

12º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2021

SISTEMA DE RECOMENDAÇÃO DE MÚSICAS BASEADO NA CORRELAÇÃO ENTRE ÁUDIOS E ARQUIVOS DIGITAIS

ITANU VAIMACA ROMERO MARTINEZ¹, LUCIANO BERNARDES DE PAULA²

¹ Graduando em Análise e Desenvolvimento de Sistemas, Bolsista PIBIFSP, IFSP, Câmpus Bragança Paulista, itanu.romero@aluno.ifsp.edu.br

² Docente, IFSP, Câmpus Bragança Paulista, lbernardes@ifsp.edu.br
Área de conhecimento (Tabela CNPq): 1.03.03.04-9 Sistemas de Informação

RESUMO: Neste artigo é apresentado um projeto de pesquisa o qual buscou uma forma alternativa de se produzir sistemas de recomendação de músicas. Esse tipo de sistema de recomendação, geralmente, considera informações abstratas como *tags* (marcações que indicam características) em sua recomendação. No caso das músicas digitais, o seu gênero, artista e outras informações constam em metadados, que são considerados ao processar uma recomendação ao usuário. Isso faz com que o algoritmo de recomendação dependa de que esses dados existam e/ou estejam fiéis com a realidade. São poucos os serviços de *streaming* que oferecem recomendações baseadas somente no dado, sem o uso de *tags*, que podem trazer vieses para os resultados, uma vez que, em alguns casos, a definição de um gênero específico para uma música não é tarefa simples. Com a utilização de um cálculo chamado de Similaridade de Hamming, o qual realiza uma comparação bit a bit (menor parte de cada arquivo digital) foram feitos testes para verificar a semelhança de cada arquivo de forma pura através de seu código binário, assim deixando de depender de interação humana.

PALAVRAS-CHAVE: algoritmo; arquivos; hamming; binário, semelhança;

Music recommendation system based on the correlation between audios and digital files

ABSTRACT: In this paper is presented a research project which sought an alternative way to produce music recommendation systems. This type of recommendation system, generally, considers abstract information like tags (which indicate some characteristics) to perform recommendation. In case of digital music, its genre, artist and other information are presented in the metadata, which are considered to process the recommendation to the user. This causes the algorithm of recommendation to rely on this data and that they are true to reality. There are few streaming services that offer recommendations based just on data, without using tags, which may have biased results, since, in some cases, the definition of a specific genre to a music is not a simple task. Using the Hamming Similarity, which performs a bit to bit comparison (the smallest part of a digital file) tests were done to verify the similarity of each file in its binary code, not relying on human interaction.

KEYWORDS: algorithm; files; hamming; binary, similarity;

INTRODUÇÃO

Atualmente a forma mais comum de se escutar música é através de serviços de *streaming*, tais como Spotify, Deezer, entre outros (Tecnundo, 2012). Esse tipo de serviço passou por um crescimento recente expressivo em números de usuários (GIMENES, 2020) e, com o intuito de oferecer um serviço personalizado para seu usuário, esses sistemas oferecem recomendações.

Os algoritmos de recomendação podem oferecer músicas da preferência do usuário a partir de músicas que ele tenha consumido. Utilizando os metadados dos arquivos, é possível identificar gênero musical, artista, localidade, álbum, entre outras informações. Essas informações são usadas como referência para a descoberta de outras músicas que possam ser de interesse do usuário.

Essa forma de se recomendar é muito utilizada, porém possui limitações. Os metadados são inseridos por pessoas e, mesmo que sejam especialistas, pode ser que exista a confusão ao definir o gênero de uma música, por exemplo, o que acaba impactando diretamente o funcionamento do algoritmo de recomendação. Considerando essa limitação, o projeto apresentado aqui estudou a possibilidade de se criar um sistema de recomendação baseado no conteúdo do dado digital do arquivo de áudio, ou seja, de seu código binário, sem depender da opinião humana.

MATERIAL E MÉTODOS

Existem diversas formas de representar áudio de forma digital. Para este estudo, decidiu-se utilizar os formatos de arquivo .MP3, .WAV e .MID (ou .MIDI), e a escolha foi feita com base na análise disponibilizada por Pereira (2017) e Tanaka, Barbosa e Kimura (2017), que disponibilizaram seus estudos com base na diferenciação e características dos arquivos digitais de áudio.

A base de dados utilizada no projeto foi a plataforma gratuita e on-line Jamendo (<https://www.jamendo.com/>), a qual disponibiliza músicas livres de direitos autorais. Desses arquivos foi feita a leitura de seu padrão binário, e para as análises de similaridade usou-se a Similaridade de Hamming (PAULA, 2011)(TEKNOMO, 2015).

Similaridade de Hamming

Como é apresentado em PAULA, L. B. (2011) e TEKNOMO (2015), a similaridade de Hamming é calculada para vetores binários, os quais possuem posições com valores binários. É então realizada uma comparação utilizando o método Ou Exclusivo (OLIVEIRA, 2017) entre os dois vetores. O número de 0s no vetor resultante são contados, ou seja, quanto mais similares os vetores binários, considerando o valor e a posição dos bits, maior o número de 0s resultantes da operação. Ao final, o valor total de 0s é dividido pela quantidade de bits dos vetores, calculando-se a semelhança.

Software desenvolvido

Para o projeto, foi desenvolvido um *software* capaz de realizar a leitura de dois arquivos, dividindo o conteúdo binário em blocos e realizando a comparação de acordo com a similaridade de Hamming, resultando assim em uma porcentagem de semelhança (ROMERO, 2020).

O sistema foi desenvolvido em linguagem C, o qual lê os arquivos de áudio na forma de leitura binária. O software faz a leitura de blocos de 32 bits das duas músicas, e a similaridade de Hamming é calculada para esses dois blocos, realizando a comparação bit a bit, que resulta em 0 caso os bits analisados sejam iguais, e 1 caso não, incrementando um contador em uma unidade a cada resultado.

Assim, para descobrir a porcentagem de similaridade entre os arquivos, pode-se calcular o total de comparações, que no caso de um bloco de 32 bits são 32 comparações, com o acumulador de zeros ou uns. Caso se encontre 16 resultados do tipo 0, por exemplo, há 50% de semelhança entre esses blocos. É importante destacar que devido os vetores serem binários, a similaridade de dois vetores com valores aleatórios tende a resultar em cerca de 50% de similaridade. Então, é possível concluir que para considerar dois arquivos similares, a porcentagem de similaridade deve partir de 50%.

Metodologia de Testes

Foram feitos testes iniciais, balizadores, com dois arquivos iguais, após isso com dois arquivos que possuem grande similaridade, como duas versões da mesma música, e, para finalizar a etapa inicial, duas músicas aleatórias, que deveria resultar em uma similaridade menor.

Os testes com MP3 resultaram para as três análises resultaram em 100% para arquivos iguais (o que era esperado para a confirmação da corretude do código desenvolvido), 76.02% para duas versões da mesma música e 76.007% para duas músicas aleatórias. Para a extensão WAV, além de ter uma análise demasiada extensa do ponto de vista de tempo, por conta do tamanho do arquivo, a comparação entregou resultados contrários, na qual as versões de uma mesma música entregaram uma porcentagem de semelhança menor que das músicas aleatórias. Os resultados foram 100% para arquivos iguais, 76.74% para versões da mesma música e 76.93% para músicas aleatórias. O arquivo MID entregou resultados otimistas, mantendo o 100% de similaridade para os arquivos iguais, e as análises com duas versões da mesma música e com músicas aleatórias foram de 78.67% e 61.27%, respectivamente. Após esses testes, para as 3 extensões de arquivos escolhidos, foi escolhido seguir o formato MID, já que foi o que representou melhor esses testes iniciais.

RESULTADOS E DISCUSSÃO

Após inicializado os testes com arquivos MIDI criados originalmente nesse formato, constatou-se que músicas da mesma banda tinham um maior grau de similaridade que outras duplas de arquivos analisados, e a partir desses resultados, foi possível continuar os testes. Por conta da limitação da diversidade de músicas criadas diretamente em MIDI e também para se adaptar melhor ao mercado atual, os próximos testes foram feitos como descrito a seguir.

Primeiro se fez o *download* da música pelo banco de dados Jamendo/Outras plataformas no formato MP3, logo depois, o arquivo era convertido para a extensão MIDI, após isso foram realizados os testes para cada dupla de arquivos.

Nesta rodada de testes, foram realizadas cerca de 30 comparações aleatórias, nas quais se buscavam 2 arquivos MIDI (provenientes de MP3) os quais serviam de base para a análise, resultando na similaridade entre os arquivos. Assim, para detectar padrões esses foram utilizados como abaixo:

- Três versões de uma mesma música, sendo sua versão estúdio (original), sua versão remixada por outro artista (remix) e sua versão gravada ao vivo (live).
- Uma música seguindo os padrões da música anterior (do gênero Eletrônica).
- Músicas de artistas diferentes, porém que seguem o mesmo gênero.
- Músicas de gêneros distintos para atenuar as diferenças.

Entregando assim os resultados apresentados na Figura 1, que são gráficos os quais cada uma das barras representa uma comparação entre 2 arquivos, que são identificados pelo seu gênero.

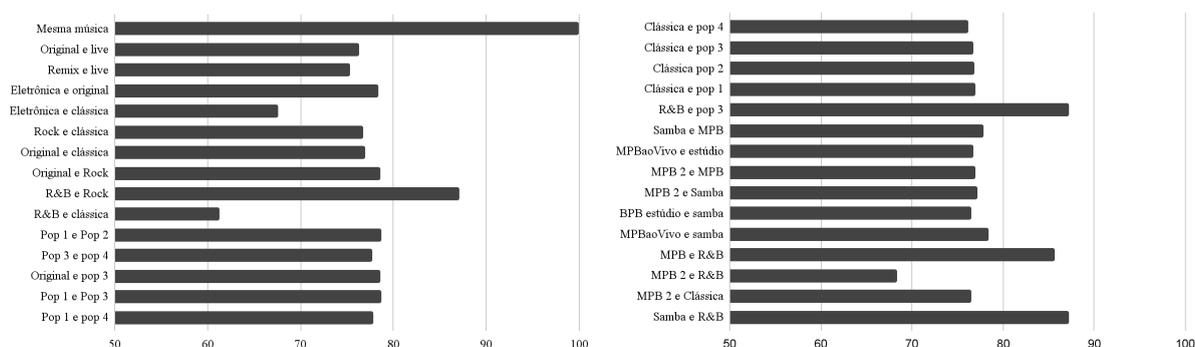


FIGURA 1. Resultados da comparação entre músicas em porcentagem.

Como é possível observar, em média encontrou-se o valor de cerca de 78%, apenas alguns se destacam positiva ou negativamente. Em alguns casos indo contra a percepção humana, o que indicou que deveriam ser feitos outros testes, agora fazendo tentativas em relação à percepção direta do autor deste trabalho, com o intuito de buscar padrões. Nesses testes, o autor deveria adicionar um valor percentual esperado (em sua opinião), e logo após rodar o algoritmo com os dois arquivos de áudio.

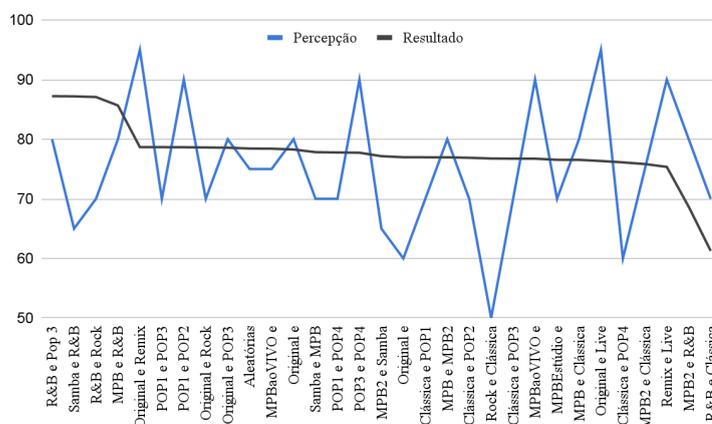


FIGURA 2. Testes realizados comparados com a percepção do autor (de 50% a 100%).

Da forma como foi ordenado, é perceptível que os resultados mantinham um padrão, próximos ao eixo de 76% a 78%, por conta disso, decidiu-se encontrar se isso não era apenas uma coincidência.

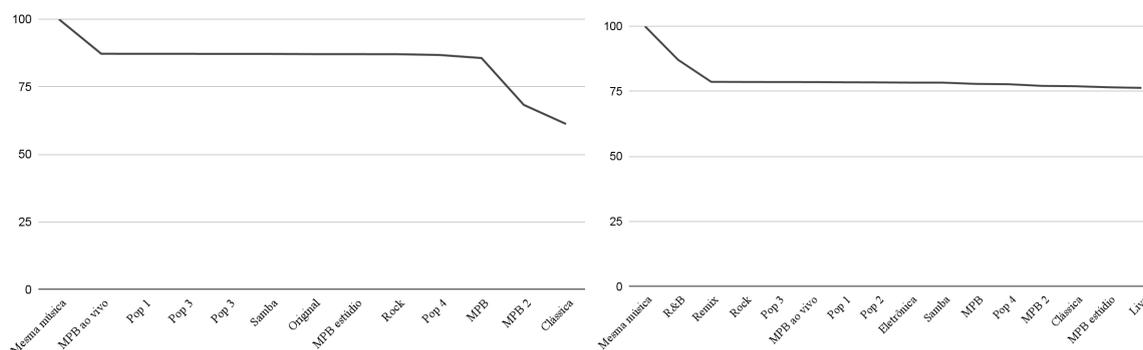


FIGURA 3. Testes realizados com apenas uma música sendo comparada com outras dentro do conjunto separado. Sendo o gráfico à esquerda a música base sendo a R&B e o gráfico à direita, com a música base Original (versão original da música utilizada).

É apresentado claramente que existe uma certa faixa de resultados onde independente de qual arquivo for comparado, o arquivo possui uma maior probabilidade de entregar resultados dentro dessa faixa, como pode ser visto na Figura 4. No caso da música R&B, é claro detectar que essa faixa está em torno de 86 a 89 por cento, enquanto na música Original a faixa está entre 76 e 78 por cento de similaridade. Isso impacta diretamente na suposição de que é possível realizar um sistema de recomendação com base nestes aspectos, já que a diferença entre a maior parte dos resultados das comparações indicariam uma distância muito sutil uma da outra, impossibilitando a estratégia.

É provável que a forma como são dispostos os dados dentro do arquivo com uma determinada extensão, que possuem um padrão, e que mesmo mudando a melodia e ritmo das músicas, esse padrão continua sendo seguido o que acaba resultando em similaridades maiores de 50% independente de quão diferentes músicas são diferentes de outras, lembrando o que foi dito anteriormente, que o valor que demonstra grande aleatoriedade nos dados seria o valor de 50% de similaridade.

Deslocamento de dados

Outra possibilidade que foi identificada e impacta nos resultados seria o deslocamento de dados, o qual apresenta um problema que pode existir em uma análise linear. Sendo dois vetores possuindo os trechos binários iguais, por exemplo 1010 no vetor A e 1010 no vetor B. Se no vetor A esse mesmo trecho seja deslocado, por exemplo, em um ou mais bits, o resultado será prejudicado, pois sofrerá alteração. No caso de arquivos de áudio, onde talvez um atraso de 1 segundo em uma música (uma versão mais lenta, por exemplo) poderia teoricamente modificar o resultado. Foram feitos testes saltando blocos do primeiro arquivo de áudio e comparando com o segundo arquivo. Os resultados não sofreram grandes modificações, seguindo a entrega da mesma faixa de similaridade mesmo após ignorar diversos blocos de dados do arquivo, como pode ser visto na figura a seguir, onde a similaridade baixa cerca de 2% mesmo após ignorar cerca de 5 mil blocos de 32 bits no primeiro arquivo. Portanto, isso indica que, existindo o problema do deslocamento, seria preciso uma análise mais profunda da sequência binária envolvida.

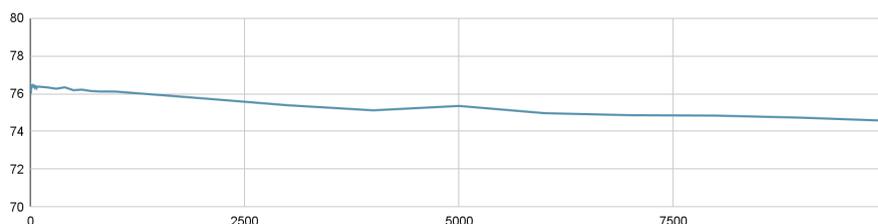


FIGURA 4. Testes realizados para comprovar o impacto do deslocamento de dados com porcentagem no eixo vertical (70% a 80%) e blocos ignorados no primeiro arquivo na horizontal.

CONCLUSÕES

Dados os resultados dos testes deste projeto, é possível concluir que, apesar de existir similaridade entre arquivos de áudio, porém, essa similaridade não se mostra suficiente para um sistema de recomendação. Dessa forma, a partir dos resultados obtidos neste projeto, não é possível afirmar que um algoritmo de recomendação poderia ser criado com base nesse tipo de análise. Sendo assim, como trabalho futuro, será explorada a geração de similaridade entre arquivos de áudios por meio da análise do seu espectrograma musical, uma imagem que representa a música através de cores em um plano cartesiano, como visto em COSTA (2013).

AGRADECIMENTOS

Omitido para não favorecer a identificação dos autores e do campus de origem do trabalho.

REFERÊNCIAS

COSTA, Yandre Maldonado e Gomes da. Reconhecimento de gêneros musicais utilizando espectrogramas com combinação de classificadores. 2013. Disponível em: <<https://acervodigital.ufpr.br/handle/1884/34886>>. Acesso em 08 set 2021.

GIMENES, Diego. Veja. Brasileiros passaram a consumir mais músicas por streaming de 2020. 7 dez 2020. Disponível em <<https://veja.abril.com.br/economia/brasileiros-passaram-a-consumir-mais-musicas-por-streaming-em-2020/>>. Acesso em 12 abr 2021.

TANAKA, A. Y.; BARBOSA, E. R.; KIMURA, R. S. Y. Análise de Qualidade de Áudio Objetiva e Subjetiva em vários formatos digitais. 2017. 144p. Trabalho de conclusão de curso (Engenharia elétrica). Universidade tecnológica federal do Paraná, câmpus Curitiba. Curitiba, 2017.

PANDEI, Harshita. Hamming Code In Computer Network. 2020. Disponível em: <<https://www.geeksforgeeks.org/hamming-code-in-computer-network/>>. Acesso em 29 abr 2021.

OLIVEIRA, Rodolpho Chrispim de. Aprenda sobre as Portas lógicas XOR. 2017. Disponível em: <<https://www.embarcados.com.br/aprenda-sobre-as-portas-logicas-xor/>>. Acesso em 16 ago 2021.

PAULA, Luciano Bernardes de. Utilização de funções LSH para busca conceitual baseada em ontologias. 2011. 54 de 134 p. Tese de doutorado em engenharia da computação - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Departamento de Computação e Automação Industrial, Campinas, 2011.

PEREIRA, Dimitri. Tecmundo. Quais são as diferenças entre os formatos de áudio WAV, MP3, AAC e FLAC? .08 ago 2017. Disponível em <<https://www.tecmundo.com.br/software/120043-diferencas-entre-formatos-audio-wav-mp3-aac-flac-htm>>. Acesso em 09 jan 2021.

ROMERO, Itanu. ICMusicBinary. 2020. Disponível em: <<https://github.com/ItanuRomero/ICMusicBinary>>. Acesso em 03/05/2021.

TEKNOMO, Kardi. Similarity Measurement. 2015. <<https://people.revoledu.com/kardi/tutorial/Similarity/HammingDistance.html>>. Acesso em 08 set 2021.