



Universidade Estadual de Maringá
Centro de Tecnologia
Departamento de Informática



**Uma Ferramenta para Exploração de
Relacionamentos de Artistas em Redes de Músicas**

William Rodrigues da Silva

TCC-2020



**Universidade Estadual de Maringá
Centro de Tecnologia
Departamento de Informática**



Uma Ferramenta para Exploração de Relacionamentos de Artistas em Redes de Músicas

William Rodrigues da Silva

TCC-2020

Trabalho de Conclusão de Curso apresentado à Universidade Estadual de Maringá, como parte dos requisitos necessários à obtenção do título de Bacharel em Informática.

Orientador: Prof. Dr. Marcos Aurelio Domingues

Maringá - Paraná
2021

WILLIAM RODRIGUES DA SILVA

**UMA FERRAMENTA PARA EXPLORAÇÃO DE RELACIONAMENTOS DE
ARTISTAS EM REDES DE MÚSICAS**

Trabalho de Conclusão de Curso apresentado
à Universidade Estadual de Maringá, como
parte dos requisitos necessários à obtenção do
título de Bacharel em Informática.

Prof. Dr. Marcos Aurelio Domingues

Orientador

Departamento de Informática - DIN, CTC

Prof. Dr. Daniel Kikuti

Departamento de Informática - DIN, CTC

Prof. Dr. Wagner Igarashi

Departamento de Informática - DIN, CTC

Maringá - Paraná
2021

Com gratidão, dedico este trabalho a Deus,
familiares, professores e amigos.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder a oportunidade e capacidade para a realização deste trabalho, cuidando de cada detalhe para que a conclusão deste fosse possível.

Aos familiares, especialmente, aos meus pais que sempre demonstraram dedicação, apoio e cuidado em toda minha trajetória. A minha irmã e cunhado, que também sempre estiveram dispostos a ajudar.

À Universidade Estadual de Maringá e todo seu corpo docente pela excelente qualidade oferecida no ensino.

Grato ao meu professor orientador Marcos Domingues, pela confiança depositada. Por todo o incentivo e pela sempre disposição em ajudar, ações que, de fato, serviram como inspiração para a realização deste trabalho. Sua colaboração foi de extrema importância neste processo.

Aos professores Daniel Kikuti e Wagner Igarashi, por aceitarem o convite para participação da banca examinadora.

À todos os amigos pelo companheirismo e apoio demonstrados. E pela colaboração na etapa de avaliação do trabalho.

RESUMO

A música é um dos principais elementos culturais de uma sociedade, sendo considerada como um fenômeno social, capaz de manter e ditar tradições em sociedades ao longo do tempo. Desta forma, a música não é apenas uma forma de entretenimento, mas está diretamente ligada com a formação de um indivíduo. Já economicamente, destaca-se o crescimento no uso de plataformas de *streaming*, responsáveis pelo crescimento no setor fonográfico. Estas plataformas se mantêm em contínua evolução, buscando melhorar a apresentação do conteúdo musical para os seus usuários. Neste contexto, este trabalho visa contribuir com o aprimoramento das plataformas de *streaming* por meio da proposta e desenvolvimento de uma ferramenta que apresente possíveis soluções para problemas e limitações ocorrentes em *players* de música, facilitando a interação do usuário com artistas e músicas, e viabilizando a busca de relações entre músicas e artistas de seu interesse. Essas relações são construídas e exibidas para o usuário a partir de uma rede de artistas e músicas criada pela ferramenta. O método utilizado para exibir relações de similaridade entre artistas, foi baseado na criação de grafos que representam tal similaridade, onde cada artista aparece ligado a um conjunto de artistas similares. Desta forma, a rede de artistas e músicas criada pela ferramenta pode ser explorada pelo usuário por meio dos grafos de cada artista, como também a partir de um sistema de recomendação de álbuns implementado na ferramenta. A ferramenta foi avaliada com usuários reais e os resultados mostraram a utilidade da mesma.

Palavras-chave: Recuperação de informações musicais. *Players* de música. Similaridade de artistas de música. Grafos. Redes de artistas e músicas. Sistemas de recomendação.

ABSTRACT

Music is one of the main cultural elements of society, being considered as a social phenomenon, able to maintain and dictate traditions in societies over time. In this way, music is not only a form of entertainment, but is directly linked to the formation of an individual. Economically, the growth in the use of streaming platforms stands out, being responsible for the growth in the phonographic sector. These platforms are constantly evolving, seeking to improve the presentation of musical content to its users. In this context, this work aims to contribute to the improvement of streaming platforms through the proposal and development of a tool that presents possible solutions to problems and limitations occurring in music players, facilitating user interaction with artists and songs, and, moreover, enabling search for relationships between songs and artists of interest. These relationships are built and displayed to the user from a network of artists and songs created by the tool. The method used to show similarity relationships between artists was based on the creation of graphs that represent such similarity, where each artist appears linked to a set of similar artists. In this way, the network of artists and songs created by the tool can be explored by the user through the graphs of each artist, as well as through an album recommendation system implemented in the tool. The tool was evaluated with real users and the results showed its usefulness.

Key-words: Music information retrieval. Music players. Similarity of music artists. Graphs. Networks of artists and music. Recommendation systems.

LISTA DE ILUSTRAÇÕES

Figura 1 – Dimensões da Pesquisa Musical	14
Figura 2 – Pontes de Königsberg	19
Figura 3 – Problema das pontes de Königsberg representado por um grafo	20
Figura 4 – Interface de exploração de discos (álbuns) no Musicream	29
Figura 5 – Grafo de músicas e player do sistema FM4 Soundpark	30
Figura 6 – Interface do sistema MusicRainbow	31
Figura 7 – Visão geral do sistema Musipath	33
Figura 8 – Fluxo de recuperação de informações de uma música	36
Figura 9 – Diagrama de classes do sistema Musipath	37
Figura 10 – Exemplos de interfaces do usuário em um dispositivo móvel	38
Figura 11 – Grafo representando a similaridade entre artistas	39
Figura 12 – Pseudocódigo para a função de criação do Grafo de um artista	40
Figura 13 – Grafo de um artista gerado pelo sistema Musipath	41
Figura 14 – Cálculo de similaridade de cosseno	42
Figura 15 – Organização de vetores para aplicação da similaridade de cosseno...	43
Figura 16 – Interface inicial do sistema Musipath	44
Figura 17 – Disposição das informações de um artista no sistema Musipath	45
Figura 18 – Interação e atualização do grafo de artistas do sistema Musipath	45
Figura 19 – Player e disposição das informações de um álbum	46
Figura 20 – Lista de álbuns recomendados	46
Figura 21 – Fluxograma de interações do sistema Musipath	47

LISTA DE TABELAS

Tabela 1 – Resultados da pesquisa de avaliação do sistema	48
---	----

SUMÁRIO

1. INTRODUÇÃO	11
1.1. Descrição da Problemática.....	11
1.2. Justificativa e Motivação.....	12
1.3. Objetivos.....	12
2. FUNDAMENTAÇÃO.....	14
2.1. Recuperação de Informações Musicais.....	14
2.2. Similaridade de Artistas de Música.....	17
2.3. Grafos.....	19
2.4. Sistemas de Recomendação.....	22
2.4.1. Definição	23
2.4.2. Principais técnicas utilizadas em sistemas de recomendação	25
3. TRABALHOS RELACIONADOS	29
4. DESENVOLVIMENTO	33
4.1. Pré-processamento dos Dados	34
4.2. API de Comunicação do Sistema <i>Musipath</i>	37
4.3. Interface com o Usuário.....	38
4.4. Criação do Grafo de Artistas	39
4.5. Sistema de Recomendação.....	41
5. AVALIAÇÃO E RESULTADOS	44
6. CONCLUSÃO E TRABALHOS FUTUROS.....	50
REFERÊNCIAS	51

1. INTRODUÇÃO

A música é um dos principais elementos culturais de uma sociedade. Gregory, A. H. (1997) considera a música um fenômeno social, sendo capaz de manter e ditar tradições em sociedades ao longo do tempo. Desta forma, a música não é apenas uma forma de entretenimento, mas está diretamente ligada com a formação de um indivíduo. De acordo com North, A. C. et al. (2000), por meio da música crianças mudam a forma de pensar, agir e estabelecer relações com amigos, colegas, modificando a sua rede social.

Já economicamente, destaca-se o crescimento no uso de plataformas de *streaming* como Spotify¹, Apple Music², Pandora³, dentre outras plataformas, responsáveis pelo crescimento no setor fonográfico. Segundo relatório publicado em 2019 pela Federação Internacional de Indústria Fonográfica (IFPI⁴ - *International Federation of the Phonographic Industry*), o comércio no setor de mídias digitais apresentou um aumento na receita de 21,1%, gerando um lucro de US \$11,2 bilhões. Já o mercado brasileiro, impulsionado pelo mercado mundial, também apresentou um crescimento de 15,4%, sendo então considerado o maior mercado na América do Sul e o 10º país em escala mundial no consumo de mídias digitais.

Mesmo com um crescimento no uso de plataformas de *streaming*, estas se mantêm em contínua evolução, buscando melhorar a apresentação do conteúdo musical para os seus usuários. Neste contexto, este trabalho visa contribuir com o aprimoramento das plataformas de *streaming* por meio da proposta e desenvolvimento de uma ferramenta que apresente possíveis soluções para problemas e limitações ocorrentes em *players* de música, facilitando a interação do usuário com artistas e músicas, e viabilizando a busca de relações entre músicas e artistas de seu interesse.

1.1. Descrição da Problemática

Os *players* de música podem ser vistos como um importante mecanismo de interação e exibição de artistas e músicas com usuários em uma plataforma de

¹ <https://www.spotify.com>

² <https://www.apple.com/br/music>

³ <https://www.pandora.com>

⁴ <http://www.ifpi.org>

streaming. Esse tipo de mecanismo apresenta aos usuários listagens de artistas e músicas que possibilitam a interação com estes. Contudo, nem sempre o conteúdo é apresentado ao usuário de forma intuitiva, o que prejudica também o interesse do usuário no uso da plataforma. Como exemplo, em um álbum de músicas, a identificação de artistas e músicas similares ao interesse do usuário pode não ser apresentada, ou não ser informado ao usuário de maneira clara os motivos para a existência de similaridade entre músicas e artistas. É importante ressaltar também que as plataformas de *streaming* em sua maioria são ofertadas estabelecendo uma taxa de uso, gerando assim um fator limitante para o acesso às mesmas.

1.2. Justificativa e Motivação

Tendo em vista o importante papel cultural da música em uma sociedade e, como já mencionado, a função de *players* de músicas como ferramentas para interação de usuários com músicas e seus artistas, este trabalho tem como motivação a contribuição no desenvolvimento desta área. Além disso, a forma com que *players* de música atuais exibem informações para seus usuários, apresenta problemas e limitações. Neste contexto, este trabalho se justifica na intenção de propor possíveis soluções para alguns problemas e limitações dos *players* de música.

1.3. Objetivos

Este trabalho tem como objetivo principal a recuperação de informações musicais e o desenvolvimento de um sistema para visualização e interação com redes de artistas e músicas, adicionando novas funcionalidades a um *player* de música. A representação da rede de artistas e músicas por meio de grafos é inspirada no conceito apresentado por Costa, D. et al. (2009).

Desta forma torna-se necessário o desenvolvimento de sistemas distintos, um que possa realizar a tarefa de recuperação de informações musicais através de métodos de *crawler*, e outro que possa, a partir destas informações coletadas, formar e exibir uma rede de artistas e músicas para o usuário final. A visualização desta rede será dada por meio de um grafo que relaciona artistas e seus similares. Além disso, será permitido ao usuário a visualização dos álbuns de músicas de cada artista, juntamente com um sistema de recomendação para álbuns similares ao interesse do

usuário. Este segundo sistema deverá conter também um *player* de música, que permita que a partir de uma música explorada na rede de músicas da ferramenta, esta possa ser ouvida pelo usuário através da reprodução de um áudio extraído do site Youtube⁵. Esta ferramenta visa oferecer de forma simples e gratuita a interação e descoberta de artistas e músicas do interesse do usuário.

Esta monografia está organizada da seguinte forma. Na Seção 1 foram apresentadas as motivações, justificativas e objetivos para este trabalho. Na Seção 2 é apresentada a fundamentação teórica, abordando os conceitos que envolveram o desenvolvimento do trabalho. Na seção 3 são apresentados alguns exemplos de trabalhos que já existem na literatura, relacionados com o tema deste trabalho, mencionando as principais diferenças e vantagens encontradas quando relacionados com este trabalho. Na Seção 4 é apresentado o processo de implementação do sistema proposto no trabalho e suas funcionalidades. Já na Seção 5, são apresentados os resultados obtidos com o desenvolvimento do trabalho, explicando o funcionamento do sistema final, bem como os resultados obtidos em uma avaliação realizada com usuários reais. Por fim, na Seção 6 são apresentadas as conclusões e trabalhos futuros.

⁵ <https://www.youtube.com>

2. FUNDAMENTAÇÃO

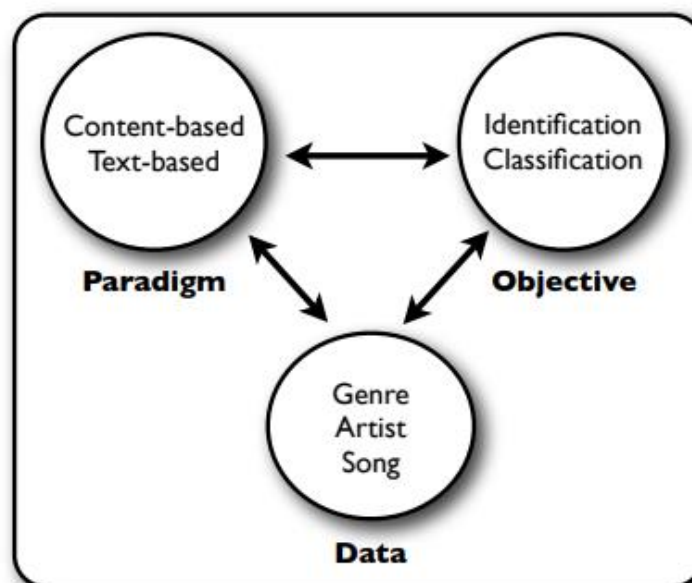
As seções a seguir apresentam conceitos técnicos e científicos existentes na literatura, e que foram adotados na fundamentação e desenvolvimento deste trabalho, partindo de uma área mais geral, com a apresentação da recuperação de informações musicais, e apresentando posteriormente áreas mais específicas, como a de similaridade de artistas de músicas, a de conceitos de grafos em sistemas computacionais e a de sistemas de recomendação.

2.1. Recuperação de Informações Musicais

A recuperação de informações musicais é uma nova e crescente área de pesquisa que envolve diversos campos do conhecimento, reunindo esforços para o desenvolvimento de técnicas de pesquisa baseadas em conteúdo, com objetivo de tornar mundialmente acessível uma ampla quantidade de músicas.

Marshall, B. (2016) apresenta a pesquisa musical como sendo a interseção de três dimensões: dimensão de paradigma, dimensão de objetivo e dimensão de dados, como exemplificado na Figura 1.

Figura 1 – Dimensões da Pesquisa Musical



Fonte: Marshall, B., Purdue University (2016).

A dimensão de paradigma descreve a música com base no seu conteúdo, como, a letra de uma música, seu estilo e rótulos associados. A dimensão de objetivo é responsável pela identificação e classificação de músicas através de métodos de aprendizagem de máquina. A dimensão de dados, por sua vez, representa a fonte de dados para pesquisas de informações musicais.

O crescente volume no compartilhamento de músicas na forma digital no século XXI, retrata um importante marco para a indústria do entretenimento, bem como para com o relacionamento desta com seus usuários. Diante de tal demanda, surgiu a necessidade do investimento no setor de pesquisa e desenvolvimento com ênfase no processamento e recuperação de informações musicais. Neste contexto, o desenvolvimento de técnicas de recomendação baseadas em conteúdo tornou-se muito importante, com o objetivo de proporcionar aos usuários uma melhor experiência no uso de sistemas através da sugestão de novos conteúdos musicais. Considera-se também que a categorização de áudios de acordo com suas propriedades é de grande importância para alcançar este objetivo, pois possibilita o agrupamento e a sugestão de músicas.

Existem diversas aplicações para recuperação de informações musicais baseadas em conteúdo nos âmbitos não só de sistemas de recomendação, mas também de reconhecimento de instrumentos, transcrição automática e categorização de músicas. Typke, R. et al. (2014) citam como exemplo:

- Consulta através do canto: Em lojas de música, ocorre com frequência a situação de um cliente conhecer apenas uma parte de uma música, mas não conhecer seu título, álbum ou até mesmo seu artista, dificultando no momento da venda. Nem sempre os vendedores possuem conhecimento sobre um vasto campo musical a ponto de conseguir identificar a música em questão apenas por uma melodia cantada pelo cliente, desta forma torna-se interessante o uso de uma ferramenta computacional que possa realizar essa identificação e a sugestão de álbuns;
- Busca de partituras semelhantes: Uma ferramenta de busca que consegue obter partituras musicais semelhantes a uma determinada partitura pode ajudar musicólogos no estudo e descoberta de influências existentes entre compositores. Esta ferramenta pode ajudar também músicos na busca de trabalhos semelhantes a trabalhos de outros artistas ou até mesmo de

trabalhos próprios anteriores. Esta tarefa foi realizada por musicólogos até então de forma manual, e se um sistema computacional pudesse executar essa tarefa, dados seriam obtidos com mais facilidade e rapidez, contribuindo com o desenvolvimento da área;

- Reconhecimento de problemas com direitos autorais: Um sistema computacional que pode reconhecer problemas de direitos autorais ajudaria na resolução e reconhecimento deste tipo de problema. Desta forma, artistas, empresas e outros sistemas poderiam reconhecer com mais facilidade questões de direitos autorais, verificando por exemplo, se alguém está os plagiando, ou até mesmo se trabalhos próprios correm o risco de acusação de plágio.

Além destas aplicações, podem-se citar aplicações de mecanismos de busca baseados em conteúdo, que funcionam especificamente com gravações de áudio:

- Identificação de uma música em reprodução: É possível identificar a música que está sendo reproduzida no rádio ou em algum tipo de alto-falante utilizando apenas um celular, através de uma ferramenta que habilita o microfone do celular para realizar a gravação de um trecho daquela faixa em reprodução e identificar então, a música exata que está sendo tocada;
- Busca de vídeo através de áudio: A busca de um vídeo baseada em conteúdo pode ser feita com mais eficiência analisando o conteúdo de áudio;
- Reconhecimento de sons suspeitos: Equipamentos de segurança e vigilância podem através de sistemas computacionais reconhecer e buscar sons suspeitos em suas gravações;
- Busca em biblioteca de áudio: Estações de rádio, cinemas, e produtores de conteúdo em geral podem utilizar uma ferramenta com um mecanismo de pesquisa que encontra efeitos sonoros e demais faixas de áudio que correspondam a uma determinada busca por descrição.

Neste trabalho, a primeira etapa de desenvolvimento baseou-se na recuperação de informações musicais, com o objetivo de preencher a base de dados do sistema com informações sobre artistas de música, seus álbuns e músicas. Com o

levantamento de informações musicais para a base de dados, tornou-se possível a construção de uma rede de artistas e músicas utilizada pela ferramenta desenvolvida.

2.2. Similaridade de Artistas de Música

O consumo de mídias no formato digital tem crescido cada vez mais nos últimos anos. Para a indústria musical isso não foi diferente, pois mídias físicas e estações de rádio AM/FM, que eram meios de acesso à música anteriormente, por exemplo, vêm sendo substituídas cada vez mais por plataformas digitais de *streaming*.

Conforme citado na seção anterior, o investimento na área de pesquisa com ênfase na recuperação de informações musicais se tornou necessário, tendo como objetivo viabilizar o acesso a uma vasta quantidade de informações sobre músicas e seus artistas.

Entretanto, ainda existem desafios que envolvem o uso dessas informações, uma vez que plataformas de *streaming* necessitam desenvolver sistemas de recomendação capazes de encontrar similaridade entre artistas de músicas. A similaridade de artistas de música pode ser de grande utilidade para a geração de recomendações que reflitam os interesses de um usuário.

Segundo Marshal, B. (2016), artistas de música podem comumente serem classificados com mais de um gênero musical, isso se deve ao fato de que artistas podem ter inspirações em mais de um gênero, e de igual forma podem produzir diferentes conteúdos, baseados em estilos musicais distintos. Isso torna-se um desafio quando há a necessidade de encontrar similaridade entre artistas de música. Outro desafio neste âmbito encontra-se no reconhecimento de similaridade de artistas para obras realizadas através de colaborações. As colaborações tornaram-se populares, e baseiam-se na apresentação de uma música com a participação de mais de um artista ou grupo musical. Ocorre que nem sempre um usuário tem interesse em todos os artistas envolvidos, mas por vezes apresenta interesse em apenas um dos artistas envolvidos na colaboração.

Desta forma, a tarefa de encontrar similaridade entre artistas de músicas deve ser realizada em um trabalho conjunto orientado por usuários, tendo em vista que esta similaridade pode ser subjetiva.

Allik, A. et al. (2018) citam alguns conjuntos de dados e serviços que dispõe informações musicais e informações sobre similaridade de artistas de música de acesso público:

- MusicBrainz⁶: Uma enciclopédia de metadados de publicações musicais construída a partir de contribuições de usuários e aberta ao público;
- Dbpedia⁷: Um serviço realizado através do esforço em comunidade para extração de conteúdos estruturados de projetos da Wikimedia. Pode ser utilizado inclusive para consulta de categorias de artistas de música;
- AcousticBrainz⁸: Uma fonte de dados também construída através do esforço conjunto da comunidade para a disposição de informações musicais. Inclui informações musicais como descrições de gêneros, humor, tons, escalas, dentre outras;
- Million Song Dataset⁹: Uma coleção de recursos de áudio e metadados criada com objetivo de incentivar e colaborar com a pesquisa sobre informações musicais;
- Last.FM: Uma rede social musical que coleta informações musicais sobre seus usuários e gostos musicais, disponibilizando dados sobre artistas, álbuns, músicas e similaridade entre artistas. Essa similaridade é extraída a partir dos gostos do usuário e de *tags* que os próprios usuários preenchem para cada artista.

O conceito de similaridade de artistas e a recuperação de informações neste sentido, foram utilizadas neste trabalho para a implementação de funcionalidades que permitem ao usuário a exploração de uma rede de artistas similares que representam seu estilo musical.

⁶ <https://musicbrainz.org>

⁷ <https://dbpedia.org>

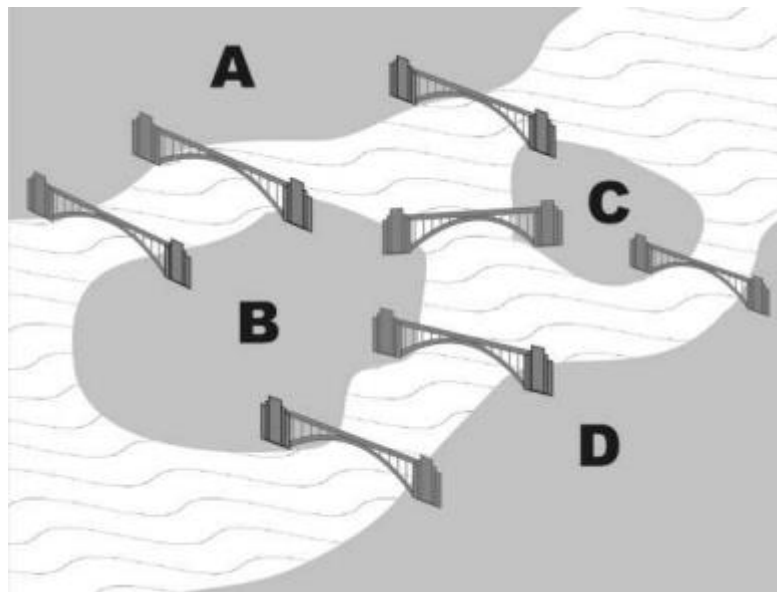
⁸ <https://acousticbrainz.org>

⁹ <https://labrosa.ee.columbia.edu/millionsong>

2.3. Grafos

Historicamente, a origem da teoria de grafos está associada com o problema das pontes de Königsberg, apresentado pelo matemático Leonhard Euler em 1773 no artigo que é considerado como o primeiro artigo relacionado a este tema. O problema descrevia como cenário uma cidade que possuía um rio, contendo duas ilhas neste, e sete pontes que interligavam as ilhas às margens do rio, e interligavam também uma ilha a outra. A Figura 2 ilustra o cenário descrito no problema. O desafio proposto no problema por Euler era realizar um trajeto pela cidade, seguindo as seguintes restrições: O trajeto deveria iniciar e terminar no mesmo ponto, e toda ponte deveria ser incluída no trajeto exatamente uma vez.

Figura 2 – Pontes de Königsberg



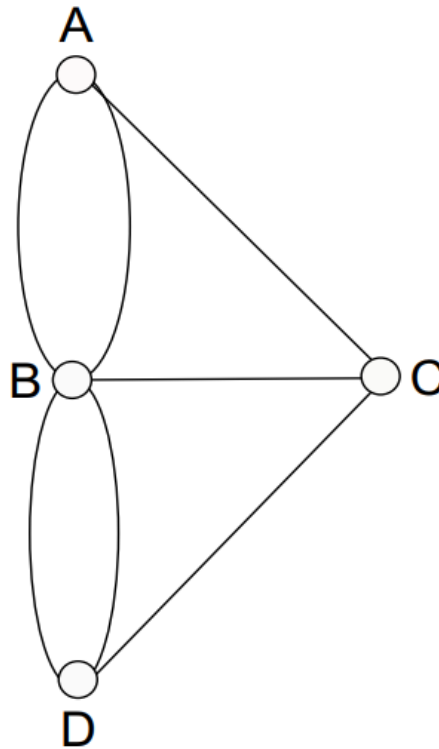
Fonte: Teoria dos grafos e Aplicações, Universidade de Aveiro, 2004.

A modelagem do problema pode ser feita a partir de um grafo onde cada porção de terra do cenário apresentado é representado por um ponto, e as pontes representadas por linhas, como exposto na Figura 3. Euler resolveu o problema e estipulou ainda a solução para um caso mais geral.

Com a contribuição de Euler na resolução do problema, posteriormente diversos outros importantes problemas foram modelados. O desenvolvimento da pesquisa sobre a teoria de grafos auxiliou a modelagem e resolução de problemas interdisciplinares como expostos por Roberts, F. S. (1984) e Netto, P. O. B. (1979) por

exemplo, a estruturação e organização de serviços em grandes centros urbanos como circuitos elétricos, distribuição de serviços como coleta de lixo e entrega de correspondências, rotas de estradas e a distribuição de água e luz. Outra possibilidade é a modelagem de relações, como de trabalho ou amizade.

Figura 3 – Problema das pontes de Königsberg representado por um grafo



Fonte: Elaborada pelo autor, 2021.

Bollobás, B. (1998) define formalmente um grafo da seguinte maneira: Um grafo G é definido por um par de conjuntos disjuntos (V, E) , onde V é um conjunto finito não vazio de vértices (ou nós) e E é um conjunto finito de arestas. Sendo G um grafo, então $V = V(G)$ é o conjunto de vértices de G , e $E = E(G)$ é o conjunto de arestas de G . Diz-se que uma aresta $\{x, y\}$, liga os vértices x e y , e é denotada por xy . Desta forma, xy e yx significam exatamente a mesma aresta. Se $xy \in E(G)$, então x e y são considerados vértices adjacentes, ou vizinhos, de G . Duas arestas são consideradas adjacentes se tiverem o mesmo vértice final em comum.

Na área da ciência da computação, os grafos são considerados como uma excelente ferramenta, utilizada na modelagem de diversos tipos de relações. Assim, eles podem ser úteis na representação de diferentes problemas do mundo real.

Singh, R. P. e Vandana (2014) citam alguns campos de aplicação da teoria de grafos na área da ciência da computação:

- Engenharia de software: Grafos podem auxiliar no processo de captura de requisitos, projeto de sistemas e descoberta de relação entre módulos de um sistema;
- Projetos de redes: Em projetos de redes a teoria de grafos é muito útil na busca e arranjo de nós da rede, encontrar o menor caminho em uma rede, analisar o tráfego na rede para segurança, modelagem de comunicação, dentre outras aplicações;
- Estrutura de dados: Os grafos oferecem uma eficiente organização de dados e otimização do tempo de recuperação de dados. Como exemplo, tem-se a aplicação de busca de árvores de custo mínimo;
- Processamento de imagens: Técnicas de processamento de imagens podem ser otimizadas utilizando a teoria de grafos, assim como o cálculo de alinhamento de imagens e a definição de transformações de distância utilizando algoritmos de caminho mínimo;
- Sistemas operacionais: Problemas como a alocação de recursos em sistemas operacionais podem ser resolvidos com a ajuda de grafos. Problemas como a distribuição de tarefas podem também ser auxiliados com o conceito de coloração de grafos, por exemplo;
- Projetos de websites: Websites podem ser projetados a partir da modelagem de grafos, tendo cada página representada por um vértice, e os links entre páginas representados pelas arestas.

Neste trabalho, o conceito de teoria de grafos foi aplicado na representação da relação existente na similaridade entre artistas de música. Essa relação pode ser representada através de um grafo $G(V, E)$ simples, isto é, um grafo que não possui laços (arestas com os dois extremos ligados a um mesmo vértice) nem arestas paralelas (arestas diferentes, mas incidentes aos mesmos dois vértices), onde o conjunto de vértices V representa artistas de música, e o conjunto de arestas E representa a relação de similaridade entre tais artistas.

2.4. Sistemas de Recomendação

De acordo com Adomavicius e Tuzhilin (2005), o surgimento de sistemas de recomendação como uma área de pesquisa independente ocorreu nos anos 90, em razão da necessidade que os pesquisadores encontraram na resolução de problemas de recomendação que dependiam expressamente de sua estrutura de classificação. Devido a sua importância, esta área vem sendo muito estudada e aplicada em seu contexto. Conforme Schafer, Konstan e Riedl (2001), atualmente já existem diversos sites que utilizam sistemas de recomendação, como sites de *e-commerce*. Segundo Sundermann (2015), um dos sites mais citados na literatura é o *e-commerce* amazon.com. Navegando pelo site, quando o cliente clica sobre um produto à venda, são exibidas as informações sobre aquele produto específico, e são exibidas também, recomendações como “Os clientes também visualizaram estes produtos” e “Produtos relacionados a esse item” seguida de uma lista de itens correspondentes a estas recomendações.

De acordo com Ricci, Rokach e Shapira (2011), existem importantes motivos que justificam uma demanda cada vez maior de aplicações que usam sistemas de recomendação:

- Aumento no número de itens vendidos: Podendo ser considerada como uma das funções mais importantes para sistemas comerciais, com o uso de sistemas de recomendação, conjuntos adicionais de itens podem ser vendidos quando comparados com aqueles vendidos normalmente sem a aplicação de qualquer tipo de recomendação. Isso se justifica pelo fato de que o usuário pode encontrar com mais facilidade, produtos que ele gosta e necessita. Esta função pode ser útil de igual forma até mesmo para sistemas não comerciais, mesmo não havendo um custo, os objetivos podem ser semelhantes quanto a seleção de itens por parte do usuário. Pode-se dizer que em geral o objetivo de um provedor de serviços é aumentar sua taxa de aceitação, ou seja, o número de usuários que se interessam por uma recomendação e consomem aquele conteúdo;
- Maior diversidade nas vendas: Outra importante função de sistemas de recomendação é prover ao usuário opções de itens que dificilmente seriam encontrados sem a utilização de recomendações. Isso ocorre pois nem sempre

as métricas de recomendação são baseadas apenas na popularidade de um item, mas estão também voltadas para os interesses do usuário;

- Aumento na satisfação do usuário: A aplicação de sistemas de recomendação fornece ao usuário uma melhor experiência no uso do sistema. O usuário poderá encontrar recomendações interessantes e relevantes que somadas a um bom projeto de interface, aumentam a avaliação subjetiva deste. Consequentemente a satisfação do usuário aumenta, junto com a probabilidade de que as recomendações sejam aceitas;
- Aumento na fidelidade dos usuários: Um usuário tende a se tornar fiel a um sistema que tem a capacidade de reconhecê-lo através de interações feitas anteriormente, como avaliações feitas sobre itens no sistema. Desta forma o sistema consegue representar melhor os interesses do usuário, gerando um sentimento de importância para este. Esta relação tende a tornar-se cada vez mais expressiva, pois, quanto maior o tempo de um usuário no sistema, maior será o conhecimento adquirido sobre os interesses deste;
- Melhor entendimento sobre os interesses do usuário: Outra funcionalidade dos sistemas de recomendação encontra-se na descrição e entendimento dos interesses do usuário, que podem ser coletadas explícita ou implicitamente pelo sistema. Desta forma o provedor de serviços consegue ter uma melhor gestão sobre seus objetivos com base neste conhecimento. Por exemplo, uma loja que consegue gerir com mais precisão sua produção e gestão de estoque, derivada de uma análise de dados apurados por um sistema de recomendação.

2.4.1. Definição

Conforme Sundermann (2015), sistemas de recomendação são ferramentas utilizadas com o intuito de encontrar e fornecer ao usuário itens de seu interesse, dentro de um conjunto de possíveis itens. Ricci, Rokach e Shapira (2011) ainda definem sistemas de recomendação como uma tecnologia empregada no filtro de informações, que podem ser utilizadas para recomendação de itens através da criação de um *ranking*, e ainda prever de acordo com os interesses do usuário, avaliações de itens.

Sistemas de recomendação tipicamente expressam sua funcionalidade de recomendação fundamentados na relação existente entre usuário e item, onde o

usuário corresponde ao alvo de uma recomendação e o item compreende as opções que podem ser sugeridas para este, sendo assim definida a recomendação como uma lista ordenada de itens que correspondem ao interesse do usuário (RICCI; ROKACH; SHAPIRA, 2011; BOBADILLA et al., 2013).

Adomavicius e Tuzhilin (2005) descrevem que o problema de recomendação pode ser reduzido ao problema de classificar qual a relevância de um item que ainda não foi exposto ao usuário. Indicam ainda que essa classificação deve ser estimada com base em classificações fornecidas pelo próprio usuário para outros itens similares. Deste modo, podendo-se estimar classificações para itens ainda não expostos aos usuários, pode-se também realizar recomendações com base nestas classificações, recomendando itens com as melhores médias. O problema de recomendação pode ser descrito formalmente como:

“Seja U o conjunto de todos os usuários e I o conjunto de todos os itens que podem ser recomendados. Seja r a função de utilidade que mede o quão útil um item i é para o usuário u , isto é, $r : U \times I \rightarrow R$, em que R é um conjunto ordenado, por exemplo, inteiros não negativos ou números reais dentro de um determinado intervalo. Então para cada usuário $u \in U$, o objetivo é encontrar um item $i \in I$ que maximiza a utilidade do usuário, isto é, que seja mais interessante a ele” (ADOMAVICIUS; TUZHILIN, 2005 apud SUNDERMANN, 2015, p. 12):

$$\forall u \in U, i'_u = \arg \max_{i \in I} r(u, i)$$

De acordo com Sundermann (2015), a função de utilidade calcula o interesse de um usuário sobre um item específico, tal cálculo é feito baseado em avaliações realizadas de forma implícita ou explícita.

Segundo Schafer et al. (2007), a forma de avaliação explícita ocorre quando a avaliação é solicitada ao usuário e registrada no sistema, podendo ser classificada como:

3. Numérica: Quando o usuário confere aos itens valores numéricos, dentro de um intervalo definido, correspondente a sua avaliação;
4. Ordinal: Quando o usuário seleciona um termo, dentre termos predefinidos, que mais se aproxima e representa sua avaliação sobre determinado item;

5. Binária: Avaliação obtida através de uma pergunta para o usuário, que responde se determinado item é considerado útil ou não;
6. Unária: Quando o usuário informa explicitamente suas áreas de interesse.

Por outro lado, na forma implícita, Claypool et al. (2001) explica que a avaliação de interesse do usuário ocorre baseada no comportamento do mesmo durante a utilização do sistema. Ricci, Rokach e Shapira (2011) inteiram ainda com um exemplo de avaliação implícita, quando o usuário seleciona determinado item, e o sistema pode então inferir o interesse do usuário por este item. Tal avaliação pode ser ainda mais precisa pela observação do tempo de retenção do usuário, isto é, o tempo de permanência na visualização ou uso do item.

2.4.2. Principais técnicas utilizadas em sistemas de recomendação

Os sistemas de recomendação são descritos na literatura em três subclasses, de acordo com a abordagem de seus algoritmos: filtragem colaborativa, filtragem baseada em conteúdo e abordagens híbridas (ADOMAVICIUS et al., 2005; BOBADILLA et al., 2013; PAZZANI, 1999; RICCI; ROKACH; SHAPIRA, 2011).

Pazzani (1999) explica sobre a filtragem colaborativa, que busca encontrar correlações entre os usuários de um sistema. O objetivo dessa abordagem é permitir que o sistema de recomendação sugira itens para um usuário em específico, com base nos itens avaliados anteriormente por outros usuários que apresentam correlação com o usuário alvo. Para Adomavicius e Tuzhilin (2005), este método é formalmente descrito como:

A utilidade $r(u, i)$ de um item i para o usuário u é estimado com base nas utilidades $r(u_j, i)$ que foram atribuídas ao item i por outros usuários $u_j \in U$ que são “semelhantes” ao usuário u .

Breese, Heckerman e Kadie (1998), Desrosiers e Karypis (2011) e Bobadilla et al. (2013) classificam ainda a abordagem de filtragem colaborativa em dois grupos:

- Baseado em memória: Esse grupo de algoritmos utiliza heurísticas para realizar recomendações e estimar avaliações, tendo como base um conjunto de avaliações de usuários realizadas anteriormente;
- Baseado em modelo: Esse grupo de algoritmos utiliza as avaliações realizadas anteriormente por usuários de forma indireta, tendo em vista que nos algoritmos baseados em modelo essas informações são utilizadas inicialmente para a construção de um modelo, que é então utilizado na previsão de avaliações. Esse modelo é construído através de técnicas de aprendizado de máquina, por exemplo: Sistemas fuzzy, algoritmos genéticos, classificadores bayesianos e redes neurais.

Por outro lado, a filtragem baseada em conteúdo não utiliza dados de correlação entre usuários, mas considera apenas o perfil do usuário para gerar recomendações, tendo como base as avaliações que o usuário fez anteriormente sobre outros itens (BALABANOVIĆ; SHOHAM, 1997).

Segundo Bobadilla et al. (2013), a filtragem baseada em conteúdo utiliza uma comparação entre o conteúdo dos itens que serão recomendados, e o conteúdo de itens já avaliados positivamente pelo usuário para gerar recomendações. Para Adomavicius e Tuzhilin (2005), uma descrição formal da filtragem baseada em conteúdo pode ser feita como:

A utilidade $r(u, i)$ de um item i para o usuário u é calculada com base nas utilidades $r(u, ij)$ que foram atribuídas pelo usuário u nos itens $ij \in I$ que são “semelhantes” ao item i .

Entretanto, algumas limitações são notadas e comentadas na literatura a respeito das duas abordagens citadas. Para Adomavicius e Tuzhilin (2005), Balabanović e Shoham (1997), Bobadilla et al. (2013), Pazzani (1999), sistemas que utilizam a abordagem de filtragem colaborativa por exemplo, mesmo conseguindo gerar boas recomendações, apresentam um problema nomeado como “cold-start”. Trata-se de um problema que ocorre por razão de uma quantidade diminuta de informações. A seguir são citadas três situações que comumente ocasionam este problema:

- Item novo: Esta situação ocorre na inserção de um novo item no sistema, como este inicialmente não possui avaliações, o sistema pode não o recomendar, tendo em vista que a recomendação é feita baseada nas avaliações do item;
- Usuário novo: Na abordagem de filtragem colaborativa o sistema utiliza correlações entre usuários com base em suas preferências, entretanto o perfil de um usuário novo não é bem definido para o sistema nesse sentido, uma vez que este usuário ainda não realizou avaliações;
- Comunidade nova: Atrelada às duas primeiras situações, esta ocorre quando o sistema de recomendação é novo e ainda não possui informações suficientes para gerar recomendações aceitáveis.

Já para sistemas de recomendação que utilizam a filtragem baseada em conteúdo, podem ser citados dois principais problemas:

- Superespecialização: Como a filtragem baseada em conteúdo também realiza recomendações a partir da análise do perfil do usuário, pode ocorrer uma tendência nas recomendações, tornando-se limitadas apenas a itens similares;
- Análise limitada do conteúdo: Este problema provém da dificuldade de se obter informações seguras para recomendações sobre os itens de forma automática.

Tendo em vista as limitações apresentadas nas duas primeiras abordagens, a abordagem híbrida é exposta como uma possível alternativa, sendo uma combinação da abordagem de filtragem colaborativa com a baseada em conteúdo. De acordo com Adomavicius e Tuzhilin (2005), existem diferentes maneiras de combinar essas abordagens, como a aplicação das duas abordagens separadamente, mas utilizando a combinação de suas previsões para as recomendações, incorporando características de uma abordagem na outra, ou até mesmo criando um modelo que unifica e incorpora as duas abordagens.

Alguns conceitos de sistemas de recomendação foram aplicados na implementação de funcionalidades neste trabalho, tendo como objetivo a produção de recomendações de álbuns para os usuários do sistema proposto. A recomendação de álbuns foi implementada de forma simplificada, utilizando conceitos da recomendação baseada em conteúdo, levando em consideração os rótulos de cada álbum para

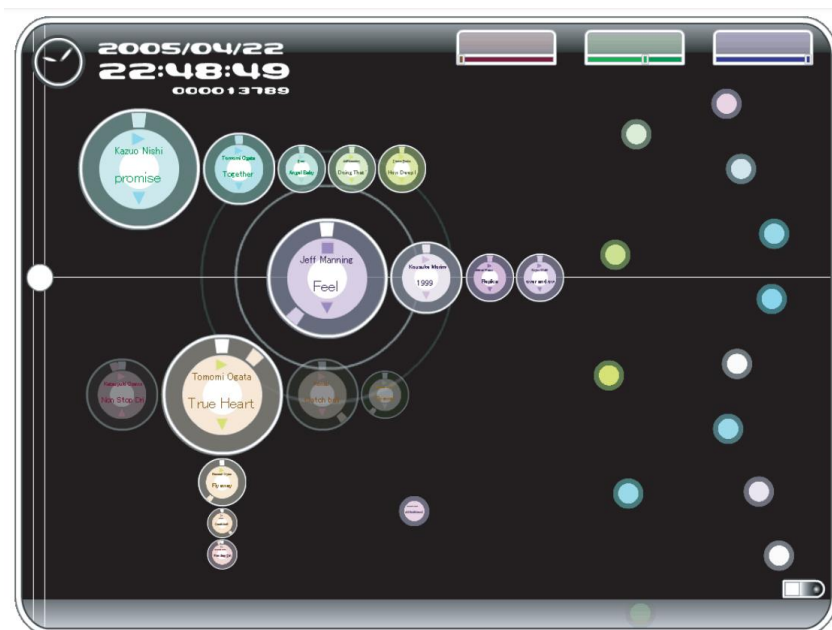
realizar a comparação de similaridade entre estes. A Subseção 4.5 descreve com mais detalhes o algoritmo implementado para a geração de recomendações de álbuns.

3. TRABALHOS RELACIONADOS

Com o desenvolvimento da área de recuperação de informações musicais, coleções de dados sobre música tornaram-se cada vez maiores com o passar dos anos. Neste contexto, verificou-se a necessidade do desenvolvimento de sistemas que fossem capazes de selecionar e exibir conjuntos de dados musicais para o usuário de maneira coerente. Nesta seção são apresentados alguns exemplos de trabalhos relacionados, encontrados na literatura, desenvolvidos neste intuito, e por fim, são discutidas as diferenças observadas com relação à proposta do presente trabalho.

O *Musicream* é um sistema proposto por Goto, M. e Goto, T. (2009), que tem como fundamento a proposta de uma interface de exibição e reprodução de músicas incomum, pois tenta desfazer o estereótipo de interfaces que exibem listas compostas por títulos de músicas e seus artistas. A ideia é proporcionar ao usuário uma experiência de descoberta de álbuns de música de forma inesperada, por meio de *playlists* criadas de acordo com o gosto do usuário. Conforme exposto na Figura 4, os discos na interface representam os álbuns de música, que são reproduzidos de forma sequencial. Através de uma busca inicial por um álbum, o sistema oferece outras opções de álbuns similares que o usuário pode anexar à reprodução.

Figura 4 – Interface de exploração de discos (álbuns) no *Musicream*



Fonte: Goto, M.; Goto, T. Musicream, 2009.

Outro exemplo é o sistema *FM4 Soundpark Player* proposto por Gasser, M. e Flexer, A. (2009), que por outro lado, utiliza como chave inicial de busca uma música específica inserida pelo usuário. Partindo desta música inicial, o sistema cria um grafo, como exposto na Figura 5, que representa sugestões de músicas similares, sugerindo até cinco músicas que coincidem com o conteúdo da música buscada. O usuário tem então a possibilidade de interagir com o grafo criado, e sempre que um dos vértices do grafo é selecionado, representando determinada música, o grafo é atualizado, determinando o vértice selecionado como central e realizando novas sugestões, partindo deste vértice central. Desta maneira o usuário consegue interagir com sugestões musicais e ouvir cada música através de um *player* de música acoplado no sistema.

Figura 5 – Grafo de músicas e *player* do sistema *FM4 Soundpark*

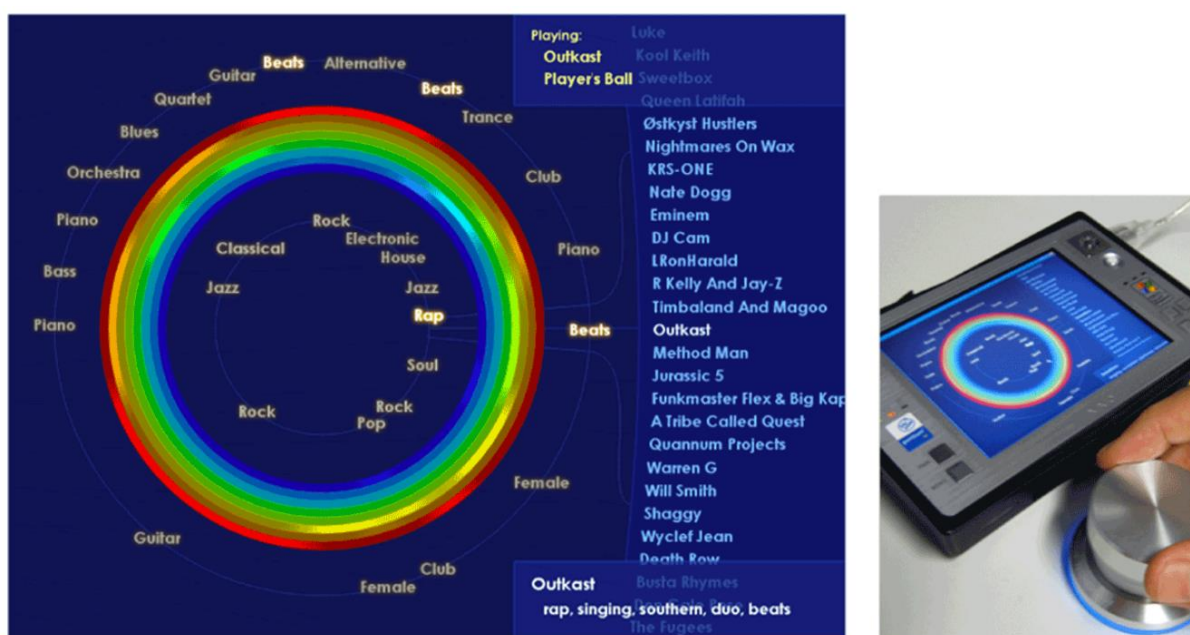


Fonte: Gasser, M.; Flexer, A. FM4 Soundpark, 2009.

Outra abordagem de interface é apresentada no sistema *MusicRainbow* proposto por Pampalk, E. e Goto, M. (2006). O sistema tem como objetivo proporcionar uma experiência de descoberta de novos artistas para o usuário. A disposição dos artistas é mapeada através de uma interface que segue a ideia de um

arco-íris em formato circular, onde cada cor do arco-íris representa um estilo de música diferente. O sistema conta com um pré-processamento de dados que separa os artistas de acordo com seus rótulos, extraídos de páginas web. Desta forma o sistema pode resumir e dividir artistas em subgrupos de acordo com os rótulos mais representativos, como é exposto na Figura 6. A interação do usuário é feita de uma forma simplificada, onde o usuário pode rotacionar o arco-íris e selecionar os rótulos de sua preferência, obtendo uma lista de artistas correspondentes.

Figura 6 – Interface do sistema *MusicRainbow*



Fonte: Pampalk, E.; Goto, M. MusicRainbow, 2006.

Em síntese, podem-se descrever os sistemas propostos nestes trabalhos, a partir dos métodos utilizados na geração de conteúdo para o usuário. No primeiro sistema citado (*Musiccream*), a proposta de busca de informações musicais é dada a partir de um álbum inicial informado pelo usuário, e o sistema gera um fluxo de conteúdo sugerindo novos álbuns similares. Já no segundo sistema (*FM4 Soundpark*), a busca inicial é realizada a partir de uma música informada pelo usuário, e o conteúdo gerado é baseado na sugestão de músicas similares. Por fim, no terceiro sistema (*MusicRainbow*), a recomendação de artistas é realizada a partir da comparação dos rótulos de cada um.

O sistema desenvolvido neste trabalho, tem um objetivo análogo de permitir ao usuário uma experiência de interação e descoberta de novos artistas, álbuns e

músicas. Entretanto, a interface proposta não restringe a exploração do conteúdo apenas a álbuns ou apenas a artistas, mas propõe uma exploração que pode ser feita por ambos os caminhos. O usuário pode interagir e descobrir novos artistas a partir de um grafo interativo que representa a conexão entre artistas similares, mas também consegue visualizar os álbuns de cada artista, recebendo recomendações de álbuns similares que pertencem a outros artistas. Desta forma, o usuário tem mais possibilidades na exploração da rede de artistas e músicas. O sistema desenvolvido também propõe uma interface de interação simplificada e intuitiva, que permita uma experiência de interação e descoberta mais fluida.

4.1. Pré-processamento dos Dados

O pré-processamento de dados compreendeu a etapa inicial de desenvolvimento do sistema. Nesta seção será apresentada a implementação do subsistema *Musipath Crawler*, que tem como principal objetivo a recuperação de informações musicais para construção da base de dados do sistema *Musipath*.

A necessidade de recuperação de informações musicais neste contexto pode ser descrita especificamente como a obtenção de dados sobre artistas de música, álbuns, músicas e existência de similaridade entre artistas.

O sistema foi desenvolvido sob a linguagem de programação Java, aplicando o conceito de orientação a objetos, e seus métodos para recuperação de informações musicais foram separados em duas classes principais, de acordo com a fonte da qual as informações são extraídas, nomeadas como *CrawlerAPI* e *CrawlerWeb*.

Para explicar o funcionamento da classe *CrawlerAPI* é preciso entender primeiramente o conceito de *Application Programming Interface* (API), que pode ser definido como um conjunto de rotinas implementadas por determinado sistema, com o objetivo de oferecer funcionalidades a serem utilizadas por outros sistemas. Desta forma uma API permite a comunicação e integração entre sistemas, sem haver a necessidade de que um sistema conheça a implementação do outro.

A classe *CrawlerAPI* tem como fonte de extração de dados a API disponibilizada pelo site Last.fm¹⁰, uma plataforma musical colaborativa, que permite que seus usuários classifiquem artistas, álbuns e músicas através de rótulos, utilizando essas informações para definir dimensões semânticas no contexto musical. A API disponibilizada pelo site Last.fm oferece métodos de recuperação de informação musical, por exemplo, busca por artistas, álbuns e músicas. Um exemplo de informação recuperada por meio desta API são os relacionamentos de similaridade entre artistas. Uma vez realizada a requisição para obter informações de um determinado artista, tem-se como resposta uma lista de artistas similares a este, uma informação de grande importância na construção da base de dados do sistema *Musipath*. Desta forma, a classe *CrawlerAPI* pode realizar requisições à API do site Last.fm, que retorna respostas no formato JSON, contendo informações úteis para construção dos objetos que representam artistas, álbuns e músicas no sistema.

¹⁰ <https://www.last.fm>

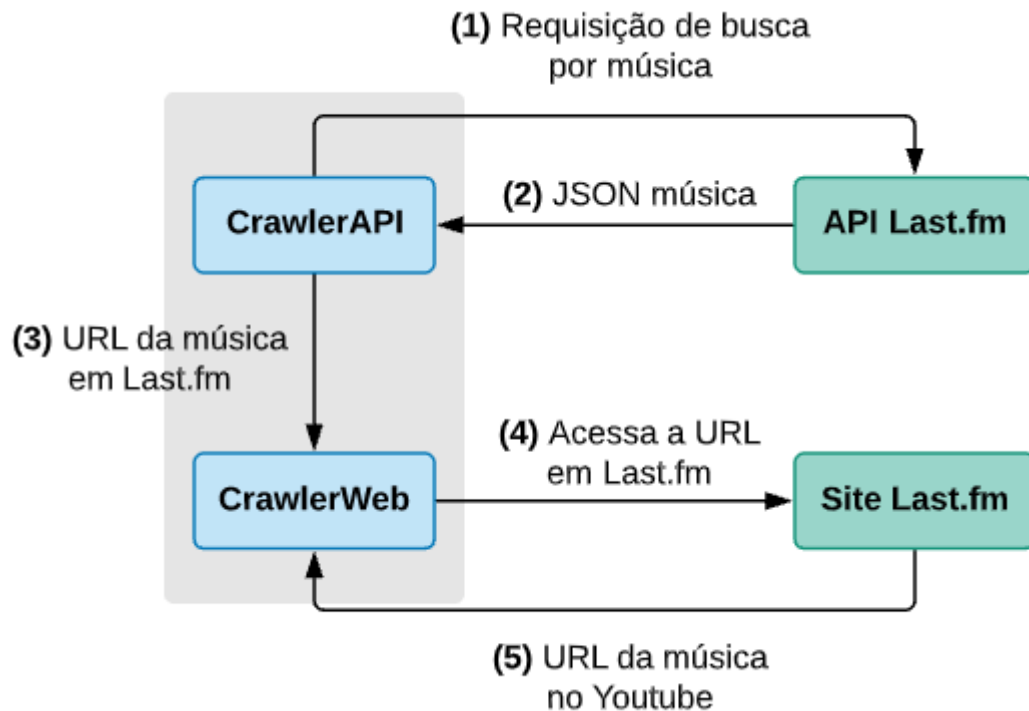
Entretanto, nem todas as informações necessárias para criação destes objetos puderam ser obtidas utilizando a API do site Last.fm, como a URL de uma música no site do Youtube, a imagem de capa de um álbum, dentre outras. Neste contexto, a criação de métodos para obtenção dessas informações se tornou necessária.

A classe *CrawlerWeb* encapsula os métodos implementados com a finalidade de obter o restante das informações necessárias, porém ausentes na recuperação de informações da classe *CrawlerAPI*. A fonte de extração de informações dessa classe é o próprio site Last.fm, desta forma os métodos implementados acessam páginas web do site, e realizam a busca de dados na estrutura da página para recuperação de informações.

É importante ressaltar que alguns métodos presentes na classe *CrawlerWeb*, necessitam de informações extraídas previamente pela classe *CrawlerAPI*. Para exemplificar, pode-se citar o contexto de recuperação de informações de uma música.

Como ilustrado na Figura 8, a recuperação de informações de uma música se inicia através de uma requisição feita pela classe *CrawlerAPI* à API do site Last.fm. A API retorna uma resposta no formato JSON contendo as informações de uma música, como: Nome da música e URL da música no site Last.fm. Essa informação é então processada e registrada no sistema. Entretanto, o sistema *Musipath* precisa também da URL da música no Youtube, que não é fornecida pela API do Last.fm. Para isso, a classe *CrawlerWeb* utiliza a URL da música no site Last.fm já coletada, e realiza uma busca sobre o conteúdo desta página com a finalidade de obter a URL da música no Youtube. Desta maneira, o objeto que representa uma música pode ser criado a partir destas informações extraídas em conjunto.

Figura 8 – Fluxo de recuperação de informações de uma música

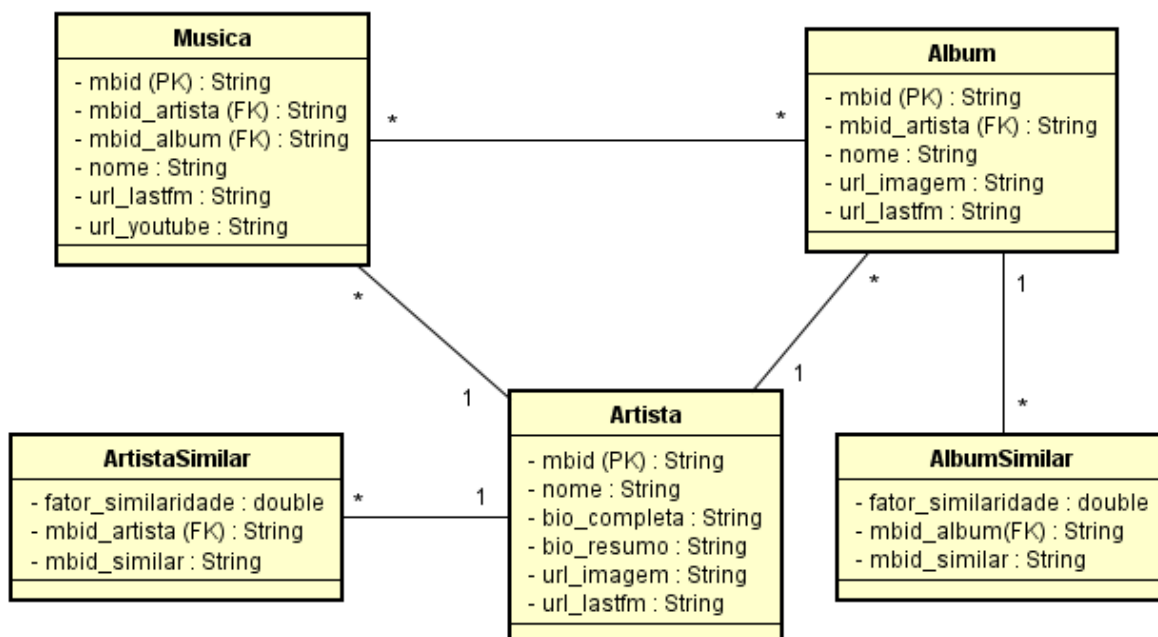


Fonte: Elaborada pelo autor, 2021.

Outra funcionalidade dos métodos implementados na classe *CrawlerWeb* é a recuperação de informações como a lista de álbuns de determinado artista, a lista de músicas de determinado álbum e a lista de artistas representados por uma categoria, tendo em vista que essas informações são necessárias para a construção da base de dados do sistema *Musipath*, mas não são informados pela API do site Last.fm.

Com a união dos métodos implementados nas duas classes citadas é possível reunir as informações recuperadas e construir os objetos e suas relações propostas para base de dados do sistema, como ilustra a modelagem na Figura 9.

Figura 9 – Diagrama de classes do sistema *Musipath*



Fonte: Elaborada pelo autor, 2021.

4.2. API de Comunicação do Sistema *Musipath*

Com o desenvolvimento do primeiro sistema (*Musipath Crawler*) e suas funcionalidades, descritas nas subseções anteriores, o banco de dados do sistema *Musipath* pode ser preenchido com informações sobre artistas, álbuns e músicas. Através dos métodos de processamento dessas informações também é possível estabelecer relações de similaridade entre artistas e álbuns de música.

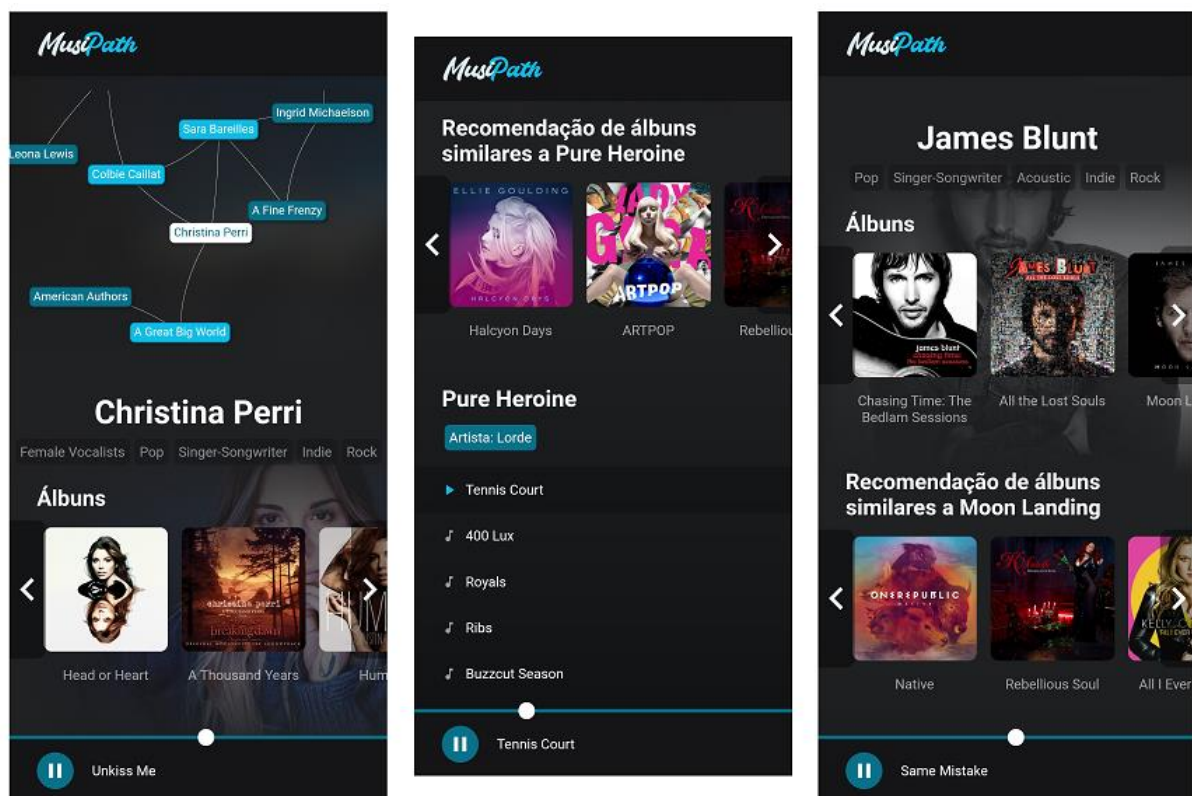
O segundo sistema desenvolvido foi nomeado como *Musipath API*. Desenvolvido sobre a linguagem de programação *Java* com auxílio do framework *Spring Boot*. O principal objetivo desse sistema foi estabelecer um meio de comunicação entre as informações coletadas e registradas no banco de dados, e o sistema Web de interface com o usuário. Como exposto anteriormente, uma API é composta por um conjunto de rotinas implementadas para permitir a comunicação e integração entre sistemas distintos. O acesso a essas rotinas pode ser feito através de *endpoints*, ou pontos de extremidade, os terminais de conexão entre o serviço disponibilizado pela API e o sistema que está consumindo esse serviço. Para exemplificar, pode-se citar o *endpoint* de acesso às informações de um artista, que pode ser feito através de uma requisição que segue o modelo: `http://{endereço da`

aplicação}/api/artistas/buscarpornome/{nome}. Neste caso, o sistema de interface do usuário pode fazer uma requisição à API, obtendo como resposta as informações que compõem um artista, no formato JSON.

4.3. Interface com o Usuário

O terceiro sistema desenvolvido foi o sistema *Musipath Frontend*, um sistema Web desenvolvido com o uso da ferramenta *React*, uma biblioteca *Javascript* declarativa, eficiente na criação de interfaces de usuário. Esta biblioteca permite a criação de interfaces através da criação de componentes. A finalidade deste terceiro sistema é disponibilizar as funcionalidades propostas inicialmente, através de interfaces para o usuário.

Figura 10 – Exemplos de interfaces do usuário em um dispositivo móvel



Fonte: Sistema *Musipath*, 2021.

O sistema Web foi desenvolvido com a proposta de responsividade, isto é, o sistema tem a capacidade de se adaptar e exibir as informações para o usuário em

diferentes tipos de dispositivos e suas resoluções de tela. A Figura 10 exibe exemplos de como a interface é modelada em um dispositivo móvel, por exemplo.

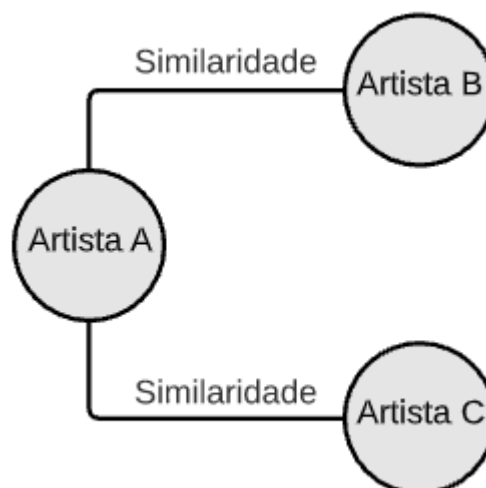
A construção da interface baseada em componentes é uma funcionalidade oferecida pela ferramenta *React*, que permite ao usuário uma interação mais fluida com o sistema, pois os componentes podem ser instanciados e atualizados dinamicamente, sem a necessidade do carregamento de novas páginas. Neste caso, foram implementados diferentes componentes para comporem a interface, cada um responsável por exibir um conjunto de informações, por exemplo, o grafo de similaridade de um artista, informações sobre um artista, informações sobre um álbum, o *player* de música, recomendação de álbuns, dentre outros.

4.4. Criação do Grafo de Artistas

Um dos objetivos do sistema *Musipath* é oferecer ao usuário a experiência de exploração de uma rede de artistas e músicas. Essa exploração pode levar à descoberta de novos artistas, álbuns e músicas que correspondem ao estilo musical do usuário. Uma das ferramentas utilizadas para gerar essa interação, foi a criação de uma rede de artistas baseada na similaridade existente entre eles.

Para representar a rede de artistas, foi utilizada a estrutura de dados de grafo, já introduzida na Subseção 2.3. A estrutura de um grafo permite a representação de relacionamentos, e foi utilizada aqui para representar a similaridade entre os artistas.

Figura 11 – Grafo representando a similaridade entre artistas



Fonte: Elaborada pelo autor, 2021.

Como ilustrado na Figura 11, cada artista pode ser representado na rede como um vértice e as arestas, por sua vez, representam a existência de similaridade entre artistas. De acordo com o exemplo de grafo apresentado pode-se concluir que o artista A é similar aos artistas B e C.

Para poder exibir a rede de artistas na interface do usuário, foi adotada a estratégia de geração do grafo individual para cada artista. Na prática isso significa que o grafo exibido para o usuário é gerado a partir de um artista inicial, e recursivamente os artistas similares vão sendo adicionados ao grafo, até que o limite de profundidade de exploração definido seja atingido. A profundidade representa o nível de recursão pelo qual o grafo foi gerado, assim, quanto maior a profundidade, maior será a complexidade do grafo. Desta maneira, para cada artista buscado ou selecionado pelo usuário na interface, um grafo é gerado e apresentado, partindo do artista inicial selecionado.

Figura 12 – Pseudocódigo para a função de criação do Grafo de um artista

Algoritmo 1 *Algoritmo recursivo para gerar o Grafo de similaridade de um Artista*

Entrada *Grafo, Artista, Profundidade Atual, Profundidade Limite*

Saída *Grafo*

```

1: Função Gerar Grafo
2:   se Profundidade Atual  $\geq$  Profundidade Limite então:
3:     Novo nó  $\leftarrow$  Criar Nó (Artista)
4:     Grafo  $\leftarrow$  Adicionar nó (Grafo, Novo nó)
5:     retornar Grafo
6:   fim se
7:   Profundidade Atual  $\leftarrow$  Profundidade Atual + 1
8:   Novo nó  $\leftarrow$  Criar Nó (Artista)
9:   Grafo  $\leftarrow$  Adicionar nó (Grafo, Novo nó)
10:  para cada Artista Similar a Artista faça:
11:    Nova aresta  $\leftarrow$  Criar Aresta (Artista, Artista Similar)
12:    Grafo  $\leftarrow$  Adicionar aresta (Grafo, Nova aresta)
13:    Grafo  $\leftarrow$  Gerar Grafo (Grafo, Artista Similar, Profundidade Atual, Profundidade Limite)
14:  fim para
15:  retornar Grafo

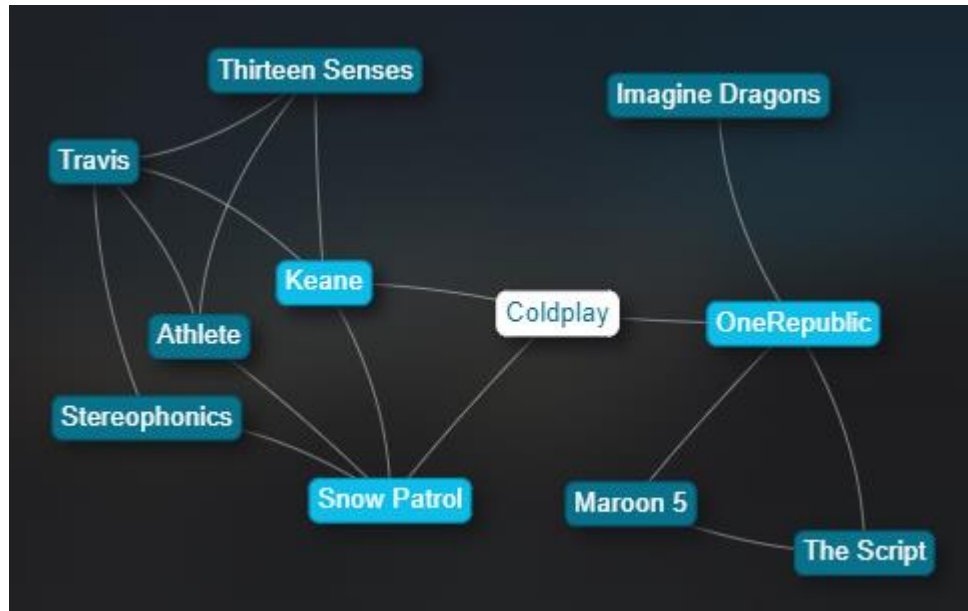
```

Fonte: Elaborada pelo autor, 2021.

A Figura 12 exibe um pseudocódigo que representa a ideia da função recursiva de geração de grafo implementada. A função recebe como argumentos um artista

inicial e a profundidade limite, e devolve como resposta o grafo de similaridade gerado para o artista inicial.

Figura 13 – Grafo de um artista gerado pelo sistema *Musipath*



Fonte: Sistema *Musipath*, 2021.

É apresentado como exemplo na Figura 13, um grafo gerado pelo sistema, tendo como artista inicial “*Coldplay*”. As cores adotadas na representação gráfica do grafo diferenciam o nível de similaridade entre os artistas. O artista inicial é representado pela cor branca, os artistas que estão diretamente associados a ele, representando uma relação direta de similaridade, são representados por um tom de azul mais claro, e os demais artistas, representados por um tom de azul mais escuro.

4.5. Sistema de Recomendação

Por meio da criação do grafo de similaridade de artistas, o sistema possibilita ao usuário a interação e descoberta de artistas similares ao seu estilo musical. Por outro lado, o sistema também propõe um mecanismo de recomendação de álbuns. Assim, o usuário pode encontrar outros caminhos na descoberta de um conteúdo musical que corresponde ao seu estilo. Por consequência, a descoberta de novos álbuns de música, também permite ao usuário o contato com novos artistas, os autores de cada álbum.

O sistema de recomendação adotado para realizar as recomendações de álbuns é baseado no conteúdo de cada álbum. Cada álbum possui um conjunto de *tags* associado a ele. Esse conjunto de *tags* tem como objetivo descrever o conteúdo do álbum através de palavras chaves ou pequenas frases. As *tags* podem referenciar o gênero de um álbum, instrumentos musicais utilizados, uma época, influências musicais, dentre outros aspectos. Esse conjunto de *tags* que descreve cada álbum foi utilizado para encontrar a similaridade entre álbuns, e assim, realizar recomendações.

Para determinar o fator de similaridade entre dois álbuns foi adotado o cálculo de similaridade de cosseno, apresentado na Figura 14.

Figura 14 – Cálculo de similaridade de cosseno

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}$$

Fonte: Neo4j - The Cosine Similarity algorithm, 2021.

Para a aplicação do cálculo de similaridade de cosseno, se faz necessário uma pré-organização de dados. Um exemplo desta organização é apresentado na Figura 15. Inicialmente, cada álbum é representado por um vetor em que cada posição é preenchida com uma *tag* do álbum. Então, para cada um dos álbuns é criado um novo vetor de tamanho N, sendo N a soma de *tags* presentes nos vetores iniciais dos dois álbuns, excluindo repetições. Desta forma tem-se dois vetores de tamanhos iguais. Cada posição deste novo vetor representa uma *tag*, e será preenchido com 1 caso a *tag* correspondente pertença ao conjunto de *tags* do álbum, e 0 caso contrário. O cálculo de similaridade de cosseno pode então ser aplicado sobre esses dois novos vetores, obtendo-se um fator normalizado de similaridade entre os dois álbuns. Assim, conclui-se que os dois álbuns são similares quanto mais este fator tende ao valor 1.

Figura 15 – Organização de vetores para aplicação da similaridade de cosseno

Álbum A			Álbum B				
pop	rock	80s	blues	80s	dance	live	pop

	pop	rock	80s	blues	dance	live
Álbum A	1	1	1	0	0	0

	pop	rock	80s	blues	dance	live
Álbum B	1	0	1	1	1	1

Fonte: Elaborada pelo autor, 2021.

Com a implementação do algoritmo de similaridade de cosseno, pode-se obter o fator de similaridade existente entre dois álbuns. Para concluir o sistema de recomendação de álbuns, foi realizado um processo de aplicação deste algoritmo em todos os álbuns do banco de dados do sistema, de forma que cada álbum pudesse ser comparado com todos os demais. Assim, definiu-se para cada álbum, um conjunto de álbuns que tem o maior fator de similaridade com este, sendo esse o conjunto de álbuns que será recomendado para o usuário.

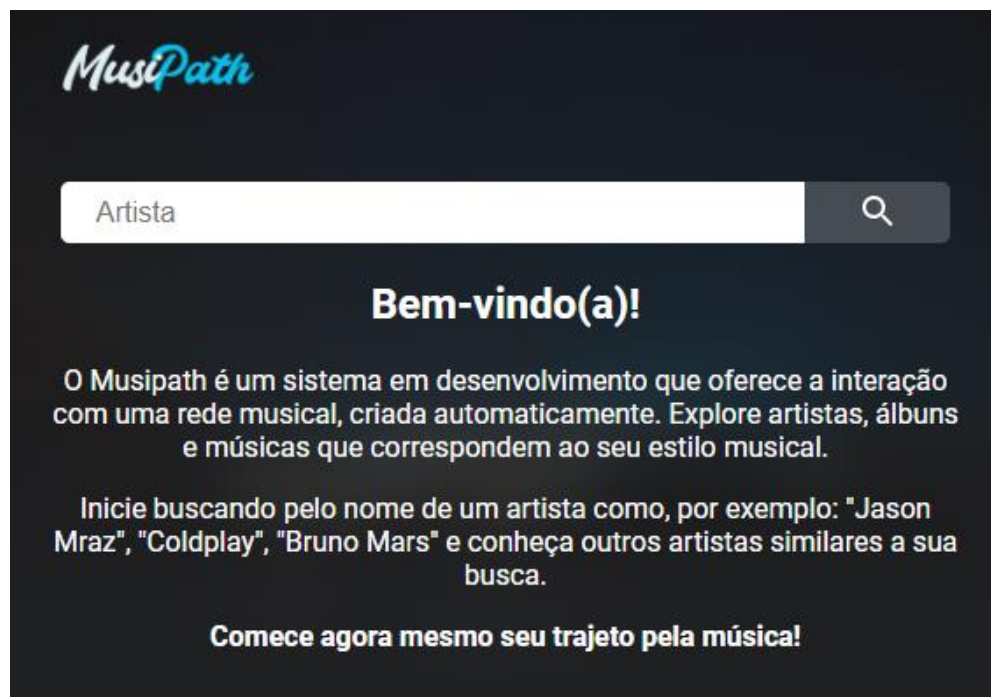
Na próxima seção são apresentados os resultados de uma avaliação realizada com o sistema *Musipath*.

5. AVALIAÇÃO E RESULTADOS

Nesta seção, será apresentado inicialmente o funcionamento da interface do sistema *Musipath*, descrevendo os fluxos de interação que podem ser tomados pelo usuário. Em seguida, serão apresentados os resultados obtidos a partir de uma avaliação do sistema com usuários reais.

A interface inicial apresenta ao usuário uma visão geral da proposta do sistema e suas funcionalidades, como ilustrado na Figura 16, e possui um campo de texto, no qual o usuário pode inserir o nome de um artista para realizar uma busca inicial.

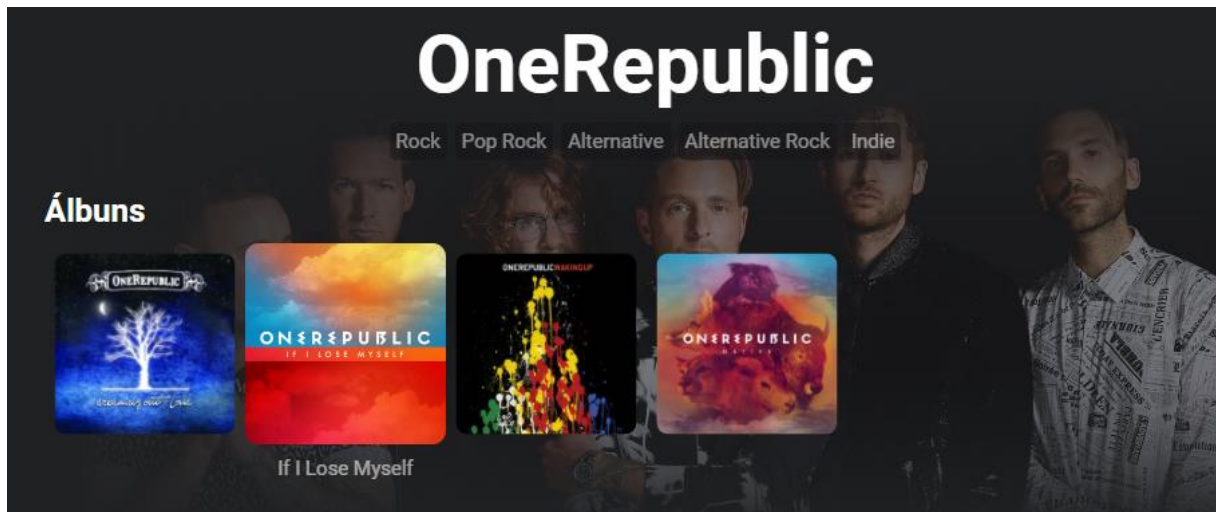
Figura 16 – Interface inicial do sistema *Musipath*



Fonte: Sistema *Musipath*, 2021.

Após realizar a busca por um artista, o sistema exibe o grafo de similaridade formado para o artista, como o grafo exibido na Figura 13, e logo abaixo as informações do artista, como, nome, tags e álbuns de música. A Figura 17 exibe a disposição das informações para o artista “OneRepublic”.

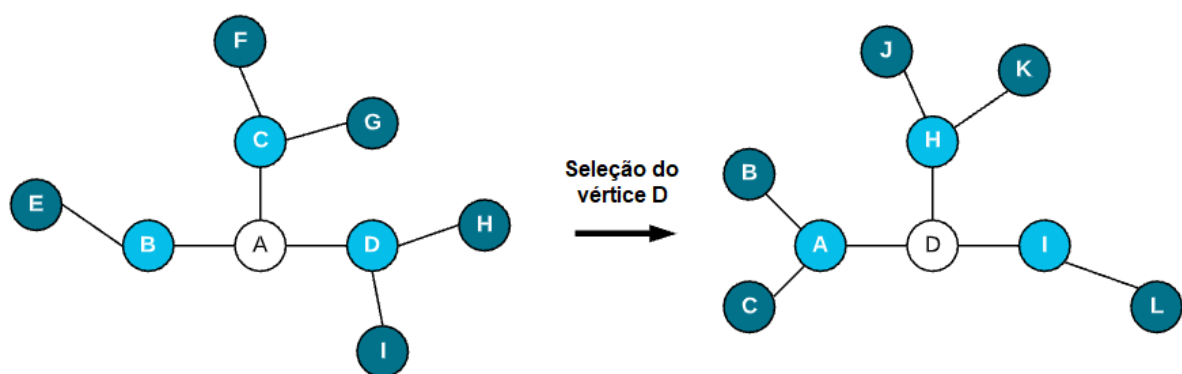
Figura 17 – Disposição das informações de um artista no sistema *Musipath*



Fonte: Sistema *Musipath*, 2021.

O grafo de similaridade oferece uma interação dinâmica, isto é, o usuário pode selecionar qualquer artista presente no grafo, e então, o grafo é atualizado tendo como base o artista selecionado. A Figura 18 exibe um exemplo de atualização do grafo de artistas. A atualização do primeiro para o segundo estado neste exemplo, ocorre quando o vértice “D” é selecionado. Também são exibidas as informações do artista selecionado no grafo, de forma análoga a ação de busca por artista.

Figura 18 – Interação e atualização do grafo de artistas do sistema *Musipath*



Fonte: William Rodrigues da Silva, 2021.

Além de interagir com o grafo de artistas, o usuário também pode selecionar os álbuns de cada artista para visualizar mais informações e ouvir a lista de músicas que o compõe. A Figura 19 expõe como exemplo a disposição das informações para o álbum “O Melhor de Tim Maia”.

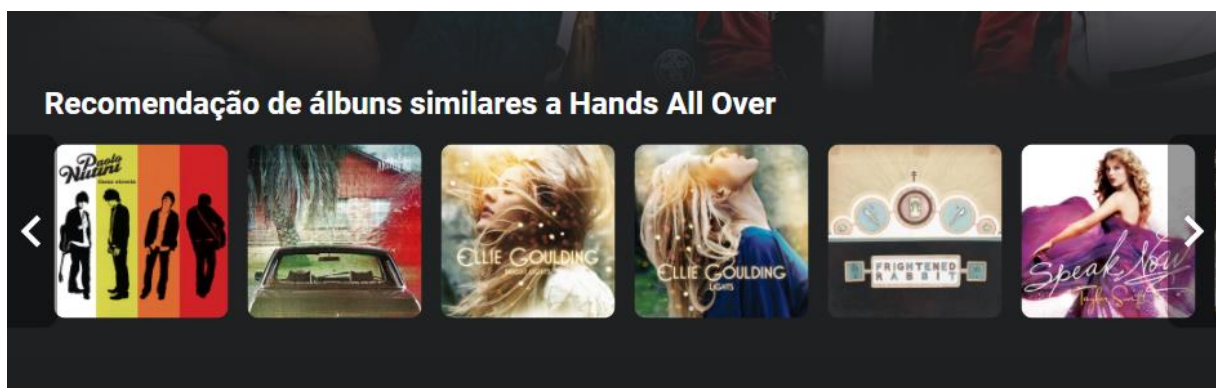
Figura 19 – Player e disposição das informações de um álbum



Fonte: Sistema *Musipath*, 2021.

Ao selecionar um álbum, o sistema também exibe uma lista de álbuns similares como recomendação, disposta de forma semelhante a lista de álbuns de um artista, como exposto na Figura 20.

Figura 20 – Lista de álbuns recomendados

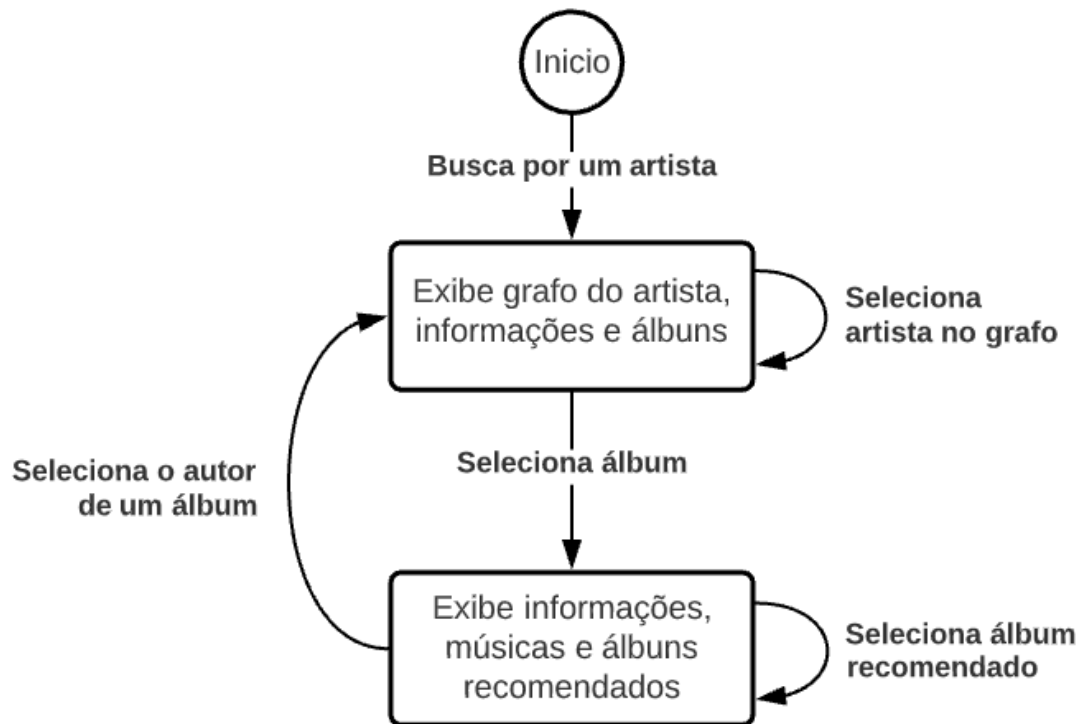


Fonte: Sistema *Musipath*, 2021.

Dentre as informações exibidas para um álbum, como ilustra a Figura 19, está o artista que é o autor deste, e o usuário consegue clicar sobre ele para visualizar

suas informações, o sistema por sua vez, realiza o mesmo mecanismo de uma busca, exibindo o grafo, álbuns e demais informações sobre o artista.

Figura 21 – Fluxograma de interações do sistema *Musipath*



Fonte: Elaborada pelo autor, 2021.

Assim, o sistema *Musipath* possibilita diversas maneiras pelas quais o usuário pode interagir e descobrir novos artistas e álbuns que correspondem a seu estilo musical. A Figura 21 ilustra um fluxograma de interações onde cada estado representa as ações do sistema, e cada transição representa as interações do usuário.

Para fins de avaliação, o sistema foi disponibilizado de forma online, e assim algumas pessoas voluntárias puderam utilizá-lo e avaliá-lo por meio de um formulário com perguntas e respostas. Os resultados desta pesquisa com perguntas e respostas são exibidos na Tabela 1.

Tabela 1 – Resultados da pesquisa de avaliação do sistema

Questões	Respostas					
	Ótima	Boa	Razoável	Ruim	Péssima	Total
“Como você classifica a compreensão do funcionamento do sistema como um todo?”	10	5	0	0	0	15
“Como você classifica a facilidade de uso do sistema?”	9	6	0	0	0	15
“Como você classifica a capacidade do sistema quanto a descoberta de artistas ou músicas que correspondem ao seu estilo?”	3	10	2	0	0	15
“Como você classifica a visibilidade das recomendações de álbuns? (Você percebeu que existem recomendações?)”	9	6	0	0	0	15
“Como você classifica a utilidade das recomendações de álbuns? (Você achou a recomendação útil?)”	4	11	0	0	0	15

Fonte: Elaborada pelo autor, 2021.

A pesquisa foi realizada com a colaboração de 15 pessoas, dentre as quais 40% são do sexo masculino e 60% feminino. Além disso, 73,3% informaram que pertencem a faixa etária entre 15 e 30 anos, 20% entre 31 e 45 anos, e 6,7% entre 46 e 60 anos.

Considerando as respostas apresentadas pelos usuários nesta pesquisa, pode-se concluir que 66,7% classificam a compreensão do funcionamento do sistema como ótima, e 33,3% como boa. Quanto a facilidade de uso do sistema 60% dos usuários consideram como boa, e 40% como ótima. Além disso, 66,7% dos usuários consideram como boa a capacidade do sistema na descoberta de artistas e músicas

que correspondem ao seu estilo, 20% como ótima, e 13,3% como razoável. Quanto à visibilidade das recomendações de álbuns, 60% dos usuários consideram como ótima, e 40% como boa. Por fim, 73,3% dos usuários classificaram a utilidade das recomendações de álbum útil como boa, e 26,7% como ótima.

6. CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho compreendeu o estudo e desenvolvimento de uma ferramenta, chamada *Musipath*, que realiza a recuperação de informações musicais e permite a interação e exploração de uma rede de artistas e músicas. A ferramenta permite que usuários explorem a similaridade existente entre artistas de músicas, descobrindo novos artistas, álbuns e músicas de seu interesse. Além disso, a ferramenta desenvolvida propõe uma interface de fácil entendimento e utilização por parte do usuário.

A elaboração da ferramenta como um todo foi dividida na implementação de três subsistemas responsáveis por diferentes tarefas, como a recuperação de informações musicais, a criação de uma rede de artistas e músicas, e a criação da interface com o usuário.

Por meio dos resultados obtidos em uma avaliação realizada com usuários reais, pode-se concluir que a ferramenta apresenta suas funcionalidades de forma descomplicada para o usuário, sendo útil na exploração da rede musical criada, e também, efetiva na exploração de novos artistas e músicas que correspondem aos interesses do usuário.

Os principais desafios encontrados na elaboração do trabalho estão relacionados à pesquisa e estudo das tecnologias aplicadas no desenvolvimento da ferramenta, tendo em vista que os sistemas implementados abordam funcionalidades tanto do contexto *back-end* de desenvolvimento, como também do contexto *front-end*.

Para trabalhos futuros propõem-se a implementação de novas funcionalidades na ferramenta, como a criação de *playlists* de música do usuário; e a possibilidade de favoritar músicas, álbuns e artistas do interesse do usuário, criando-se também uma funcionalidade para visualização destas *playlists*, músicas, álbuns e artistas favoritos.

REFERÊNCIAS

- ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **IEEE transactions on knowledge and data engineering**, IEEE, v. 17, n. 6, p. 734–749, 2005.
- ALLIK, A.; THALMANN, F.; SANDLER, M. **MusicLynx: Exploring Music Through Artist Similarity Graphs**, Queen Mary University of London, 2018.
- BALABANOVIĆ, M.; SHOHAM, Y. Fab: content-based, collaborative recommendation. **Communications of the ACM**, ACM, v. 40, n. 3, p. 66–72, 1997.
- BOLLOBÁS, B. **Modern Graph Theory**, Springer-Verlag New York, 1998. p. 1-2.
- BOBADILLA, J. et al. Recommender systems survey. **Knowledge-based systems**, Elsevier, v. 46, p. 109–132, 2013.
- BREESE, J. S.; HECKERMAN, D.; KADIE, C. Empirical analysis of predictive algorithms for collaborative filtering. In: MORGAN KAUFMANN PUBLISHERS INC. **Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence**. [S.l.], 1998. p. 43–52.
- CLAYPOOL, M. et al. Implicit interest indicators. In: ACM. **Proceedings of the 6th international conference on Intelligent user interfaces**. [S.l.], 2001. p. 33–40.
- COSTA, D.; SARMENTO, L.; GOUYON F. RAMA: **An Interactive Artist Network Visualization Tool**. Late-break/demo at ISMIR 2009.
- DESROSIERS, C.; KARYPIS, G. A comprehensive survey of neighborhood-based recommendation methods. **Recommender systems handbook**, Springer, p. 107–144, 2011.

GASSER, M.; FLEXER, A. **FM4 Soundpark: Audio-based music recommendation in everyday use**. In Proceedings of the 6th Sound and Music Computing Conference (SMC), 2009.

GOTO, M.; GOTO, T. **Musicream: Integrated music-listening interface for active, flexible, and unexpected encounters with musical pieces**. *IPSJ (Information Processing Society of Japan)*, 2009.

GREGORY, A. H. **The roles of music in society: The ethnomusicological perspective**. [S.l.]: Oxford University Press, 1997.

MARSHALL, B. **Music Artist Similarity Aggregation**, Purdue University, 2016.

NETTO, P. O. B. **Teoria e Modelos de Grafos**. Editora Edgard Blucher Ltda, 1979.

NORTH, A. C.; HARGREAVES, D. J.; O'NEILL, S. A. **The importance of music to adolescents**. *British Journal of Educational Psychology*, Wiley Online Library, v. 70, n. 2, p. 255–272, 2000.

PAMPALK, E.; GOTO, M. **MusicRainbow: A new user interface to discover artists using audiobased similarity and web-based labeling**. In Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR), 2006.

PAZZANI, M. J. A framework for collaborative, content-based and demographic filtering. **Artificial intelligence review**, Springer, v. 13, n. 5-6, p. 393–408, 1999.

ROBERTS, F. S. **Applied Combinatorics**. Rutgers University, 1984.

SCHAFER, J. B.; KONSTAN, J. A.; RIEDL, J. E-commerce recommendation applications. In: **Applications of data mining to electronic commerce**. [S.l.]: Springer, 2001. p. 115–153.

SUNDERMANN, C. V. **Extração de informação contextual utilizando mineração de textos para sistemas de recomendação sensíveis ao contexto**. Dissertação

(Mestrado) — Instituto de Ciências Matemáticas e de Computação (ICMC) - USP, São Carlos, 2015.

RICCI, F.; ROKACH, L.; SHAPIRA, B. Introduction to recommender systems handbook. In: **Recommender systems handbook**. [S.l.]: Springer, 2011. p. 1–35.

SINGH, R. P.; VANDANA. **Application of Graph Theory in Computer Science and Engineering**, International Journal of Computer Applications, Volume 104 – No.1, p.10-12, 2014.

TYPKE, R.; WIERING, F.; VELTKAMP, R. C. **A survey of Music Information Retrieval Systems**. Universiteit Utrecht, 2014.