

# Um Middleware para o Gerenciamento de Dados Musicais com Utilização de Banco de Dados Relacional

William R. Ibiapina<sup>1</sup>, Leonardo O. Moreira<sup>1</sup>

<sup>1</sup>Instituto Universidade Virtual (UFC VIRTUAL)  
Universidade Federal do Ceará – Fortaleza, CE – Brasil

william.eti@alu.ufc.br, leoomoreira@virtual.ufc.br

**Abstract.** *Upon the advent of online music repositories, digital distribution of musical data has notably increased, and studies in Music Information Retrieval (MIR) have been carried out to optimize data storage and retrieval. Querying for musical data is quite challenging, since some of that data require mathematical processing that relational databases do not provide. This paper introduces an initial design of middleware for musical data management using relational databases, through an architecture that allows communication with technologies capable of processing non-textual data.*

**Resumo.** *Com o advento de repositórios de música online, a distribuição digital de dados musicais teve um aumento notável, e estudos em Music Information Retrieval (MIR) vêm sendo conduzidos para otimizar o processo de armazenamento e recuperação desses dados. A consulta por dados musicais constitui um desafio, uma vez que alguns desses dados exigem processamento matemático que bancos de dados relacionais não providenciam. Este artigo apresenta um projeto inicial de middleware para o gerenciamento de dados musicais utilizando banco de dados relacional, através de uma arquitetura que permita a comunicação com tecnologias capazes de processar dados não-textuais.*

## 1. Introdução

Com o crescimento das aplicações multimídia nos últimos anos, a quantidade e variedade de dados a serem processados ou manipulados também aumentou [Lo and Tsai 2009]. Além de informações textuais e numéricas, dados binários, como imagens, áudio e vídeo, vêm sendo amplamente utilizados. Dentre essas aplicações, há os repositórios de músicas *online*, que atraíram interesse da indústria musical e da comunidade científica para estudo de técnicas de armazenamento e recuperação de dados musicais [Shen et al. 2006], objetivando prover suporte a classificação e reconhecimento de dados musicais com base em parâmetros específicos.

Tal estudo deu origem a uma ciência interdisciplinar para recuperação e extração de dados musicais, chamada *Music Information Retrieval* (MIR) [Downie 2005]. As pesquisas nesse campo podem valer-se de conhecimentos específicos de diversas áreas, como psicologia, estatística e aprendizagem de máquina, além da música propriamente dita. Frutos dessas pesquisas incluem categorização automática [Wang et al. 2012], reconhecimento de instrumentos [Han et al. 2017] e sistemas de recomendação [Knees and Schedl 2013], podendo-se destacar dentre estes as aplicações da Spotify, Last.fm, Amazon, Google e Pandora [McFee et al. 2012].

Os Banco de Dados (BD) Relacionais têm sido muito utilizados como modelos de armazenamento e recuperação de dados em vários sistemas, como comerciais [Mrozek et al. 2016], redes sociais, sistemas geográficos e científicos. O modelo relacional, adotado pelos Banco de Dados Relacionais, emprega os conceitos de entidade e relação [Date 2003] em estruturas de tabela, sendo a entidade identificada por seus dados em uma tabela e a relação definida pela associação entre registros de tabela [Elmasri and Navathe 2015].

A união entre os conceitos de MIR e BD deu origem aos estudos sobre bancos de dados musicais, permitindo que usuários recuperem dados por meio de parâmetros musicais. As consultas a esses BD são geralmente feitas por descrições textuais, de acordo com os rótulos dados aos objetos musicais, ou por trechos da música, para recuperar informações como o gênero musical ou a presença de um instrumento específico. Esses métodos são chamados de consultas por texto e consultas por conteúdo, respectivamente [Cui et al. 2007].

Este artigo aborda uma extensão do trabalho descrito em [Ibiapina et al. 2017]. O objetivo principal deste presente trabalho é apresentar um projeto inicial de *middleware* para suportar o armazenamento e recuperação de dados musicais em banco de dados relacional. Para a realização do objetivo principal, alguns objetivos específicos foram elencados: i) estudar as técnicas ou tecnologias para armazenamento e recuperação de dados musicais; ii) especificar um conjunto de características necessárias para o armazenamento e recuperação de dados musicais; iii) projetar os requisitos de *middleware* para suportar o armazenamento e recuperação de dados musicais em banco de dados relacional; iv) desenvolver um protótipo funcional do banco de dados relacional e do *middleware* projetado; e v) avaliar o protótipo funcional do *middleware* sob a perspectiva de desempenho.

O restante do artigo está organizado da seguinte forma. A seção 2 apresenta os conceitos teóricos que são necessários para uma melhor compreensão do trabalho. Já a seção 3 comenta e discute os trabalhos relacionados. A seção 4 apresenta o projeto e as características do *middleware* proposto. A seção 5 apresenta a metodologia de avaliação e discute os resultados obtidos. Por fim, a seção 6 conclui o trabalho e apresenta os trabalhos futuros.

## **2. Referencial Teórico**

Esta seção apresenta os conceitos teóricos que são necessários para um melhor entendimento do trabalho. Assim, são comentados os seguintes assuntos: Banco de Dados Relacional, *Music Information Retrieval* (MIR) e Tecnologias para Manipulação de Dados Musicais.

### **2.1. Banco de Dados (BD) Relacional**

Um Banco de Dados Relacional é uma estratégia de persistência e recuperação de dados que utiliza o modelo relacional, representando um banco de dados como uma coleção de relações [Elmasri and Navathe 2015]. Uma relação pode ser vista como uma tabela de valores composta por linhas. Uma linha em uma tabela de valores é chamada de tupla. Cada coluna em uma tabela de valores é chamada de atributo. Um atributo possui domínio associado onde se define o tipo de dados que o atributo pode armazenar. Segundo Date (2003), um Banco de Dados pode ser visualizado como um repositório para uma

coleção de dados que possui uma organização cuja estrutura facilita o armazenamento e recuperação destes dados.

Um dos principais motivos para o sucesso dos Banco de Dados Relacionais é a utilização do *Structured Query Language* (SQL) como linguagem de consulta para banco de dados [Elmasri and Navathe 2015]. O SQL pode ser considerada uma linguagem completa para banco de dados, pois abrange instruções tanto para criação das estruturas de banco de dados quanto para manipulação dos dados nas estruturas de banco de dados [Elmasri and Navathe 2015].

## 2.2. Music Information Retrieval (MIR)

Muitas aplicações, dentre elas Last.fm e Spotify, exigem o gerenciamento de dados musicais. Neste sentido, ações de interpretar e fazer relações entre os dados musicais, por exemplo, interpretação de gênero musical e similaridade entre músicas, são essenciais. Tais ações parecem ser um desafio na área de *Music Information Retrieval* (MIR) [Café and Barros 2018]. Segundo Mrozek et al. (2016), MIR é considerada uma ciência que possui um caráter interdisciplinar que objetiva a recuperação ou extração de informações em uma música. Diversas pesquisas e aplicações que envolvem o uso de MIR estão sendo propostas para diversos contextos, como atividades de relaxamento, atividades físicas, estudo, entretenimento etc. [Café and Barros 2018].

A potencial aplicabilidade desse campo de pesquisa é bastante vasto. Desde encontrar uma música a partir de um trecho que vem à sua cabeça durante o café da manhã [Downie 2005] até os exemplos científicos e mercadológicos supracitados. Sistemas de MIR também têm potencial para gerar receita, além de notável melhoria de acessibilidade. Por exemplo, a Vivendi Universal, em meados dos anos 2000, comprou o domínio MP3.com, muito popular como distribuidor de arquivos MP3, por 372 milhões de dólares [Downie 2005]. A pronta acessibilidade do acervo desse domínio traz grande benefício para entusiastas, músicos, estudantes etc. [Downie 2005].

## 2.3. Tecnologias para Manipulação de Dados Musicais

Em Ibiapina et al. (2017) foi feita uma pesquisa por tecnologias para execução, extração e manipulação de objetos musicais, objetivando a obtenção de dados musicais por conteúdo do arquivo de áudio. Assim, foram elencadas quatro tecnologias de código aberto.

O *Essentia*<sup>1</sup> é uma biblioteca que implementa funcionalidades de entrada e saída de áudio, processamento de sinais digitais, caracterização estatística de dados e vários descritores de alto nível. Já o *Marsyas*<sup>2</sup> é um *framework* voltado para prototipação/experimentação rápida, síntese e análise de áudio, projetado para prover uma arquitetura flexível, geral e extensa que facilite a experimentação de algoritmos e proporcione alto desempenho.

O *Madmom*<sup>3</sup> é uma biblioteca de processamento digital de áudio criada para facilidade de uso, prototipação rápida de fluxos de trabalho para processamento digital,

---

<sup>1</sup>Essentia. Disponível em: <http://essentia.upf.edu/documentation/>. Acesso em: 20 de novembro de 2018.

<sup>2</sup>Marsyas. Disponível em: <http://marsyas.info/>. Acesso em: 20 de novembro de 2018.

<sup>3</sup>Madmom. <https://madmom.readthedocs.io/en/latest/>. Acesso em: 20 de novembro de 2018.

conversão simples de um fluxo de trabalho para um programa executável e independência de softwares externos. Por fim, o LibROSA<sup>4</sup> é uma biblioteca para análise de áudio, projetado para estimativa de tempo, separação de instrumentos de percussão, harmonia, processamento de áudio no domínio do tempo e segmentação estrutural, dentre outras possibilidades.

Os dados musicais obtíveis por tais tecnologias incluem:

- **frequência:** número de ciclos (ou oscilações) por unidade de tempo;
- **andamento (*tempo*):** velocidade do compasso sonoro, geralmente medido em batidas por minuto, ou *beats per minute* (bpm);
- **duração:** valor no tempo que ilustra a duração total de um objeto musical;
- **gênero musical:** classificação dada a objetos musicais a partir da presença de elementos musicais em comum;
- **ritmo:** valor das notas musicais de acordo com a intensidade e o andamento, determinando a duração de cada som no objeto musical e também a duração dos silêncios;
- **diferenciação de componentes:** por exemplo, diferenciação entre harmonia e percussão.

A Tabela 1 apresenta as tecnologias que foram elencadas em Ibiapina et al. (2017) e detalha quais dados musicais podem ser obtidos por cada tecnologia.

**Tabela 1. Visão geral das tecnologias e quais dados podem ser obtidos**

<b>Tecnologia</b>	<b>Frequência</b>	<b>Andamento</b>	<b>Duração</b>	<b>Gênero</b>	<b>Ritmo</b>	<b>Diferenciação</b>
Essentia	Sim	Sim	Não	Não	Não	Não
Marsyas	Sim	Não	Sim	Sim	Sim	Não
Madmom	Sim	Não	Não	Não	Não	Não
LibROSA	Sim	Sim	Sim	Não	Não	Sim

**Adaptado de Ibiapina et al. (2017)**

### 3. Trabalhos Relacionados

Foi feita uma pesquisa na ACM Digital Library<sup>5</sup> e na IEEE Xplore Digital Library<sup>6</sup>, usando a string de busca “music information retrieval” e passando por duas filtragens; a primeira, por ano de publicação para artigos publicados a partir de 2013; a segunda, com base na metodologia e nas palavras-chave, as quais incluíram “music information retrieval”, “database” e atributos musicais como frequência fundamental (“fundamental frequency”) ou ritmo (“rhythm”). O resultado obtido totalizou cinco artigos.

Pugin e Zitellini (2013) escreveram sobre distribuição de instrumentos e funcionalidades de busca de notação musical para o Swiss RISM (*Swiss Répertoire International des Sources Musicales*, ou Diretório Internacional Suíço de Fontes Musicais), usando uma

<sup>4</sup>LibROSA. Disponível em: <https://librosa.github.io/>. Acesso em: 20 de novembro de 2018.

<sup>5</sup>ACM Digital Library. Disponível em: <http://dl.acm.org/>. Acesso em: 23 de novembro de 2018.

<sup>6</sup>IEEE Xplore Digital Library. Disponível em: <http://ieeexplore.ieee.org/>. Acesso em: 23 de novembro de 2018.

ferramenta de MIR chamada Themefinder. Sua implementação adiciona a possibilidade de operar em tom, métrica e duração, de forma separada ou combinada, mas carece de funcionalidades de busca por texto. Outros trabalhos relacionados incluem distribuição de instrumental e vocal e busca de *incipit*<sup>7</sup>.

Xing et al. (2013) trabalham com terapia musical baseada em reconhecimento de emoções musicais, construindo um acervo de emoções e características, comparando algoritmos de aprendizado de máquina, como SVM (*Support Vector Machine*, ou Máquina de Vetores de Suporte) e retropropagação, e analisando pressão sanguínea e batimentos cardíacos para medir flutuação emocional. Porém, embora faça uso de tecnologias similares às abordadas neste artigo para extrair as características necessárias, sua aplicabilidade é limitada à medicina.

Yang et al. (2014) trazem um algoritmo de detecção de frequência fundamental chamado BaNa, que combina relação harmônica e análise de cepstrum.<sup>8</sup> Sua principal característica é a resiliência a ruído, comparável a vários algoritmos similares, como as funções de autocorrelação e correlação cruzada. Todavia, apesar da vasta aplicabilidade, este algoritmo foca a detecção de um atributo musical específico. Sua implementação gira em torno da classificação de emoções baseada na fala, e está disponível no *website* do Wireless Communication & Networking Group<sup>9</sup>.

Prockup et al. (2014) focam seu trabalho na análise de ritmo e melodia para representar padrões musicais. Para tanto, fazem uso de modelos probabilísticos e aplicam algoritmos de redução de dimensionalidade para possibilitar o uso de algoritmos mais complexos em aplicações específicas, como aprendizado de máquina. Contudo, como apontado pelos próprios autores, sua análise peca pela ausência de características mais significativas, como magnitude e fase, que poderiam implicar resultados mais detalhados e precisos.

Bayle et al. (2017) apresentam o SATIN (*Set of Audio Tags and Identifiers Normalized*, ou Conjunto de Identificadores e Tags de Áudio Normalizados), uma base de dados criada para facilitar a reutilizabilidade e comparação entre algoritmos de MIR. Ele ajuda, por exemplo, a economizar tempo computacional, acessar as trilhas, garantir consistência entre várias bases de dados. O SATIN pode também distribuir gêneros musicais em uma nuvem de palavras, ordenar as trilhas por ano de publicação em um grafo e extrair a distribuição geográfica das trilhas através do código ISRC. Entretanto, essa base de dados trabalha primordialmente com bancos de consulta por conteúdo. O código-fonte da API está disponível no GitHub<sup>10</sup>.

Ibiapina et al. (2017) focam no estudo para elaboração de estratégias de adaptação de um banco de dados relacional, propondo uma arquitetura de sistema para que seja possível suportar o armazenamento e recuperação de dados musicais em um banco de

---

<sup>7</sup>Incipit é o nome dado à sequência inicial de notas.

<sup>8</sup>Cepstrum é o resultado da transformada inversa de Fourier do logaritmo do espectro estimado de um sinal.

<sup>9</sup>Wireless Communication & Networking Group. Disponível em: [http://www2.ece.rochester.edu/projects/wcng/project\\_bridge.html](http://www2.ece.rochester.edu/projects/wcng/project_bridge.html). Acesso em: 23 de novembro de 2018.

<sup>10</sup>GitHub. Disponível em: <https://github.com/ybayle/SATIN>. Acesso em: 23 de novembro de 2018.

dado relacional. Para isso, seu trabalho se voltou à análise de tecnologias de manipulação de dados musicais e apresentou uma proposta de arquitetura de sistema que integre essas tecnologias e um SGBD relacional. Porém, não houve etapas de implementação e avaliação das soluções encontradas e propostas.

**Tabela 2. Trabalhos relacionados e suas características**

<b>Trabalho</b>	<b>MIR</b>	<b>BD</b>	<b>Tipos de Consulta</b>	<b>Implementação</b>
Pugin e Zitellini (2013)	Sim	Sim	Conteúdo	Não
Xing et al. (2013)	Sim	Sim	Texto	Não
Yang et al. (2014)	Sim	Não	Conteúdo	Sim
Prockup et al. (2014)	Sim	Não	Conteúdo	Não
Bayle et al. (2017)	Sim	Sim	Texto	Sim
Ibiapina et al. (2017)	Sim	Sim	Conteúdo e Texto	Não
<b>Nossa Abordagem</b>	<b>Sim</b>	<b>Sim</b>	<b>Conteúdo e Texto</b>	<b>Sim</b>

A Tabela 2 destaca os trabalhos relacionados discutidos e traz uma taxonomia para estabelecer uma comparação entre as características de cada trabalho. Como mostrado, a abordagem proposta neste artigo é a única que contempla elementos de MIR, usa bancos de dados relacionais para atender a certos tipos de consultas, apresenta uma implementação para validar a abordagem proposta e executa consultas de dados musicais por texto e conteúdo.

#### **4. O Middleware**

Esta seção apresenta o projeto da arquitetura de sistema e a modelagem do Banco de Dados Relacional para o *middleware* discutido neste trabalho. Segundo Craig (2002) o termo *middleware* pode ser adotado quando uma tecnologia permite a integração de elementos (dados e programas), que podem ser distribuídos e/ou integrados. Assim, a proposta discutida neste trabalho pode realizar a integração de tecnologias para manipulação de dados musicais, possibilitando o armazenamento de dados musicais em banco de dados relacional no intuito de facilitar e agilizar a busca de tais dados. Esses dados musicais gerenciados pela solução proposta podem ser consumidos ou utilizados por usuários e outros programas. A Figura 1 demonstra os detalhes da arquitetura de sistema do *middleware* e destaca os componentes que foram especificados. A arquitetura de sistema adotada para o *middleware* proposto foi apresentada e detalhada em [Ibiapina et al. 2017].

A arquitetura de sistema é dividida em cinco estruturas principais, que se comunicam como apresentado na Figura 1. O usuário final ou programas que se servem da solução proposta efetuam uma consulta de determinados dados musicais através da Interface com o Usuário (1), que repassa a consulta para o Processador de Consultas (2). Este redireciona a consulta dependendo da natureza dos dados musicais requeridos: consultas textuais são repassadas para os Arquivos de Metadados (5), enquanto consultas por conteúdo são levadas às Tecnologias para Manipulação de Dados Musicais (3), as quais são responsáveis pelo processamento de dados não-textuais. As tecnologias a serem utilizadas dependerão das consultas requeridas pelo usuário e serão dinamicamente escolhidas pelo próprio *middleware*, com base em variáveis como desempenho, processamento e consumo de memória.

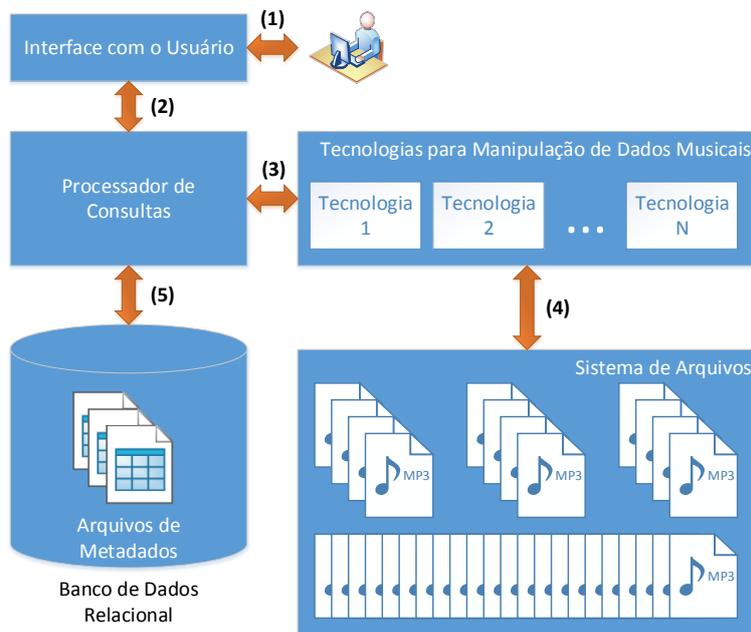


Figura 1. Arquitetura do Sistema [Ibiapina et al. 2017]

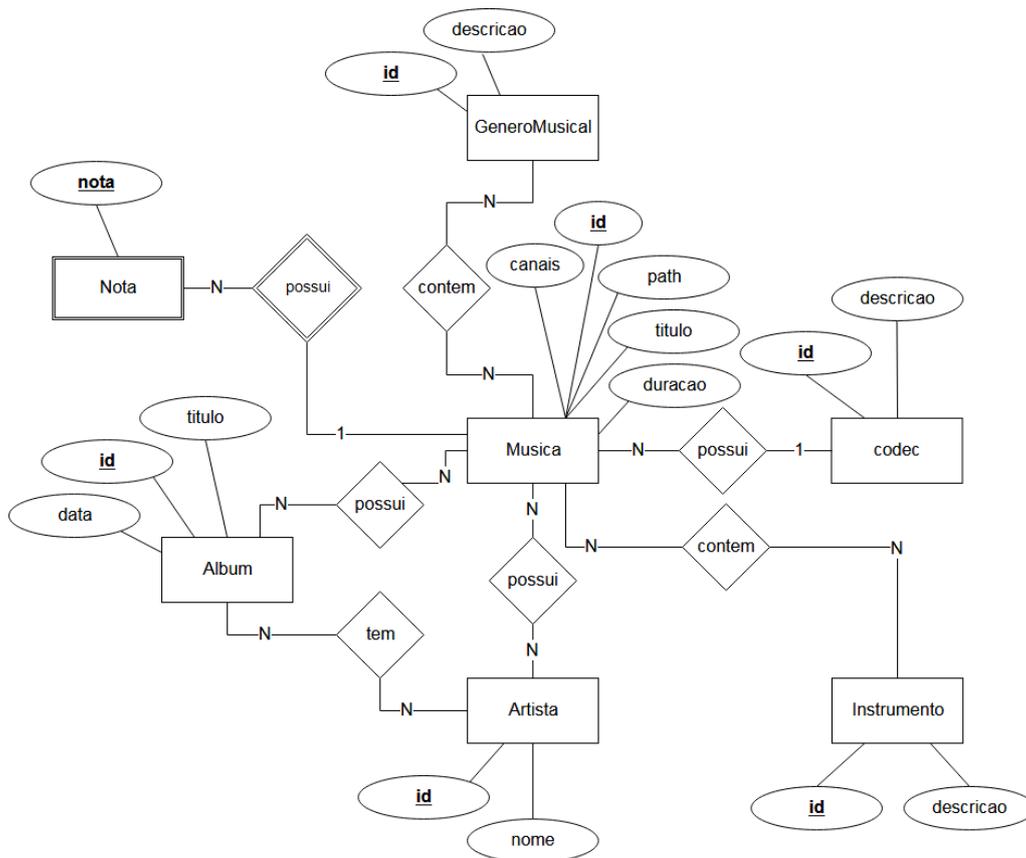


Figura 2. Modelo Entidade-Relacionamento

Para tanto, as tecnologias, neste artigo exemplificadas pelo LibROSA, acessam o Sistema de Arquivos (4) para processamento das consultas por conteúdo. Além disso, os Arquivos de Metadados são alimentados em tempo de armazenamento das músicas, em que os dados musicais textuais e numéricos são pré-processados e armazenados para serem posteriormente recuperados pelos bancos de dados relacionais, como o PostgreSQL. Os metadados pré-processados e armazenados obedecem à estrutura mostrada na Figura 2, que apresenta dados textuais e numéricos, assim como alguns dados de conteúdo mais comumente utilizados como critérios de classificação musical, o que abre um cenário de pré-processamento que será explicado mais adiante.

## 5. Avaliação

Para avaliar o projeto inicial do *middleware* proposto, foi escolhida a tecnologia do LibROSA, pela gama de dados musicais que ela pode extrair e pela praticidade de implementação, graças à linguagem Python, predominante no seu código-fonte. Para conduzir os testes preliminares, foram elaboradas as seguintes consultas:

- (Q1) Quais músicas são lentas e duram no máximo três minutos?
- (Q2) Quais músicas são rápidas e apresentam percussão?
- (Q3) Quais músicas não apresentam percussão e duram no mínimo dois minutos?

A partir das possibilidades oferecidas pela tecnologia e das consultas preliminares, as métricas escolhidas foram as seguintes:

- A **duração** da música é medida pelo LibROSA em segundos, com precisão de até centésimos de segundo.
- O **andamento** é uma classificação subjetiva baseada no tempo musical, medido em batidas por minuto (bpm). Os termos classificativos geralmente são em italiano, mas nesta avaliação será feita uma classificação mais simplificada para fins de referência: músicas com tempo inferior a 80 bpm serão rotuladas como lentas, músicas com tempo superior a 112 bpm serão consideradas rápidas, e músicas com tempo entre esses dois valores serão tidas como moderadas.
- A **presença ou ausência de instrumentos de percussão** é feita com base na potência em decibel (dB). A medição é realizada por meio de uma diferenciação de componentes, que resulta em duas variáveis de série temporal distintas, uma contendo os componentes percussivos e a outra com os componentes harmônicos. Desta é calculada a potência sonora máxima, e os componentes percussivos são rotulados como ausentes caso essa potência seja inferior a 30 dB, ou presentes em caso contrário.

### 5.1. Implementação

A implementação teve como objetivo de mostrar o uso da tecnologia, das métricas adotadas e dos algoritmos construídos para a extração de tais métricas. Para tanto, foram criados um código-fonte simples em Python, com uso extensivo das funções do LibROSA, e um banco de dados relacional gerenciado no SGBD PostgreSQL 11<sup>11</sup>, através da ferramenta pgAdmin 4<sup>12</sup>. O código-fonte carrega arquivos de áudio através da árvore de

<sup>11</sup>PostgreSQL. Disponível em: <https://www.postgresql.org/>. Acesso em: 23 de novembro de 2018.

<sup>12</sup>pgAdmin. Disponível em: <https://www.pgadmin.org/>. Acesso em: 23 de novembro de 2018.

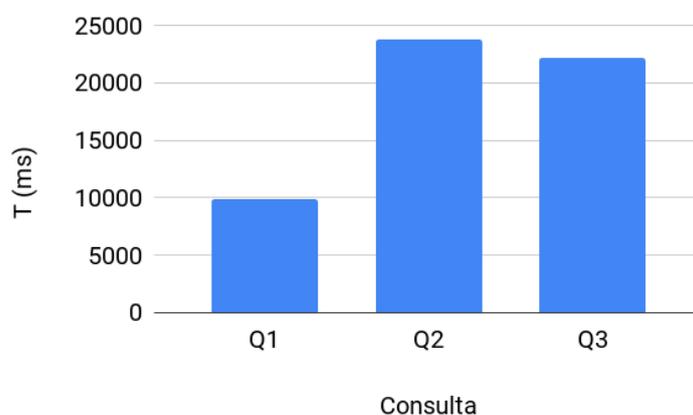
diretórios de cada arquivo, guarda-os em variáveis de série temporal e extrai deles as três métricas escolhidas.

Para o espaço amostral utilizado pelo código-fonte e pelo banco de dados, foram escolhidas para a avaliação 25 músicas, dentre músicas rápidas, lentas e moderadas, com durações variando entre 1 minuto e 5,5 minutos, e componentes percussivos presentes em algumas e ausentes em outras. Os testes foram conduzidos em um *laptop* HP EliteBook 8470p, com processador Intel Core i7-3720QM de 2,6GHz, 8 GB de memória RAM e sistema operacional Windows 7 Ultimate x64. O ambiente foi centralizado de modo a oferecer melhor controle sobre as variáveis presentes na análise.

## 5.2. Análise

A análise dos resultados foi feita em dois cenários distintos. As Figuras 3 e 4 mostram graficamente o tempo médio de resposta ( $T$ ), em relação ao espaço amostral utilizado, para cada consulta.

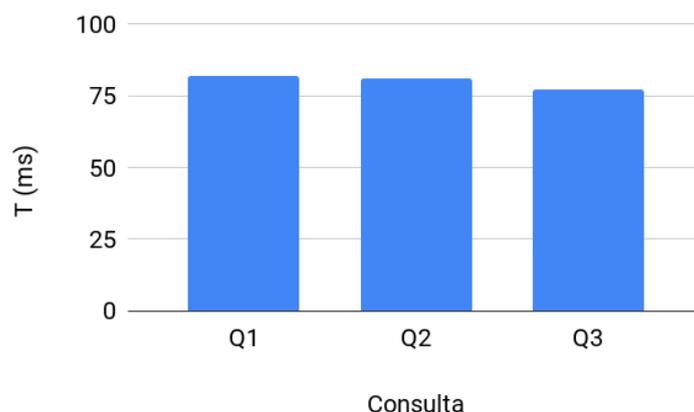
No primeiro cenário experimental, denominado Cenário 1, o LibROSA é acionado sempre que há uma consulta por conteúdo. A cada consulta, o sistema de arquivos é acessado, as métricas requisitadas são extraídas dos arquivos de áudio correspondentes e enviadas diretamente ao usuário pelo processador de consultas. As métricas não são armazenadas nos arquivos de metadados neste cenário, e são recuperadas individualmente. Devido a isso, os arquivos musicais são processados pelo *middleware* em cada consulta, aumentando consideravelmente o tempo de resposta. Além disso, a presença de elementos percussivos é a métrica mais custosa, resultando em outro aumento notável nesse tempo.



**Figura 3. Tempo médio de resposta  $T$  (em ms) de cada consulta no Cenário 1**

No segundo cenário experimental, denominado Cenário 2, a extração das métricas é feita em tempo de *upload* dos arquivos de áudio. À medida que os arquivos são carregados no sistema, o LibROSA já é acionado para extrair métricas específicas e armazená-las nos arquivos de metadados. Dessa forma, em eventuais consultas que as requisitem, essas métricas pré-processadas serão acessadas da mesma forma que métricas textuais convencionais, o que permite que sejam recuperadas em lotes.

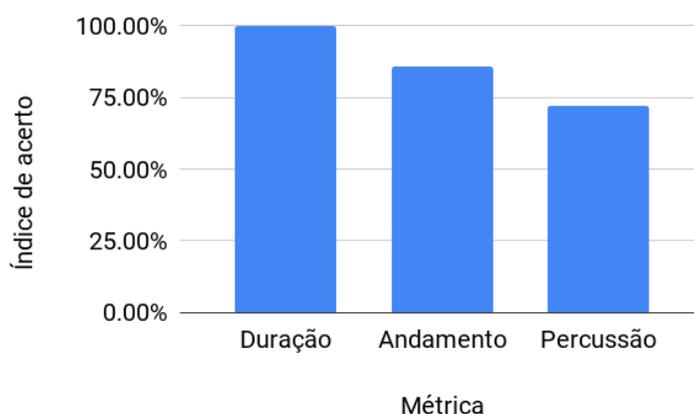
A Tabela 3 resume os tempos de respostas obtidos, em ambos os cenários, após a execução das consultas Q1, Q2 e Q3 para as 25 músicas que foram especificadas como



**Figura 4. Tempo médio de resposta  $T$  (em ms) de cada consulta no Cenário 2**

amostra. Os tempos médios de resposta mostrados nas Figuras 3 e 4 resultam do cálculo das médias aritméticas dos tempos de resposta dos arquivos musicais para as respectivas consultas nos respectivos cenários.

Após a avaliação do *middleware* proposto, foi feita uma segunda análise, de forma manual, objetivando medir a precisão da extração das métricas adotadas. Essa análise resultou num segundo conjunto de dados, que foi comparado aos dados extraídos pela tecnologia.



**Figura 5. Índice de acerto da extração das métricas**

Como mostrado na Figura 5, a duração obteve máximo acerto, pois é um dado mais simples de se extrair do arquivo de áudio, o que se verifica também pelo tempo de resposta da extração dessa métrica, o mais curto das três. O andamento apresentou índice de 85,6%, com a análise manual sendo feita com auxílio de um metrônomo *online*<sup>13</sup>, e a tecnologia apresentando estimativas diferentes para a velocidade ou o compasso<sup>14</sup> de

<sup>13</sup>Este metrônomo está disponível em: <https://www.imusic-school.com/en/tools/online-metronome/>

<sup>14</sup>O compasso foi medido externamente aos testes pelo metrônomo online, e auxiliou na medição da precisão da estimativa da tecnologia estudada.

**Tabela 3. Tempo de resposta por música para cada consulta**

Música	Cenário 1			Cenário 2		
	Q1 (ms)	Q2 (ms)	Q3 (ms)	Q1 (ms)	Q2 (ms)	Q3 (ms)
M1	6650	13529	12694	82	81	77
M2	10939	24689	23113	82	81	77
M3	3405	7819	7300	82	81	77
M4	15413	39665	36960	82	81	77
M5	10617	26653	24827	82	81	77
M6	11682	26839	25046	82	81	77
M7	15003	37706	35131	82	81	77
M8	6234	14981	13955	82	81	77
M9	12671	28755	26880	82	81	77
M10	7863	19166	17862	82	81	77
M11	12782	29018	27153	82	81	77
M12	12771	32473	30266	82	81	77
M13	6187	14053	13167	82	81	77
M14	11831	29351	27287	82	81	77
M15	15790	40234	37495	82	81	77
M16	9553	23003	21428	82	81	77
M17	12455	28724	26833	82	81	77
M18	7061	16839	15688	82	81	77
M19	4979	11746	10952	82	81	77
M20	9524	23545	21950	82	81	77
M21	8813	21625	20119	82	81	77
M22	10068	24702	22999	82	81	77
M23	6019	14455	13470	82	81	77
M24	10821	24822	23203	82	81	77
M25	8299	20112	18703	82	81	77

algumas músicas. A presença de elementos percussivos foi a métrica com menor índice, de apenas 72%, pois a diferenciação de componentes não é perfeita e a extração baseada na potência sonora é bastante rudimentar, apresentando razoável margem de erro em virtude das potências sonoras diferentes de cada arquivo musical. Os arquivos musicais sem percussão foram os que apresentaram maior inconsistência na leitura do andamento, exatamente devido à falta de um instrumento percussivo com som mais alto para marcar as batidas.

## 6. Conclusão

Este artigo apresentou um projeto inicial de *middleware* para suportar o armazenamento e recuperação de dados musicais em banco de dados relacional. Assim, foram explora-

dos e estudados conceitos teóricos sobre Banco de Dados Relacional e MIR. Além disso, também foi feita a exploração por tecnologias que dão suporte a extração de dados musicais em objetos musicais. O projeto inicial do *middleware* proposto foi comentado e detalhado, destacando os aspectos de arquitetura de sistema e projeto de Banco de Dados Relacional para manipulação de dados musicais dos objetos musicais em metadados. Por fim, foi feita um protótipo de implementação para avaliar o projeto inicial de *middleware* proposto no contexto de consultas por conteúdo e por texto servindo-se de um Banco de Dados Relacional.

Como trabalhos futuros, pretende-se: (i) realizar uma implementação completa do *middleware* proposto, já que neste artigo foi feito apenas um protótipo de implementação; (ii) avaliar o *middleware* proposto com uma variedade maior de objetos musicais e bases de dados maiores, como bancos de músicas abertos disponíveis na Web; (iii) adicionar novas tecnologias de manipulação de dados musicais para enriquecer o número de métricas que podem ser obtidas, assim resolvendo um número maior de consultas a dados musicais; (iv) fazer análises comparativas entre o uso de bancos de dados relacionais e bancos de dados em grafos, como Neo4j; (v) medir o armazenamento em disco nos diferentes cenários; (vi) propor critérios de busca em banco de dados e modelos de interface com o usuário; (vii) analisar mais detalhadamente o desempenho em relação ao espaço amostral, volume de consultas e diferentes tecnologias; (viii) comparar o *middleware* proposto com sistemas similares; (ix) estender este trabalho para a curadoria de aplicações ou outros conteúdos de mídia; e (x) propor um sistema de *benchmarking* com base na arquitetura de sistema estudada.

### Agradecimentos

Este trabalho foi realizado com recursos do projeto **Uma Estratégia para Persistência de Dados Musicais em Banco de Dados Relacional**, conforme Edital 03/17 (PIBIC 2017/2018), da Universidade Federal do Ceará (UFC). Agradecemos também ao Grupo de Pesquisa em Computação em Nuvem e Sistemas (**IBITURUNA**) pela disponibilização da aplicação em sua infraestrutura.

### Referências

- Bayle, Y., Hanna, P., and Robine, M. (2017). Satin: A persistent musical database for music information retrieval. In *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*, CBMI '17, pages 2:1–2:5, New York, NY, USA. ACM.
- Café, L. M. A. and Barros, C. M. (2018). Abordagens metodológicas as pesquisas sobre organização da informação musical. *Revista Brasileira de Biblioteconomia e Documentação (RBBDD)*, 14(3):304–323.
- Craig, I. and Serain, D. (2002). *Middleware and Enterprise Application Integration: The Architecture of e-Business Solutions*. Practitioner Series. Springer London.
- Cui, B., Liu, L., Pu, C., Shen, J., and Tan, K.-L. (2007). Quest: Querying music databases by acoustic and textual features. In *Proceedings of the 15th ACM International Conference on Multimedia*, MM '07, pages 1055–1064, New York, NY, USA. ACM.
- Date, C. (2003). *An Introduction to Database Systems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 8 edition.

- Downie, J. S. (2005). Music information retrieval. *Annual Review of Information Science and Technology*, 37(1):295–340.
- Elmasri, R. and Navathe, S. B. (2015). *Fundamentals of Database Systems*. Pearson, 7th edition.
- Han, Y., Kim, J., Lee, K., Han, Y., Kim, J., and Lee, K. (2017). Deep convolutional neural networks for predominant instrument recognition in polyphonic music. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, 25(1):208–221.
- Ibiapina, W. R., Maia, J. G. R., Neto, E. T. L., and Moreira, L. O. (2017). Uma proposta para o gerenciamento de dados musicais utilizando banco de dados relacional. *Revista Sistemas e Mídias Digitais (RSMD)*, 2(2).
- Knees, P. and Schedl, M. (2013). A survey of music similarity and recommendation from music context data. *ACM Trans. Multimedia Comput. Commun. Appl.*, 10(1):2:1–2:21.
- Lo, Y.-l. and Tsai, L.-y. (2009). Real-valued feature indexing for music databases. In *Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication*, ICUIMC '09, pages 550–556, New York, NY, USA. ACM.
- McFee, B., Bertin-Mahieux, T., Ellis, D. P., and Lanckriet, G. R. (2012). The million song dataset challenge. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12 Companion, pages 909–916, New York, NY, USA. ACM.
- Mrozek, D., Socha, B., Kozielski, S., and Małysiak-Mrozek, B. (2016). An efficient and flexible scanning of databases of protein secondary structures. *Journal of Intelligent Information Systems*, 46(1):213–233.
- Prockup, M., Scott, J., and Kim, Y. E. (2014). Representing musical patterns via the rhythmic style histogram feature. In *Proceedings of the 22nd ACM International Conference on Multimedia*, MM '14, pages 1057–1060, New York, NY, USA. ACM.
- Pugin, L. and Zitellini, R. (2013). Instrument distribution and music notation search for enhancing bibliographic music score retrieval. In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '13, pages 195–198, New York, NY, USA. ACM.
- Shen, J., Cui, B., Shepherd, J., and Tan, K.-L. (2006). Towards efficient automated singer identification in large music databases. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '06, pages 59–66, New York, NY, USA. ACM.
- Wang, J.-C., Yang, Y.-H., Chang, K., Wang, H.-M., and Jeng, S.-K. (2012). Exploring the relationship between categorical and dimensional emotion semantics of music. In *Proceedings of the Second International ACM Workshop on Music Information Retrieval with User-centered and Multimodal Strategies*, MIRUM '12, pages 63–68, New York, NY, USA. ACM.
- Xing, B., Zhang, K., Zhang, L., Lua, E. K., and Sun, S. (2013). Human-centric music medical therapy exploration system. In *Proceedings of the 2013 ACM SIGCOMM Workshop on Future Human-centric Multimedia Networking*, FhMN '13, pages 3–8, New York, NY, USA. ACM.

Yang, N., Ba, H., Cai, W., Demirkol, I., and Heinzelman, W. (2014). Bana: A noise resilient fundamental frequency detection algorithm for speech and music. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, 22(12):1833–1848.