

Similaridades entre músicas de uma playlist criada por humanos influenciam na sua popularidade?

Andryw Marques

Laboratório de Sistemas Distribuídos (LSD)
Universidade Federal de Campina Grande (UFCG)
Campina Grande, PB, Brasil

Resumo—O grande volume de música disponível torna importante a pesquisa e uso de recomendadores musicais e geradores automáticos de playlists. Porém, existem vários critérios para gerar playlists de forma automática. Para inferir qual desses critérios é o mais relevante, foi realizado um estudo para descobrir qual dos seguintes fatores, similaridade acústica entre músicas, similaridade entre tags das músicas, similaridade entre artistas das músicas e a popularidade atual das músicas, afetam a popularidade de playlists criadas de forma não-automática por humanos. Porém, notou-se que estes fatores pouco influenciam a popularidade das playlists.

I. INTRODUÇÃO

O volume de música disponível cresceu muito nos últimos anos. A internet, áudio no formato digital, tocadores de mp3 e redes sociais ajudaram nessa disseminação da música. Essa disponibilidade se dá tanto do aumento das bibliotecas musicais pessoais quanto da presença de serviços de streaming, como o Spotify¹, que possui mais de 15 milhões de músicas disponíveis. Apesar de ser algo importante para o usuário, isto acarretou em um problema: devido a grande quantidade de música disponível, o usuário perde muito tempo escolhendo as músicas que ele deseja ouvir. É neste contexto que entram os recomendadores musicais automáticos.

Recomendadores musicais têm como objetivo recomendar músicas de acordo com algum critério. Existe um tipo específico de recomendador musical denominado gerador automático de playlists. Playlists são conjuntos de músicas para um fim². O gerador não só recomenda uma música isolada, mas um conjunto onde tipicamente as músicas do conjunto possuem algo em comum. Em muitas situações é preferível a recomendação de playlists em vez da recomendação individual de músicas, como em músicas para uma festa, viagem, gravar CD para presentear, entre outras.

Existem várias abordagens para a geração de playlists, como por exemplo: geração por similaridade acústica, onde as músicas são selecionadas de acordo com similaridades dos atributos acústicos (timbre, pitch, tempo); similaridade entre metadados, onde as músicas são selecionadas de acordo com similaridades nos metadados, geralmente provenientes

de tags, como gênero (rock, pop, rap), humor (depressivo, relaxante); similaridade entre artistas, onde os artistas das músicas são similares; e geração por filtragem colaborativa, onde as músicas são escolhidas baseadas no histórico das músicas de outras pessoas com gosto similar ao do usuário que requisitou a playlist.

Na área de criação não-automática de playlists, as pessoas criam constantemente playlists desta maneira. Sites como 8tracks.com e artofthemix.com são redes sociais que permitem a criação e compartilhamento destas playlists. É possível que algumas das similaridades apontadas acima (acústica, por metadados e por artistas) estejam presentes nas playlists criadas não-automaticamente (manualmente)? Se sim, essas similaridades influenciam na popularidade das playlists? O objetivo deste trabalho é descobrir que fatores influenciam a popularidade de uma playlist criada manualmente e isto pode ajudar na escolha da abordagem de criação automática de playlists. Até onde pudemos determinar, não há estudos cobrindo essa questão.

Para descobrir isso, foram coletadas playlists criadas manualmente no site 8tracks, calculados valores de similaridades acústica, entre metadados e entre artistas das músicas de cada playlist e, junto com uma métrica de popularidade atual de cada música da playlist (também conhecida como hotness), foram verificados se estes quatro fatores influenciam a popularidade das playlists. Para isso foi utilizado um modelo de regressão multivariada.

Os resultados apontaram que os fatores estudados são responsáveis por uma influência pequena, ou quase nenhuma, da popularidade. É possível que haja outros fatores além dos estudados que influenciem a popularidade das playlists.

II. TRABALHOS RELACIONADOS

A questão da criação manual de playlists foi abordada em outros trabalhos. O artigo [1] levantou fatores que influenciam na criação de playlists, a partir de dados de entrevistas. Entre outras conclusões, ele apontou os principais critérios que os usuários utilizam para a criação de suas playlists: gênero, evento ou atividade, romance e humor. Outra contribuição foi evidenciar que há situações onde a ordem das músicas das playlists importa e outras onde não importa, muitas vezes as músicas sendo tocadas aleatoriamente.

¹<http://www.spotify.com/>

²O artigo [1] define playlist como conjunto de músicas para uso pessoal, e mix como conjunto de músicas com tema bem definido, onde geralmente a sequência das músicas importa. No artigo proposto ambos os conceitos foram atribuídos à playlist

Na área de recomendadores musicais, foram levantados trabalhos que utilizam diferentes critérios de recomendação. Recomendação musical baseada em valores acústicos podem ser vistos em [2] e [3]; recomendações baseadas em tags podem ser vistos em [4]; baseadas na similaridade de artistas também em [4]; baseadas em filtragem colaborativa ([5]) e até mesmo recomendadores híbridos ([6] e [7]), utilizando mais de um critério de recomendação.

Na questão da avaliação dos diferentes tipos de geradores de playlists, foram encontrados 2 tipos de avaliação. O primeiro é a avaliação a partir de respostas dos usuários às músicas recomendadas por recomendadores onde já se sabe previamente o critério de recomendação. O artigo [8] propõe a criação de recomendadores musicais baseados em medidas referentes ao conteúdo acústico musical e depois faz uma comparação destes com outros recomendadores baseados em gênero, filtragem colaborativa, análise musicológica, recomendação humana e randomicidade. Foi apontado neste artigo que o recomendador baseado na acústica recomenda músicas tão bem quanto os outros recomendadores do estado da arte. Já o artigo [4] compara os resultados do recomendador Genius com um recomendador baseado em similaridade de artistas e com um baseado em similaridade acústica, mostrando que o Genius, que é baseado em filtragem colaborativa, tem melhores recomendações que os outros. O segundo tipo de avaliação é criar um modelo a partir de um conjunto de playlists e verificar se as recomendações do modelo são parecidas com outro conjunto de playlist ([9] e [5]).

Diferente destes tipos de avaliação, o que vai ser proposto neste artigo é verificar se, em um conjunto de playlists criadas por humanos, existe algum fator que influencie a popularidade delas (playlists).

III. METODOLOGIA

A. Objetivo

O objetivo desta análise é verificar se algum destes 4 fatores: similaridade acústica entre músicas, similaridade entre as tags das músicas, similaridade entre os artistas de cada música e hotness das músicas, influenciam na popularidade de playlists. Para isso, foram calculadas medidas de centralidade (média e mediana) e uma medida de dispersão (desvio-padrão) da distribuição de cada fator dentro das playlists. Assim, foi investigado se há uma relação entre a popularidade e medidas de centralidade e variação da distribuição de cada fator dentro das playlists.

B. Contexto

As playlists foram coletadas do site 8tracks.com. O 8tracks é uma rede social que provê criação, compartilhamento e votação de playlists. O usuário pode acessar o site tanto em computadores quanto em smartphones.

C. Unidades experimentais - Coleta e Seleção

1) *Informações sobre músicas:* Para calcular as diversas similaridades, foram utilizadas informações sobre as músicas.

Estas informações foram coletadas do Million Song Dataset (MSD)³. Os seguintes dados foram coletados:

- Valores sobre o timbre da música: para calcular a similaridade acústica, foram coletados os valores do timbre de cada música. Estes valores são provenientes do projeto echonest.com e disponibilizados pelo MSD. Cada música é dividida em segmentos e esses segmentos são valorados de acordo com 12 dimensões relacionadas ao timbre da música, como o volume médio do segmento, o "brightness" da música, etc.⁴.
- Conjunto de artistas similares de um artista: estes valores são provenientes do projeto echonest.com e disponibilizados pelo MSD.
- Conjunto de tags que caracterizam uma música e o peso de cada tag sobre a mesma: estes dados são provenientes do site Last.fm e disponibilizados pelo MSD. As tags são adicionadas por pessoas e há a possibilidade de mais de uma pessoa adicionar a mesma tag. Assim, o site provê um peso para cada tag, que varia de 0 a 100, onde o peso 100 é atribuído à tag que mais pessoas marcaram⁵.
- Hotness da música: o echonest.com disponibiliza um valor, entre 0 e 1, sobre o quanto a música está popular no momento.

2) *Unidades experimentais:* As unidades experimentais são as playlists do site 8tracks.com. Cada playlist possui um número de *likes*, atribuídos pelos usuários, e um número de *listens*, que é incrementado quando um usuário escuta todas as músicas da playlist. Foi definida como "popularidade" a razão entre o número de *likes* e o número de *listens*.

Foi coletado um grupo randômico de 500 playlists, com número de *listens* maior que 20, onde mais de 50% das músicas das playlists estavam presentes no MSD. 500 playlists pois, quanto mais dados, maior probabilidade de encontrar significância onde não há. Foram feitos estudos com 1000 playlists e ocorreu isto. Mais de 20 *listens* para evitar que playlists as quais usuários dão *likes* sem terem escutado, ou que, com poucas execuções, sejam consideradas populares.

D. Variáveis, Métricas e Aplicação

Foram utilizadas 4 métricas:

1) Similaridade acústica

Os valores coletados do timbre das músicas foram modelados para um Single Gaussian Model e então calculada a distância entre duas músicas segundo a Kullback-Leibler divergence entre Single Gaussian Models. [2]

2) Similaridade entre tags

O last.fm provê uma lista de tags e de seus pesos para cada música. Então, foi calculada a distância euclidiana para o valor de cada peso. Se por acaso uma tag estivesse na música A e não estivesse na música B, o valor do peso da tag na música B seria 0.

³<http://labrosa.ee.columbia.edu/millionsong/>

⁴http://docs.echonest.com.s3-website-us-east-1.amazonaws.com/_static/AnalyzeDocumentation.pdf

⁵<http://www.lastfm.com.br/api/intro>

3) Similaridade entre artistas

O MSD provê uma lista de artistas similares para cada artista. Assim, foi calculado o Jaccard Index entre os dois conjuntos de artistas similares de duas músicas.

4) Popularidade da música

Cada música possui um valor para sua popularidade, proveniente do echonest.com. Este valor é calculado baseado na atividade dos artistas nos sites analisados pelo echonest.

Os valores das similaridades acústica, entre tags e entre artistas foram calculados entre duas músicas consecutivas da playlist, gerando 3 conjuntos de similaridades, que junto com a popularidade de cada faixa, formam uma tupla de 4 conjuntos para cada playlist. Para cada conjunto foram calculados a média, a mediana e o desvio padrão para serem analisados na regressão, totalizando 12 variáveis independentes.

E. Hipóteses

Após a definição das variáveis, foram definidas 12 hipóteses:

H1-0: a média da similaridade acústica entre músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H2-0: a média da similaridade entre artistas de músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H3-0: a média da similaridade entre tags de músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H4-0: a média da popularidade atual das músicas de uma playlist não influencia a popularidade da mesma;

H5-0: a mediana da similaridade acústica entre músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H6-0: a mediana da similaridade entre artistas de músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H7-0: a mediana da similaridade entre tags de músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H8-0: a mediana da popularidade atual das músicas de uma playlist não influencia a popularidade da mesma;

H9-0: o desvio padrão da similaridade acústica entre músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H10-0: o desvio padrão da similaridade entre artistas de músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H11-0: o desvio padrão da similaridade entre tags de músicas consecutivas de uma playlist não influencia a popularidade da mesma;

H12-0: o desvio padrão da popularidade atual das músicas de uma playlist não influencia a popularidade da mesma.

F. Design

Inicialmente foi realizado um estudo de correlação, para verificar o quanto as variáveis independentes estavam relacionadas com a variável dependente. Notou-se que a correlação era fraca, para todas as variáveis. Próximo passo foi a construção do modelo de regressão multivariada.

G. Ameaças à validade

Apontamos algumas ameaças à validade do trabalho:

- As playlists foram coletadas de um só local. Talvez seja um problema generalizar os resultados;
- As tags e a similaridade entre artistas não são entidades objetivas, mas são valores atribuídos por humanos. Assim, calcular a similaridade entre tags das músicas e similaridade entre os artistas das músicas vai depender dessas entidades subjetivas.
- Falta de investigação mais profunda sobre que transformação nos dados utilizar: processar dados musicais nem sempre é trivial. Entre alguns problemas, foram encontrados: atributos das playlists e das músicas (artista, nome) com nomes errados; artista e nome da música colocados juntos como nome da música; uma música possuir várias versões; entre outros problemas. Foi utilizado Levenshtein Distance para minimizar estes problemas.
- Não levar em conta outros fatores para popularidade da playlist, como influência do criador: pode ser que outros fatores, como o número de followers de um usuário, possam influenciar em sua popularidade.
- A forma de calcular cada tipo de similaridade pode não representar significativamente o conceito da similaridade.
- Impossibilidade de coletar amostra de playlists com quase todas as músicas com informações no MSD: nem todas as músicas das playlists possuíam atributos no MSD.

IV. RESULTADOS

Tabela I
ANÁLISE DE CORRELAÇÃO - VARIÁVEIS INDEPENDENTES E POPULARIDADE (MÉTODO DE SPEARMAN)

Variável Independente	Coefficiente de correlação
Média similaridade entre artistas	0.15
Média similaridade acústica	-0.10
Média similaridade entre tags	-0.13
Média popularidade atual	-0.039
Mediana similaridade entre artistas	0.019
Mediana similaridade acústica	-0.16
Mediana similaridade entre tags	-0.11
Mediana popularidade atual	-0.0073
Desvio padrão similaridade entre artistas	0.25
Desvio padrão similaridade acústica	0.049
Desvio padrão similaridade entre tags	0.084
Desvio padrão popularidade atual	0.14

O estudo da correlação (Tabela I) foi feito utilizando o método de Spearman. O estudo aponta que as correlações das variáveis independentes com a popularidade são correlações fracas.

Tabela II
MODELO II DA REGRESSÃO

Variável Dependente	
$\sqrt{\text{popularidade}}$	
Variável Independente	Coefficiente
Média similaridade entre artistas	-5.7e-01
Média similaridade acústica	-2.3e-06
Desvio padrão similaridade entre artistas	7.8e-01
Desvio padrão similaridade acústica	2.1e-06
Desvio padrão similaridade entre tags	1.3e-04
Desvio padrão popularidade atual	4.2e-01

Tabela III
VALORES DOS MODELOS DE REGRESSÃO

R-squared	R-squared ajustado	P-value	Modelo
0.088	0.066	8.3e-06	Modelo I
0.092	0.081	1.3e-08	Modelo II

Tabela IV
TESTE DE NORMALIDADE PARA OS DOIS MODELOS (SHAPIRO-WILK TEST)

P-value
< 2.2e-16

Próxima etapa foi realizar a construção dos modelos de regressão. Utilizando todas as variáveis (Modelo I), o modelo explica cerca de 6,6% da variação da popularidade (R-squared ajustado - Tabela III). Foi realizado o teste Shapiro-Wilk no Modelo I para verificar a normalidade dos resíduos (Tabela IV). Porém, o modelo não passou no teste da normalidade, sendo evidenciado esse resultado também no Gráfico Normal Q-Q Plot (Figura 1). Com isso, não foi possível analisar os resultados dos Testes T e F do modelo.

Para tentar alcançar a normalidade dos resíduos, foram feitas algumas alterações no Modelo I, resultando no Modelo II. A Tabela II mostra quais as variáveis utilizadas neste modelo, incluindo a transformação da variável popularidade (raiz-quadrada). Está presente também na Tabela II os valores dos coeficientes encontrados. Este modelo explica cerca de 8,1% da variação da popularidade (R-squared ajustado - Tabela III), porém os resíduos também não vêm de uma população normal (Tabela IV). O Gráfico Normal Q-Q Plot do Modelo II é bem parecido com o do Modelo I (Figura 1). Assim, também não foi possível analisar os resultados dos Testes T e F do modelo.

Com estes resultados (correlação fraca e baixa explicação da variação da popularidade) pode-se afirmar que as variáveis estudadas possuem uma influência muito baixa na popularidade das playlists criadas por humanos. Possivelmente existem outros fatores que podem influenciar na popularidade e que não estão sendo levados em consideração, como influência do criador da playlist, a visibilidade que a playlist possui no site, entre outros.

V. CONCLUSÕES E TRABALHOS FUTUROS

Muitos fatores podem influenciar a popularidade de uma playlist. Similaridades acústica, entre tags, entre artistas, po-

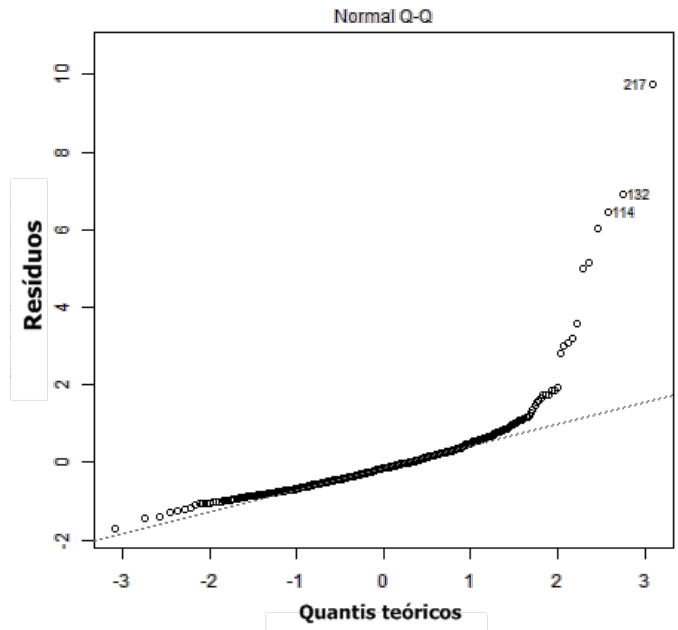


Figura 1. Normal Q-Q Plot dos resíduos do Modelo I

pularidade das músicas, músicas que seus vizinhos escutam, entre outros. Porém, os modelos gerados a partir dos fatores estudados neste trabalho explicam pouco a variação da popularidade. Como contribuição do trabalho, pode-se concluir que os fatores estudados não influenciam a popularidade das playlists criadas manualmente.

Outra contribuição do trabalho foi servir como base para trabalhos futuros. É importante listar e entender outros fatores que possam influenciar a popularidade das playlists criadas manualmente. A forma de cálculo das similaridades também pode ser analisada futuramente, já que não há um consenso na literatura de como calcular similaridades entre músicas.

REFERÊNCIAS

- [1] S. J. Cunningham, D. Bainbridge, and A. Falconer, "“more of an art than a science”: Supporting the creation of playlists and mixes,” in *International Society for Music Information Retrieval Conference (ISMIR)*, PAGES =.
- [2] D. Bogdanov, J. Serrà, N. Wack, P. Herrera, and X. Serra, "Unifying low-level and high-level music similarity measures,” *IEEE Transactions on Multimedia*, vol. 13, pp. 687–701, 08/2011 2011.
- [3] E. Pampalk and M. Gasser, "An implementation of a simple playlist generator based on audio similarity measures and user feedback,” in *International Society for Music Information Retrieval Conference (ISMIR)*, 2006, pp. 389–390.
- [4] L. Barrington, R. Oda, and G. R. G. Lanckriet, "Smarter than genius? human evaluation of music recommender systems,” in *International Society for Music Information Retrieval Conference (ISMIR)*, 2009, pp. 357–362.
- [5] N. Aizenberg, Y. Koren, and O. Somekh, "Build your own music recommender by modeling internet radio streams,” in *Proceedings of the 21st international conference on World Wide Web*, 2012, pp. 1–10.
- [6] F. Maillet, D. Eck, G. Desjardins, and P. Lamere, "Steerable playlist generation by learning song similarity from radio station playlists,” in *International Society for Music Information Retrieval Conference (ISMIR)*, 2009, pp. 345–350.
- [7] D. Bogdanov and P. Herrera, "How much metadata do we need in music recommendation? a subjective evaluation using preference sets,” in *International Society for Music Information Retrieval Conference (ISMIR)*, 2011.

- [8] T. Magno and C. Sable, "A comparison of signal based music recommendation to genre labels, collaborative filtering, musicological analysis, human recommendation and random baseline," in *International Society for Music Information Retrieval Conference (ISMIR)*, 2008, pp. 161–166.
- [9] B. McFee and G. R. G. Lanckriet, "The natural language of playlists," in *International Society for Music Information Retrieval Conference (ISMIR)*, 2011, pp. 537–542.