

# Redes de colaboração musical envolvendo os artistas mais reproduzidos no Spotify

Victor Ramos de Albuquerque Cabral  
Centro Federal de Educação Tecnológica de Minas Gerais  
Divinópolis, Brasil  
vramoscabral2020@gmail.com

**Resumo**—Este estudo tem como iniciativa analisar colaborações entre os principais artistas da indústria musical, utilizando como base de dados os artistas mais reproduzidos na plataforma de streaming Spotify e todas as músicas que eles fizeram em colaboração. Dentro da Teoria dos Grafos, o algoritmo principal utiliza o clique - subconjunto de um grafo onde todos vértices possuem arestas que os conectam - para representar uma rede de colaborações entre artistas (vértices) que se conectam entre si caso haja o registro de pelo menos uma música em parceria. Ao determinar qual é o tamanho do clique com maior número de vértices, é possível identificar os artistas mais envolvidos em parcerias musicais. O objetivo deste trabalho é identificar em qual gênero musical a interconectividade artística é mais frequente, fundamentando-a como um elemento influente na popularidade dos artistas.

**Palavras-Chave**—artistas, colaborações, musical, clique, grafo

## I. INTRODUÇÃO

A distribuição da música no mundo passou pelas mídias físicas - Disco de vinil, Fita K7, Compact Disc -, e pelo rádio, até chegar ao ponto atual onde a predominância pertence às plataformas digitais de streaming. Devido ao avanço da internet e das mídias digitais, o streaming de música se tornou a principal ferramenta da indústria musical para lançar novidades e armazenar uma extensa biblioteca que abrange a história da música gravada. Os artistas compartilham suas músicas nas plataformas de streaming e usam a internet como principal meio de divulgação. A construção de uma presença online é crucial para os artistas, pois o declínio das mídias físicas diminuiu a receita proveniente das vendas de álbuns.

Nesse contexto, muitos artistas conseguiram atingir o sucesso de maneira individual. Porém alguns artistas necessitam de colaborações para manterem sua popularidade e relevância nos rankings de reprodução musical, por exemplo, os Djs. Uma parte significativa das canções lançadas por Djs, principalmente os Djs que estiveram nos dados que foram objeto desse estudo, envolve colaboração com outros artistas, por eles assumirem a função de produzir a música e precisar de um outro artista para gravar a voz. Além disso, muitos artistas populares participam frequentemente de músicas em colaboração, por influência de gravadora ou por vontade própria.

O uso de grafos como estrutura de dados neste estudo oferece uma abordagem alternativa para identificar grupos no âmbito musical com registros de colaborações frequentes e forte conexão, além de proporcionar meios para distinguir

esses grupos e gerar resultados que podem ser analisados detalhadamente sob diferentes perspectivas. O algoritmo ao fazer isso, elementa os nomes que mais praticam a interconectividade artística permitindo uma análise do cenário musical com exemplos concretos.

## II. TRABALHOS RELACIONADOS

O uso da programação e da Teoria de Grafos aplicado no contexto musical trata-se de uma abordagem que já foi utilizada em trabalhos anteriores, em destaque nessa seção do artigo estão dois trabalhos: "Detecting Collaboration Profiles in Success-based Music Genre Networks" [1] por estudantes da Universidade Federal de Minas Gerais, que detecta perfis de colaboração dentro de redes de gêneros populares, e "Network Analysis of the Spotify Artist Collaboration Graph" [2] de Tobin South, que também utiliza o Spotify como base para identificar redes de colaboração entre artistas.

## III. METODOLOGIA

Primeiramente, foi feita uma coleta dos dados que foram utilizados nesse estudo na API do Spotify, um serviço que permite a obtenção de informações sobre a plataforma, incluindo estatísticas a respeito dos artistas, músicas e álbuns, e também estatísticas individuais de cada usuário que está acessando o servidor da API com a própria conta. Nessa coleta de dados, foram considerados os artistas mais reproduzidos do Spotify de acordo com o ranking do site ChartMasters [3]. Na execução do algoritmo de coleta, foi acessada a lista de todas as músicas de cada artista, e foi feita uma filtragem que entregou aos dados finais apenas as músicas que possuíam a participação de um ou mais artistas. Depois da coleta de dados, foram reunidas as músicas de 150 artistas em um arquivo .csv. Para essa implementação, foi utilizada uma biblioteca específica da linguagem Python chamada *Spotipy* [4], que contém as funções que fazem o acesso e coleta de informações da API.

Com os dados obtidos, a segunda parte foi a execução do arquivo *grafo.py*. No primeiro momento da execução, esse código pega um dataframe que foi gerado no arquivo *dataframe.py*. O código nele implementado pega os dados do .csv único obtido na coleta e transforma ele no dataframe. Dataframe é uma estrutura de dados bidimensional, de linhas e colunas. No caso dessa implementação, o dataframe gerado possui 150 linhas e em cada uma dessas linhas possuindo duas colunas, uma coluna com o nome do artista, e a outra coluna

com a lista de músicas desse artista filtradas por colaborações. O código abre o .csv com os dados e faz uma leitura completa do arquivo. Foi feita uma separação entre os 150 artistas do arquivo: "- artista; Nome do artista", para indicar o começo de uma nova linha no dataframe e o início da lista das músicas do artista a partir da próxima linha, e "- fimartista;", para indicar o final da lista de músicas do artista. Em cada linha do arquivo que possui informações de música, contém o nome da música, artistas envolvidos, e o ID da música de acordo com o registro do Spotify. Cada linha do arquivo foi iterada, a primeira linha possui um número que representa a quantidade de artistas, as linhas contendo "- artista; Nome do artista" e "- fimartista;" marcando uma nova linha do dataframe e o fim da segunda coluna, e as outras linhas sendo as músicas, tiveram os dados separados em música, lista de artistas e id. As informações do dataframe de resultado foram salvas em um arquivo csv como forma de verificação. Nesse processo foi usada a biblioteca *Pandas* [5] de Python, para implementar as funções de tratamento do Dataframe. O primeiro pseudocódigo a seguir demonstra como é a implementação básica desse algoritmo secundário, que gera a base pro algoritmo principal.

---

**Algorithm 1** Gerador de DataFrame: 'dataframe.py'

---

```

1: Gerador-DataFrame(full.csv, dataframe)
2: dataframe ← Novo-DataFrame(nome, musicas)
3: main_arch ← full.csv
4: listamusic ← [ ]                                ▷ lista vazia
5: listart ← [ ]
6: i ← 0                                            ▷ Contador de linhas
7: Abra main_arch como arq_csv
8: leitor ← ler_csv(arq_csv)
9: for cada linha em leitor do                ▷ Separa o delimitador do
   nome do artista
10:  if '- artista;' está em linha then
11:    a ← separa(linha)
12:    artist ← a[1]
13:    listart.adiciona(artist)
14:  else if '- fimartista;' está em linha then
15:    if dataframe.vazio() then
16:      dataframe[0] ← [listart[0], listamusic[0]]
17:    else
18:      dataframe[0] ← [listart[i], listamusic[i]]
19:    end if
20:  else
21:    n ← separa(linha)
22:    musica ← n[0]
23:    artistas ← n[1]
24:    id ← n[2]
25:    nova_musica ← NovaMusica(musica, artistas, id)
26:    listamusic.adiciona(nova_musica)
27:    i += 1
28:  end if
29: end for
30: Fechar main_arch
31: Retornar dataframe

```

---

Com o dataframe pronto, o código com a implementação das funções do grafo desempenha o seu papel de algoritmo principal desse trabalho. Foram utilizadas duas bibliotecas: a *NetworkX* [6], que gera o grafo e possui as funções específicas, e a *Matplotlib* [7], que é responsável por plotar o grafo em um ambiente interativo. Primeiro é criado o Grafo, e todos os vértices desse grafo são preenchidos com os artistas da coluna nome do dataframe. Depois é percorrida a lista de músicas de cada artista do dataframe para a formação das arestas. Em cada música é verificado se o artista é diferente do que está sendo iterado e também se ele faz parte dos nomes do dataframe, em caso positivo, é criada uma aresta com os nomes dos dois artistas e atribuição de peso 1, peso que aumentará a cada ocorrência da mesma aresta. Após isso, é criado um grafo não-direcionado para ser realizada a busca pelo clique máximo, com a maior quantidade de vértices. A decisão em determinar o clique máximo é um Problema NP-Completo, portanto foi descartada a possibilidade de trabalhar com um clique único, logo, o algoritmo lista todos os cliques e depois os ordena por quantidade de vértice, e então printa no terminal 10 cliques máximos. O segundo pseudocódigo demonstra esse processo.

---

**Algorithm 2** Criar Grafo e Achar Cliques 'grafo.py'

---

```

procedure CRIAR-GRACO(dataframe)
2:  G ← Networkx.novografo()
   for cada linha em dataframe do
4:    vértice ← linha[nome]
    G.add_vértice(vértice)
6:    for cada musica em linha[musicas] do    ▷ Separa
   os artistas da música
    arts ← musica[artistas].split()
8:    for cada art em arts do
    if art ≠ vértice e art em dataframe then
10:      if G.aresta_existe() then
        G[vértice][art].peso += 1
12:      else
        G.add_aresta(vértice, art, 1)    ▷ Peso
        inicial é 1
14:      end if
    end if
16:  end for
18:  end for
   Retornar G
20: end procedure
procedure ACHAR-CLIQUE-MÁXIMO(G)
22:  cliques ← lista(nx.find_cliques(G))
    cliques_ordenados ← sorted(cliques, key =  $\lambda x$  :
        len(x), reverse)    ▷ Decrescentemente
24:  for cada i, clique em enumerate(cliques_ordenados[
        10], start = 1) do
    Imprimir(Clique: {clique}, vértices: {len(clique)})
26:  end for
end procedure

```

---

#### IV. RESULTADOS

Esse trabalho gerou um grafo interativo por meio da biblioteca *Matplotlib* [7] para mostrar todas as ligações considerando todos os artistas da lista dos mais populares no Spotify. E na análise de cliques, foi identificados dois grupos recorrentes e fortemente conectados. O primeiro grupo envolve os artistas: J Balvin, Ozuna, Nicky Jam, Farruko, Daddy Yankee, Maluma, Myke Towers, Bad Bunny, Rauw Alejandro, Anuel AA. Esse grupo é formado por artistas latinos, em sua maioria porto-riquenhos, e o gênero predominante nessa rede é o Hip-Hop/Rap. Um indicativo da grande relevância da música latina e da grande conexão que existe entre os artistas envolvidos.

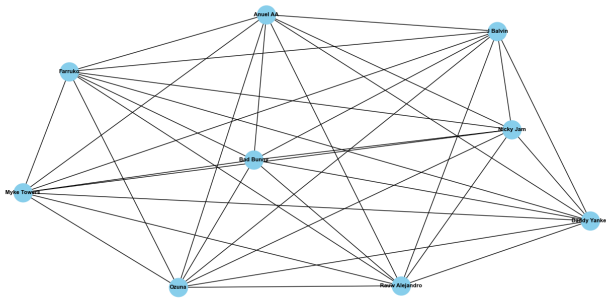


Fig. 1. Representação do clique do primeiro grupo.

No outro grupo, há a presença dos seguintes artistas: *Lil Wayne, Future, Young Thug, Nicki Minaj, DJ Khaled, Migos, 21 Savage, DaBaby, Travis Scott*, em suas variações de clique participaram *Drake, Post Malone, YoungBoy Never Broke Again e Lil Uzi Vert*. Novamente, foi identificado mais um grupo de contribuições musicais entre rappers, e o DJ Khaled na função de produção musical. Esse grupo proveniente dos Estados Unidos da América, é mais um indicador do impacto desse país na produção de músicas populares e da forte influência cultural que eles exercem globalmente. Também foi nos Estados Unidos onde se originou o Rap e o Hip-Hop. O Hip-Hop além de um estilo musical, é também um movimento cultural que influencia até hoje a cultura pop e a moda.

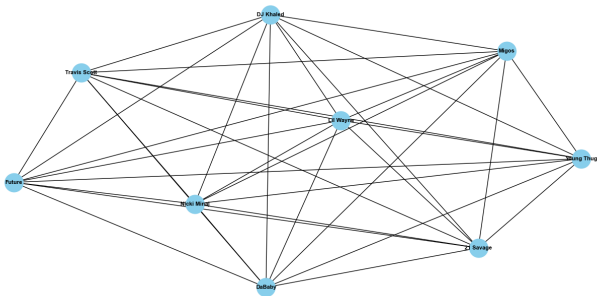


Fig. 2. Representação do clique do segundo grupo.

Os resultados e metodologia apresentados nesse artigo foram extraídos com a execução de códigos disponíveis no

repositório do GitHub "Trabalho Final AEDS II". [8]

#### V. CONCLUSÃO

Este trabalho, fundamentado na Teoria dos Grafos e na implementação dos cliques, revelou resultados indicativos de que o gênero Hip-Hop/Rap, dentro da base de dados que abrange os principais artistas reproduzidos no Spotify dos mais variados gêneros, incluindo Pop, Rock, Dance, entre outros, é o gênero musical onde a interconectividade artística é mais proeminente. O Hip-Hop se mostra desde o princípio do movimento cultural como um ambiente propício para a realização de parcerias musicais, contribuindo com a formação de grupos de colaboração musical. Pode-se dizer que dentro dessa realidade, o Hip-Hop apresenta um forte network musical.

As conexões geradas pelo uso de Grafos facilitam a análise de redes não só de colaborações musicais, como também em networks do mundo empresarial que engloba parcerias entre fornecedores, empresários, entre outros, também na análise de relações entre usuários de redes sociais como amigos, seguidores, produtores de conteúdo, e questões menos subjetivas como Redes de Transporte permitindo melhor planejamento logístico, e Redes de Infraestrutura Urbana, como a distribuição da água e dos cabos elétricos.

Portanto, esse estudo explorou o conceito de interconectividade artística como uma ferramenta de influência no mercado musical e também na carreira dos próprios artistas, permitindo que eles consigam renovar seu público por meio dessas parcerias e conquistar mais audiência consolidando suas carreiras e seus nomes na história da música. Além da visão abrangente das colaborações musicais, esse estudo destacou a importância de outros fatores como a cultura, gênero, mídia digital e localização para a formação desses networks musicais. Espera-se que esse estudo possa contribuir para a realização de novas pesquisas sobre a dinâmica da indústria musical.

#### REFERENCES

- [1] G. P. Oliveira, M. O. Silva, Danilo B. S., Anisio Lacerda, M. M. Moro, "Detecting Collaboration Profiles in Success-based Music Genre Networks." [https://program.ismir2020.net/static/final\\_papers/275.pdf](https://program.ismir2020.net/static/final_papers/275.pdf)
- [2] Tobin South, "Network Analysis of the Spotify Artist Collaboration Graph", [https://vrs.amsl.org.au/wp-content/uploads/sites/84/2018/04/tobin\\_south\\_vrs-report.pdf](https://vrs.amsl.org.au/wp-content/uploads/sites/84/2018/04/tobin_south_vrs-report.pdf)
- [3] ChartMasters, "Most streamed artists on Spotify (daily update) - ChartMasters," ChartMasters, Nov. 23, 2023. <https://chartmasters.org/most-streamed-artists-ever-on-spotify/>
- [4] "Spotipy 2.0 documentation" <https://spotipy.readthedocs.io/>
- [5] "Pandas 2.1.3 documentation" <https://pandas.pydata.org/docs/>
- [6] "NetworkX documentation." <https://networkx.org/>
- [7] "Matplotlib 3.8.2 documentation." <https://matplotlib.org/stable/index.html>
- [8] CABRAL, V. R. A., "Trabalho Final AEDS II," GitHub, 2023, [Online]. <https://github.com/vramoscabral/Trabalho-Final-AEDS-II>