.2</b>	<b>Disrupção</b>	<b>6</b>
2.2.1	Disrupção Musical	6
<b>3</b>	<b>REFERENCIAL TEÓRICO</b>	<b>7</b>
<b>3.1</b>	<b>Representação da Similaridade</b>	<b>7</b>
3.1.1	MFCC	7
3.1.2	Atributo Convnet	7
3.1.3	Similaridade entre músicas através do RBF kernel	8
<b>3.2</b>	<b>Índice de Disrupção</b>	<b>8</b>
<b>3.3</b>	<b>Redes de Alta Ordem</b>	<b>9</b>
3.3.1	Análise da Influência de um nó Disruptivo na trajetória de um usuário	10
<b>4</b>	<b>O CONJUNTO DE DADOS</b>	<b>11</b>
<b>4.1</b>	<b>Music4All</b>	<b>11</b>
<b>5</b>	<b>METODOLOGIA</b>	<b>13</b>
<b>5.1</b>	<b>Tratamento dos dados</b>	<b>13</b>
<b>5.2</b>	<b>Extração os Atributos das Músicas</b>	<b>13</b>
5.2.1	Extração do MFCC	13
5.2.2	Extração do Atributo Convnet	13
<b>5.3</b>	<b>Calculando a Similaridade entre Músicas</b>	<b>14</b>
<b>5.4</b>	<b>Criação da Rede de Similaridades</b>	<b>14</b>
5.4.1	Cálculo da Disrupção	14
<b>5.5</b>	<b>Análise Musicológica da Disrupção</b>	<b>15</b>
<b>6</b>	<b>EXPERIMENTOS</b>	<b>16</b>
<b>6.1</b>	<b>Tratamento dos dados</b>	<b>16</b>
<b>6.2</b>	<b>Extração dos Atributos</b>	<b>16</b>
<b>6.3</b>	<b>Geração da Rede e Cálculo da Disrupção</b>	<b>17</b>
<b>7</b>	<b>RESULTADOS</b>	<b>19</b>
7.0.1	Disrupção no Conjunto de Dados	19

---

7.0.2	Álbuns mais disruptivos . . . . .	22
7.0.3	Gêneros Disruptivos e Consolidadores . . . . .	23
7.0.3.1	Experimento 1 . . . . .	23
7.0.3.2	Experimento 2 . . . . .	25
7.0.4	Disrupção ao longo das décadas . . . . .	28
7.0.5	Disrupção na trajetória dos usuários . . . . .	28
<b>8</b>	<b>CONCLUSÃO . . . . .</b>	<b>31</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>32</b>

# 1 Introdução

A musicologia é o estudo científico ou mesmo a ciência da música, em que se estuda as amplas perspectivas históricas, antropológicas e estético-poéticas, abrangendo tanto questões técnico-operativas como filosóficas da música. Um campo específico da musicologia aborda a influência musical e como ela afeta o desenvolvimento de gêneros musicais. Seu papel tem sido debatido por estudiosos e críticos, mas somente recentemente através de uma abordagem orientada por dados [1].

Ao analisar a influência musical através de redes complexas é possível tirar conclusões quantitativas de como as influências modelam a forma que artistas e grupos produzem música, provendo informações valiosas das tendências e inovações músico-culturais.

Recentemente surgiram trabalhos baseados em uma métrica de rede proposta por Funk & Owen-Smith [2] para medir as influências desestabilizadoras e consolidadoras, chamada *disrupção*. Originalmente foi utilizada para analisar as mudanças tecnológicas no domínio das patentes para demonstrar de forma geral que um objeto disruptivo é aquele que cria um novo fluxo de eventos, mudando a maneira tradicional de como um contexto geralmente funciona, mostrando assim quais eram as patentes inovadoras ao longo do tempo.

No contexto musical, dado o máximo que nosso conhecimento possibilita, a *disrupção* foi primeiramente utilizada por Figueiredo e Andrade (2019) [3] através de uma abordagem Bayesiana da métrica para identificar artistas musicais disruptivos, utilizando-se dos seus metadados textuais, obtidos do conjunto de dados do All Music Guide. Este estudo forneceu evidências de que a métrica da *disrupção* pode prover informações importantes no domínio musical. Mais recentemente em 2020 a *disrupção* foi utilizada no contexto musical para identificar músicas disruptivas em uma rede de similaridade de músicas [4], para o conjunto de dados *Forró em Vinil* [5]. Esta rede foi construída por meio de atributos extraídos do áudio das composições, provendo evidência do potencial musicológico da métrica em redes de similaridade baseadas em metadados sonoros.

Motivado pelas recentes demonstrações da utilidade da *disrupção* [2, 6, 3, 4] o objetivo deste trabalho consiste em explorar o potencial musicológico da métrica para se obter informações das tendências e inovações musicais. Analisando sobre a ótica da *disrupção* uma rede de influências representada pela similaridade entre músicas. Diferente do modelo de Falcão et al. [4], esta rede irá abranger vários gêneros musicais.

## 1.0.1 Objetivos

Com intuito de observar o efeito da *disrupção* nas tendências musicais, e como as músicas disruptivas influenciam na forma que os usuários navegam por esta rede, será explorado através da métrica:

- Quais são os artistas e álbuns mais disruptivos.
- Quais consequências nas tendências musicais ao longo do tempo podem ser percebidas.
- Como músicas disruptivas impactam na sequência de músicas ouvidas por um usuário.

## 2 Trabalhos Relacionados

### 2.1 Estudo da influência e da inovação

Analisar e modelar como se dá a influência musical é um tópico já explorado de diversas formas no contexto da recuperação de informação musical, apesar de não ser um tópico popular.

Vários autores abordaram a influência no domínio musical utilizando modelagens obtidas através dos metadados das letras [7]. Outros autores como Mauch et al. utilizam do timbre, tom e da informação harmônica para analisar a influência musical [8].

A abordagem de Shalit et al. [1] em específico propõe um modelo de tópico dinâmico para representar a influência da música ao longo do tempo usando metadados e áudio. Suas descobertas utilizaram as músicas das paradas da Billboard para apontar correlações entre a influência e inovação apenas durante alguns curtos períodos: no início dos anos 70 e em meados dos anos 90.

### 2.2 Disrupção

Este trabalho é inspirado pela métrica de rede proposta por Funk & Owen-Smith [2] para estudar as mudanças tecnológicas de um contexto. O índice de disrupção é descrito em detalhes na seção 3.2.

Outras abordagens expandem as informações obtidas no contexto original como o que foi feito por Wu et al. [6]. Que adicionou ao conjunto de dados original repositórios de software e artigos científicos e concluiu que inovações estão mais associadas a times pequenos ao passo que times maiores consolidavam mais conhecimento.

#### 2.2.1 Disrupção Musical

Alguns trabalhos validam o índice em contextos diferentes do original, em especial os trabalhos feitos por Figueiredo e Andrade (2019) e Falcão et al. [3, 4], exploram o potencial musicológico da métrica, em que foram criadas redes de similaridade musical utilizando metadados textuais e sonoros respectivamente.

## 3 Referencial Teórico

### 3.1 Representação da Similaridade

Para estimar a similaridade entre músicas é necessário uma forma de comparação entre elas, logo é necessário uma representação das músicas que possibilite a sua análise.

#### 3.1.1 MFCC

Uma das formas utilizadas para representar as músicas é o coeficiente cepstral de frequência mel (MFCC). Atributos baseados no MFCC capturam propriedades do timbre do sinal [9].

O MFCC foi originalmente desenvolvido para reconhecimento automatizado de voz, e sua representação de áudio costuma ser robusta, fazendo com que seja costumeiramente utilizada em várias tarefas de recuperação de informação musical [10], como classificação de gênero [11], recomendação de músicas [12] e de similaridade entre áudios [13].

Espera-se que este atributo capturará eventos relevantes do áudio contido no conjunto de dados. Já que diversos casos de inflexões musicais (em que aconteceu um evento disruptivo) aconteceram devido um artista ou grupo adicionando um novo instrumento ou técnica a um gênero ou música.

No trabalho de Falcão et al. em que é construída uma rede de similaridade musical a métrica foi utilizada para representar as músicas do conjunto de dados do Forró em Vinil [5], que contém apenas o gênero do Forró. A partir desta rede foi possível identificar inflexões musicais de quando novos instrumentos eram utilizados nas músicas. Em nosso trabalho visa-se explorar se a métrica consegue ser representativa para um conjunto de dados multigênero, contendo músicas de diferentes épocas.

#### 3.1.2 Atributo Convnet

No trabalho de Choi et al. [10] é proposto uma outra representação para uma música. Sendo este um vetor de atributos gerado através de uma rede neural convolucional (convnet).

Esta rede é previamente treinada para uma tarefa de etiquetamento de músicas, e em seguida, após o treino e atualização de seus pesos é utilizada como extrator de atributos para outras tarefas.

Para gerar a representação de interesse para este trabalho dá-se como entrada o MFCC de uma música e então concatena-se os mapas de atributos (feature maps) de várias

camadas desta rede convolucional treinada, então constrói-se a representação da música a ser utilizada.

Nos experimentos do trabalho de Choi et al. [10] esta representação se provou melhor que o MFCC em todos os seus experimentos.

Outra motivação para utilizar esta representação além da possível performance superior ao MFCC é de que o conjunto de dados utilizado para treinamento da tarefa original do trabalho de Choi et al. foi o do Million Song Dataset [14], que contém diversos gêneros musicais de diferentes épocas, uma característica similar ao conjunto de dados utilizado neste trabalho. Logo espera-se que os pesos treinados em um conjunto de dados parecido reflitam em uma representação mais robusta e fiel das músicas.

### 3.1.3 Similaridade entre músicas através do RBF kernel

O Kernel RBF modela a distância vetorial, em que quanto maior a distância menor o seu valor. Sua definição matemática [15] atribui a si mesma uma interpretação de similaridade [16] (os valores que variam de 0 a 1, inversamente proporcionais à distância vetorial). Dado seu potencial em modelar a similaridade, além das evidências em favor desta abordagem utilizada por outros autores [4], opta-se por usar o Kernel RBF como medida de similaridade para pares de vetores de atributos.

## 3.2 Índice de Disrupção

A métrica, um conceito chave para este trabalho, foi proposta por Funk & Owen-Smith [2].

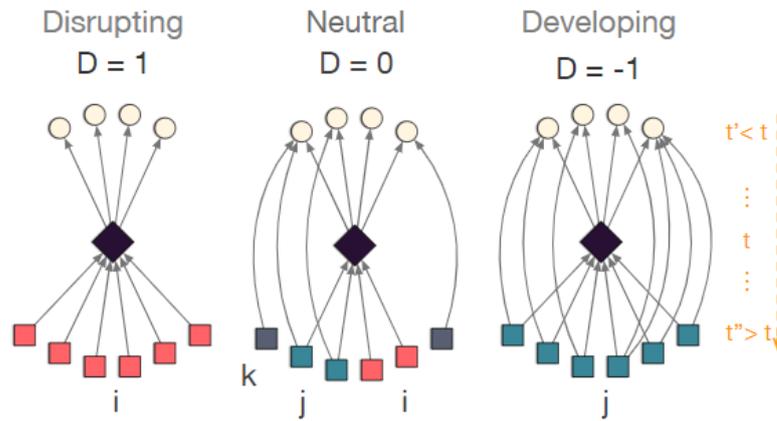


Figura 1 – Exemplo da disrupção em rede de Figueiredo e Andrade (2019) [3]. Os losangos são os trabalhos centrais, suas referências anteriores são os círculos e os trabalhos posteriores são os quadrados. Os nós  $i$  (em vermelho) referenciam o trabalho central, mas nenhum de seus predecessores; nós  $j$  (em azul) referenciam ambos o nó focal e os seus predecessores, e os nós (em cinza) vinculam apenas aos predecessores do nó focal

A figura mostra uma rede de influência e exemplifica a disrupção do ponto de vista do objeto central, em forma de diamante. As setas representam a influência ou as citações. Nesta rede, um objeto inovador influente será um nó com um alto grau de entrada. Em contraste, um nó disruptivo ( $D = 1$ ) é isolado e por isso autossuficiente em comparação com os seus antecessores. [3] Para formalizar a métrica para este trabalho, chama-se o nó focal em forma de diamante de,  $m$  (para música). Existem  $n_j$  nós que se referem à música  $m$  e tem pelo menos uma referência a um de seus antecessores (músicas pelos quais foi influenciado). Os nós  $n_i$  referenciam a música  $a$ , mas nenhuma de suas antecessoras. Há também  $n_k$  nós que não fazem referência a  $m$ , mas referenciam pelo menos uma de suas antecessoras. O índice de interrupção de Funk e Owen-Smith [2], aqui chamado de  $D$ , é medido como:

$$D = \frac{n_i - n_j}{n_i + n_j + n_k}$$

Em que  $D$  varia de -1 a 1. O extremo negativo captura a ideia de um trabalho em desenvolvimento, um que é muito citado em conjunto com suas influências (i.e.,  $n_i = 0$  e  $n_k = 0$ ). Já o extremo positivo captura a disrupção (i.e.,  $n_j = 0$  e  $n_k = 0$ ).

### 3.3 Redes de Alta Ordem

Tendo posse da rede de influências e tendo calculado o índice de disrupção para cada música. Para que seja possível analisar as sequências de músicas ouvidas pelos usuários sobre a ótica da disrupção é necessário uma representação fidedigna das relações subjacentes de como ocorre esta sequência de músicas.

Como representar estes dados sequenciais é um passo importante já que as representações de rede convencionais que implicitamente assumem a propriedade Markov (dependência de primeira ordem) podem rapidamente se tornar limitantes [17].

Essa suposição sustenta que quando movimentos são simulados na rede, o próximo movimento depende apenas do nó atual, desconsiderando o fato de que o movimento pode depender de várias etapas anteriores.

No entanto, os dados derivados de muitos sistemas complexos podem mostrar dependências de até quinta ordem. Nesses casos, a suposição de simplificação excessiva da representação de rede de primeira ordem pode levar a resultados imprecisos.

Para resolver este problema, espera-se representar as sequências de música dos usuários utilizando uma rede de ordem superior (HON) [17] que pode descobrir e incorporar ordens de dependências de variáveis em uma representação de rede. A HON pode ser aplicada a uma ampla variedade de tarefas, como caminhada aleatória, agrupamento e classificação, e como demonstrado por Xu et al. [17] ao utilizá-la como base para os experimentos, a HON produz resultados mais precisos.

### 3.3.1 Análise da Influência de um nó Disruptivo na trajetória de um usuário

Para analisar a influência de uma música disruptiva na trajetória de um usuário busque-se modelar a sequência de músicas que um usuário ouve como uma sequência de músicas ouvidas em uma sessão. É analisado a partir do conjunto de dados quanto tempo os usuários permanecem ouvindo músicas, obtendo-se a noção de quanto tempo dura a sessão de um usuário.

A partir desta modelagem pode-se gerar uma rede através dos caminhos percorridos pelas sequências de músicas ouvidas durante as sessões dos usuários. Permitindo assim calcular métricas de centralidade de rede, que medem a influência de um nó nesta rede, como PageRank [18] e Intermediação (Betweenness) para as músicas (nós) que possuem a disruptão calculada, visando entender qual a influência de uma música disruptiva na trajetória do usuário.

## 4 O conjunto de Dados

### 4.1 Music4All

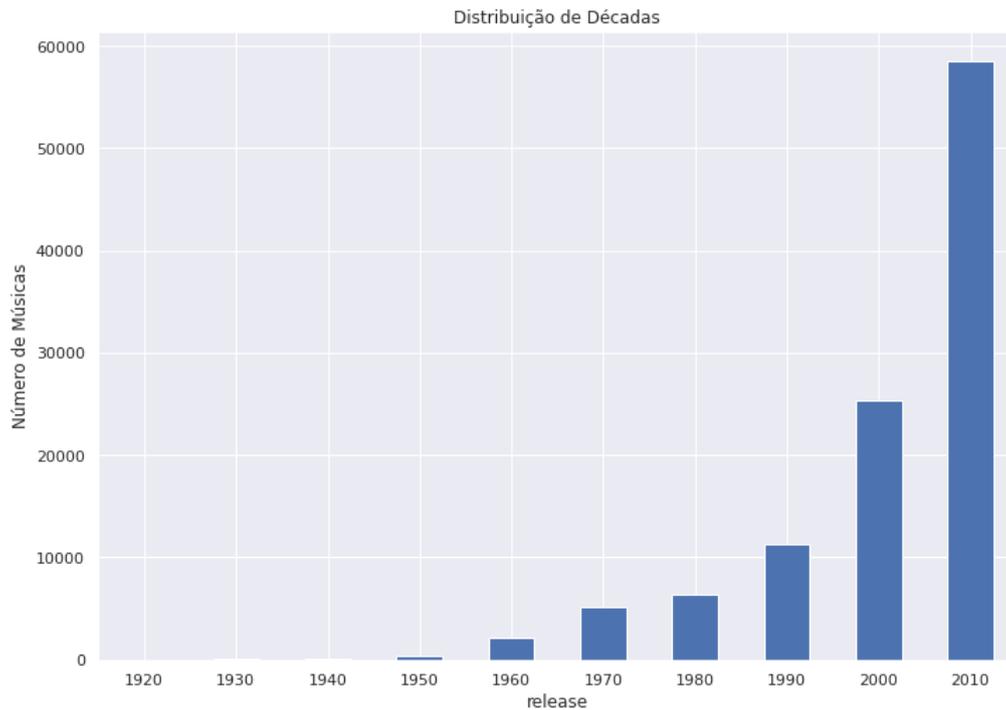


Figura 2 – Histograma com o número de músicas ao longo das décadas.

A campo de recuperação de informação musical (MIR) possui um grande volume pesquisas, apesar disso faltam bancos de dados musicais que suportem esses trabalhos, ou seja, bancos de dados que atendam a alguns requisitos bastante desejáveis para o desenvolvimento de pesquisas, tais como: grande quantidade de músicas, disponibilidade de amostras dos áudios, uma grande diversidade nas amostras e o histórico dos usuários que ouvem essas músicas.

O conjunto de dados do Music4All contém 16.269 artistas, 38.363 álbuns, 109.269 músicas. Para cada uma das músicas têm-se os metadados, tags, informações dos gêneros das músicas, clipes de áudio de 30 segundos, letras, somando ao todo 16 metadados extraídos através da API do Spotify. Além disso, é disponibilizado o histórico de músicas ouvidas por 15.602 usuários anônimos.

Analisando o conjunto de dados foi possível perceber que apesar da grande quantidade de gêneros e músicas disponíveis, os gêneros mais frequentes são relacionados às músicas

ocidentais. Os gêneros mais populares são Rock, Pop, Eletrônica, Rock Alternativo e Indie Rock respectivamente, compondo 29% de todos os gêneros disponíveis. Vale ressaltar que uma mesma música pode ter vários gêneros associados a ela neste conjunto de dados.

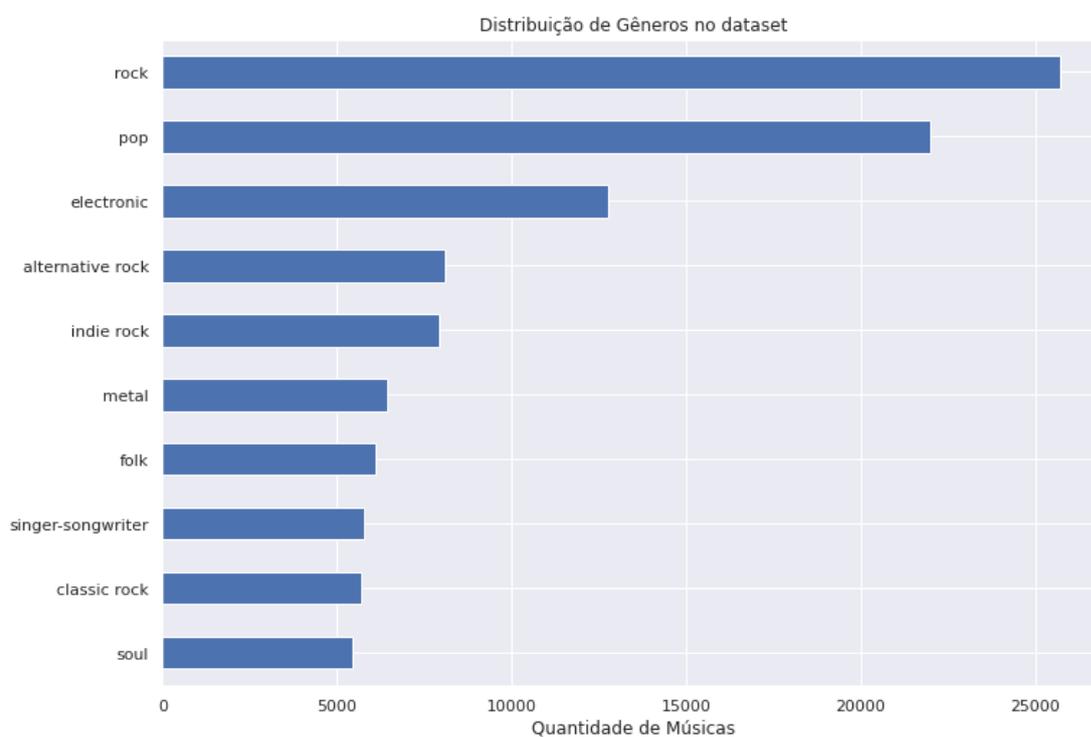


Figura 3 – Distribuição de gêneros do conjunto de dados

# 5 Metodologia

## 5.1 Tratamento dos dados

O primeiro passo para este trabalho resume-se em corrigir as entradas errôneas do conjunto de dados, já que músicas com ano de lançamento errado podem ser consideradas disruptivas quando este não é o caso.

Além disso busca-se remover músicas cuja amostra de áudio seja inválida, como no caso do áudio estar vazio, já que ao calcular a distância entre dois vetores de atributos que representam músicas vazias a similaridade entre eles vai ser muito alta e influenciará nos resultados.

## 5.2 Extração os Atributos das Músicas

Devido o tamanho massivo do conjunto de dados e com intuito de realizar os experimentos dentro do prazo disponível, foi feita a escolha de trabalhar com uma parte deste conjunto, já que realizar os cálculos das similaridades e da construção da rede tornou-se um processo computacionalmente restritivo.

### 5.2.1 Extração do MFCC

Para poder fazer as análises a seguir, é necessário extrair o MFCC das músicas para poder calcular qual é a similaridade entre elas. A extração e processamento é feita utilizando o Librosa [19].

Em trabalhos anteriores na área, o MFCC foi utilizado com atributo linha de base para diversas tarefas, sendo este extraído utilizando as médias e desvios-padrão de 20 coeficientes ceptrais de frequência mel (MFCCs) e suas derivadas de primeira e segunda ordem [10, 4], essa é a mesma abordagem utilizada para criar o atributo do MFCC de cada amostra de 30 segundos do conjunto de dados.

### 5.2.2 Extração do Atributo Convnet

A extração do atributo Convnet se dá de forma facilitada, utilizando o código proveniente do repositório original do artigo do Choi et al. [20]. Sendo assim possível extrair representações das músicas como vetores de atributos.

Atributos Extraídos		
Quantidade de Músicas Utilizada	30000	20000
Nós restantes na rede	26638	16876
Perda	11.3%	15.2%

Tabela 1 – Perda de informação ao se criar a rede.

### 5.3 Calculando a Similaridade entre Músicas

Dado o potencial do RBF Kernel para modelar a similaridade entre vetores de atributos o mesmo é utilizado para gerar uma matriz  $N \times N$  em que  $N$  é o número de músicas que se deseja calcular a similaridade, representando qual é a distância (ou similaridade) de cada música para cada outra música. Permitindo encontrar qual é a similaridade de qualquer música  $i$  para outra música  $j$  acessando a matriz na posição  $i, j$ .

### 5.4 Criação da Rede de Similaridades

Dado que possuímos a similaridade entre músicas, podemos representar as músicas como nós em uma rede e a relação de influência delas em arestas, criando-se assim uma rede de influências, que no futuro será utilizada para calcular a disrupção das músicas.

Esta é uma rede direcionada conectando as músicas similares, no entanto a escolha de criar uma aresta entre duas músicas depende do limite de similaridade escolhido para gerar esta relação.

Não há um consenso de como gerar esta relação arbitrária, uma formulação possivelmente válida é de que as músicas de um mesmo artista, ou de um mesmo álbum são uma boa base para determinar similaridade, já que essas músicas costumemente vão possuir instrumentos, vozes e produção similar. Neste trabalho a decisão de criar uma aresta será baseada na similaridade média das músicas contidas em um álbum, ou seja, caso a similaridade entre duas músicas supere esse limiar, então uma aresta irá conectar essas duas músicas.

Um problema inerente na criação desta rede é a perda de informação, já que podem existir músicas que não são similares a nenhuma outra, neste caso o nó desta música não é adicionado à rede. Uma sumarização desta perda pode ser vista na Tabela 1.

#### 5.4.1 Cálculo da Disrupção

De posse da rede de influências e da data de lançamento das músicas sabe-se qual música influenciou outra e qual a ordem de precedência entre elas, por isso itera-se sobre a rede e calcula-se o índice de disrupção para cada nó.

## 5.5 Análise Musicológica da Disrupção

Com o índice de disrupção é possível entender como é a distribuição da métrica no conjunto de dados e também quais foram os artistas, álbuns mais disruptivos, analisando possíveis tendências e informações contextuais que podem ser extraídas através dos resultados da métrica.

É possível modelar o histórico de músicas que os usuários ouviram em sequências de músicas em uma sessão. Neste trabalho é considerada a noção de sessão como uma sequência de músicas que um usuário ouviu dado um intervalo de tempo.

Através das sequências de músicas que os usuários ouvem, ou de forma análoga, os caminhos que estes usuários percorrem, é possível transformar esta representação em uma representação de rede, em que nós influentes influenciam na trajetória de músicas que um usuário ouve. Para analisar a influência dos nós nesta rede serão calculadas as métricas de PageRank e Intermediação (Betweenness) e será feita a correlação destas com a disrupção das músicas, no intuito de revelar ou não uma correlação entre uma música disruptiva e sua influência nas trajetórias dos usuários.

Será testado adicionalmente se esta rede do histórico dos usuários suporta uma representação de alta ordem e se esta pode representar melhor as trajetórias dos usuários.

O código para este trabalho se encontra em: Quantificando a Disrupção em Redes de Similaridade Musicais

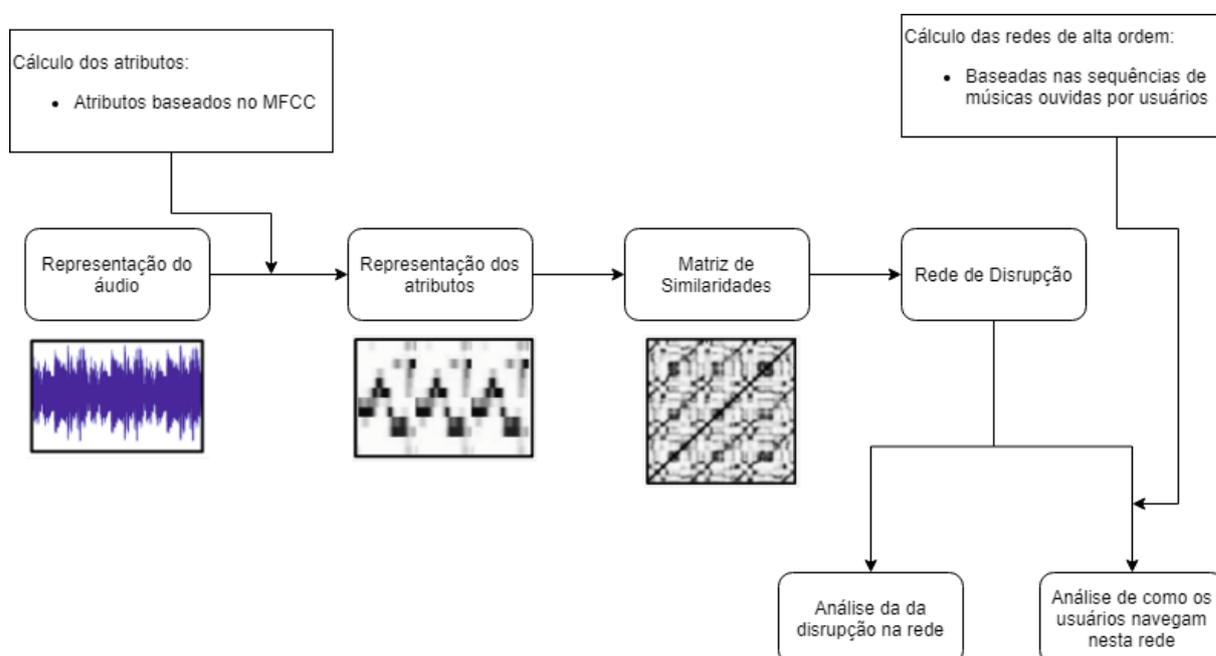


Figura 4 – Fluxo de Trabalho esperado

## 6 Experimentos

### 6.1 Tratamento dos dados

Foi corrigido manualmente as músicas cujo ano de lançamento estava incorreto, como músicas com datas anteriores a 1900. Assim como foi utilizado o software SoX [21] para identificação e remoção do conjunto de dados as músicas com áudios vazios.

Foram identificadas 91 amostras de música com áudios vazios, as mesmas foram removidas das etapas posteriores de processamento e análise dos dados.

### 6.2 Extração dos Atributos

Ao se extrair o MFCC das músicas, foi obtido 109.269 vetores (que serão referenciados como vetores de atributos), cada um contendo 120 elementos que representam a informação contida no áudio.

No caso do atributo da convnet, o mesmo processo foi realizado para as 109.269 músicas, gerando vetores de atributos, cada um contendo 160 elementos representando a informação contida no áudio.

Atributos Extraídos		
Atributo	Quantidade	Número de Atributos
MFCC	109.269	120
Convnet Feature	109.269	160

Devido o número massivo de dados os mesmo não cabiam em memória, por isso foram utilizadas matrizes esparsas e descarregamento para o disco que as computações fossem possíveis. No entanto mesmo tendo as limitações de memória resolvidas, o custo computacional  $O(2)$  para computar as similaridades foi proibitivo, por isso resolveu-se amostrar apenas 30.000 músicas do conjunto de dados para realizar os experimentos e análises.

O conjunto de dados possui muito mais músicas nas décadas de 2000 e 2010 como pode ser visto na Figura 2. Por isso dois subconjuntos foram feitos para fazer os experimentos, um visando representar o conjunto de dados total, já o outro visando ter uma melhor representação da história da música no século 20.

**Experimento 1:** Amostragem aleatória de 30.000 elementos do conjunto total

Esse experimento resulta em uma melhor representação dos gêneros musicais do conjunto todo, há a presença de 625 gêneros, ao passo que o conjunto total possui 853 gêneros. No entanto artistas sofrem de uma representação limitada, já que em média há

somente 3.22 de músicas por artista neste subconjunto. Quando é feita a criação da rede de disrupção há uma perda maior de nós se comparado ao experimento 2.

**Experimento 2:** Amostragem dos primeiros 30.000 elementos ordenados por data de lançamento da música.

Tendo em vista as limitações do experimento 1 foi proposto obter as 30.000 músicas ordenadas por ano de lançamento, de forma a conter todas as músicas durante um período, essa foi a representação que apresentou resultados mais detalhados, já que em média foi possível obter uma quantidade de 6.5 músicas por artista, um número próximo à quantidade de músicas em álbum musical [22]. Porém este conjunto contém apenas 460 dos 853 gêneros do conjunto total.

Desta forma é possível abordar a disrupção tendo em vista uma gama maior de gêneros presentes no conjunto de dados (experimento 1) e melhor analisar artistas e álbuns disruptivos do século 20 a partir do experimento 2.

Tipos de experimentos		
	Anos das Músicas	
	Início	Fim
<b>Experimento 1</b>	1928	2019
<b>Experimento 2</b>	1923	2002

Assim foi feito o cálculo da matriz de similaridades através do RBF Kernel utilizando ambos os atributos descritos anteriormente.

No entanto o MFCC se provou incapaz de gerar uma boa representação para o dado, mesmo com valores altos de gamma (permitindo que distâncias maiores fossem consideradas similares), as matrizes geradas com o MFCC foram efetivamente esparsas, já que poucos elementos eram realmente similares. Ao se tentar gerar as Redes de Disrupção com essas matrizes essas redes também eram esparsas e compostas de poucas músicas e gêneros.

Experimento	Músicas de Entrada	Feature	Gamma	Nós	Arestas
-	30.000	MFCC	0.0005	9234	106636
-	30.000	MFCC	0.1	1046	754
<b>1</b>	30.000	Convnet	0.1	26119	4462095
<b>2</b>	30.000	Convnet	0.1	26638	4710644

Levando em conta superioridade do atributo da convnet em relação ao MFCC todos os experimentos seguintes foram feitos utilizando somente ele.

### 6.3 Geração da Rede e Cálculo da Disrupção

Uma diferença marcante entre as redes dos dois experimentos é que a rede gerada no experimento 2 apresenta um valor significativamente maior de grau ponderado médio,

significando a maior diversidade em suas relações de influência se comparada a rede do experimento 1.

	<b>Experimento 1</b>	<b>Experimento 2</b>
<b>Grau Ponderado Médio</b>	40,909	176,839
<b>Grau Médio</b>	133,388	110,304

Tabela 2 – Informações sobre as redes de disrupção para ambos os experimentos.

## 7 Resultados

### 7.0.1 Disrupção no Conjunto de Dados

Calculada a métrica da disrupção para os experimentos 1 e 2 foi possível observar qual a sua distribuição nos subconjuntos de dados.

	Porcentagem dos Dados	
	Experimento 1	Experimento 2
<b>Disrupção acima de 0.8</b>	5.7%	4.9%
<b>Disrupção abaixo de 0.1</b>	93.1%	93.9%

Tabela 3 – Disrupção para os experimentos 1 e 2.

Pode-se perceber que para ambos os experimentos a disrupção é algo incomum nos subconjuntos, em que por volta de 5% das músicas dos conjuntos são efetivamente disruptivas.

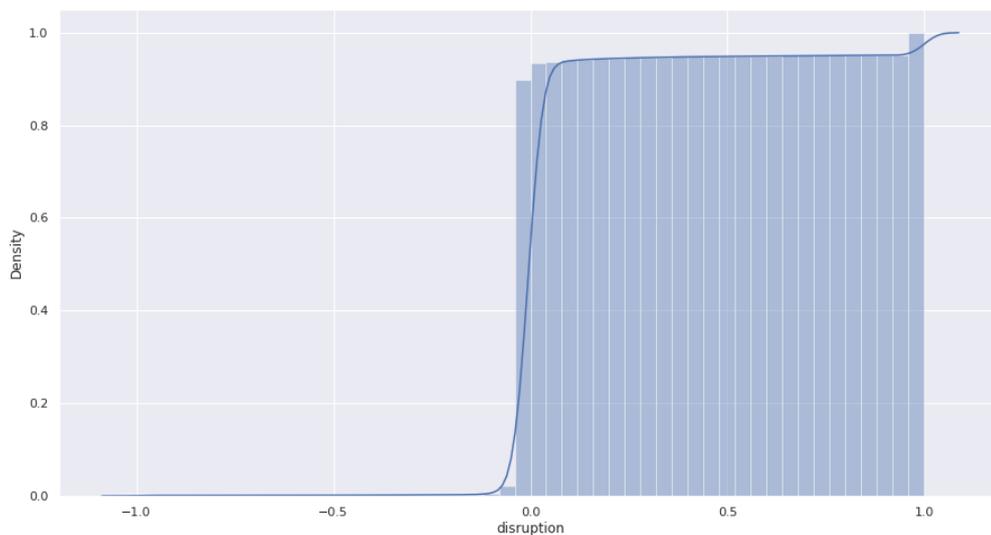


Figura 5 – Disrupção nos subconjuntos de dados

Utilizando dos resultados do experimento 2 foi possível calcular quais foram os artistas e álbuns mais disruptivos no conjunto de dados no intervalo de 1923 à 2002.

Uma das formas para analisar os artistas mais disruptivos foi através da **disrupção média de um artista**, calculada somando a disrupção de todas as suas músicas e dividindo pelo número de músicas que o mesmo possui. Essa métrica favorece artistas

que foram disruptivos ou inovadores em seu contexto. A partir dela foi possível capturar músicos e bandas que alteraram o curso de seus e outros gêneros, não tendo correlação direta com sua popularidade.

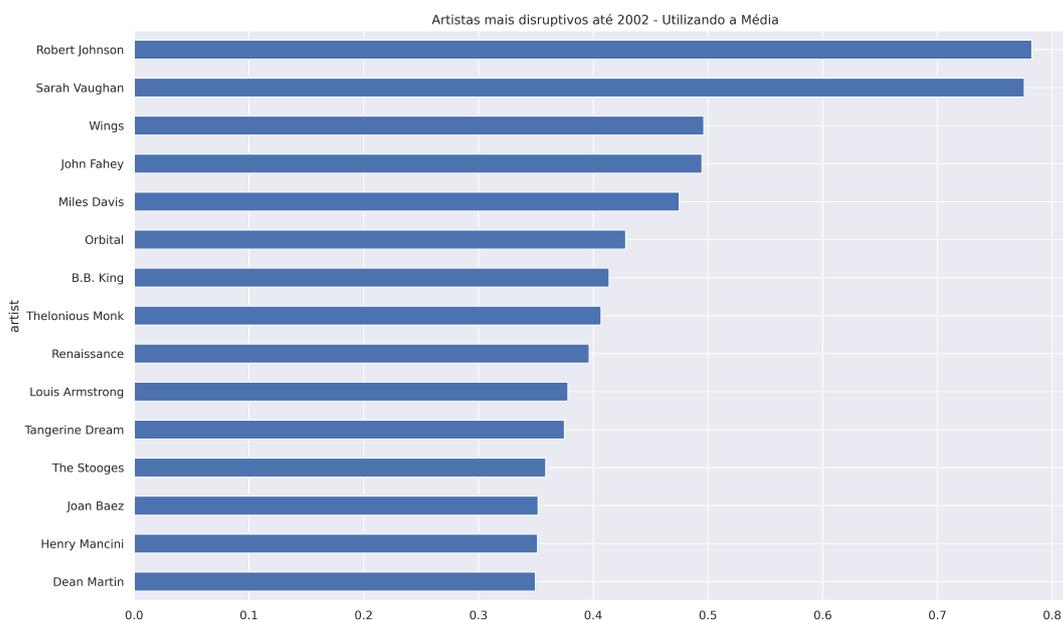


Figura 6 – Artistas mais disruptivos até 2002 - Utilizando a Media

Por exemplo Robert Johnson é frequentemente citado como "o maior cantor de blues de todos os tempos" e até mesmo o músico mais importante do século 20 [23]. Ele definiu o som do Delta Blues e consolidou o blues de 12 compassos. Porém em vida o mesmo não obteve sucesso comercial, somente após sua morte suas músicas foram descobertas e serviram de grande inspiração para artistas do rock como Led Zeppelin, os Rolling Stones e do folk como Bob Dylan [24].

Conhecido somente pelos aficionados do rock e do blues e esquecido durante muitos anos, o fato de Johnson aparecer como o artista mais disruptivo demonstra o poder da disrupção em descrever a trajetória entre gêneros musicais somente através do áudio desses artistas.

Há ainda na lista músicos como Miles Davis, considerado um dos mais influentes músicos do século XX, já que esteve na vanguarda de quase todos os desenvolvimentos do jazz desde a Segunda Guerra Mundial até a década de 90. Outra musicista que aparece na lista é a Sarah Vaughan, uma das primeiras vocalistas a incorporar o fraseio do bebop, uma corrente mais influente do Jazz. Já Henry Mancini é eternamente lembrado como sendo um dos compositores mais conhecidos de trilhas sonoras para a televisão e o cinema,

devido as suas famosas composições para filmes como música tema da pantera cor de rosa e Moon River.

Através da disrupção média é possível descobrir artistas que não só foram precursores em seus próprios gêneros, como tiveram influência intergênero musical.

A segunda forma de analisar a disrupção foi a partir do **somatório da disrupção** das músicas de um artista, porém desta vez sem dividi-lo pela quantidade de músicas do mesmo. Essa métrica favorece artistas que tiveram várias músicas disruptivas dentre as músicas que ele lançou, porém ela é enviesada para artistas famosos que possuem músicas disruptivas, já que os mesmos tem uma maior representatividade no conjunto de dados por terem mais músicas ou álbuns.

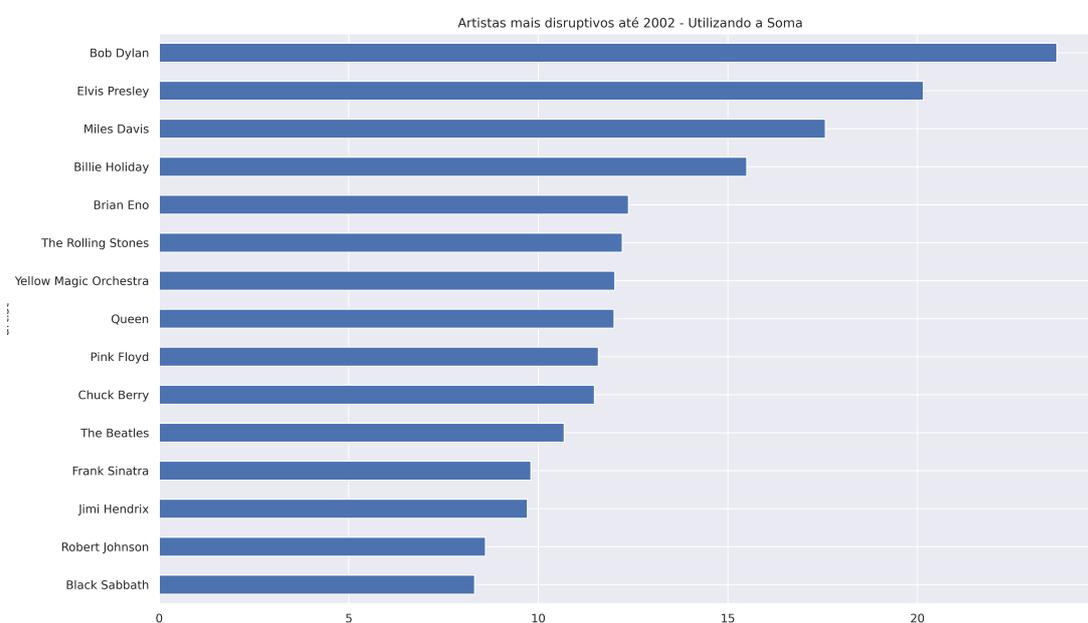


Figura 7 – Artistas mais disruptivos até 2002 - Utilizando a Soma

É possível ver artistas que foram bastante influentes e disruptivos, como Bob Dylan que influenciou muitos artistas e bandas como os Beatles e os Rolling Stones.

Já a banda Black Sabbath é usualmente citada como um dos pioneiros do Heavy Metal. [25]. O grupo ajudou a definir o gênero com lançamentos como Paranoid (1970), Master of Reality (1971) e Heaven and Hell (1980), tendo vendido mais de 70 milhões de cópias durante sua carreira. Eles foram definidos pela MTV como "a maior banda de Heavy Metal de todos os tempos".

O modelo também é capaz de capturar bandas que criaram seus gêneros musicais a partir da mixagem de gêneros existentes e a partir deste se distinguem fundamentalmente

dos gêneros dos quais se inspiraram. A banda Yellow Magic Orchestra (YMO), uma banda japonesa de música eletrônica, surgiu como um projeto de estúdio para criar um álbum que fundisse o exotismo orientalista com a música eletrônica moderna. Os mesmos já alcançaram o segundo lugar na lista dos 100 melhores músicos do Japão da HMV, sendo considerados inovadores no campo da música eletrônica popular.

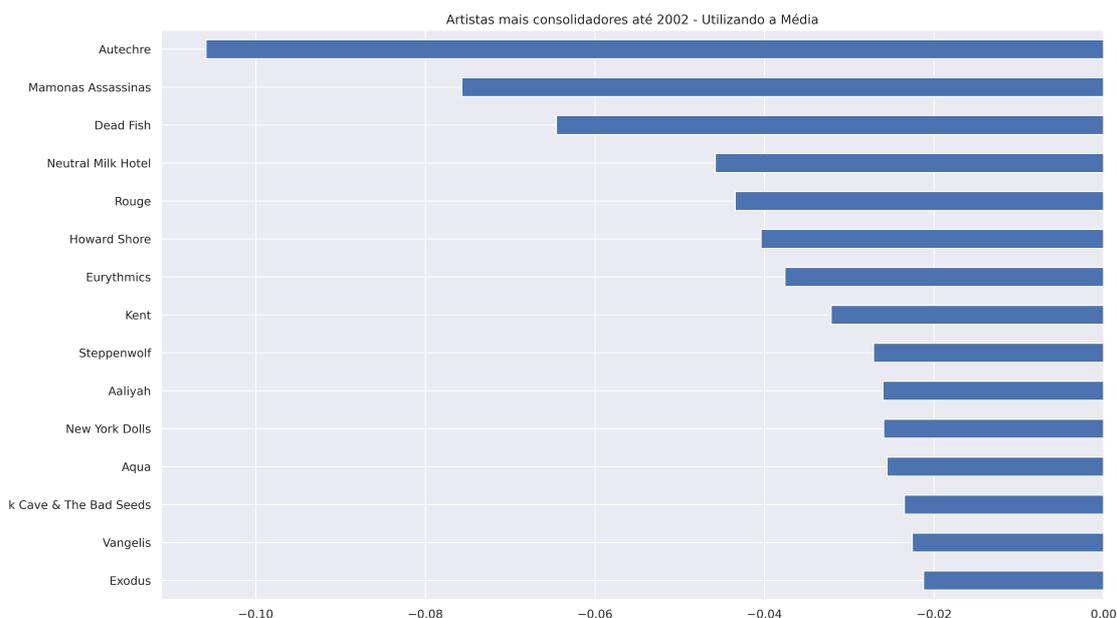


Figura 8 – Artistas mais consolidadores até 2002 - Utilizando a Média

Através dos artistas mais consolidadores em média é possível observar um ponto fraco do modelo. A banda Mamonas Assassinas foi extremamente influente nos anos 90 e quebrou os padrões da época através do seu rock cômico [26]. O foco no humor das letras os diferenciava, no entanto sua sonoridade musical se dava muitas vezes da apropriação de gêneros existentes em suas músicas como pop rock, sertanejo, brega, entre outros [27]. Logo a banda aparenta estar consolidando outros gêneros, quando a mesma está utilizando desses para criar o seu próprio.

### 7.0.2 Álbuns mais disruptivos

Um álbum é considerado um aglomerado de 6 ou mais músicas com mais de 30 minutos de duração [22], buscando entender a relevância semântica que álbuns tem para o contexto de um gênero ou época, na tabela abaixo são mostrados os álbuns mais disruptivos em média até 2002.

Os resultados se referem apenas aos dados do experimento 2, já que o experimento 1 possui poucas músicas por artista/álbum o que diminui a significância dos seus resultados.

Álbum	Artista	Ano de Lançamento
Elvis Presley	Elvis	1956
Robert Johnson	King Of The Delta Blues	1936
Joan Baez	Joan Baez	1960
B.B. King	Live At The Regal	1965
Miles Davis	Birth Of The Cool	1957
Pink Floyd	The Piper at the Gates of Dawn	1967
The Rolling Stones	12 X 5	1964

Em primeiro lugar aparece o álbum do Elvis Presley, sendo este o segundo álbum de estúdio de músico. O álbum passou quatro semanas no topo da parada Billboard Top Pop Albums em 1956, o segundo álbum de rock and roll a liderar a parada [28].

Já King of the Delta Blues é um compilado de músicas de Robert Johnson, é considerado um dos maiores e mais influentes lançamentos do blues.

Live At The Regal é um álbum considerado como um dos mais influentes álbuns de blues já gravados e foi ranqueado como álbum número 141 pela Revista Rolling Stones em sua lista dos melhores 500 álbuns da historia [29].

The Piper at the Gates of Dawn é o álbum de estreia da banda britânica de rock Pink Floyd, lançado em agosto de 1967, sendo teste considerado um dos pioneiros do art rock.

Pode-se perceber que a disrupção média de um álbum refere-se diretamente a álbuns muito influentes, como o do Elvis, ou precursores em seus respectivos estilos, como o primeiro álbum do Pink Floyd.

### 7.0.3 Gêneros Disruptivos e Consolidadores

Outra informação possível de ser observada é se gêneros musicais tem maior correlação com a disrupção, a correlação é feita através do coeficiente de correlação tau de Kendall.

#### 7.0.3.1 Experimento 1

Os dados do experimento 1 utilizam as informações do gêneros até 2019. Todas as correlações apresentadas na Figura 9 demonstram relações muito fracas (abaixo de 0.1) entre um gênero e a disrupção. Sendo a maior correlação observada de  $\tau = 0.07$  para o gênero de Rock Indie.

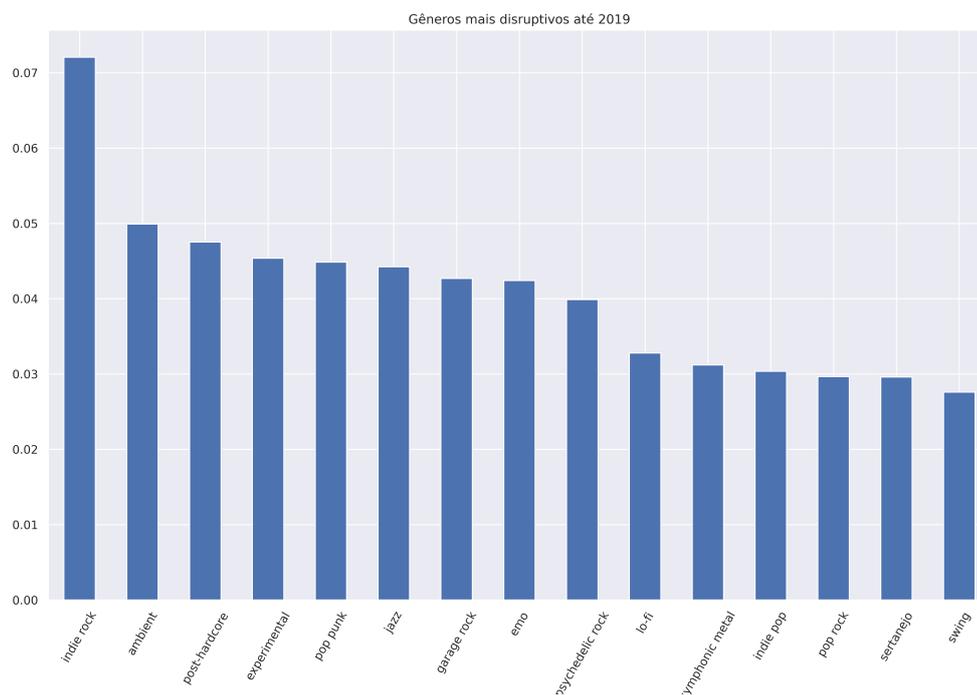


Figura 9 – Gêneros mais Disruptivos - Até 2019 - Experimento 1

O mesmo comportamento de correlações muito fracas aparece para os gêneros mais consolidadores, sendo a menor correlação de  $\tau = -0.08$  atribuída aos gêneros de Rock Clássico e Pop.

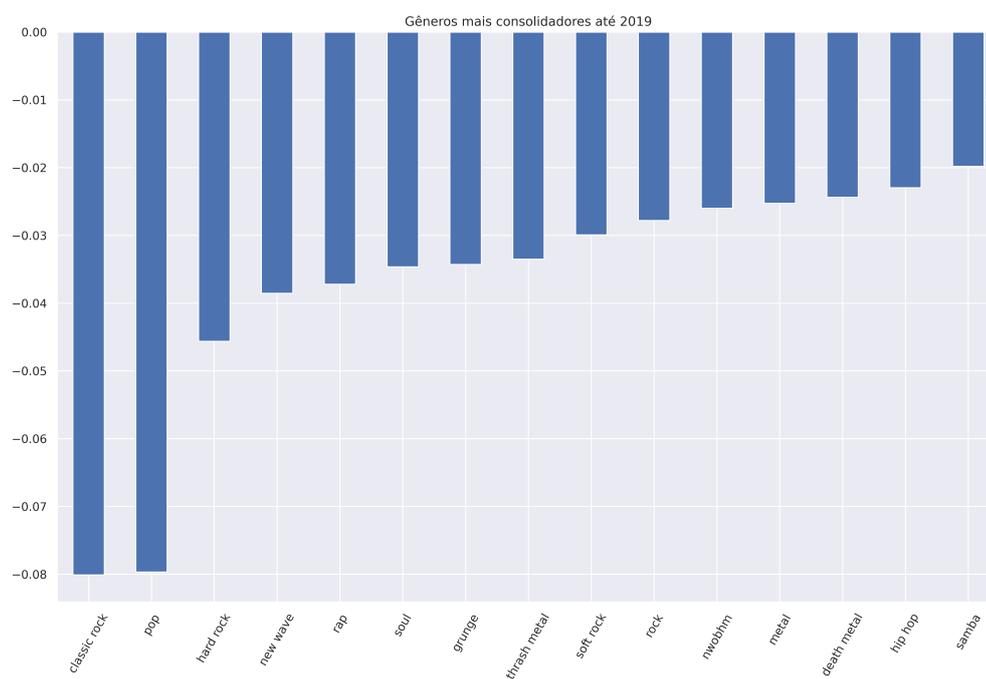


Figura 10 – Gêneros mais Consolidadores - Até 2019 - Experimento 1

### 7.0.3.2 Experimento 2

De forma similar aos resultados do experimento 1, nenhum gênero apresenta correlação sequer fraca com a métrica da disrupção, sendo a maior correlação de  $\tau = 0.06$  atribuída ao gênero de Rock Alternativo.

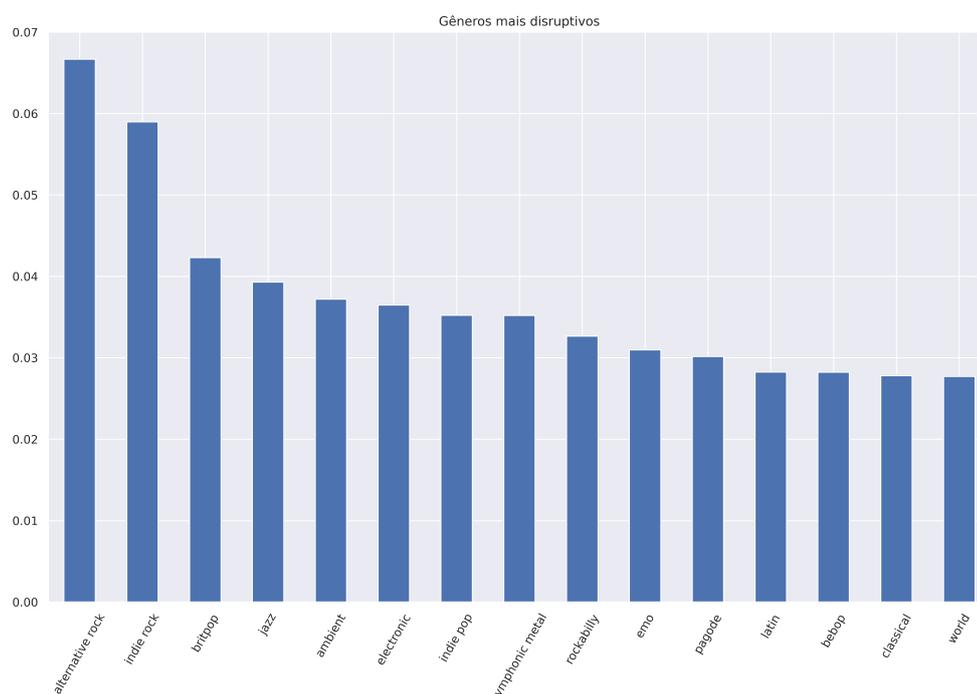


Figura 11 – Gêneros mais Disruptivos - Até 2002 - Experimento 2

A única correlação fraca entre gênero e disruptividade aparece nos resultados do gráfico da Figura 12, em que o Rock Clássico apresenta uma correlação negativa fraca de  $\tau = -0.12$  com a disruptividade, dando indícios que este é um gênero musical mais consolidador.

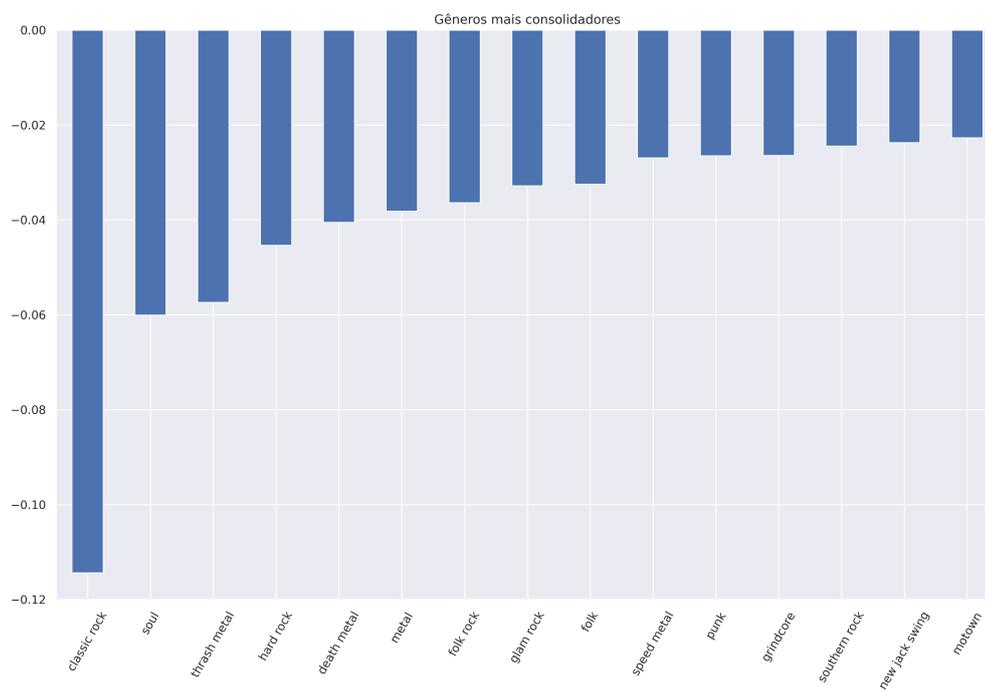


Figura 12 – Gêneros mais Consolidadores - Até 2002 - Experimento 2

### 7.0.4 Disrupção ao longo das décadas

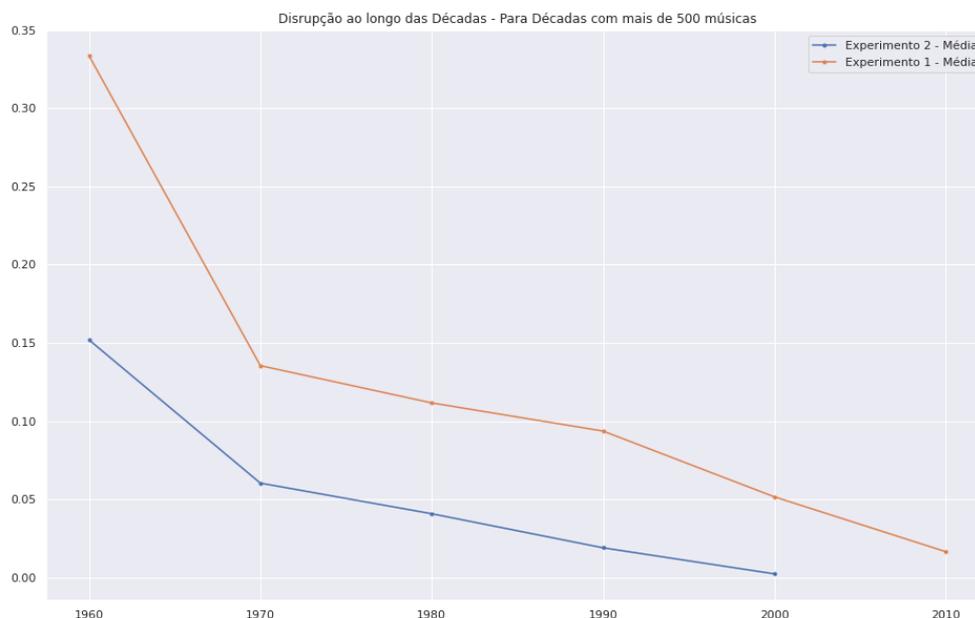


Figura 13 – Disrupção durante as décadas - Experimento 1 e 2

Para observar como foi a disrupção média das décadas, fez-se o somatório da disrupção para cada década e dividiu-se pela quantidade de músicas desta década, obtendo-se assim uma noção de como uma década foi em média disruptiva. Apenas décadas com mais de 500 músicas foram escolhidas. Os resultados da Figura 13 acima mostram disrupção média das décadas de ambos experimentos 1 e 2, de forma a comparar a visão geral que estes modelos podem dar.

Uma característica que pode-se perceber é que a disrupção ao longo das décadas vem diminuindo ao passar das mesmas. De forma que a década de 60 se encontra como a mais disruptiva para ambos os experimentos.

### 7.0.5 Disrupção na trajetória dos usuários

Para analisar o efeito da disrupção na trajetória dos usuários foi necessário transformar a trajetória dos usuários em uma rede. Foram selecionadas sessões dos usuários, em que nestas os mesmos ouviram uma sequência de músicas. Essa sequência foi transformada numa rede e a partir desta foi analisada a influência dos nós disruptivos através do cálculo de métricas de centralidade de rede.

A primeira tarefa foi de entender qual é o tempo de sessão de um usuário, para realizá-la foi calculada a diferença de tempo entre cada ação para cada usuário no conjunto de dados,

em que uma ação é quando o usuário ouve uma música. Analisando esta distribuição foi visto que 90% dos usuários mantém uma sessão de até no máximo 101 minutos. Sabendo quanto dura uma sessão, foi possível transformar o histórico dos usuários sequências de músicas que duraram até 101 minutos.

Somente foram adicionadas à rede sessões em que todas as suas músicas estavam contidas no subconjunto de dados utilizado, das quais já havia sido calculado a métrica da disrupção. Então foi feita construída uma rede contendo 13.241 nós (representando as músicas) e 30.277 arestas. Para o cálculo dos caminhos e construção da rede foi utilizada a biblioteca PathPy [30].

<b>Caminhos</b>	<b>Tamanho dos caminhos</b>
1	4174
2	2246
3	1372
4	923
5	657
6	511
7	396
8	303
9	235
10	179
11	116
12	115
13	66
14	54
15	50
16	54
17	26
18	24
19	21
20	21

Tabela 4 – Tamanho dos caminhos da rede de sequências de músicas.

Foi possível perceber que a maior parte das sessões consiste em sequências de 1 até 20 músicas, indicando que usuários ouvem pequenas sequências de músicas durante suas sessões.

Foi feita a modelagem do histórico dos usuários como uma rede alta ordem e feito o teste que dá um indicativo para saber se a mesma comporta essa representação de alta ordem. O teste é feito da seguinte forma:

1. Calcula-se a diferença  $d$  entre o grau de liberdade do modelo de segunda e de primeira ordem.

2. Calcula-se o teste de razão de verossimilhança

$$x = -2 \cdot (\log(\text{hon\_1.likelihood}) - \log(\text{hon\_2.likelihood}))$$

3. Calcula-se o p-valor como  $1 - \text{cdf}(x, d)$ , em que a *cdf* é a função de distribuição cumulativa de uma distribuição qui-quadrado.

O teste teve resultado 1, **indicando que a rede suporta uma representação de segunda ordem**, ou seja, indicando que ouvir uma música pode depender de 2 ou mais músicas anteriores. Porém devido aos problemas computacionais de calcular as métricas de centralidade de rede na rede de alta ordem, foi utilizada a representação de primeira ordem para os experimentos seguintes.

Foram calculadas as métricas de PageRank e Betweenness para todos os nós, tendo assim a noção de influência destes nós na trajetória que os usuários escutam. Para analisar a correlação dessas métricas de influência dos nós e a disrupção foi utilizada o Coeficiente de correlação tau de Kendall,  $\tau$ , já que são necessárias menos suposições sobre os dados, a métrica aborda problema de empates, e permite descobrir tanto relações lineares quanto não lineares [31].

Os resultados mostram correlações muito fracas negativas entre a disrupção e as métricas de centralidade de rede no histórico dos usuários.

Correlação	PageRank	Betweenness
Disrupção	-0.035728	-0.067913

## 8 Conclusão

Neste trabalho foi possível obter uma formulação para analisar a disrupção no contexto da música, expandindo os modelos anteriores de disrupção em redes de similaridade musical. Mostrando sua utilidade em extrair informações semânticas de redes multigênero geradas através dos metadados sonoros.

Dada uma representação suficientemente boa das músicas a disrupção pode ser utilizada para analisar artistas, músicas e álbuns disruptivos ao longo da história, mesmo em conjuntos de dados contendo vários gêneros e décadas. Valorizando e reforçando a contribuição dos músicos para o cenário como um todo.

Foi possível perceber correlações negativas muito fracas entre as músicas disruptivas e a sua influência na trajetória de músicas que um usuário ouve, podendo concluir que apesar destas músicas mudarem o contexto musical ao longo dos anos, elas tem pouca relação com a sequência de músicas ouvidas por um usuário.

# Referências

- [1] SHALIT, U.; WEINSHALL, D.; CHECHIK, G. Modeling musical influence with topic models. In: . [S.l.]: JMLR.org, 2013. (ICML'13).
- [2] FUNK, R. J.; OWEN-SMITH, J. A dynamic network measure of technological change. *Management Science*, mar. 2016. ISSN 0025-1909, 1526-5501.
- [3] FIGUEIREDO, F.; ANDRADE, N. Quantifying Disruptive Influence in the AllMusic Guide. ISMIR, Delft, The Netherlands, nov. 2019. Disponível em: <<https://doi.org/10.5281/zenodo.3527940>>.
- [4] FALCÃO, F. et al. Measuring disruption in song similarity networks. In: . [S.l.]: ISMIR, 2020.
- [5] FELIPE. *Forró em Vinil Dataset*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/4079503/export/json>>.
- [6] WU LINGFEI, W. D. E. J. A. Large teams develop and small teams disrupt science and technology. *Nature*, v. 566, n. 7744, p. 378–382, fev. 2019. ISSN 1476-4687. Disponível em: <<https://doi.org/10.1038/s41586-019-0941-9>>.
- [7] ATHERTON, J.; KANESHIRO, B. I Said it First: Topological Analysis of Lyrical Influence Networks. In: *Proceedings of the 17th International Society for Music Information Retrieval Conference*. New York City, United States: ISMIR, 2016. Disponível em: <<https://doi.org/10.5281/zenodo.1418047>>.
- [8] MAUCH, M. et al. The evolution of popular music: Usa 1960–2010. *Royal Society Open Science*, Royal Society, v. 2, n. 5.
- [9] MLLER, M. *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2015. 177 p. ISBN 3319219448.
- [10] CHOI, K. et al. Transfer learning for music classification and regression tasks. 2017.
- [11] KOUR, G. Music genre classification using mfcc, svm and bpnn. *International Journal of Computer Applications*, v. 112, n. 6, p. 3.
- [12] HAN, H. et al. Music recommendation based on feature similarity. In: *2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*. [S.l.: s.n.], 2018. p. 650–654.

- [13] JENSEN, J. H. et al. Evaluation of mfcc estimation techniques for music similarity. In: *2006 14th European Signal Processing Conference*. [S.l.: s.n.], 2006. p. 1–5. ISSN 2219-5491.
- [14] BERTIN-MAHIEUX, T. et al. The million song dataset. In: *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*. [S.l.: s.n.], 2011.
- [15] Kai-Min Chung et al. Radius margin bounds for support vector machines with the rbf kernel. In: *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02*. [S.l.: s.n.], 2002. v. 2, p. 893–897 vol.2.
- [16] VERT, J.; TSUDA, K.; SCHÖLKOPF, B. A primer on kernel methods. In: MAX-PLANCK-GESELLSCHAFT. *Kernel Methods in Computational Biology*. Cambridge, MA, USA: MIT Press, 2004. p. 35–70.
- [17] XU, J.; WICKRAMARATHNE, T. L.; CHAWLA, N. V. Representing higher-order dependencies in networks. *Science Advances*, American Association for the Advancement of Science, v. 2, n. 5, p. e1600028, May 2016. ISSN 2375-2548.
- [18] PAGE, L. et al. *The PageRank Citation Ranking: Bringing Order to the Web*. [S.l.], 1999. Previous number = SIDL-WP-1999-0120. Disponível em: <<http://ilpubs.stanford.edu:8090/422/>>.
- [19] MCFEE, B. et al. librosa: Audio and music signal analysis in python. In: . [s.n.], 2015. p. 18–24. Disponível em: <[https://conference.scipy.org/proceedings/scipy2015/brian\\_mcfee.html](https://conference.scipy.org/proceedings/scipy2015/brian_mcfee.html)>.
- [20] CHOI, K. *transfer\_earning\_music*. [s.n.], 2021. Disponível em : <[https://github.com/keunwoochoi/transfer\\_earning\\_music](https://github.com/keunwoochoi/transfer_earning_music)>.
- [21] Disponível em: <<http://sox.sourceforge.net/>>.
- [22] Disponível em: <<https://support.tunecore.com/hc/en-ca/articles/115006689928-What-is-the-difference-between-a-Single-an-EP-and-an-Album->>.
- [23] Disponível em: <<https://www.allmusic.com/artist/robert-johnson-mn0000832288/discography>>.
- [24] 2020. Disponível em: <<https://www.ploddings.com/blog/who-did-robert-johnson-influence/>>.
- [25] Disponível em: <<https://www.allmusic.com/artist/black-sabbath-mn0000771438/biography>>.

- [26] Disponível em: <<https://cultura.estadao.com.br/noticias/musica,existencia-dos-mamonas-assassinas-foi-essencial-nos-anos-90-,1811416>>.
- [27] Disponível em: <<https://super.abril.com.br/cultura/mamonas-cometa-loucura/>>.
- [28] Disponível em: <<https://www.billboard.com/articles/columns/rock/6835191/happy-birthday-elvis-first-time-listener-guide>>.
- [29] STONE, R.; STONE, R. *500 Greatest Albums of All Time*. 2009. Disponível em: <<https://www.rollingstone.com/music/music-lists/500-greatest-albums-of-all-time-156826/>>.
- [30] SCHOLTES, I. When is a network a network? multi-order graphical model selection in pathways and temporal networks. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 1037–1046, Aug 2017. ArXiv: 1702.05499.
- [31] KENDALL, M. G. A new measure of rank correlation. *Biometrika*, Biometrika Trust, v. 30, n. 1/2, p. 81–93, 1938. ISSN 00063444. Disponível em: <<http://www.jstor.org/stable/2332226>>.
- [32] BRYAN, N. J.; WANG, G. Musical influence network analysis and rank of sample-based music. *Oral Session*, p. 6, 2011.
- [33] MÜLLER, M. *Information Retrieval for Music and Motion*. Berlin, Heidelberg: Springer-Verlag, 2007. ISBN 3540740473.
- [34] MLLER, M. *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2015. ISBN 3319219448.