

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

GUILHERME GREGIANIN TESTA

**Uma Abordagem Híbrida para Recomendação  
de Parceiros em Ambientes Virtuais  
Colaborativos de Composição Musical**

Dissertação apresentada como requisito parcial  
para a obtenção do grau de Mestre em Ciência  
da Computação

Prof. Dr. Marcelo Soares Pimenta  
Orientador

Porto Alegre, dezembro de 2013.

## CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Testa, Guilherme Gregianin

Uma Abordagem Híbrida para Recomendação de Parceiros em Ambientes Virtuais Colaborativos de Composição Musical [manuscrito] / Guilherme Gregianin Testa. – 2013.

65 f.:il.

Orientador: Marcelo Soares Pimenta.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Porto Alegre, BR – RS, 2013.

1. Sistemas de recomendação. 2. Recomendação de parceiros. 3. Ambientes colaborativos de composição musical. I. Pimenta, Marcelo Soares. II. Uma Abordagem Híbrida para Recomendação de Parceiros em Ambientes Virtuais Colaborativos de Composição Musical.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos Alexandre Netto

Vice-Reitor: Prof. Rui Vicente Oppermann

Pró-Reitor de Pós-Graduação: Prof. Vladimir Pinheiro do Nascimento

Diretor do Instituto de Informática: Prof. Luís da Cunha Lamb

Coordenador do PPGC: Prof. Luigi Carro

Bibliotecário-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

## **AGRADECIMENTOS**

Gostaria de agradecer meu orientador, Marcelo Soares Pimenta, e a todos os envolvidos direta ou indiretamente por minha educação e instrução nestes mais de vinte anos de estudos. Em especial, meus pais, Jacir e Maria Oliva, pelo constante incentivo e dedicação; meus irmãos Francisco, Lourenço, Henrique, Maurício e Fernando, pelo estímulo, disponibilidade para discutir assuntos aleatórios e por não atrapalharem demasiadamente meus momentos de concentração; aos meus amigos, por fornecerem uma via de escape à lucidez e realidade não acadêmicas; e, por fim, à minha namorada, Raquel, pelo seu amor, carinho e companhia constantes, além de seu apoio e consultorias diversas.

# SUMÁRIO

<b>LISTA DE ABREVIATURAS E SIGLAS.....</b>	<b>6</b>
<b>LISTA DE FIGURAS.....</b>	<b>7</b>
<b>LISTA DE TABELAS.....</b>	<b>8</b>
<b>RESUMO.....</b>	<b>9</b>
<b>ABSTRACT.....</b>	<b>10</b>
<b>1 INTRODUÇÃO.....</b>	<b>11</b>
<b>2 TRABALHOS RELACIONADOS.....</b>	<b>13</b>
2.1 Indaba.....	13
2.2 Kompoz.....	14
2.3 CODES.....	15
2.4 Considerações sobre os trabalhos relacionados.....	16
<b>3 SISTEMAS DE RECOMENDAÇÃO.....</b>	<b>19</b>
3.1 Abordagens de sistemas de recomendação.....	19
3.1.1 Filtragem colaborativa.....	20
3.1.2 Filtragem baseada em conteúdo.....	21
3.1.3 Recomendação híbrida.....	21
3.2 Análise comparativa e principais problemas.....	22
3.3 Recomendação de pessoas e parceiros.....	24
3.3.1 Recomendação de pessoas.....	24
3.3.2 Recomendação de parceiros.....	25
<b>4 COLABORAÇÃO NA COMPOSIÇÃO MUSICAL: PESQUISA COM USUÁRIOS DE WEBSITES DE COMPOSIÇÃO COLABORATIVA.....</b>	<b>27</b>
4.1 Motivação.....	27
4.2 Metodologia.....	29
4.3 Resultados.....	30
4.3.1 Relatos de experiência.....	33
4.4 Considerações finais sobre a pesquisa.....	34
<b>5 ABORDAGEM DE RECOMENDAÇÃO DE PARCEIROS.....</b>	<b>36</b>
5.1 Representação musical.....	36
5.2 Perfil musical.....	38
5.2.1 Perfil descritivo.....	38
5.2.2 Composições próprias.....	39
5.2.3 Colaborações.....	40
5.2.4 Avaliações.....	40
5.2.5 Perfil numérico.....	41
5.3 Agrupamento.....	41
5.3.1 Matriz de similaridade.....	42
5.4 Seleção e recomendação.....	43
5.5 Considerações sobre a abordagem.....	45

<b>6</b>	<b>ACOPLANDO RECOMENDAÇÃO DE PARCEIROS AO CODES.....</b>	<b>46</b>
6.1	Gêneros musicais no CODES .....	46
6.2	Perfil musical .....	47
6.3	Arquitetura e banco de dados .....	47
6.4	Interface com o usuário .....	50
6.5	Considerações finais .....	51
<b>7</b>	<b>CONCLUSÃO .....</b>	<b>52</b>
7.1	Trabalhos futuros .....	53
	<b>REFERÊNCIAS.....</b>	<b>55</b>
	<b>ANEXO A &lt;QUESTIONÁRIO DE PESQUISA&gt;.....</b>	<b>60</b>
	<b>ANEXO B &lt;DEPOIMENTOS DOS ENTREVISTADOS&gt; .....</b>	<b>63</b>
	<b>ANEXO C &lt;ARTIGO PUBLICADO&gt;.....</b>	<b>66</b>

## LISTA DE ABREVIATURAS E SIGLAS

BD	Banco de Dados
CMP	<i>Cooperative Music Prototyping</i>
CSCW	<i>Computer Supported Cooperative Work</i>
CODES	<i>Cooperative Music Prototyping Design</i>
FMOL	<i>F@ust Music On Line</i>
HCI	<i>Human-Computer Interaction</i>
HTML	<i>HyperText Markup Language</i>
IR	<i>Information Retrieval</i>
kNN	<i>K-Nearest-Neighbors</i>
LCM	Laboratório de Computação e Música da UFRGS
MP	<i>Music Prototype</i>
MVC	<i>Model-View-Controller</i>
MIR	<i>Music Information Retrieval</i>
MXML	<i>Macromedia Flex Markup Language</i>
UFRGS	Universidade Federal do Rio Grande do Sul
POI	<i>Points of Interest</i>
PRES	<i>Personalized Recommender System</i>
SBCM	Simpósio Brasileiro de Computação Musical
SQL	<i>Structured Query Language</i>
TF-IDF	<i>Term Frequency - Inverse Document Frequency</i>
XML	<i>Extensible Markup Language</i>

## LISTA DE FIGURAS

Figura 2.1: Tela de edição de uma sessão no Indaba .....	14
Figura 2.2: Tela de controle de faixas de um projeto no Kompoz .....	15
Figura 2.3: Tela da área de edição de MPs .....	16
Figura 4.1: Distribuição dos <i>websites</i> utilizados pelos respondentes à pesquisa .....	30
Figura 4.2: Experiência musical relatada pelos entrevistados .....	31
Figura 4.3: Aceitabilidade de colaborações dos usuários conforme seu tempo de utilização em websites de composição colaborativa .....	33
Figura 5.1: Composição do perfil musical de um usuário .....	38
Figura 5.2: Grafo gerado pelo software Graph Magics representativo da relação da Tabela 5.4. ....	43
Figura 6.1: Arquitetura do CODES estendida com o módulo de recomendação. Adaptado de Miletto (2010) .....	48
Figura 6.2: Esquema do banco de dados estendido do CODES. Adaptado de Miletto (2010) .....	49
Figura 6.3: Página de gerenciamento de MPs, com a área de recomendação de parceiros em destaque .....	50

## LISTA DE TABELAS

Tabela 3.1: Vantagens e desvantagens de cada uma das abordagens de recomendação .....	24
Tabela 4.1: Resumo e tradução das perguntas do questionário da pesquisa realizada...	29
Tabela 4.2: Distribuição da preferência dos respondentes quanto à criação de projetos próprios ou contribuição à outros usuários .....	32
Tabela 5.1: Convenções representativas para descrição da abordagem de recomendação .....	37
Tabela 5.2: Relação entre a nota atribuída pelo usuário, seu parecer associado e a influência sobre seu gênero musical .....	40
Tabela 5.3: Vetores $VM$ de cinco usuários hipotéticos .....	42
Tabela 5.4.a: Distância entre diferentes usuários em um determinado grupo .....	43
Tabela 5.4.b: Distância entre usuários e suas relações de colaboração .....	43
Tabela 5.5: Algoritmo das funções de recomendação .....	44
Tabela 6.1: Estilos musicais do CODES.....	47
Tabela 6.2: Relação de páginas da <i>interface</i> do CODES com necessidade de adaptação ou criação .....	46

## RESUMO

Ambientes virtuais colaborativos de composição musical têm apresentado expressivo crescimento, acompanhando a tendência de disponibilização de serviços de apoio a atividades colaborativas na Internet. Tais plataformas fornecem um meio para que pessoas possam compartilhar uma experiência musical em conjunto remotamente. A atividade de composição musical, entretanto, é influenciada por diversos fatores subjetivos de origem pessoal, social e cultural; e encontrar pessoas com objetivos convergentes para colaborar em uma criação musical em conjunto têm se demonstrado como uma tarefa complexa e desgastante. Neste contexto, esta dissertação apresenta uma forma automatizada para aproximar pessoas com objetivos e preferências musicais compatíveis para colaborar entre si. Utilizando técnicas da área de *sistemas de recomendação*, este trabalho propõe uma abordagem híbrida de recomendação de parceiros – baseada em filtragem colaborativa e por conteúdo –, em ambientes colaborativos de composição musical. Ao longo deste trabalho, os principais conceitos e características da abordagem são apresentados, assim como um roteiro de sua implantação em um sistema real, o CODES – um ambiente de suporte à prototipação musical voltado à leigos.

**Palavras-Chave:** Sistemas de Recomendação, CSCW, Computação Musical, Ambientes Colaborativos de Composição Musical.

# **A Hybrid Approach to Recommending Partners in Collaborative Environments for Musical Composition**

## **ABSTRACT**

Collaborative environments for musical composition have grown significantly lately, following the current tendency of providing services that support cooperative activities on the Internet. Such platforms provide a common way for people to share a musical experience remotely. However, the music composition activity is influenced by subjective factors of personal, social and cultural backgrounds. And the task of finding people with the same goals to collaborate with on a musical creation has been shown very complex and overwhelming. In this context, this dissertation explores an automated approach to approximate people with compatible objectives and musical preferences to compose together. Using techniques from the recommender systems' area, it proposes a hybrid approach, based both on collaborative filtering and content-based recommendation, for recommending partners in collaborative environments for musical composition. Throughout this work, the key concepts, principles and characteristics of the approach are presented, as well as the details of its implementation in CODES, a real environment to support music creation by novices.

**Keywords:** Recommender Systems, CSCW, Computer Music, Collaborative Environments for Music Composition.

# 1 INTRODUÇÃO

No imaginário popular, a composição musical mostra-se como uma tarefa solitária, resultante de um esforço individual de intensa iluminação criativa. Porém, existem evidências que a colaboração possui um papel fundamental no desenvolvimento do trabalho criativo humano. Farrel (2003), em seu extenso estudo sobre círculos colaborativos – grupo de pessoas que trabalha cooperativamente –, é categórico ao afirmar que a quase totalidade de artistas, escritores, compositores, cientistas e políticos veem a colaboração como parte indispensável do desenvolvimento de seus trabalhos. Diversos estudos de antropologia, sociologia e psicologia têm demonstrado a importância da criatividade coletiva para o avanço intelectual e tecnológico da sociedade humana (PALISPIS, 1996)(WINTHROP, 1991)(TUOMELA, 2000).

A atividade de composição musical de forma colaborativa expandiu significativamente com o avanço da Web 2.0, das redes sociais virtuais e a crescente oferta de serviços interativos na Internet (MCDONALD, 2003). Assim, vários desses sistemas surgiram na última década, atraindo pessoas de diversas partes do mundo para compartilhar uma experiência musical conjunta, nascendo, desta maneira, uma nova forma de criar música através do computador. Entretanto, a música, posta como arte, é influenciada por fatores subjetivos de origem pessoal, social e cultural, que contribuem para a formação da identidade de cada indivíduo (REYNOLDS, 2013). Encontrar pessoas com perfis em comum para compor em conjunto em tais ambientes, que são tipicamente multiculturais, torna-se uma tarefa complicada e desgastante.

Sistemas de recomendação surgiram como uma alternativa para lidar com desafios desta natureza, relacionados ao evento conhecido na literatura como o problema da sobrecarga de informações (*the information overload problem*), onde é necessário lidar com uma quantidade de dados muito grande durante a busca por algo objetivo (SCHICK; GORDON; HAKA, 1999). Com esta motivação, a presente dissertação apresenta um estudo sobre sistemas de recomendação e ambientes colaborativos de composição musical, além de apresentar dados coletados de uma pesquisa quantitativa com usuários de sistemas comerciais de criação musical, onde foram examinados os fatores que influenciam o relacionamento de compositores que criam música em conjunto. Dentro deste contexto, propõe-se uma abordagem para a recomendação de pessoas (parceiros) em tais ambientes, elaborada com base em diferentes áreas de pesquisa como computação musical, sistemas de recomendação, psicologia musical, além de uma pesquisa quantitativa conduzida com usuários reais de *websites* colaborativos de composição musical. Recomendação de pessoas envolve aspectos complexos, especialmente em áreas muito suscetíveis a interferências de cunho pessoal. Na música, especificamente, cada indivíduo possui preferências e inclinações muito sutis de gênero. Encontrar parceiros para trabalhar em conjunto em tais ambientes, e que

se beneficiem reciprocamente e convirjam para a satisfação mútua é o objetivo do trabalho apresentado nesta dissertação.

A abordagem de recomendação proposta é híbrida, combinando a filtragem colaborativa e filtragem baseada em conteúdo. Para atingir tal objetivo, características específicas dos usuários, suas interações e colaborações são recuperadas do sistema para criar perfis numéricos representados por vetores de gêneros musicais, de modo que possam ser agrupados, comparados e analisados juntamente com perfis descritivos para posterior recomendação. Como exemplificação da abordagem proposta, é apresentado um roteiro para a implementação de um sistema de recomendação no CODES: um ambiente real de prototipação musical colaborativa (CMP – *Cooperative Music Prototyping*) voltado a leigos (MILETTO, 2010).

Esta dissertação está organizada como segue: o Capítulo 3 apresenta alguns websites atuais de composição colaborativa, destacando os aspectos colaborativos dos mesmos; o Capítulo 3 descreve conceitos importantes sobre sistemas de recomendação em geral; o Capítulo 4 traça uma breve discussão sobre composição colaborativa e apresenta a motivação, metodologia e resultados obtidos através da pesquisa realizada com usuários de *websites* de composição colaborativa; no Capítulo 5, encontra-se detalhadamente a abordagem de recomendação de parceiros proposta; posteriormente, o Capítulo 6 descreve um roteiro para o acoplamento de um módulo de recomendação no CODES; e, finalmente, o Capítulo 7 apresenta conclusões e possíveis trabalhos futuros complementares a esta pesquisa.

## 2 TRABALHOS RELACIONADOS

Os primeiros estudos destacando o suporte computacional para a composição musical datam da década de 70, com destaque para a *League of Automatic Music Composers* (BISCHOFF; GOLD; HORTON, 1978). Em 1991, o professor de música James A. Hoffmann foi o pioneiro a relevar a importância da composição musical colaborativa através do computador (HOFFMANN, 1991). Hoffmann utilizou *softwares* musicais simples como ferramenta de apoio ao ensino de harmonia musical. Os resultados de sua experiência foram muito positivos, fato que o impulsionou a incentivar a comunidade internacional a rever o modelo de ensino tradicional de música. Seu método colaborativo, onde os alunos participam, discutem e argumentam ativamente sobre os exercícios de composição, obteve grande interesse e aceitabilidade por parte dos estudantes.

Atualmente, ambientes virtuais colaborativos de apoio à composição musical estão em perceptível expansão, acompanhando o aumento do número de comunidades virtuais e serviços *online* interativos. Tais ferramentas oferecem diversos recursos e facilidades que visam atrair tanto usuários iniciantes em música quanto aqueles experientes ou, até mesmo, músicos profissionais. Nas próximas seções deste capítulo, serão apresentados alguns destes ambientes, dando destaque para suas principais características e aspectos relacionados à colaboração.

### 2.1 Indaba

O Indaba<sup>1</sup> (Figura 2.1) foi fundado em 2005 e é a rede social musical colaborativa com o maior número de usuários atualmente, com cerca de 800.000 registrados (INDABA, 2013). Nela, o usuário pode criar uma composição utilizando uma ferramenta que permite gravar e editar vozes ou instrumentos musicais diretamente no *browser*, ou através do *upload* de arquivos de áudio gravados em outros meios. Assim inicia-se uma composição colaborativa, denominada de sessão. Uma vez aberta, é possível o envio de convites individuais para outros usuários colaborarem, ou ainda publicar a sessão para atrair parceiros desconhecidos entusiasmados com a ideia musical. Uma composição pode permanecer como privada ou pública. Todas as discussões sobre as adições que os colaboradores passam a realizar na sessão, assim como o gerenciamento da mesma são administradas através da *interface* denominada Indaba Music Session. O site ainda provê salas de bate-papo e chats específicos para cada sessão, permitindo uma troca de ideias em tempo real, além de possibilitar comentários em pontos específicos das músicas.

---

<sup>1</sup> <http://www.indabamusic.com>



Figura 2.1: Tela de edição de uma sessão no Indaba.

O Indaba é bastante popular por prover oportunidades para seus usuários comporem com músicos profissionais, pagos para este fim, além de promover eventos e campeonatos de composição e mixagem. Oferece ainda assistência de profissionais e regularmente apresenta composições do *site* para produtores reconhecidos, proporcionando assim críticas de qualidade aos seus usuários.

O site possui um sistema de recomendação de colaboradores, fornecido pela EchoNest, empresa especializada em soluções de áudio e aplicações relacionadas à música. Ao criar uma nova sessão, Indaba sugere parceiros através de um mecanismo de recomendação que analisa as composições prévias e outras características explícitas providas pelos usuários, e procura por compatibilidades (TECHCRUNCH, 2013). Entretanto, existem muito poucas referências sobre a ferramenta e seu funcionamento, e até mesmo no próprio *website* do Indaba este mecanismo não é citado.

## 2.2 Kompoz

O Kompoz<sup>2</sup> (Figura 2.2) é focado principalmente em músicos profissionais, ainda que seu funcionamento seja muito semelhante àquele do Indaba. Suas composições colaborativas são denominadas de projetos. Em um projeto, criado por um proprietário, adiciona-se uma descrição sobre a ideia da composição, objetivos e as trilhas de instrumentos que faltam. Outros usuários podem colaborar acrescentando faixas sob a aprovação do criador do projeto. Assim como no Indaba, é possível enviar convites para colaboração específicos ou publicar a composição para atrair usuários desconhecidos eventualmente interessados na ideia musical.

<sup>2</sup> <http://www.kompoz.com>

The screenshot shows the Kompoz website interface. At the top, there is a navigation bar with 'Home', 'Projects', 'Published', 'Videos & Podcasts', 'Community', and a user profile 'GuilhermeT'. Below this, the project page for 'So Many Stars' is displayed, with the URL 'kompoz.com/p/43337'. The page has tabs for 'Overview', 'Tracks', 'Discussions', 'Members', and 'Info & Settings'. The 'Tracks' tab is active, showing a table of tracks. The table has columns for 'Type', 'Track Name', 'Member', 'Added', 'Updated', 'Status', and a set of icons. The tracks are categorized into 'Mixdown', 'Guitar, Acoustic', 'Ideas', and 'Vocals'.

Type	Track Name	Member	Added	Updated	Status	Icons
<b>Mixdown</b>						
🎧	mix_gtr	[User]	1d 2h 20m	22h 8m	🎧	109 0 7
🎧	rough_mix	[User]	6d 12m	5d 21h 24m	🎧	354 1 2
🎧	mix_01	[User]	7d 1h 13m	7d 1h 13m	🎧	110 4 0
<b>Guitar, Acoustic</b>						
🎸	Acoustic Guitar Pilot full...	[User]	5d 12h 57m	5d 1h 50m	🎧 HD	9 2 0
🎸	Pilot Project Acoustic Guitar	[User]	6d 8h 17m	6d 2h 16m	🎧 HD	13 2 1
<b>Ideas</b>						
💡	So Many Stars	[User]	6d 13h 36m	3d 14h 24m	🎧	36 1 14
<b>Vocals</b>						
🎤	Backing Vocal Dry Sep	[User]	6d 11h 12m	5d 1h 45m	🎧 HD	4 2 0

Figura 2.2: Tela de controle de faixas de um projeto no Kompoz.

A troca de informações e discussão sobre os rumos da música sendo criada são essenciais para que uma composição satisfaça tanto seu proprietário quanto seus colaboradores. Em composições com alto índice de aceitação, percebe-se um número elevado de troca de mensagens entre os envolvidos na mesma. Diferentemente do Indaba, não existem referências sobre qualquer tipo de recomendação de colaboradores no Kompoz.

## 2.3 CODES

CODES é um *website* desenvolvido no LCM (Laboratório de Computação Musical), da UFRGS, concebido para *prototipação cooperativa de música* (CMP – *Cooperative Music Prototyping*). Foi projetado com foco específico em pessoas sem conhecimentos de teoria musical. Um processo de prototipação musical significa que os usuários podem rascunhar peças musicais simples, através de passos cíclicos de refinamento. O objetivo principal do sistema é incentivar leigos a criar música de maneira colaborativa, sem necessidade de conhecimentos prévios de música. Para isso, CODES oferece uma representação de alto nível e recursos de *interface* intuitivos que possibilitam a manipulação direta de padrões sonoros predefinidos que, combinados, produzem uma peça musical, denominada *protótipo musical* (MP – *Musical Prototype*) (MILETTO 2010).

No CODES, os usuários podem criar, editar, compartilhar e publicar MPs de maneira colaborativa. Para gerar um novo protótipo, basta o usuário informar o título do

mesmo e começar a compor arrastando ícones de padrões sonoros predefinidos (pequenos arquivos MP3 de quatro segundos) da biblioteca sonora para a área de edição (Figura 2.3). Estes padrões são harmônicos e com o mesmo andamento, o que garante a conformidade do resultado final.

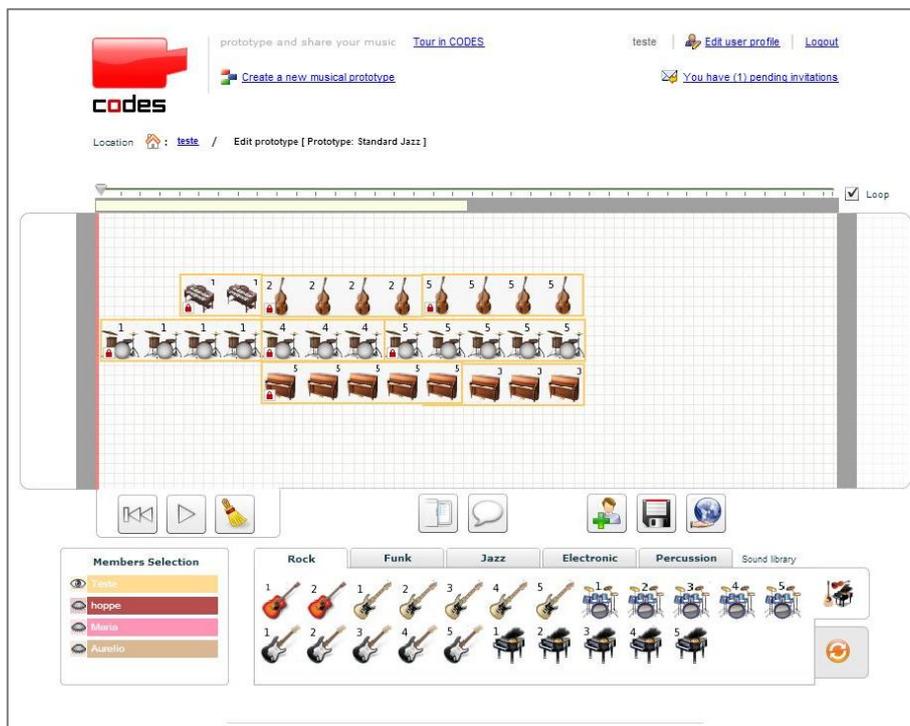


Figura 2.3: Tela da área de edição de MPs.

Na edição, usuários podem clicar no ícone de reprodução para ouvir o resultado de suas modificações a qualquer momento. Esta reprodução ocorre em *loop* e de maneira dinâmica, permitindo que o usuário escute e edite a MP de maneira concorrente. Esta é uma importante característica do CODES, já que permite ao usuário ouvir os resultados enquanto realiza modificações e novas combinações. O início de uma colaboração é realizado através de convites explícitos enviados pelo criador da MP a usuários específicos. Uma vez aceitos os convites, os colaboradores podem editar o protótipo da mesma maneira que o criador, e ter acesso às diferentes versões e à área de discussão da MP, onde os mesmos podem trocar ideias e debater sobre cada contribuição efetuada ao protótipo.

CODES atualmente não possui um mecanismo de recomendação de parceiros. Assim, um roteiro para sua implantação será proposto posteriormente, no Capítulo 6 desta dissertação, onde também serão apresentados maiores detalhes do ambiente, como sua arquitetura e estrutura de banco de dados.

## 2.4 Considerações sobre os trabalhos relacionados

Os mecanismos de suporte à colaboração dos *websites* apresentados variam sutilmente, sendo oferecidos, basicamente, fóruns especializados e locais de discussão e argumentação sobre as contribuições individuais de cada participante. A exceção é o Indaba, que possui referências sobre um mecanismo de recomendação de colaboradores.

Outros *websites* desta categoria, como MixMatchMusic<sup>3</sup>, Merge.fm<sup>4</sup> e Sound Keep<sup>5</sup>, oferecem suporte à colaboração de maneira semelhante aos já citados. Entretanto, nenhum oferece algum sistema de recomendação.

Estas plataformas possibilitam que tanto músicos profissionais quanto amadores colaborem entre si livremente. Outros *websites* são também utilizados para esse fim, porém não são exatamente de composição colaborativa, como o SoundCloud<sup>6</sup>, que possibilita a publicação e compartilhamento livre de faixas de instrumentos, vozes e sons, sem suporte direto à colaboração; além de serviços de compartilhamento de arquivos como o Google Drive<sup>7</sup> e Dropbox<sup>8</sup>, que são utilizados por pessoas que desejam colaborar entre si, mas preferem utilizar *softwares* consagrados ou profissionais para editar e gravar os sons.

Miletto (2010), em sua tese, faz uma análise detalhada sobre vários ambientes de composição semelhantes ao CODES, que são voltados a pessoas sem conhecimento musical (ou seja, que utilizam abstrações e notações mais intuitivas para que usuários leigos possam entender e compor músicas). Dentre os trabalhos estudados, sejam eles em versão *online*, quanto aplicações *desktop* específicas, estão o Daisyphone, PitchWeb, WebDrum, Public Sound Objects, EduMusical e JamSpace. Os mecanismos para suportar e fomentar a colaboração destes diferem pouco daqueles apresentados anteriormente, como a disponibilização de fóruns e listas de discussão específicas para os usuários. E nenhum deles possui algum tipo de recomendação que aproxime usuários com objetivos em comum para colaborar entre si.

Pode-se afirmar, portanto, que a recomendação de parceiros em ambientes musicais é uma área muito pouco explorada até o momento. Na literatura acadêmica, o trabalho de Wüst e Jordà, que propõe uma nova arquitetura e funcionalidades ao FMOL – *F@ust Music On-Line* (WUST; JORDÀ, 2001), um sistema de composição colaborativa *online* em tempo real, faz menção sobre a possibilidade de recomendar parceiros como um incentivo à cooperação, utilizando os perfis dos usuários para buscar similaridades. Os perfis seriam fomentados com informações explícitas inseridas pelos próprios usuários, e implícitas, inferidas pela interação dos mesmos com o sistema. Entretanto, não é descrita nenhuma abordagem efetiva de recomendação e o projeto foi, posteriormente, descontinuado.

Esta lacuna de pesquisas e, por consequência, de resultados sobre recomendação de parceiros em tais ambientes, traz dúvidas sobre quais são os fatores influentes mais importantes em uma composição colaborativa. Como descobrir e, por consequência, selecionar e recomendar os usuários mais compatíveis dentro de um universo heterogêneo e multicultural? Nos próximos dois capítulos, serão apresentados conceitos da área de sistemas de recomendação, e estudos que buscam responder estes

---

<sup>3</sup> <http://www.mixmatchmusic.com>

<sup>4</sup> <http://merge.fm>

<sup>5</sup> <http://www.soundkeep.com>

<sup>6</sup> <http://www.soundcloud.com>

<sup>7</sup> <http://drive.google.com>

<sup>8</sup> <http://www.dropbox.com>

questionamentos de maneira satisfatória para direcionar e traçar uma abordagem efetiva para este problema.

### 3 SISTEMAS DE RECOMENDAÇÃO

O termo *sistemas de recomendação* abrange um conjunto de ferramentas e técnicas que têm como objetivo primário gerar uma lista de sugestões de itens possivelmente interessantes para usuários de sistemas computacionais (JANNACH et al, 2010). Para isso, são empregadas técnicas de filtragem de dados com foco na previsão de comportamentos e preferências pessoais. Estas predições são baseadas nos perfis individuais gerados pelo sistema, e contêm julgamentos de interesses ou graus de relevância de itens previamente consumidos pelos usuários.

Ainda que a área de sistemas de recomendação seja alvo de estudos desde a década de 80, somente nos últimos anos ela foi efetivamente utilizada em larga escala. Tal acontecimento está diretamente relacionado ao problema da sobrecarga de informações (SCHICK; GORDON; HAKA, 1999). Com a crescente quantidade de dados sendo gerados e disponibilizados na Web, houve uma demanda por ferramentas que auxiliassem usuários de sistemas computacionais na busca e filtragem por conteúdos relevantes e sugestões automatizadas de objetos e produtos de interesse.

Este capítulo está organizado da seguinte maneira: a seção 3.1 apresenta as abordagens utilizadas em sistemas de recomendação atualmente, suas aplicações, e pontos positivos e negativos; a seção 3.2, explica e elenca os principais desafios dos sistemas de recomendação atuais; por fim, a seção 3.3 debate conceitos relacionados à recomendação de pessoas e parceiros.

#### 3.1 Abordagens de sistemas de recomendação

Sistemas de recomendação são implementados através de três abordagens: filtragem colaborativa (*collaborative filtering*), filtragem baseada em conteúdo (*content-based filtering*), ou então, uma combinação destas duas, denominada abordagem híbrida de recomendação (*hybrid recommendation*). A utilização de uma ou outra varia conforme o tipo de informação que será extraída do banco de dados para realizar as sugestões e o item alvo sendo recomendado. *Item* é o termo utilizado para denotar o que está sendo recomendado aos usuários, que pode variar dentre uma infinidade de produtos, como filmes (IMDB<sup>9</sup>), músicas (Last.fm<sup>10</sup>) a livros e objetos em geral (Amazon<sup>11</sup> e eBay<sup>12</sup>).

---

<sup>9</sup> <http://imdb.com>

<sup>10</sup> <http://last.fm>

<sup>11</sup> <http://amazon.com>

<sup>12</sup> <http://ebay.com>

De modo resumido, existem três tipos básicos de informações importantes para a recomendação: dados sobre o usuário, sobre os itens e informação transacional (históricos de sugestões e suas respectivas avaliações) (CAI; FRANCIS; GHEYSENS, 2009). São essas informações que são extraídas implícita ou explicitamente pelo sistema e, posteriormente, filtradas de modo que seja possível inferir comportamentos específicos sobre seus usuários.

### 3.1.1 Filtragem colaborativa

O termo *filtragem colaborativa* foi introduzido pela primeira vez em 1992, e faz referência ao processo de descoberta de padrões através do uso de técnicas que envolvem a colaboração entre os agentes ou fontes de informação, e que extrapolam sua utilização a diferentes aplicações como monitoramento de dados, exploração mineral, mercado financeiro, dentre outros (GOLDBERG ET AL, 1992)(RICCI ET AL, 2010).

No processo de recomendação, mais especificamente, filtragem colaborativa refere-se à forma de realizar previsões automáticas sobre os interesses dos usuários de um sistema por meio do processamento de informações sobre o comportamento de muitos outros indivíduos, traçando assim um perfil comportamental de um determinado grupo. Ou seja, baseia-se na premissa de que se um usuário  $a$  possui a mesma opinião de outro  $b$  sobre um determinado item  $x$ ,  $a$  possui uma maior probabilidade de ter uma opinião similar a  $b$  acerca de outro item  $y$  do que à opinião de um terceiro usuário escolhido aleatoriamente. Como esta técnica procura por padrões semelhantes de comportamento considerando o universo de todos os usuários, quanto maior forem as interações de uma pessoa com o sistema, e quanto mais usuários este possuir, melhores serão as recomendações.

A forma largamente utilizada para traçar perfis que representem o gosto dos usuários é a avaliação de itens públicos. Esta é uma maneira simples e suficiente para as pessoas fomentarem o sistema com suas opiniões e que possibilite refletir satisfatoriamente suas preferências. Esta avaliação pode ser realizada de inúmeras maneiras, porém, tipicamente, é binária: do tipo gostei, não gostei; ou classificativa, através da escala de Likert, com notas (pareceres) de um a cinco (MATELL; MICHAEL, 1971). Através destes dados, o indivíduo é disposto dentro do universo de usuários do sistema de modo que seus vizinhos possuam gostos e opiniões semelhantes. E são os itens mais bem classificados por estes vizinhos que lhe são recomendados.

Um exemplo de mensuração utilizada em filtragem colaborativa é o kNN (*k-nearest-neighbor*) (DUDANI, 1976). Neste algoritmo, o problema é modelado como uma matriz onde as linhas representam usuários e as colunas representam os itens. Assim, cada célula contém o valor que o usuário (linha) classificou o item (coluna) correspondente. Células vazias correspondem a itens aos quais não houve classificação. Desta forma, o problema resume-se em predizer os valores para estas células vazias.

Para realizar esta predição, é essencial encontrar os vizinhos mais próximos do usuário-alvo, obtendo assim uma base na qual a recomendação irá guiar-se. De modo a calcular a similaridade entre indivíduos e, portanto, encontrar aqueles mais próximos deste usuário-alvo, são utilizados diferentes cálculos. Os mais populares são a distância cosseno e a correlação de Pearson (MILLER et al, 2003), que traduzem numericamente a proximidade de dois usuários dentro do domínio. A predição então é realizada calculando uma nota média que o sistema de recomendação supõe que um usuário irá avaliar um determinado item baseando-se em seus vizinhos. A recomendação, por sua

vez, é uma lista das notas mais altas preditas pelo sistema para o usuário-alvo, dentre os itens aos quais não houve avaliação ou consumo anterior.

### 3.1.2 Filtragem baseada em conteúdo

A recomendação através da filtragem baseada em conteúdo possui suas raízes no campo da recuperação de informações (IR - *Information Retrieval*) e utiliza o histórico de itens consumidos (visitados/avaliados) pelo próprio usuário para elaborar novas sugestões de consumo (BALABANOVIC; SHOHAM, 1997). As informações utilizadas para tal são a descrição textual dos objetos e do perfil do usuário. Ou seja, um item é recomendado se o mesmo é similar a outros aos quais o usuário já tenha avaliado positivamente em um momento anterior. O raciocínio intuitivo por trás desta abordagem é: se uma pessoa gostou de algo no passado recente, ela tende a gostar de outras semelhantes no futuro.

O procedimento de recomendação varia conforme a representação dos itens envolvidos, que podem ser desestruturados: quando são armazenados em texto livre; semiestruturados: combinação de descrições livres com atributos específicos; ou estruturados: por exemplo, tabelas de atributos. Como o processamento de informações desestruturadas é mais custosa, eventualmente são utilizados algoritmos para converter um texto livre em uma informação estruturada, como o TF-IDF (*Term Frequency – Inverse Document Frequency*), que analisa o texto e elege termos de maior relevância para representa-lo (PAZZANI; BILLSUS, 2007).

Dentre os algoritmos de classificação de dados desestruturados mais utilizados podemos citar o método Bayesiano, reconhecido por seu bom desempenho em análises textuais, assim como variações do algoritmo de Rocchio (MITCHELL, 1997)(NIGAM et al, 1998). Para dados estruturados, tipicamente se utiliza árvores de decisão ou o kNN, complementado com a distância do cosseno ou distância euclidiana (HAND, 2007). Desta forma, são eleitos termos chaves para representar o gosto de um usuário de acordo com seu histórico de consumo.

Um exemplo de sistema de recomendação utilizando filtragem baseada em conteúdo é o PRES (acrônimo para *Personalized Recommender System*), um sistema para a recomendação de documentos. Os perfis de usuário são gerados através do histórico dos *feedbacks* dados pelos mesmos aos documentos analisados. Tanto os perfis de usuário quanto os documentos são representados por vetores que contém palavras-chave relevantes. A eleição das palavras representativas dos documentos é realizada com o TF-IDF. Assim, as recomendações para cada usuário são geradas através da comparação dos vetores e análise de características específicas dos documentos como novidade (se é recente ou não) e relevância (METEREN; SOMEREN, 2000).

### 3.1.3 Recomendação híbrida

Com o objetivo de oferecer sugestões mais precisas e driblar os problemas individuais enfrentados por cada uma das abordagens, foram concebidos algoritmos híbridos de recomendação combinando as duas técnicas apresentadas anteriormente. A aplicação de uma abordagem híbrida depende principalmente do domínio sendo explorado, sendo que existem inúmeras maneiras de combinar a filtragem colaborativa com a filtragem baseada em conteúdo ou, até mesmo, combinar diferentes algoritmos de uma mesma técnica.

Burke (2002), em seus estudos, identificou 53 maneiras possíveis de realizar esta junção e introduziu uma taxonomia para sistemas de recomendação híbridos,

classificando-os em sete categorias de combinação: ponderada, comutativa, mista, de recursos, de aumento de recursos, em cascata e em meta-nível. Burke procurou ainda mensurar e comparar o desempenho de diversos tipos de algoritmos híbridos e tradicionais e obteve resultados superiores com abordagens mistas, salientando o seu potencial.

Dentre os principais expoentes acadêmicos híbridos de recomendação pode-se citar o algoritmo *Content-Boosted Collaborative Filtering*, um *framework* modular que utiliza a filtragem baseada em conteúdo para aprimorar dados existentes dos usuários expressos na matriz *usuários × itens* que, por sua vez, são recomendados através de filtragem colaborativa. Devido à sua natureza modular, são mencionados ainda outros tipos de composição das duas técnicas que podem ser exploradas e trazer resultados superiores às abordagens isoladas. (MELVILLE, 2002)

O estudo de Claypool et al (1999), por sua vez, detalha uma abordagem que produz recomendações separadas através da filtragem colaborativa e baseadas em conteúdo para posteriormente, através de uma média ponderada das previsões geradas por ambas, selecionar aquelas mais adequadas ao usuário em questão.

Abordagens híbridas são especialmente interessantes quando o domínio de aplicação produz informações relevantes tanto do histórico individual de atividade e/ou consumo quanto às relações dos usuários entre si. Nestes cenários, a combinação das duas abordagens pode resolver ou ao menos minimizar as principais dificuldades dos sistemas de recomendação, como será detalhado na próxima subseção.

### 3.2 Análise comparativa e principais problemas

Sistemas de recomendação apresentam alguns problemas práticos que podem ficar mais ou menos evidentes dependendo do domínio de aplicação e abordagem utilizada. Vozalis (2003) elencou os principais desafios encontrados pelos sistemas de recomendação atualmente. Reconhecer e driblar estas dificuldades são essenciais para obter sugestões mais eficientes. São eles:

- **Qualidade das recomendações:** diz respeito à confiança dos usuários nas recomendações propostas. Sugestões mal preditas irão gerar descrenças e, conseqüentemente, levar os usuários à ignorarem futuras recomendações.
- **Esparsidade dos dados:** problema comum em vários domínios onde as avaliações públicas de itens são poucas, em relação ao montante total de itens. Isto resulta em matrizes *usuários × itens* esparsas que, por sua vez, tornam as previsões imprecisas.
- **Escalabilidade:** sistemas de recomendação realizam cálculos que dependem tanto do número de usuários quanto de itens do domínio. Algoritmos eficazes devem prover alternativas para tratar adequadamente o crescimento de dados para aplicações em grandes bases de dados.
- **Recomendações transitivas:** a correlação transitiva de usuários é uma importante forma de identificar potenciais recomendações. Algoritmos eficientes devem ser capazes de reconhecer tais relações.
- **Sinônimos:** é importante que os sistemas de recomendação consigam identificar a utilização de sinônimos para descrever um mesmo item.

- **Problema da partida fria** (*the cold-start problem*): este problema afeta novos itens ou usuários do sistema, quando não existem informações sobre os mesmos para fomentar as recomendações.
- **Problema do usuário incomum** (*the gray sheep problem*): faz referência aos usuários que fogem dos padrões de comportamento da maioria, fazendo com que as recomendações para os mesmos sejam imprecisas.

Eventualmente, os pontos relacionados acima referem-se somente a uma das abordagens e não se aplicam à outra. Às vezes são limitadas também ao domínio de aplicação. Porém, quando pertinentes, sistemas de recomendação devem procurar tratar estas questões. A filtragem baseada em conteúdo, por exemplo, é aplicada com sucesso principalmente na recomendação de itens em domínios textuais e, por isso, sofre mais com o uso de sinônimos para descrição de itens semelhantes, o que acontece em menor grau na filtragem colaborativa. Entretanto, possui duas vantagens importantes: (i) não necessita de avaliações de usuários para iniciar o processo de recomendação; e (ii) não apresenta o problema de esparsidade de dados.

As duas vantagens citadas acima se devem ao fato de que a filtragem baseada em conteúdo compara o perfil do usuário com a descrição do item, e não às avaliações de outros usuários. Portanto ela pertence a um domínio textual, onde as informações sobre os itens estão claras e disponíveis (ao contrário da filtragem colaborativa, onde necessitamos de uma série de avaliações para podermos sugerir determinado item a um usuário). A filtragem colaborativa, por sua vez, apresenta três vantagens importantes sobre a filtragem baseada em conteúdo: (i) independência sobre a natureza dos itens; (ii) habilidade para filtrar itens baseando-se em aspectos subjetivos; e (iii) capacidade de recomendar itens inesperados (BALABANOVIC; SHOHAM, 1997)(GOOD et al, 1999).

Primeiramente, como a avaliação de itens é realizada diretamente pelos usuários, ou seja, pelo ser humano, não existe restrição sobre a natureza do item a ser recomendado. Segundo, a filtragem colaborativa consegue expandir sua análise a pontos subjetivos, que são difíceis de serem analisados computacionalmente, desde que exista um retorno sobre estes quesitos no sistema. Ou seja, é possível recomendar músicas de melodia triste ou um livro bem escrito. E, por fim, a filtragem colaborativa pode recomendar itens que não são esperados pelo usuário. Isto ocorre pelo fato que são comparados os usuários em si, e não os itens. Mesmo que um usuário goste de rock, por exemplo, ele pode receber uma recomendação sobre uma música de jazz, devido ao fato que muitos usuários com perfis semelhantes ao seu tenham avaliado positivamente esta música de jazz.

Algoritmos híbridos, por sua vez, apresentam vantagens das duas abordagens, em maior ou menor grau, dependendo de sua aplicação e combinação. A principal desvantagem da abordagem híbrida é que, por envolver ambas as técnicas, ela necessita de uma quantidade de dados maior que as demais para realizar recomendações boas dentro de seu contexto e, por isso, é computacionalmente mais custosa.

A Tabela 3.1 mostra um resumo comparativo das abordagens de recomendação com as características mencionadas acima e outros pontos importantes retirados de estudos relacionados (BREESE; HECKERMAN; KADIE, 1998)(ZAIER; GODIN; FAUCHER, 2008)(SARWAR et al, 2001).

Tabela 3.1: Vantagens e desvantagens de cada uma das abordagens de recomendação

	Vantagens	Desvantagens
<b>Filtragem colaborativa</b>	<ul style="list-style-type: none"> <li>• Independência sobre a natureza dos itens;</li> <li>• Habilidade para filtrar itens baseando-se em aspectos subjetivos;</li> <li>• Capacidade de recomendar itens inesperados.</li> </ul>	<ul style="list-style-type: none"> <li>• Problema da partida fria possui maior impacto</li> <li>• Apresenta o problema de esparsidade dos dados.</li> </ul>
<b>Filtragem baseada em conteúdo</b>	<ul style="list-style-type: none"> <li>• Não necessita de avaliações de usuários para iniciar o processo de recomendação;</li> <li>• Possibilidade de personalizar as recomendações.</li> </ul>	<ul style="list-style-type: none"> <li>• O processamento textual dos itens é custoso;</li> <li>• Depende das descrições objetivas dos itens;</li> <li>• Não identifica subjetividade na preferência dos usuários;</li> </ul>
<b>Abordagem híbrida</b>	<ul style="list-style-type: none"> <li>• Recomendações mais precisas;</li> <li>• Independência sobre a natureza dos itens;</li> <li>• Habilidade para filtrar itens baseando-se em aspectos subjetivos;</li> <li>• Possibilidade de recomendar itens inesperados.</li> </ul>	<ul style="list-style-type: none"> <li>• Necessário informações de um número maior de fontes.</li> <li>• Computacionalmente mais custosa.</li> </ul>

O uso de *folksonomia* também tem sido empregado na última década para categorizar itens em ambientes colaborativos específicos. Este termo foi criado por Vander Wal (2013) na junção das palavras *folk* e *taxonomy*. Basicamente, é um sistema de classificação colaborativo que utiliza *tags* (etiquetas) para categorizar conteúdos. Por vezes esta prática também é conhecida como *social tagging*. As *tags* são utilizadas como palavras-chave relevantes para cada item do domínio, que podem ter diferentes graus de importância.

### 3.3 Recomendação de pessoas e parceiros

Estudos da Psicologia Social demonstram que, em geral, as pessoas possuem uma predisposição maior à aceitação de outras que possuam características e gostos semelhantes aos seus. Ou seja, criamos uma empatia maior àquelas pessoas de cultura e interesses pessoais parecidos aos nossos (COLEMAN; COLEMAN, 1994). Sob esta premissa, e impulsionados com a difusão das redes sociais virtuais e a consequente virtualização das relações humanas, novos desafios na área de sistemas de recomendação emergiram. Essas mudanças no estilo de vida do homem moderno trouxe uma grande quantidade de serviços que oferecem oportunidades de conhecer outras pessoas buscando um objetivo comum: de amizades a encontros românticos ou sexuais.

#### 3.3.1 Recomendação de pessoas

Independentemente do domínio de aplicação, a recomendação de pessoas difere-se em vários aspectos da recomendação de objetos comuns. A mais evidente consiste no fato de que as recomendações (e compatibilidades) acontecem de maneira bidirecional: um usuário  $x$  deve despertar interesse de um usuário  $y$  e assim como  $y$  deve interessar a

x. Caso contrário, a relação não ocorrerá. Além disso, aspectos subjetivos das relações devem ser levadas em consideração para sintonizar melhor as recomendações, assim como as rejeições.

Recomendação de pessoas têm sido aplicada com sucesso principalmente em *websites* para encontros amorosos e de novas amizades, como eHarmony<sup>13</sup>, OkCupid<sup>14</sup> e match.com<sup>15</sup>. Muitas vezes é referenciada na literatura como *social matching*: termo originado na teoria de mesmo nome e importada para sistemas computacionais. Trata-se da mesma questão de recomendação de pessoas, onde o objetivo final é aumentar as interações sociais dos usuários e também impulsionar a colaboração (TERVEEN; MCDONALD, 2005).

A prática demonstrou que as técnicas tradicionais de recomendação focam-se demasiadamente em comportamentos explícitos dos usuários, ignorando suas relações implícitas (NAYAK; ZHANG; CHEN, 2010). O trabalho de Chen et al (2009) relata a condução de um estudo com 3000 usuários onde foi realizada uma análise comparativa de quatro algoritmos muito utilizados de recomendação de pessoas, sendo dois deles utilizando filtragem baseada em conteúdo: *Content Matching* e CplusL (*Content-plus-Link*); e dois baseados em informações sobre as relações sociais dos usuários: FoF (*Friend-of-Friend*) e SONAR. Nesta análise, que foi realizada dentro do domínio específico de procura por amizades, verificou-se que os algoritmos baseados em filtragem colaborativa são mais eficazes na procura por contatos conhecidos (devido à rede de colaborações), enquanto os que utilizam filtragem baseada em conteúdo trazem bons resultados na recomendação de novas amizades. Este fato sugere que uma abordagem híbrida de recomendação talvez seja a melhor alternativa para recomendar pessoas, como também concluiu Kazienko (2006) e McDonald (2003) ao utilizar algoritmos híbridos de recomendação de pessoas utilizando como base a autodescrição dos usuários, comparação de atributos específicos do domínio, relações com os outros usuários e seu histórico comportamental no sistema.

### 3.3.2 Recomendação de parceiros

Em alguns domínios específicos, a recomendação de pessoas é denominada de recomendação de colaboradores (GUNAWARDENA; WEBER, 2009) ou parceiros. Pode-se afirmar que estes termos são, na verdade, subconjuntos de sistemas de recomendação, onde o objetivo final é aproximar usuários para uma atividade colaborativa comum. Os fatores influentes decisivos para uma recomendação interessante neste contexto variam conforme seu domínio. Considerando especificamente a composição musical colaborativa, por exemplo, fatores importantes são o objetivo pessoal de cada um no ambiente, o gosto musical de cada colaborador e sua flexibilidade em relação às ideias musicais dos parceiros.

Embora na literatura o termo *recomendação de parceiros* não seja muito difundido, alguns estudos importantes já abordam o tema com esta nomenclatura. O estudo de Wu, Sun e Tang (2003) propõe um *framework* para a recomendação de parceiros para desenvolvimento e patenteamento de invenções e produtos de maneira colaborativa. Neste sistema, perfis de usuários são criados segundo suas produções e referências e

---

<sup>13</sup> <http://www.eharmony.com>

<sup>14</sup> <http://www.okcupid.com>

<sup>15</sup> <http://www.match.com>

então são comparados através de um cálculo de similaridade para, finalmente, serem recomendados. O algoritmo ainda possui um módulo de aprendizado interativo, que refina as recomendações de acordo com o *feedback* dado pelos usuários.

Chen, Hong e Chang (2008) apresenta uma abordagem para a recomendação de parceiros em ambientes colaborativos de aprendizado *online* (*web learning*), tipo de atividade que também está se tornando popular recentemente. O estudo explora dados retirados do ambiente social virtual para elevar a compatibilidade entre parceiros de estudos e, assim, aumentar o desempenho e o resultado do aprendizado colaborativo.

Especificamente no contexto desta dissertação, recomendação de parceiros refere-se à sugestão automatizada de pessoas para trabalhar em conjunto em ambientes colaborativos de composição musical. Neste sentido, é necessário identificar quais são os fatores individuais que levam os agentes desta atividade coletiva a obter resultados satisfatórios. Diferentemente de *websites* de relacionamento pessoal, empatia, neste caso, não é essencial. Como ressaltado por Miletto (2010), é importante principalmente que os usuários tenham objetivos sintonizados e comunicação argumentativa satisfatória para discutir as ideias de cada um. Estes e outros aspectos subjetivos da colaboração em ambientes musicais serão discutidos no próximo capítulo.

## 4 COLABORAÇÃO NA COMPOSIÇÃO MUSICAL: PESQUISA COM USUÁRIOS DE WEBSITES DE COMPOSIÇÃO COLABORATIVA

Este capítulo traz uma discussão sobre fatores importantes da composição musical colaborativa e descreve uma pesquisa quantitativa conduzida com usuários de *websites* desta categoria, onde são apresentados sua motivação, com uma análise de alguns estudos sobre composição musical; sua metodologia, descrevendo como foi elaborado o questionário submetido aos entrevistados; e os resultados obtidos, ressaltando os principais fatores que impactam diretamente no relacionamento de colaboradores dentro de tais ambientes, além de depoimentos relevantes dos respondentes à pesquisa.

### 4.1 Motivação

O processo mental relacionado à atividade de composição musical ainda é objeto de grande incompreensão humana. Sua natureza abstrata faz com que não exista regra, teoria ou roteiro definido para criar música. Cada compositor pode explorar uma infinidade de combinações através de diferentes e também infinitos caminhos, tornando uma composição complexa inconscientemente irrepetível (FERNEYHOUGH, 1995). Como confirma Reynolds (2013), a música é um reflexo da personalidade de seu criador que, por sua vez é influenciada por diversos elementos de origem social, cultural e pessoal. Assim, cada criação musical é um objeto único no tempo e espaço.

Tratando-se de composições colaborativas, algumas questões surgem neste contexto abstrato, como observado no final do capítulo anterior: como aproximar pessoas para compor em conjunto? E como identificar compositores compatíveis e com objetivos convergentes? Para responder a estas perguntas é primordial entendermos os diversos fatores que impactam no relacionamento entre compositores que cooperam entre si, assim como identificarmos os objetivos comuns de cada indivíduo participante deste processo.

A complexa influência da música na vida do ser humano já motivou diversos estudos da *psicologia da música*. Este termo, às vezes referenciado como *psicologia musical*, é um ramo da psicologia que tem como objetivo explicar e entender o comportamento humano relacionado à experiência musical. Ou seja, compreende aspectos como o entendimento das funções da música sob as perspectivas pessoal, social e cultural; o papel dela como agente de expressão de sentimentos e emoções; a musicoterapia; assim como os aspectos psicológicos relacionados à composição e improvisação musicais.

A atividade de composição musical, especificamente, sempre foi um domínio de difícil mensuração, modelagem e estudo. Esta talvez seja a explicação mais razoável

para o fato que, embora exista um crescente interesse na percepção e cognição musicais, exista uma falta impressionante de pesquisas acerca do processo psicológico relacionado à composição em si, como afirma Brown (2003) e Inpett (2008).

Apesar desta falta de pesquisas, alguns indicativos relacionados à atividade de composição musical colaborativa em crianças apontam resultados muito positivos quando estes trabalham em conjunto frente à composição individual. MacDonald e Miell (2000) estabelecem uma relação direta entre amizade e comunicação à qualidade das criações musicais. Foi observado que entre amigos, há mais troca de informações e liberdade para experimentação. Como consequência, estas composições foram classificadas como de maior qualidade que aquelas criadas por pares de crianças que não se conheciam. Conclusão semelhante foi elaborada por MacDonald, Byrne e Carlton (2006) onde alguns fatores impactantes no relacionamento entre os parceiros e, conseqüentemente, nas composições sendo criadas foram apontados, como a identificação cultural e o relacionamento entre os pares. Crianças se sentem mais à vontade para experimentar e comunicar com amigos próximos e com crianças que possuem maior identificação cultural, resultando assim em composições mais interessantes.

Estes estudos, porém, focam-se principalmente na educação musical de crianças de maneira colaborativa e, portanto, não são exatamente aplicáveis em ambientes *online*, multiculturais e voltados a adultos. Assim, foi idealizada e conduzida uma pesquisa quantitativa entre usuários de serviços de composição musical colaborativa *online* que fosse capaz de refletir as motivações que levam uma colaboração nestes ambientes a um resultado satisfatório a todos os envolvidos. Os dados colhidos forneceram uma base para a concepção da abordagem de recomendação de parceiros proposta.

O principal agente motivador que impulsionou esta pesquisa inicial é a falta de referências sobre os elementos mais importantes que resultam em composições colaborativas de sucesso. Entende-se como sucesso, neste caso, um resultado satisfatório a todos os compositores. O questionário concebido teve como objetivo o entendimento de três pontos principais, que seguem:

1. Traçar um perfil do usuário típico de ambientes colaborativos de composição musical, contemplando tempo de utilização, experiência musical e os objetivos individuais que o leva à utilização de tais sistemas;
2. Entender o que espera um compositor ao buscar pessoas para trabalhar em conjunto em um projeto musical de sua autoria;
3. Compreender a influência do gosto musical dos colaboradores na composição sendo compartilhada e seu impacto nas relações interpessoais, assim como a flexibilidade e aceitabilidade dos usuários neste quesito.

Atingindo estes objetivos elementares, seria possível traçar uma abordagem de recomendação eficiente, que realmente fosse capaz de aproximar pessoas para compor em conjunto. Assim, esta pesquisa foi conduzida de maneira simples e objetiva, realizada através de um questionário acessível aos usuários de tais serviços. A motivação da pesquisa em si também foi apresentada de forma transparente ao entrevistado, deixando claros os objetivos do estudo (GUNTHER, 2006).

É importante salientar que o universo de serviços *online* de composição colaborativa é consideravelmente limitado, sendo dominado majoritariamente por dois *websites* (Indaba e Kompoz). Em contrapartida, tais ambientes contém um número significativo

de usuários, facilitando a busca por respondentes. Outro quesito importante é que, de maneira geral, os usuários interrogados de tais sistemas já possuem alguma experiência musical. Isto deve-se ao fato que todos os *websites* abrangidos na pesquisa são destinados à músicos profissionais ou amadores, e não à leigos.

## 4.2 Metodologia

Para obter um número significativo de respostas, o questionário foi disponibilizado na Internet, em inglês, de modo que fosse possível responde-lo em questão de poucos minutos. Assim, foi elaborado com oito perguntas de múltipla escolha objetivas e duas questões opcionais dissertativas, abrindo espaço para aquelas pessoas que gostariam de expressar melhor suas experiências.

A divulgação foi realizada principalmente através dos fóruns de discussão do Indaba e do Kompoz, os dois *websites* mais populares de composição colaborativa. Também foram enviados pedidos de contribuição nas páginas do Facebook do Merge.fm<sup>16</sup> e SoundKeep<sup>17</sup>. A Tabela 4.1, a seguir, traz um resumo traduzido das perguntas realizadas no questionário da pesquisa. A versão original do mesmo pode ser visualizada no Anexo A, no final desta dissertação.

Tabela 4.1: Resumo e tradução das perguntas do questionário da pesquisa realizada.

Número	Questão	Características
1	Quais <i>websites</i> de composição musical colaborativa você utiliza?	- Múltipla escolha; - Múltiplas respostas; - Resposta editável.
2	A quanto tempo você utiliza <i>websites</i> de composição colaborativa?	- Múltipla escolha; - Única resposta.
3	Qual é a sua experiência com música?	- Múltipla escolha; - Resposta única; - Resposta editável.
4	Qual fator mais lhe influenciou a utilizar <i>websites</i> de composição colaborativa?	- Múltipla escolha; - Resposta única; - Resposta editável.
5	Geralmente você cria projetos novos ou contribui para projetos existentes de outros usuários?	- Múltipla escolha; - Resposta única.
6	Em um projeto de sua criação, o quão flexível você é sobre as contribuições propostas por outros usuários?	- Múltipla escolha; - Resposta única; - Resposta editável.
7	Em relação ao seu gosto musical, indique sua preferência sobre os colaboradores de seus projetos.	- Múltipla escolha; - Resposta única; - Resposta editável.
8	Escreva sobre sua experiência compondo <i>online</i> com outras pessoas.	- Dissertativa; - Opcional.
9	Sugestões e observações sobre esta pesquisa.	- Dissertativa; - Opcional.

<sup>16</sup> <https://www.facebook.com/mergefm>

<sup>17</sup> <https://www.facebook.com/Soundkeep>

As questões de número 1 a 3 pretendiam traçar um perfil do usuário típico de *websites* de composição musical colaborativa. A primeira delas procurava especificamente saber qual ou quais os serviços de composição colaborativa o respondente participa, oferecendo a possibilidade de selecionar múltiplas respostas, assim como uma resposta editável pelo usuário. A pergunta de número 2 questionava há quanto tempo o usuário utilizava tais sistemas, enquanto a questão de número 3 complementava as duas anteriores perguntando a experiência do usuário com música em geral.

As questões de número 4 a 7 focaram especificamente a atividade de composição musical colaborativa. A pergunta 4 objetivou traçar um panorama sobre os fatores que levam as pessoas a buscar estes serviços, permitindo uma resposta editável pelo usuário. A questão 5, por sua vez, buscou captar a preferência dos mesmos como atores ativos de novas composições ou como colaboradores de projetos existentes. Na pergunta 6, procurou-se entender o quão tolerantes são os usuários em relação às contribuições de seus colaboradores. Ou seja, se procuram contribuições específicas que respeitam suas ideias iniciais, ou permitem qualquer tipo de colaboração. A última questão objetiva, de número 7, pretendia complementar a anterior e permitia aos entrevistados selecionar, dentre as respostas, o perfil desejado de seus colaboradores no que diz respeito ao seu gosto musical específico.

Para finalizar, foram ainda apresentadas as questões 8 e 9, ambas opcionais e dissertativas, para os entrevistados escreverem livremente sobre suas experiências e impressões compondo colaborativamente na Internet e acerca da pesquisa conduzida em si.

### **4.3 Resultados**

O questionário esteve disponível na Internet durante 20 dias, entre 11 de Junho e 1º de Julho de 2013 e foi respondido por 52 pessoas. É importante salientar o fato que uma parcela considerável dos entrevistados (48%) utilizam mais de uma plataforma para compor *online*, sendo Kompoz o serviço de maior popularidade, com 92% de utilização (Figura 4.1). Tal resultado, muito provavelmente, deve-se ao modo de divulgação da pesquisa e não necessariamente reflete a real quantidade de usuários ativos dos *websites* abrangidos. No Kompoz, diferentemente dos outros serviços, foi possível divulgar nos grupos de discussão específicos, que são bastante ativos, sem necessidade de moderação e bloqueio de *links* externos.

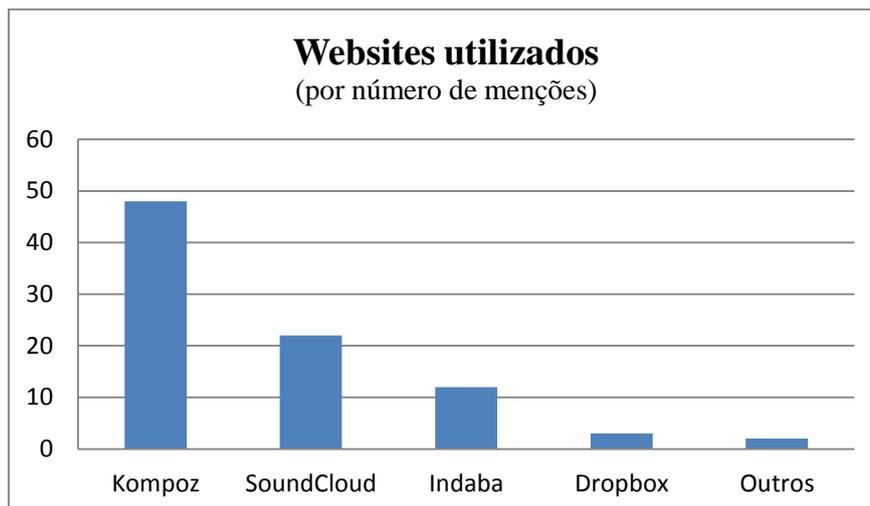


Figura 4.1: Distribuição dos *websites* utilizados pelos respondentes à pesquisa.

A maioria do público entrevistado se define com considerável experiência musical (59%), sendo que quase um quarto (23%) são músicos profissionais, conforme se observa na Figura 4.2. Tal fato pode ser um reflexo da grande quantidade de respondentes que utilizam o Kompoz, *website* focado em usuários com maiores conhecimentos de música. Outra explicação para a baixa quantidade de músicos amadores é a possível rejeição dos músicos experientes e mais frequentes, ao público novato, que talvez não contribua da maneira esperada nos projetos. Isto demonstra também a importância de oferecer ferramentas de apoio à composição colaborativa específicas a pessoas que possuem pouca ou nenhuma experiência musical, visto que os objetivos desejados por este grupo divergem consideravelmente daqueles mais experientes.



Figura 4.2: Experiência musical relatada dos entrevistados.

A maior experiência do público entrevistado também se torna evidente pelo tempo de utilização deste tipo de serviço. Enquanto apenas um usuário respondeu que iniciou a compor de forma colaborativa há menos de três meses, o que corresponde a apenas 2% do total; 89% possuem mais de dois anos de experiência com este tipo de ferramenta; e 64%, mais de três anos. Neste ponto é possível observar também que a maioria do público possui considerável experiência tanto em música em si quanto em ferramentas de composição. Mas apesar de positivo o fato que exista uma grande parcela de usuários que utilizam há muitos anos (usuários ativos de longa data), talvez exista uma necessidade de adaptar estas ferramentas de modo a atrair e manter o usuário com pouca ou nenhuma experiência musical.

Em relação à preferência do público entre contribuir com projetos de outros usuários ou iniciar novos projetos (questão 5), houve uma distribuição praticamente equilibrada das respostas, com leve predominância por colaboração em projetos de outros usuários, como mostra a Tabela 4.2.

Tabela 4.2: Distribuição da preferência dos respondentes quanto à criação de projetos próprios ou contribuição a outros usuários.

<b>Preferência dos usuários</b>	<b>Percentual dos respondentes</b>
Somente criam projetos próprios.	5,7%
Em geral criam projetos próprios, mas também contribuem eventualmente.	21,2%
Criam projetos próprios e contribuem de maneira equitativa.	32,7%
Em geral contribuem, mas eventualmente também criam projetos próprios.	30,8%
Somente contribuem em projetos de outros usuários.	9,6%

Apesar dos *websites* pesquisados possibilitarem tanto composições individuais, quanto colaborativas, a maioria dos respondentes (71%) os elege por seu caráter colaborativo. Nesse sentido, essas ferramentas acabam atraindo um perfil de usuário que costuma ser muito flexível às colaborações de desconhecidos em suas composições. Os dados levantados demonstram que 73% dos respondentes se julgam flexíveis, ou muito flexíveis quanto às contribuições de seus colaboradores, independentemente do tempo de uso da ferramenta (Figura 4.3). A partir desses dados, podemos afirmar que a própria natureza deste tipo de ferramenta trás consigo o apelo pela flexibilidade. Certamente compositores muito restritos às suas ideias preferem compor sozinhos.

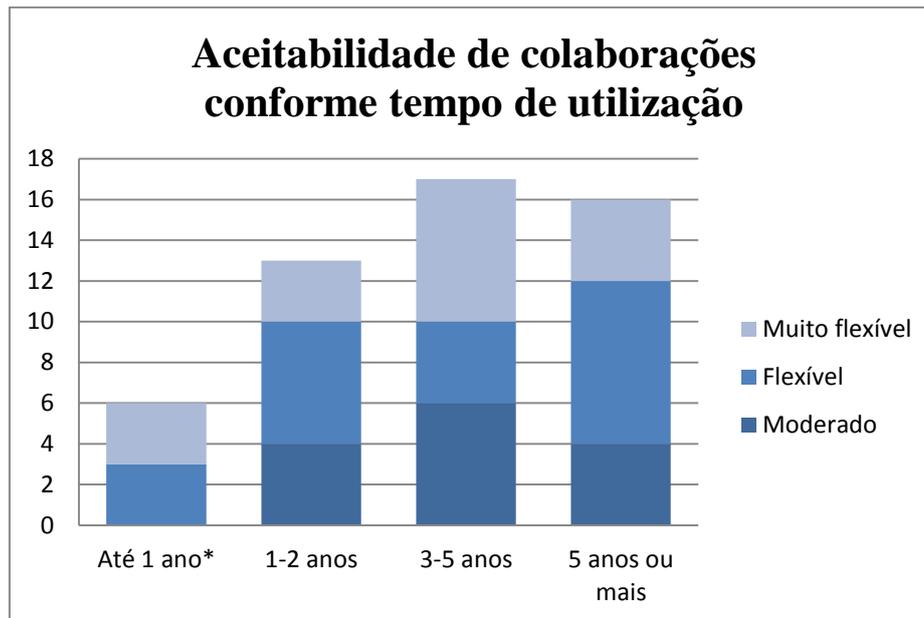


Figura 4.3: Aceitabilidade de colaborações dos usuários conforme seu tempo de utilização em *websites* de composição colaborativa.

Nas respostas acerca das preferências de gênero musical dos colaboradores, houve grande divergência de opiniões, com uma ligeira inclinação por aqueles que possuem gosto semelhante. Enquanto 50% alega preferência por colaborações de pessoas que tenham gosto similar ou muito similar; 36% faz questão de que seus parceiros possuam gosto musical diferente ou muito diferente. Os restantes 14% não possuem opção específica, mantendo-se focados no resultado da composição. Estes dados divididos tornam evidentes as necessidades de adaptação das técnicas tradicionais de sistemas de recomendação no quesito similaridade: ao recomendar parceiros, é essencial entender (implícita ou explicitamente) a preferência e tolerância de cada envolvido acerca dos gostos pessoais dos seus colaboradores. Um compositor que busca por algo inesperado e inovador, por exemplo, certamente aceitará parceiros que possuam experiências e gostos musicais diferentes.

#### 4.3.1 Relatos de experiência

Os relatos de experiência dos entrevistados evidenciaram aspectos positivos dos ambientes de composição colaborativa, como o fato de possibilitar o encontro e troca de ideias entre pessoas com diferentes perfis e características, de partes remotas do globo. Os depoimentos também foram particularmente importantes para compreender alguns pontos práticos da composição colaborativa, como a comunicação entre os colaboradores para obter um resultado interessante. A aceitação e flexibilidade em relação às colaborações também ficaram evidentes, em diferentes níveis. A seguir serão apresentadas algumas transcrições em tradução livre dos depoimentos sobre aspectos relevantes dentro do contexto da recomendação de parceiros. Uma lista com todos os relatos originais, em inglês, pode ser encontrada no Anexo B, ao final desta dissertação.

A quase totalidade dos depoimentos foram positivos e entusiásticos, como o relato do entrevistado 39, abaixo, que faz também uma boa síntese de alguns outros comentários.

Compor *online* é incrível e é uma parte importante da evolução musical no mundo de hoje. Acredito que o público em geral, e muitos músicos não

perceberam isso ainda, mas este é o futuro da música e a quebra dos gêneros musicais, onde as ideias de um homem ou mulher são colididas nas ideias de outro homem ou mulher de algum canto do mundo para criar algo singular e global que não é guiado por dinheiro [...] mas somente pela paixão verdadeira à música e por criar algo que é único.

Eu já colaborei com vários músicos locais e é difícil ser melhor que trabalhar na mesma sala que alguém. Porém às vezes trabalhar sozinho em uma ideia de outra pessoa pode ser ótimo também. Assim você não precisa lidar com egos ou ser guiado pela ideia de outras pessoas sobre como você deveria estar tocando ou compondo.

A maior parte de minhas experiências compondo online foram incríveis. Um aspecto negativo talvez seja o desafio de expressar fielmente suas ideias através de e-mails [...].

Além deste aspecto internacional, singular e livre da composição colaborativa *online* citada no depoimento acima e presente também em alguns outros comentários, um fato que merece destaque entre os relatos é como a colaboração incentiva o estudo da música, como descrito pelo entrevistado 4:

Eu cresci muito utilizando Kompoz [...]. Esta experiência também me trouxe confiança em meu trabalho e também em meu desempenho. Meus colaboradores têm me apoiado muito e também me influenciado com sua diversidade de experiências musicais.

E também pelo depoimento 49:

Colaborar online é uma ótima plataforma de aprendizado, tanto para músicos como para produtores. É também muito gratificante a nível pessoal, apesar de não ajudar a promover você como artista.

Os relatos de pontos negativos evidenciaram principalmente a dificuldade de encontrar instrumentistas específicos, como bateristas (depoimento 48) ou vocalistas (depoimento 6); além de problemas de comunicação e expressão de ideias (depoimentos 11, 46 e 51); e dificuldade em encontrar parceiros para colaborar em projetos, como no depoimento 38: “Um problema das composições online é que as pessoas vêm e vão, então nem sempre é fácil achar pessoas que colaboram em suas canções”, e no depoimento 48:

Têm sido difícil encontrar pessoas dedicadas e/ou com tempo para contribuir. É necessário muito tempo para que, daqueles que fazem requisições para participarem das minhas sessões, realmente contribuam.

Com estes relatos é possível observar que, mesmo tendo diversos fóruns e grupos específicos dentro destes ambientes, muitas vezes os usuários têm dificuldade em encontrar parceiros para compor em conjunto. E as buscas por colaboradores também exigem um esforço de tempo considerável por parte dos usuários, o que torna a tarefa desgastante. São exatamente estas dificuldades em encontrar colaboradores dispostos a contribuir em projetos específicos que a abordagem de recomendação de parceiros procura suprir.

#### **4.4 Considerações finais sobre a pesquisa**

A partir da análise dos dados resultantes da pesquisa conduzida e dos depoimentos dos usuários descritos neste capítulo, é possível identificar três pontos principais que são essenciais para a recomendação de parceiros em ambientes colaborativos virtuais de composição musical, que são:

1. **Preferência musical:** o gosto pessoal de cada usuário tem uma influência importante na colaboração musical. Assim, deve ser possível reconhecer e identificar as preferências de gênero e estilo dos usuários, de maneira implícita ou explícita, para uma recomendação mais eficaz.
2. **Flexibilidade:** de maneira geral, os usuários são bastante flexíveis quanto às contribuições de seus colaboradores. Porém, esta característica varia dentro de certos limites para cada indivíduo. Alguns usuários fazem questão de buscar parceiros com influências completamente diferentes para colaboração; outros se restringem a um ou mais gêneros. Desta forma, é necessário rever a questão de similaridade das metodologias de recomendação existentes para compreender esta flexibilização.
3. **Objetivos:** as pessoas divergem e se complementam nos objetivos individuais dentro do ambiente colaborativo. Alguns usuários preferem contribuir em projetos iniciados por outros do que criar projetos próprios; alguns ainda preferem somente iniciar projetos de sua autoria; enquanto a maioria se distribui entre estes dois extremos.

Estes pontos constituem a principal orientação para a abordagem de recomendação de parceiros detalhada no próximo capítulo.

## 5 ABORDAGEM DE RECOMENDAÇÃO DE PARCEIROS

Neste capítulo estão detalhadas cada uma das etapas que compõe a solução de recomendação de parceiros para ambientes colaborativos de composição musical. A abordagem está exposta de maneira tão genérica quanto possível, dentro do contexto da solução. Ou seja, está demonstrada de modo que seja possível adapta-la com pouco esforço a qualquer sistema já existente.

A primeira seção deste capítulo, 5.1, traz uma breve discussão acerca da classificação de músicas em gêneros musicais e um resumo das notações utilizadas durante o capítulo. As seções que seguem descrevem o processo de recomendação de parceiros em si, que é composto por três etapas principais: a primeira, detalhada na seção 5.2, aborda os procedimentos necessários para criação de um perfil musical de um usuário-alvo; a seção 5.3, por sua vez, detalha a etapa de agrupamento de usuários e os cálculos de similaridade e, finalmente, a seção 5.4 explana os algoritmos de seleção de usuários e criação efetiva da lista de recomendações.

### 5.1 Representação musical

É importante neste momento definir uma forma de representar as músicas em gêneros de modo que reflita suas características de maneira eficiente para o cálculo das recomendações. A tarefa de classificação das músicas presentes no ambiente será considerada realizada, seja através do trabalho humano ou de forma automática, com a utilização de algoritmos de classificação automática, que analisam a onda sonora para obter características musicais específicas (BURRED; LERCH, 2003)(XU et al, 2003).

Cabe salientar, neste sentido, que não existe uma classificação exaustiva de gêneros musicais. Alguns modelos de taxonomia foram propostos na literatura, procurando categorizar o maior número de gêneros possíveis, focados principalmente na automação do processo (AUCOUTURIER; PACHET, 2003)(PACHET; CAZALY, 2000)(LI; OGIHARA, 2005). Porém, diferentes ambientes utilizam organizações específicas adaptadas cada uma ao seu contexto. Existem classificações simples, não hierarquizadas, como o Music Match Jukebox, que categoriza 34 gêneros distintos; e outros organizados hierarquicamente com alguns gêneros principais (raízes), e subdivisões mais específicas, como a Amazon, que possui 23 níveis principais e centenas de subníveis (LI; OGIHARA, 2005).

Algumas formas de classificação, ainda, podem não refletir fielmente algumas características específicas, como quando se atrela exclusivamente uma música a um determinado gênero. Isso ocorre porque, na verdade, uma música pode pertencer a diferentes gêneros musicais concomitantemente. A música *A little Less Conversation*, gravada por Elvis Presley, por exemplo, pertence a gêneros como Rock e Rockabilly. Sua versão remixada pelo DJ holandês Tom Holkenborg, conhecido como Junkie XL ou

JXL, e que fez relativo sucesso, contém ainda elementos de música eletrônica, pertencendo, até certo ponto, também a este gênero.

Com estas considerações, optou-se por uma representação baseada em vetores, estrutura também utilizada em outros sistemas de recomendação, devido à naturalidade da distribuição de gêneros neste formato, e à facilidade da manipulação matemática necessária para os cálculos de recomendação. Assim, uma composição será representada por um vetor  $C$ , de dimensão  $m$ , onde  $m$  corresponde ao número de gêneros musicais do ambiente. Os valores que compõe  $C$  são percentuais de pertencimento àquele determinado gênero musical. Além da representação vetorial, uma composição pode possuir uma série de descritores textuais, ou *tags*, para representar características que não são compreendidas por gêneros específicos, mas que também podem ser utilizadas na recomendação, como, por exemplo, *vocal feminino*, *euforia* ou *dançante*. A Tabela 5.1, abaixo, traz uma síntese das convenções e notações que serão utilizadas ao longo deste capítulo para explicar a abordagem apresentada.

Tabela 5.1: Convenções representativas para descrição da abordagem de recomendação.

Convenção	Descrição
$u$	Representa um determinado usuário do sistema.
$c$	Constante inteira global que representa o limite máximo de vetores utilizados para calcular a recomendação.
$n$	Denota o número de composições realizadas por $u$ para fins de recomendação, sendo seu máximo igual a $c$ .
$n'$	Número de colaborações realizadas por $u$ para fins de recomendação, sendo seu máximo igual a $c$ .
$n''$	Número de avaliações realizadas pelo usuário para fins de recomendação.
$m$	Constante inteira global que representa o número de gêneros musicais do sistema.
$g$	Variável que faz referência a um gênero específico.
$C$	Vetor de gêneros de uma composição. Possui tamanho $m$ .
<i>descritores</i>	Lista de descritores da composição $C$ .
$G$	Representação de um grupo genérico.
$[C^1 C^2  \dots  C^n]$	Matriz de ordem $n \times m$ formada pela composição de $C^1$ até $C^n$ .
$VCP$	Vetor de Composições Próprias do usuário $u$ . Possui tamanho $m$ .
$VCL$	Vetor de Colaborações do usuário $u$ . Possui tamanho $m$ .
$VAL$	Vetor de Avaliações de $u$ . Possui tamanho $m$ .
$VM$	Vetor Musical combinado do usuário $u$ . Possui tamanho $m$ .
<i>pref</i>	Conjunto de interesses do usuário $u$ .
<i>des</i>	Conjunto de desinteresses do usuário $u$ .

<i>rej</i>	Lista de usuários rejeitados para recomendação
<i>flex</i>	Número representativo da flexibilidade de um usuário, tendo valores de 1 a 5.

## 5.2 Perfil musical

O perfil musical do usuário é aquele construído pelo sistema objetivando refletir seu gosto musical pessoal e, assim, guiar as recomendações. As informações nele contidas provêm de quatro fontes (Figura 5.1): o perfil descritivo introduzido explicitamente pelo indivíduo; as composições criadas por ele; as suas colaborações em músicas de outros usuários; e suas avaliações de composições públicas (de terceiros). Estes dados são combinados de maneira ponderada para serem posteriormente comparados a fim de identificar compatibilidades e realizar as recomendações entre usuários.

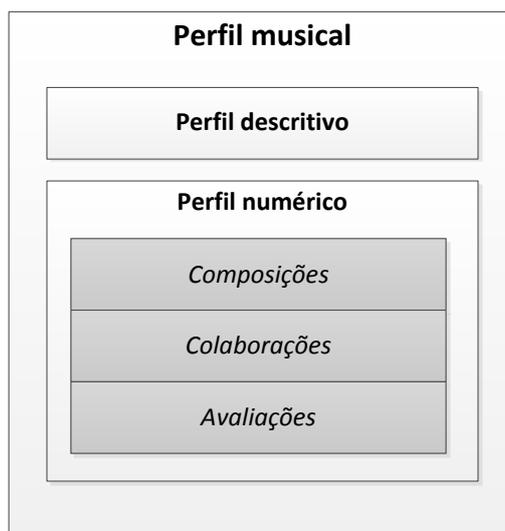


Figura 5.1: Composição do perfil musical de um usuário

Nas próximas subseções, serão detalhadas as formas de construção de cada componente deste perfil musical, assim como sua unificação.

### 5.2.1 Perfil descritivo

Como observado no capítulo anterior, quando um usuário cadastra-se em um ambiente colaborativo, seus objetivos e preferências individuais naquele sistema pode variar consideravelmente. Portanto, é importante entender de alguma maneira essas informações para realizar uma recomendação interessante. Para isso, é possível submetê-lo a um breve questionário, opcional, que seja capaz de refletir adequadamente estes pontos. Ainda que este tipo de interrogatório possa às vezes parecer intrusivo e desgastante, as informações coletadas nessa etapa serão valiosas para entender três pontos principais, detalhados a seguir, e que podem ser obtidas através de poucas perguntas.

- 1. Objetivos na colaboração:** procura identificar o interesse do usuário referente à composição colaborativa. Ou seja, se prefere criar projetos musicais próprios ou colaborar em peças já existentes. Tal informação pode

ser adquirida através de uma pergunta objetiva de múltipla escolha, classificando-o em uma das três seguintes classes de usuários: apenas cria projetos próprios; cria projetos próprios e colabora; e apenas colabora em projetos de outros usuários.

Claramente, usuários que pertencem conjuntamente à primeira ou terceira categorias são incompatíveis, sendo necessário excluir recomendações entre estes grupos.

Adicionalmente, é possível estender esta estrutura para incorporar objetivos mais restritos e específicos, além de divisão de tarefas, como seleção de instrumentistas, letristas ou compositores. Isso ajuda a evitar situações onde se recomenda um baterista a outro baterista, em *websites* de composição livre, e permite filtrar as recomendações.

2. **Preferência e rejeição de gêneros:** procura reduzir o universo de cálculo de recomendações por antever incompatibilidades de gênero musical entre os usuários. Pode ser realizada através de duas questões semelhantes, perguntando os gêneros prediletos e rejeitados para composição/colaboração, apresentando uma lista exaustiva dos principais gêneros musicais, onde o usuário pode selecionar múltiplas opções. Assim, para cada usuário  $u$ , é mantido dois conjuntos: um de preferências,  $pref$  e outro de desinteresse,  $des$ .
3. **Flexibilidade:** têm como objetivo antever a flexibilidade do usuário em relação às contribuições de outros indivíduos em composições de sua criação. Esta característica, que é disposta em uma escala com valores de um a cinco para representar desde “inflexível” (valor 1) a “muito flexível” (valor 5), pode ser adquirida através de uma questão de múltipla escolha. Esta informação irá aumentar ou diminuir o círculo de tolerância em relação às recomendações de usuários com preferências diferentes. Será representada por  $flex$ .

### 5.2.2 Composições próprias

As composições próprias, em conjunto com as colaborações em projetos de outros usuários e as avaliações de composições públicas, serão responsáveis por refinar o perfil descritivo e, assim, identificar compatibilidades maiores entre indivíduos. Os vetores que representam as últimas  $n$  criações do usuário são combinados em forma de uma matriz de ordem  $n' \times m$  para cálculo do vetor de composições próprias,  $VCP = [a_1, a_2 \dots a_m]$ , onde  $a_g \in VCP$  de  $u$  e é calculado da seguinte forma:

$$a_g = \frac{\sum_{i=1}^n b_{g,i} \times 0,9^i}{\sum_{i=1}^n 0,9^i}, \text{ onde } g = 1 \dots m \quad (1)$$

Em (1), o elemento  $a_g$  de  $VCP$  contém o somatório normalizado de cada gênero  $g$  obtido para as últimas  $n$  composições criadas pelo usuário  $u$ , sendo o elemento  $b_{g,i} \in [B^1|B^2| \dots |B^n]$ , onde  $B^1$  representa o vetor da composição mais recente até a mais antiga,  $B^n$ . Deste modo, a fórmula decreta 10% na nota relativa àquela imediatamente mais recente.

De forma a evitar cálculos muito custosos que comprometam a escalabilidade do sistema, é interessante que seja atribuído um valor fixo máximo para  $n$ . Como as composições próprias deve ter um peso considerável para traçar o perfil musical do usuário, este número também não deve ser muito baixo (AGRAWAL; SRIKANT, 1994). Assim, assume-se inicialmente  $c = \max(n) = 10$ . O vetor  $VCP$  de  $u$  é atualizado pelo módulo de recomendação cada vez que uma composição é criada ou alterada pelo usuário  $u$ .

### 5.2.3 Colaborações

As colaborações realizadas por  $u$  em criações musicais de outros usuários também influenciam no seu perfil musical, visto que uma pessoa somente contribuirá naquilo que julgar interessante. Assim, esta informação é armazenada em um vetor de gêneros denominado  $VCL$ , e calculado de maneira semelhante ao vetor de composições próprias, como segue, onde  $c_g \in VCL$  de  $u$ :

$$c_g = \frac{\sum_{i=1}^{n'} d_{g,i}}{n'}, \text{ onde } g = 1 \dots m \quad (2)$$

Em (2), cada elemento  $c_g$  de  $VCL$  armazena a média aritmética do gênero  $g$  para as  $n$  composições em que houve colaboração de  $u$ . Ou seja,  $d_{g,i}$  pertence à matriz de ordem  $n' \times m$  formada pelos vetores destas composições:  $[D^1 | D^2 | \dots | D^{n'}]$ . A diferença básica entre  $VCP$  e  $VCL$  é que, na segunda, não há o decréscimo proporcional de 10% para cada composição mais antiga..

### 5.2.4 Avaliações

As avaliações realizadas pelo usuário  $u$  de músicas publicadas no ambiente colaborativo são calculadas e armazenadas em um vetor de gêneros denominado  $VAL$ . Este parecer, realizado de acordo com a escala de Likert com valores de um a cinco, terá uma influência relativa positiva ou negativa, conforme a Tabela 5.2. Se ao usuário, por exemplo, não lhe agrada composições que possuam elementos de música eletrônica, este irá avaliar, em geral, negativamente músicas deste gênero. E esta avaliação irá refletir-se como um decréscimo na nota atribuída a este gênero em  $VAL(u)$ .

Tabela 5.2: Relação entre a nota atribuída pelo usuário, seu parecer associado e a influência sobre seu gênero musical.

Nota atribuída	Opinião associada	Influência relativa
1	Péssimo	-0,25
2	Ruim	-0,1
3	Neutro	0
4	Bom	+0,1
5	Ótimo	+0,25

Inicialmente,  $VAL$  é preenchida por valores igualmente distribuídos para cada gênero. Cada vez que  $u$  atribui um parecer a um item público, este vetor é atualizado calculando-se a influência relativa de acordo com sua avaliação conforme a Tabela 5.2. Assim, um fator de peso proporcional será atribuído dependendo de sua nota. No pior caso, a nota será decrescida de 25% e, no melhor, será acrescida pelo mesmo fator, calculados da seguinte maneira, onde  $e_g \in VAL(u)$ :

$$e_g = \sum_{j=1}^{n''} e_g \times (1 + f_{g,j} \times feedback_j), \quad \text{onde } g = 1 \dots m \quad (3)$$

Em (3),  $f_{g,j} \in F^j$  que corresponde ao gênero  $g$  relativo à  $j$ -ésima composição que o usuário atribuiu parecer até  $n''$  (número total de avaliações), sendo  $feedback_j$  o valor numérico dado a esta composição conforme a terceira coluna da Tabela 5.2. E, posteriormente normalizada utilizando-se  $d = \sum_{g=1}^m e_g$ .

### 5.2.5 Perfil numérico

De posse dos vetores  $VCP$ ,  $VCL$  e  $VAL$  de  $u$ , é possível agora unificar estas estruturas em uma única estrutura,  $VM$ , que será aquele efetivamente utilizado para calcular a similaridade entre usuários.  $VM$  de  $u$  é calculado da seguinte forma, com  $n''$  representando o número de avaliações dadas por  $u$ :

$$VM = \begin{cases} \frac{(20 - n'') \times \frac{nVCP \times n'VCL}{n + n'} + n''VAL}{20}, & \text{se } n'' < 6 \\ 0,7 \times (VCP + VCL) + 0,3 \times VAL, & \text{se } n'' \geq 6 \end{cases} \quad (4)$$

A equação (4) atribui um peso mínimo de 70% para as composições próprias do usuário e suas colaborações. Cada avaliação que o usuário realizar terá uma influência de 5% sobre a nota final até, no máximo 30% (6 avaliações). A partir deste número, o percentual de influência permanecerá constante e as avaliações continuarão a serem consideradas, porém, com impacto individual cada vez menor. Esta limitação procura manter uma proporção de influência menor para as avaliações dadas pelo usuário aos itens públicos do que às suas composições e colaborações, visto que estas últimas são consideradas mais importantes.

## 5.3 Agrupamento

Agrupamento ou *clustering* é uma maneira eficiente de reduzir o universo de cálculos necessários para identificar vizinhos de um usuário alvo, atribuindo previamente os indivíduos a grupos de interesses ou características comuns relevantes (CATALTEPE; ALTINEL, 2007). Assim, são realizados cálculos de comparação somente dentro de cada grupo, tornando o sistema mais escalável. Nesta proposta, a lista de todos os grupos corresponde à lista de todos os gêneros do ambiente colaborativo. Portanto, existem  $m$  grupos. Como existe uma pluralidade de preferências pessoais de gêneros não exclusivas, uma pessoa pode pertencer a mais de um grupo. O que define essa informação é o perfil descritivo do usuário.

Caso um determinado usuário, em sua descrição, tenha classificado interesse por *blues*, *rock* e *soul*, por exemplo, ele irá pertencer a estes três grupos. Portanto, os cálculos de similaridade serão limitados aos usuários destas categorias. Assim, só haverá uma possível recomendação (através da matriz de similaridade) entre um usuário  $u$  e  $p$  caso  $pref \cap pref' \neq \emptyset$ , onde  $pref \in u$  e  $pref' \in p$ .

### 5.3.1 Matriz de similaridade

O próximo passo para a recomendação é o cálculo de similaridade, que irá mensurar numericamente a proximidade entre usuários de um grupo. Para isso, é utilizada uma fórmula muito popular em sistemas de recomendação em geral: a *similaridade do cosseno* (ou *distância do cosseno*). Ela calcula a semelhança entre dois vetores através da medida do cosseno do ângulo formado entre eles no espaço gerado pelo produto dos mesmos. Existem outras formas de calcular a similaridade, como a *correlação de Pearson* e a *distância euclidiana*. (VOZALIS; MARGARITIS, 2003). Porém, adotou-se esta medida devido à representação dos usuários no ambiente (vetores); e por ser uma fórmula já muito utilizada em sistemas de recomendação, com bons resultados. A distância é representada por um valor real entre zero e um. Quanto mais próximo de zero, menos semelhanças existem entre os usuários; e quanto mais próximo de um, mais similaridades os mesmos possuem. Abaixo, a fórmula da distância cosseno é apresentada com  $VM \in u$  e  $VM' \in p$ .

$$dist(u, p) = dist(VM, VM') = \frac{\sum_{g=1}^m a_g \times b_g}{\sqrt{\sum_{g=1}^m (a_g)^2} \times \sqrt{\sum_{g=1}^m (b_g)^2}} \quad (5)$$

Onde em (5), a distância entre dois usuários  $u$  e  $p$  é calculada utilizando seus vetores musicais, com  $a_g \in VM$  e  $b_g \in VM'$ , para cada gênero  $g$ . Assim, a matriz de correlação de usuários para um determinado grupo  $G$  vai sendo preenchida. É importante salientar que, diferentemente dos sistemas de recomendação de itens ou objetos comuns através da filtragem colaborativa, onde neste momento seria gerada uma matriz *usuários*  $\times$  *itens*, na recomendação de pessoas gera-se uma matriz quadrada simétrica *usuários*  $\times$  *usuários*.

Tabela 5.3: vetores  $VM$  de cinco usuários hipotéticos.

<b>u1</b>	<b>u2</b>	<b>u3</b>	<b>u4</b>	<b>u5</b>
0,5	0,1	0,4	0,2	0,1
0,3	0,3	0,3	0,3	0,4
0,2	0,2	0	0,5	0,2
0	0	0,3	0	0,1
0	0,4	0	0	0,2

Para exemplificar este processo, considere cinco usuários hipotéticos pertencentes a um determinado grupo  $G = [u1, u2, u3, u4, u5]$ , em um ambiente com cinco gêneros musicais. Os vetores musicais,  $VM$ , que representam os perfis destes usuários estão elencados na Tabela 4.3. Desta forma, a matriz  $MU$  de  $G$  que fornece as informações sobre as similaridades entre estes usuários, calculadas conforme (5), está apresentada na Tabela 5.4.a. Alternativamente, é possível aproveitar a mesma estrutura para incluir os relacionamentos entre estes indivíduos, como na Tabela 5.4.b, para cálculos de relacionamento mais rápidos.

Tabela 5.4.a: Distância entre diferentes usuários em um determinado grupo.

	u1	u2	u3	u4	u5
u1		0,53	0,81	0,76	0,67
u2	0,53		0,41	0,62	0,90
u3	0,81	0,41		0,47	0,64
u4	0,76	0,62	0,47		0,76
u5	0,67	0,90	0,64	0,76	

Tabela 5.4.b: Distância entre usuários e suas relações de colaboração.

	u1	u2	u3	u4	u5
u1		0,53	0,81	0,76	0,67
u2	0		0,41	0,62	0,90
u3	0	1		0,47	0,64
u4	1	0	0		0,76
u5	0	0	0	1	

Apesar dos valores mais altos representarem uma maior similaridade, não há limite mínimo para seleção e recomendação de usuários. Para uma melhor visualização e entendimento, as distâncias e relações entre os usuários podem ser representadas através de um grafo desconexo, como ilustrado na Figura 5.2, onde os nodos representam os usuários e, as ligações, as colaborações. A proximidade entre os nodos refletem as semelhanças calculadas e apresentadas na Tabela 5.4: quanto mais próximos, mais semelhanças possuem em comum.

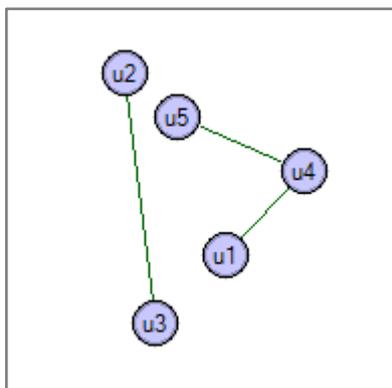


Figura 5.2: Grafo gerado pelo *software* Graph Magics<sup>18</sup> representativo da relação da Tabela 5.4.

## 5.4 Seleção e recomendação

A geração de uma lista de sugestões, denominada *recom*, contendo referências a outros usuários potencialmente interessados em colaborar em uma composição de  $u$ , é realizada através de um algoritmo, detalhado na Tabela 5.5. Resumidamente, o procedimento percorre as matrizes relativas aos grupos de gêneros aos quais pertence  $u$ , e visita os usuários mais similares. Caso o usuário visitado é compatível, segundo seu perfil descritivo, o mesmo é adicionado à lista de recomendações. Dependendo de sua flexibilidade, usuários de outros grupos também podem ser adicionados, através de transitividade. Quanto mais flexível, mais frequentes e mais tolerantes serão as recomendações a  $u$ .

<sup>18</sup> <http://www.graph-magics.com/>

Tabela 5.5: Algoritmo das funções de recomendação.

<p><b>Função <i>criaListaRecomendação</i>(usuário <i>u</i>):</b></p> <ol style="list-style-type: none"> <li>1. Seleciona grupo <math>G</math>, onde <math>G \in u.pref</math>; <math>iter \leftarrow 0</math>;</li> <li>2. Seleciona usuário <math>p \in G.MU</math>, com a maior similaridade a <math>u</math>, desde que <math>p \notin u.rej</math>;       <ol style="list-style-type: none"> <li>a. Se <math>p \in u.recom</math>, volta a 2;</li> <li>b. Se <math>p.classe \cap u.classe \neq \emptyset</math> então <math>u.rej \leftarrow p</math> e volta a 2;</li> <li>c. Se <math>p \in u.colaboradores</math>, volta a 2;</li> <li>d. <math>u.recom \leftarrow p</math>; <math>incrementa(iter)</math>;</li> </ol> </li> <li>3. Se <math>u.recom.tamanho \geq 20</math>, retorna <math>u.recom</math>; Senão, Caso:       <ol style="list-style-type: none"> <li>a. <math>u.flex = 1</math>: retorna à etapa 1.</li> <li>b. <math>u.flex = 2</math> e <math>count = 5</math>: <math>u.recom \leftarrow encontraParceiro(u, 2, 0.8)</math>;</li> <li>c. <math>u.flex = 3</math> e <math>count = 4</math>: <math>u.recom \leftarrow encontraParceiro(u, 3, 0.7)</math>;</li> <li>d. <math>u.flex = 4</math> e <math>count = 3</math>: <math>u.recom \leftarrow encontraParceiro(u, 4, 0.6)</math>;</li> <li>e. <math>u.flex = 5</math> e <math>count = 2</math>: <math>u.recom \leftarrow encontraParceiro(u, 5, 0.5)</math>;</li> </ol> </li> <li>4. Se <math>u.recom \geq 20</math>, retorna <math>u.recom</math>; senão, volta a 1.</li> </ol>
<p><b>Função <i>encontraParceiro</i>(usuário <i>u</i>, int nível, real similaridade):</b></p> <ol style="list-style-type: none"> <li>1. Seleciona usuário <math>p \in u.colaboradores</math>;</li> <li>2. Seleciona usuário <math>q \in p.colaboradores</math> e <math>q \notin u.colaboradores</math>;</li> <li>3. Se nível = 2 então       <ol style="list-style-type: none"> <li>a. Se <math>dist(u, q) \geq similaridade</math> retorna <math>q</math>;</li> </ol>       Senão       <ol style="list-style-type: none"> <li>b. <math>decrementa(nível)</math>; retorna <math>encontraParceiro(p, nível, similaridade)</math>;</li> </ol> </li> </ol>

Considerando que a segunda função *encontraParceiro*, apesar de recursiva, irá percorrer no máximo cinco níveis, nota-se que o crescimento do algoritmo é afetado principalmente pela primeira função: *criaListaRecomendação*, que necessita percorrer as  $n$  matrizes de gênero e comparar todos os usuários de cada uma delas. As matrizes, por sua vez, não possuem limites de entradas, que correspondem aos usuários que pertencem ao gênero respectivo. Assim, a complexidade do algoritmo é expressa pelo número de comparações da matriz de gênero com o maior número de usuários ( $num$ ), e dada por  $O(num^2)$  (LEVITIN, 2007).

Neste algoritmo, uma lista com vinte recomendações é calculada para um usuário-alvo (número que pode ser variável). As sugestões são baseadas no perfil musical global do usuário, mas podem ser calculadas também a partir de uma composição específica. Neste caso, as sugestões podem ser mais focadas em uma determinada criação, aumentando as possibilidades de uma recomendação de sucesso.

Quando o usuário acata uma recomendação dada, é enviado um convite ao colaborador recomendado com informações que podem ser editadas para melhor descrever a ideia do projeto/composição. Caso a pessoa convidada recuse o convite, a lista de rejeições de ambos são atualizadas para que esta sugestão não conste futuramente na lista de possíveis recomendações.

## 5.5 Considerações sobre a abordagem

Na seção 3.2 desta dissertação, foi apresentada uma lista com características importantes e problemas comuns de sistemas de recomendação em geral. A seguir, serão detalhadas as formas que a abordagem trata alguns dos pontos relevados:

- **Esparsidade dos dados:** os vetores musicais dos usuários possuem todos o mesmo padrão, com soma total igual a um, e os cálculos de similaridade compõem matrizes com valores completos dentro de cada grupo, não havendo, portanto, o problema de esparsidade de dados típico de matrizes *usuários × itens*.
- **Escalabilidade:** foram tomadas algumas medidas para reduzir a quantidade de cálculos para gerar os perfis musicais pensando na escalabilidade do sistema. Além disso, as etapas de construção da lista de recomendações podem ser realizadas no lado do servidor e mantidas no banco de dados de maneira independente do cliente. Assim, a lista é buscada e apresentada como uma informação estática.
- **Recomendações transitivas:** o algoritmo proposto faz extenso uso da transitividade das relações dos usuários para enriquecer a lista de recomendações.
- **Problema da partida fria:** quando um usuário novo se cadastrar no sistema, seu perfil musical irá refletir exatamente as informações sobre preferências de gêneros elencadas em seu perfil descritivo (inseridas por ele). Assim, gera-se uma aproximação inicial real do seu gosto musical.
- **Problema do usuário incomum:** por utilizar uma abordagem híbrida, as recomendações não são geradas a partir de uma classe de comportamentos dos usuários, mas sim, sobre suas preferências. Ou seja, todos os usuários inevitavelmente irão pertencer a algum grupo no sistema, e receberão recomendações calculadas dentro deste. É possível que a similaridade não seja muito alta para usuários que fogem de certos padrões, porém, receberão, de qualquer forma, recomendações que mais lhe aproximam.

No próximo capítulo, será demonstrado um roteiro para a implementação da abordagem em um ambiente de composição colaborativa real, onde alguns detalhes expostos serão explorados em mais detalhes.

## 6 ACOPLANDO RECOMENDAÇÃO DE PARCEIROS AO CODES

Este capítulo apresenta um roteiro para acoplar um sistema de recomendação de parceiros no CODES, ambiente colaborativo de composição musical voltado a leigos introduzido na seção 2.3. CODES foi escolhido por ter sido desenvolvido no LCM (Laboratório de Computação e Música) da UFRGS e, conseqüentemente, ter inspirado e servido de base para este trabalho, que faz parte de uma pesquisa no mesmo laboratório. A implementação definitiva de um sistema de recomendação no CODES não foi realizada devido à uma migração de plataforma que está atualmente em curso. Assim, sugere-se um roteiro para realizar tal acoplamento.

Primeiramente, foi preciso avaliar as informações presentes no sistema que poderiam ser utilizadas na abordagem proposta, assim como aquelas que deveriam ser incluídas para completar o processo de recomendação. CODES apresenta características típicas de ambientes colaborativos de composição musical: controle de versionamento; avaliação de itens públicos; histórico de colaborações; informações sobre gêneros musicais das composições; e espaço voltado à discussão do processo de prototipação pelos colaboradores. Da análise destas suas características e de posse do seu esquema de banco de dados, foi possível eleger as necessidades efetivas de modificação e alteração, que incluem os perfis descritivos dos usuários; o controle sobre as recomendações e suas estruturas de dados; as matrizes de similaridades; o módulo de recomendação em si; e as *interfaces* com o usuário.

### 6.1 Gêneros musicais no CODES

A recuperação de informações sobre os gêneros musicais das MPs é relativamente simples, visto que elas estão explícitas no sistema. Em sua versão atual, CODES possui padrões sonoros predefinidos de quatro gêneros musicais: funk, jazz, rock e eletrônica (Tabela 6.1), que são referenciados como *estilos musicais* no ambiente. Assim, de acordo com a nomenclatura definida no Capítulo 5, teremos  $m = 4$  e, conseqüentemente, quatro matrizes de similaridade. De forma semelhante, todas as composições serão representadas por vetores de quatro dimensões, que fazem referência a estes quatro gêneros.

Tabela 6.1: Estilos musicais do CODES.

Estilo musical	$g$
Funk	1
Jazz	2
Rock	3
Eletrônica	4

Para definir a distribuição de gêneros de uma determinada MP, é possível utilizar os padrões sonoros da biblioteca que compõe o protótipo em questão, uma vez que estes já estão distribuídos nos quatro gêneros do ambiente. No protótipo ilustrado na Figura 6.1, por exemplo, cada ícone representa um padrão de quatro segundos atribuído a um gênero específico. Neste caso, contabilizamos 12 padrões de funk e 16 padrões de jazz. Assim, podemos representar seu vetor de gêneros com estes valores normalizados, que resultaria em  $C = [0.4, 0.6, 0, 0]$ . Ou seja, o valor 0.4 para funk; 0.6 para jazz; 0 para rock; e 0 para eletrônica.

Em ambientes mais complexos, onde as informações sobre gêneros não estão explícitas, a geração do vetor de gêneros torna-se mais difícil. Dependendo do caso, talvez seja necessário utilizar classificadores automáticos para obter informações semânticas a partir da análise da onda sonora de cada composição (como descrito no início da seção 5.1). Isto ocorre em casos onde os usuários fazem *upload* de faixas de áudio em formatos *crus*, como WAV ou MP3.

## 6.2 Perfil musical

Os perfis musicais dos usuários devem ser criados conforme a definição de gêneros, seguindo os cálculos apresentados no Capítulo 5. Portanto, para cada usuário, devem ser calculados os vetores  $VCP$ ,  $VCL$ ,  $VAL$  e  $VM$ . Os cálculos envolvidos na composição destes vetores são relativamente simples, e as informações para realiza-los já estão presentes no banco de dados do CODES. Assim, se faz necessário somente estender a estrutura de dados dos usuários em si.

$VCP$  é calculado sobre os últimos  $n$  protótipos criados pelo usuário em questão, até o limite de  $c = 10$ ;  $VCL$ , por sua vez, é inferido através das últimas colaborações do usuário, até o mesmo limite;  $VAL$  armazena o vetor respectivo à ponderação das avaliações do usuário à MPs já publicadas; e  $VM$  é o perfil musical numérico calculado com base nos demais vetores. De maneira complementar, CODES não possui um perfil descritivo dos usuários que reflita suas preferências, desinteresses e flexibilidade (*pref*, *des* e *flex*, respectivamente). Portanto, é necessário também incorporar tais informações ao perfil do usuário, questionando-o explicitamente no momento da criação de sua conta.

## 6.3 Arquitetura e banco de dados

CODES é implementado através do clássico modelo cliente-servidor para aplicações *web*. Devido a necessidades relacionadas a usabilidade e flexibilidade de interação, o

sistema utiliza Adobe® Action Script e arquivos .swf embutidos em HTML no lado cliente. No lado servidor, CODES aplica o Modelo-Visão-Controlador (*Model-View-Controller* - MVC), que separa a lógica de aplicação dos dados e da *interface*. A parte do modelo, implementado com Apache e PHP, conecta-se ao banco de dados MySQL que armazena toda a informação da aplicação. O controlador, por sua vez, gerencia a comunicação entre os dados e as regras de negócio do modelo. Para isso, CODES utiliza Adobe® MXML, uma linguagem baseada em XML para desenhar *interfaces* de usuário para aplicações Adobe® Flex.

O módulo de recomendação pode ser acoplado ao controlador de maneira assíncrona ao uso do sistema, para não comprometer a experiência do usuário. Assim, as atualizações das matrizes de similaridade e dos vetores relacionados à recomendação ocorrem de maneira independente. Os vetores de recomendação dos perfis de usuário podem ser atualizados em tempo real, conforme as respectivas atividades (avaliações de itens públicos, criação de uma nova MP, etc.). Em períodos específicos, as matrizes de similaridade são recalculadas, baseando-se nos perfis numéricos dos usuários, atualizando, assim, as informações sobre as recomendações. A Figura 6.1 ilustra a arquitetura do CODES incluindo o módulo de recomendação, que interage com o sistema somente através do banco de dados.

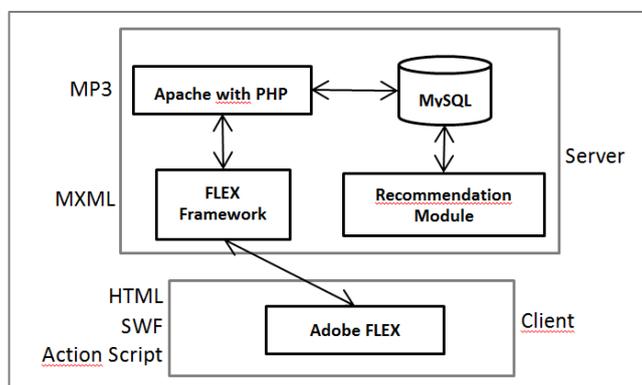


Figura 6.1: Arquitetura do CODES estendida com o módulo de recomendação. Adaptado de Miletto (2010)

Existem algumas mudanças necessárias no banco de dados para incorporar as estruturas relacionadas à recomendação. Dentre elas, estão incluídos os vetores representativos específicos de cada usuário, *VCP*, *VCL*, *VAL* e *VM*, que terão todos quatro dimensões. CODES também não possui um perfil descritivo, portanto, se faz necessário também estender a estrutura de dados dos usuários para manter as informações de preferência, desinteresse, rejeição e flexibilidade, representados por *pref*, *des*, *rej* e *flex*. A Figura 6.2, abaixo, destaca as alterações identificadas como necessárias a serem realizadas no banco de dados para implementar a abordagem apropriadamente (em azul). Este esquema do banco de dados do ambiente CODES foi retirado de Miletto (2010).

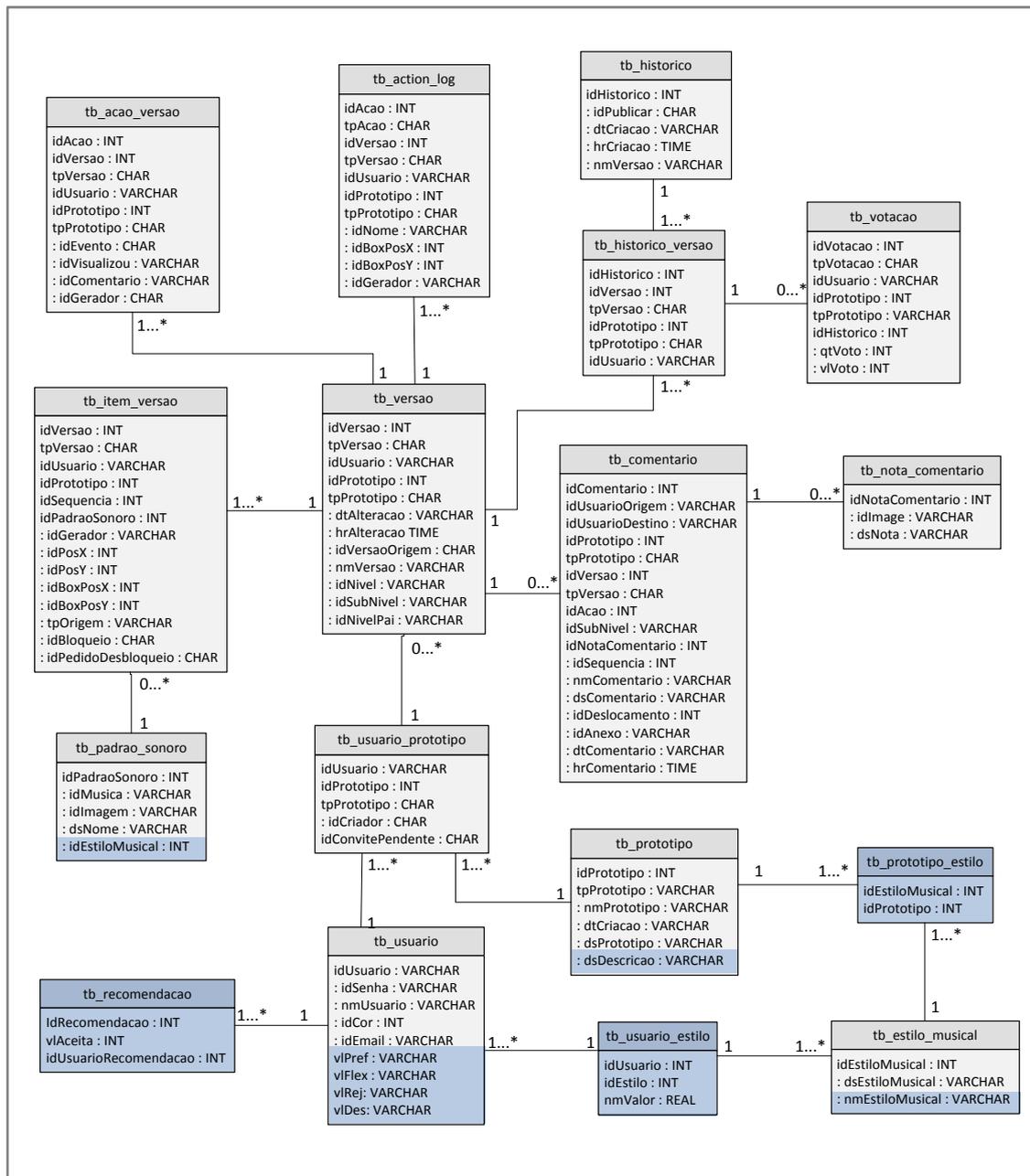


Figura 6.2: Esquema do banco de dados estendido do CODES. Adaptado de Miletto (2010)

Dentre as mudanças ilustradas na Figura 6.2, incluem-se as relações entre as tabelas de estilo musical (*tb\_estilo\_musical*) e usuários (*tb\_usuarios*). Esta alteração é necessária pois um usuário pode pertencer a vários estilos, ou seja, estar presente em diversas matrizes de gêneros musicais. Além desta, três inclusões de tabelas também devem ser incluídas: *tb\_recomendacao*, que armazena informações sobre uma recomendação específica (um usuário pode ter inúmeras recomendações); *tb\_usuario\_estilo*, que unifica as tabelas referentes ao usuário e estilo; e *tb\_prototipo\_estilo*, que unifica as tabelas de protótipos musicais e estilo musical.

Houve necessidade de adaptações da tabela *tb\_usuario*, que foi estendida para incluir as informações referentes ao perfil descritivo do usuário: preferência, flexibilidade e rejeição. A tabela *tb\_estilo musical* recebeu um atributo a mais com o

nome do gênero correspondente e, por fim, a tabela *tb\_padrao\_sonoro* foi estendida para incluir uma referência ao estilo musical de cada padrão.

#### 6.4 Interface com o usuário

Com o esquema do banco de dados definido para o cálculo das recomendações de cada usuário, o próximo passo consistiu em identificar as páginas da *interface* com necessidade de alteração ou inserção no sistema. Estas alterações visam permitir um acesso fácil ao convite de parceiros recomendados, apresentando listas de sugestões em diferentes páginas, assim como uma área específica para o gerenciamento das recomendações. A Tabela 6.2 elenca estas necessidades.

Tabela 6.2: Relação de páginas da *interface* do CODES com necessidade de adaptação ou criação.

Título da página	Tipo	Descrição da necessidade
Página inicial	Adaptação	Inclusão de seção com resumo de recomendações, caso visitante esteja logado.
Criação de usuário	Adaptação	Inclusão de seção para incorporar o perfil descritivo do usuário.
Criação de MP	Adaptação	Inclusão de seção com resumo de recomendações.
Gerenciamento de MPs	Adaptação	Inclusão de seção com resumo de recomendações.
Gerenciamento de versões de uma MP	Adaptação	Inclusão de seção com resumo de recomendações.
Edição de MPs	Adaptação	Inclusão de seção com resumo de recomendações.
Gerenciamento de recomendações	Criação	Página de gerenciamento das recomendações atuais e histórico de recomendações.

Para efeito ilustrativo, a Figura 6.3 mostra uma tela de gerenciamento de protótipos musicais de um usuário hipotético, com destaque à lista de sugestões de parceiros. Neste caso são apresentadas apenas as primeiras recomendações. Para uma lista mais completa e com informações sobre cada perfil, é possível clicar sobre o link “*mais recomendações*”, abaixo da lista sugerida, que irá redirecionar à página de gerenciamento de recomendações.

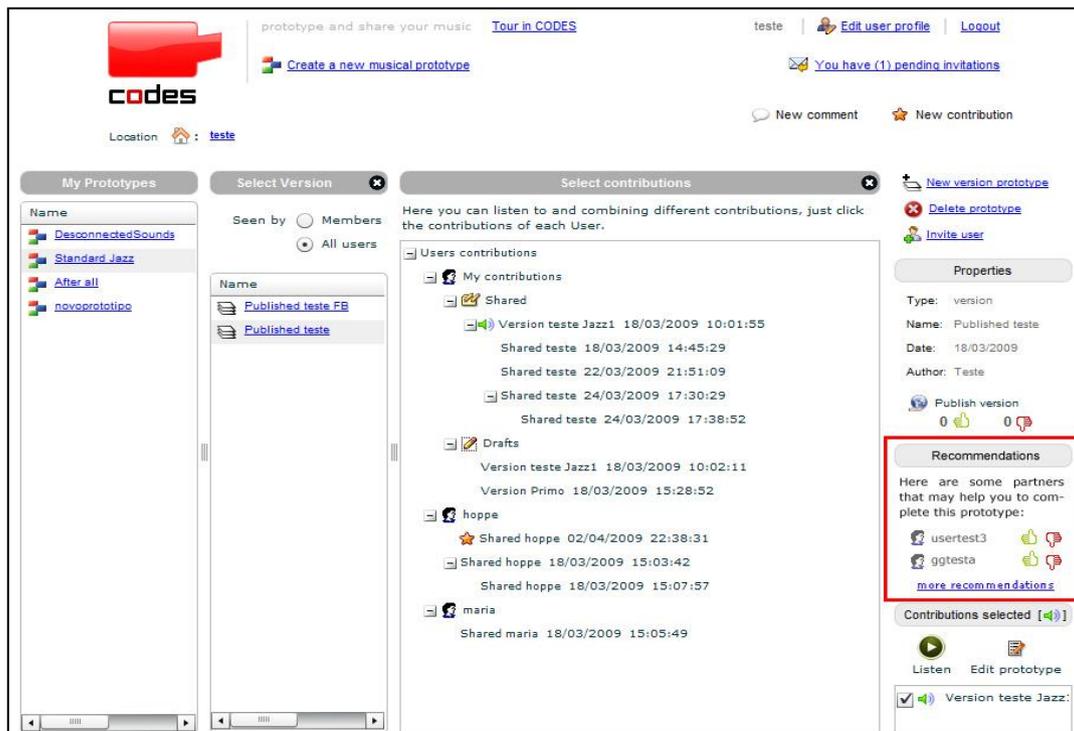


Figura 6.3: Página de gerenciamento de MPs, com a área de recomendação de parceiros em destaque.

## 6.5 Considerações finais

Este capítulo apresentou um roteiro para acoplar a abordagem de recomendação de parceiros proposta no CODES, detalhando os principais pontos com necessidade de alteração em sua versão original. Esta apresentação objetivou ilustrar os procedimentos e aspectos a serem levados em consideração na implementação de um módulo de recomendação em sistemas reais. Tais alterações não são numerosas nem complexas, e se considerarmos os pontos positivos trazidos pelas recomendações à atividade de composição musical colaborativa, estes custos tornam-se relativamente ainda menores.

A impossibilidade da implementação da abordagem conforme descrição deste capítulo tornou inviável uma avaliação final com os usuários do CODES. Esta avaliação poderia mensurar aspectos qualitativos das recomendações para posterior comparação (com outras abordagens e sistemas). Esta implementação e avaliação está prevista para ser realizada assim que o processo de migração de plataforma do CODES esteja finalizado.

## 7 CONCLUSÃO

O presente trabalho apresenta uma abordagem de recomendação de parceiros para sistemas colaborativos de composição musical *online*. Tais ambientes possuem um número significativo de usuários cadastrados atualmente, e encontrar parceiros com objetivos e preferências compatíveis tem demonstrado ser uma tarefa difícil e penosa. Durante este trabalho foram pesquisados e analisados vários *websites* desta categoria, e foram identificadas poucas referências sobre mecanismos de recomendação de parceiros que auxiliam usuários a encontrarem pessoas para cooperar entre si em projetos musicais colaborativos.

De modo a identificar melhor questões específicas de tais ambientes que influenciam nas relações entre seus usuários, foi conduzida uma pesquisa quantitativa com 52 pessoas que utilizam *websites* de composição colaborativa, realizada através de um questionário *online*. O questionário focou-se em pontos como os objetivos específicos de cada pessoa no ambiente, a tolerância acerca das contribuições de parceiros e como suas preferências musicais impactam em um projeto colaborativo. Os resultados demonstraram que os usuários são, em média, flexíveis quanto às contribuições de seus parceiros. Porém, a influência das preferências pessoais de gêneros musicais são por vezes bastante diferentes.

Assim, foi concebida uma abordagem híbrida – utilizando filtragem baseada em conteúdo e filtragem colaborativa – para a recomendação de parceiros capaz de identificar compatibilidades e divergências acerca das preferências de gênero e objetivos no ambiente. Para isso, foram utilizadas recomendações transitivas e matrizes de similaridade para cada gênero específico, onde são selecionados usuários com maior semelhança para recomendação, com graus de flexibilidade variáveis para cada usuário. A implantação da abordagem foi ilustrada através de seu acoplamento no CODES, um ambiente colaborativo de prototipação musical voltado a leigos.

As principais contribuições advindas deste trabalho são:

- A reunião e análise de estudos acerca da atividade de composição musical de maneira colaborativa, destacando a importância das relações entre os diversos colaboradores para um resultado positivo.
- A pesquisa e discussão sobre sistemas de recomendação e suas metodologias, e como estes podem ser utilizados para sugerir parceiros em ambientes colaborativos de composição musical.
- A análise das relações entre colaboradores de *websites* de composição, destacando a importância dos objetivos pessoais dos parceiros, suas preferências musicais pessoais, flexibilidade e tolerância, obtidas através da pesquisa conduzida com usuários reais.

- A proposta de uma abordagem de recomendação de parceiros para ambientes de composição colaborativa, capaz de abranger os aspectos identificados na pesquisa realizada.

As principais limitações deste trabalho são:

- A falta de resultados concretos sobre as recomendações da proposta, advindas de sua implementação em um sistema real, com usuários colaborando entre si a partir das recomendações, e de entrevistas com os mesmos sobre suas experiências. Tal experimento se mostra necessário para comprovar a eficácia da abordagem.
- A limitação de estudos sobre escalabilidade e confiabilidade na abordagem. Apesar de ter sido dada alguma atenção a aspectos específicos que impactam na escalabilidade da proposta, não houve uma exploração maior sobre seu comportamento em situações mais extremas e com grande quantidade de informações.

Durante este trabalho de pesquisa, foram publicados dois trabalhos relacionados a este tema. Primeiramente, o capítulo de Pimenta et al (2013) sobre o sistema CODES, que descreve a recomendação de parceiros como um dos mecanismos de suporte a comunidades de criação musical *online*; e o artigo publicado no SBCM (Simpósio Brasileiro de Computação Musical), sobre uma abordagem de recomendação de parceiros em ambientes de composição musical colaborativa (TESTA et al, 2013). Este último é transcrito na íntegra no Anexo C, ao final deste documento.

## 7.1 Trabalhos futuros

A seguir, estão elencados alguns pontos e ideias identificados durante o desenvolvimento deste trabalho que podem complementar a abordagem de recomendação, ou questões específicas da proposta em si que podem ser exploradas adicionalmente. As duas primeiras estão diretamente relacionadas às limitações observadas anteriormente.

- Implementação em um sistema real e validação da abordagem com usuários reais, obtendo assim resultados concretos sobre a qualidade das recomendações.
- Estudo do comportamento da abordagem em uma aplicação de larga escala, e seu impacto sobre as recomendações.
- Explorar o uso de folksonomia para indexar informações sobre usuários e composições, de forma complementar à recomendação de parceiros.
- Aprimorar a abordagem incorporando informações sobre as experiências de colaboração após as recomendações, mantendo um histórico de cada usuário como colaborador. Assim, é possível ajustar as recomendações baseando-se na reputação dos usuários como colaboradores.
- Possibilidade de utilizar POIs (*Points of Interest* – Pontos de Interesse), que permitem aos usuários marcarem pontos específicos interessantes nas composições para aperfeiçoar as recomendações, de maneira semelhante ao trabalho de Dias, Wives e Roesler (2012), que propõe uma abordagem para utilização de POIs em um sistema de recomendação de vídeos.

- Adaptar a abordagem para torna-la mais genérica, sendo possível utiliza-la em ambientes colaborativos criativos em geral;

## REFERÊNCIAS

- AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules. In: 20th INT. CONFERENCE OF VERY LARGE DATA BASES, ... **proceedings**, [S.l.]: [s.n], 1994. p. 487-499.
- AUCOUTURIER, J. J.; PACHET, F. Representing musical genre: a state of the art. **Journal of New Music Research**. [S.l.] 2003. p. 83-93.
- BALABANOVIĆ, M.; SHOHAM, Y. Fab: content-based, collaborative recommendation. **Communications of the ACM**, v. 40, n. 3, 1997. p. 66-72.
- BREESE, J. S.; HECKERMAN, D.; KADIE, C. Empirical analysis of predictive algorithms for collaborative filtering. In: 14th CONFERENCE ON UNCERTAINTY IN ARTIFICIAL INTELLIGENCE... **proceedings**. [S.l.]: Morgan Kaufmann Publishers Inc., 1998. p. 43-52.
- BROWN, A. R.. **Music composition and the computer: an examination of the work practices of five experienced composers**. Tese de doutorado, School of Music, The University of Queensland. 2003.
- BURKE, R. Hybrid recommender systems: Survey and experiments. **User modeling and user-adapted interaction**, v. 12, n. 4, [S.l.], 2002. p. 331-370.
- BURRED, J. J.; LERCH, A. A hierarchical approach to automatic musical genre classification. In: INTERNATIONAL CONFERENCE ON DIGITAL AUDIO EFFECTS, 2003... **proceedings**. [S.l.]: [S.N], 2003.
- CAI, J.; FRANCIS J.; GHEYSENS, S. Creating a Hybrid Music Recommendation System from Content and Social-Based Algorithms. In: **GSET, 2009...proceedings**. Governor's School of Engineering and Technology. New Jersey: Governor's School of Engineering and Technology. 2009.
- ÇATALTEPE, Z; ALTINEL, B. Music recommendation based on adaptive feature and user grouping. In: 22nd international symposium on Computer and information sciences, ISICIS 2007...**proceedings**. [S.l.]: IEEE, 2007. p. 1-6.
- CHEN, C. M.; HONG, C. M.; CHANG, C. C. Mining interactive social network for recommending appropriate learning partners in a Web-based cooperative learning environment. In: CONFERENCE ON CYBERNETICS AND INTELLIGENT SYSTEMS, 2008...**proceedings**. [S.l.]: IEEE, 2008. p. 642-647.
- CHEN, H. C.; CHEN, A. L. A music recommendation system based on music data grouping and user interests. **CIKM**. Vol 1.[S.l.]: 2001, p. 231-238.

CHEN, J.; GEYER, W.; DUGAN, C.; MULLER, M.; GUY, I. Make new friends, but keep the old: recommending people on social networking sites. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2009. ...**proceedings**. [S.l.]: ACM, 2009. p. 201-210.

CLAYPOOL, M.; GOKHALE, A.; MIRANDA, T.; MURNIKOV, P.; NETES, D.; SARTIN, M. Combining content-based and collaborative filters in an online newspaper. In: ACM SIGIR WORKSHOP ON RECOMMENDER SYSTEMS... **proceedings**. [S.l.]: [s.n.], 1999.

COLEMAN, J. S.; COLEMAN, J. S. **Foundations of social theory**. Harvard University Press, 1994.

DIAS, A. S.; WIVES, L. K.; ROESLER, V. POI Enhanced Video Recommender System using Collaboration and Social Networks. In: WEBIST, 2012...**proceedings** [S.l.:s.n.] 2012. p. 717-722.

DUDANI, S A. The distance-weighted k-nearest-neighbor rule. **Systems, Man and Cybernetics, IEEE Transactions on**, n. 4, p. 325-327, 1976.

FARREL, M. P. **Collaborative Circles: friendship dynamics and creative work**. [S.l.]: University Of Chicago Press . 2003. p. 328.

FERNEYHOUGH, B. **Collected writings**. Routledge, 1995.

GOLDBERG, D.; NICHOLS, D.; Oki, B. M.; TERRY, D. Using collaborative filtering to weave an information tapestry. **Communications of the ACM**, New York, v. 35, n. 12, p. 61-70, December, 1992.

GOOD, N.; SCHAFFER, J. B.; KONSTAN, J. A.; BORCHERS, A.; SARWAR, B.; HERLOCKER, J.; RIEDLE, J. Combining collaborative filtering with personal agents for better recommendations. In: AAAI/IAAI, 1999.... **proceedings**. [S.l.]: [s.n.], 1999. p. 439-446.

GUNAWARDENA, S; WEBER, R. Discovering Patterns of Collaboration for Recommendation. In: FLAIRS CONFERENCE, 2009...**proceedings**. [S.l.]: [s.n.], 2009

HAND, D. J. Principles of data mining. **Drug safety**, v. 30, n. 7, p. 621-622, 2007.

HOFFMANN, J. A. Computer-aided collaborative music instruction. **Harvard educational review**, v. 61, n. 3, p. 270-279, 1991.

INDABAMUSIC.COM. **About**. Disponível em <https://www.indabamusic.com/about>. Acessado em novembro de 2013.

IMPETT, J. **Making a mark: the psychology of composition**. Oxford handbook of music psychology. Oxford: Oxford Library of Psychology, 2008. p. 403-412.

JANNACH, D.; ZANKER, M.; FELFERNIG, A.; FRIEDRICH, G. **Recommender systems: an introduction**. [S.l.]: Cambridge University Press. 2010.

KAZIENKO, P.; MUSIAŁ, K.. **Recommendation framework for online social networks**. Advances in Web Intelligence and Data Mining. Berlim: Springer, 2006. p. 111-120.

BISCHOFF, J.; GOLD, R.; HORTON, J. Music for an interactive network of microcomputers. **Computer Music Journal**, v. 2, n. 3, 1978. p. 24-29.

- KIM, K.; LEE, D.; YOON, T.; LEE J. A music recommendation system based on personal preference analysis. In: INTERNATIONAL CONFERENCE ON THE APPLICATIONS OF DIGITAL INFORMATION AND WEB TECHNOLOGIES, 2008. ..**proceedings** [S.l.]: [s.n], 2008. p. 102-106.
- LEVITIN, A. **Introduction to the design & analysis of algorithms**. [S.l.]: Pearson Addison-Wesley, 2007.
- LI, T.; OGIHARA, M. Music genre classification with taxonomy. INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 2005. ...**proceedings**. [S.l.]: IEEE, 2005. p. 197-200.
- MACDONALD, R. A.; MIELL, D. Creativity and music education: The impact of social variables. **International Journal of Music Education**, n. 1. 2000. p. 58-68.
- MACDONALD, R.; BYRNE, C.; CARLTON, L. Creativity and flow in musical composition: An empirical investigation. **Psychology of Music**, v. 34, n. 3, p. 292-306, 2006.
- MATELL, M. S.; JACOBY, J. Is there an optimal number of alternatives for Likert scale items? I. Reliability and validity. **Educational and psychological measurement**, 1971.
- MCDONALD, D. W. Recommending collaboration with social networks: a comparative evaluation. SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2003...**proceedings**. [S.l.]: ACM, 2003. p. 593-600.
- MELVILLE, P.; MOONEY, R. J.; NAGARAJAN, R. Content-boosted collaborative filtering for improved recommendations. In: AAAI/IAAI, 2002...**proceedings**. [S.l.]: S.n.] 2002. p. 187-192.
- MELVILLE, P.; SINDHWANI, V. **Recommender Systems**. Encyclopedia of Machine Learning, Disponível em <http://www.prem-melville.com/publications/recommender-systems-eml2010.pdf>. Acesso em: julho de 2012.
- METEREN, R. V.; SOMEREN, M. V. Using content-based filtering for recommendation. In: MACHINE LEARNING IN THE NEW INFORMATION, 2000...**proceedings**. [S.l.]: [s.n], 2000.
- MILETTO, E. M. **CODES: an Interactive Novice-oriented Web-based Environment for Cooperative Musical Prototyping**. Tese de Doutorado. PPGC, Instituto de Informática, UFRGS. 2010.
- MILETTO, E. M.; FLORES, L. V.; RUTILY, J.; PIMENTA, M. S. CODES: supporting awareness in a web-based environment for collective music prototyping. In: BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS , 2006. ..**proceedings**. New York: ACM, 2006. p. 50-55.
- MILLER, B. N.; ALBERT, I.; LAM, S. K.; KONSTAN, J. A.; RIEDL, J. MovieLens unplugged: experiences with an occasionally connected recommender system. In: 8th INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES, 2003...**proceedings**. [S.l.]: ACM, 2003. p. 263-266.
- MITCHELL, T. M. Machine learning. 1997. **Burr Ridge, IL**: McGraw Hill, v. 45, 1997.

- NAYAK, R.; ZHANG, M.; CHEN, L. A Social Matching System for an Online Dating Network: A Preliminary Study. IEEE INTERNATIONAL CONFERENCE ON DATA MINING WORKSHOPS (ICDMW), 2010...**proceedings**. [S.l.]: IEEE, 2010. p. 352-357.
- NIGAM, K.; McCALLUM, A.; THRUN, S.; MITCHELL, T. Learning to classify text from labeled and unlabeled documents. **AAAI/IAAI**, v. 792. [S.l.:S.n.], 1998.
- PACHET, F.; CAZALY, D. A taxonomy of musical genres. **RIAO**. [S.l.:S.n.] 2000. p. 1238-1245.
- PALISPIS, E. S. **Introduction to Sociology and Anthropology**. Manila: Rex Printing Company, Inc.1996. p. 313.
- PAZZANI, M. J.; BILLSUS, D. Content-based recommendation systems. **The adaptive web**. Berlin: Springer, 2007. p. 325-341.
- PIMENTA, M. S.; MILETTO, E. M.; KELLER, D.; FLORES, L. V.; TESTA, G. G. Technological Support for Online Communities Focusing on Music Creation: Adopting Collaboration, Flexibility, and Multiculturality from Brazilian Creativity Styles. **Cases on Web 2.0 in Developing Countries: Studies on Implementation, Application, and Use**. IGI Global. p. 283-312. 2013.
- PIMENTA, M.; MILETTO, E.; FLORES, L. V.; HOPPE, A. **Cooperative Mechanisms for Networked Music**. Future Generation Computer Systems archive. Volume 27, Issue 1, January, 2011. Amsterdam, The Netherlands. p. 100-108. 2011.
- REYNOLDS, R. **Form and Method: Composing Music: The Rothschild Essays**. Routledge, 2013.
- RICCI, F.; ROKACH, L.; SHAPIRA, B.; KANTOR, P. B. **Recommender Systems Handbook**. [S.l.]: Springer. 2010. p. 872p.
- SARWAR, B.; KARYPIS, G.; KONSTAN, J.; RIEDL, J. Item-based collaborative filtering recommendation algorithms. In: 10th INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, 2001. ..**proceedings**. [S.l.]: ACM, 2001. p. 285-295.
- SCHICK, A. G.; GORDON, L A.; HAKA, S. Information overload: A temporal approach. **Accounting, Organizations and Society**, Volume 15, Issue 3, p 199-220, 1999.
- TECHCRUNCH.COM. **Indaba**: a music collaboration hub that can build a band for you. Disponível em <http://techcrunch.com/2009/01/13/indaba-a-music-collaboration-hub-that-can-build-your-band-for-you/>. Acessado em novembro de 2013.
- TERVEEN, L.; MCDONALD, D. W. Social matching: A framework and research agenda. **ACM transactions on computer-human interaction (TOCHI)**, v. 12, n. 3, p. 401-434, 2005.
- TESTA, G. G.; PIMENTA, M. S.; KELLER, D.; LAZZARINI, V. A Hybrid Approach to Recommend Partners in Collaborative Musical Environments. In: 14° SIMPOSIO BRASILEIRO DE COMPUTAÇÃO MUSICAL – SBCM, 2013...**anais**. Brasília: [S.n.], 2013.
- TUOMELA, R. **Cooperation: A Philosophical Study**. Netherlands: Kluwer, 2000. p. 431.

VANDER,W. **Folksonomy**. Disponível em <http://vanderwal.net/folksonomy.html>. Acessado em Outubro de 2013.

VOZALIS, E; MARGARITIS, K. G. Analysis of recommender systems algorithms. In: 6th HELLENIC EUROPEAN CONFERENCE ON COMPUTER MATHEMATICS AND ITS APPLICATIONS, 2003. ...**proceedings**. Athens, Greece: [s.n.], 2003.

WINTHROP, R. H. **Dictionary of Concepts in Cultural Anthropology**. New York: Greenwood Press, 1991.

WU, S.; SUN, J.; TANG, J. Patent partner recommendation in enterprise social networks. In: 6th ACM INTERNATIONAL CONFERENCE ON WEB SEARCH AN DATA MINING, 2013...**proceedings**. [S.l.]: ACM, 2013. p. 43-52.

WÜST, O.; JORDÀ, S. Architectural overview of a system for collaborative music composition over the web. In: INTERNATIONAL COMPUTER MUSIC CONFERENCE, 2001... **proceedings**. [S.l.]: [s.n.], 2001. p. 298-301.

XU, C.; MADDAGE, N. C.; SHAO, X.; CAO, F.; TIAN, Q. Musical genre classification using support vector machines. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 2003. ...**proceedings**. [S.l.]: IEEE, 2003. p. 429-432.

ZAIER, Z.; GODIN, R.; FAUCHER, L. Evaluating recommender systems. In: AXMEDIS, 2008...**proceedings**. [S.l.]: IEEE, 2008. p. 211-217.

## ANEXO A <QUESTIONÁRIO DE PESQUISA>

Este anexo apresenta o questionário de pesquisa original submetido aos entrevistados.

### Collaborative music composition

This is a study about the motivations of users of collaborative music composition websites. This web form will contribute in an academic research about music collaboration and music composition been conducted in the Institute of Informatics of the Federal University of Rio Grande do Sul, Brazil.

No personal data will be asked. If you have any doubts or suggestions before answering it, you can write to [gqtesta@inf.ufrgs.br](mailto:gqtesta@inf.ufrgs.br).

Please, answer this survey in a honest way, and thank you for your participation.

\* Required

**1 - Which website of collaborative music composition do you use? \***

- Indaba
- Kompoz
- Merge.fm
- SoundKeep
- SoundCloud
- Other:

**2 - How long have you been using websites of collaborative composition? \***

- 0 - 3 months
- 4-12 months
- 1 - 2 years
- 3 - 5 years
- More than 5 years

**3 - What is your experience with music? \***

- Professional
- Occasional player, with large experience
- Occasional player, with some experience
- Just as a hobby, with no experience
- Other:

**4 - Which factor have influenced you the most to use a collaborative music composition website? \***

- Just fun.
- Receive feedback of my compositions.
- The chance to create real music with people from all around the world.
- Find musicians to play together.
- Other:

**5 - Usually you create new projects or contribute to existing projects from other users? \***

- I only create new projects.
- Generally I create projects, but I also contribute sometimes.
- Both situations, equally.
- Generally I contribute in existing projects, but I also create my own sometimes.
- I only contribute in existing projects.

**6 - When other users collaborate on your project, how flexible are you about the contributions? \***

- Very flexible. Any kind of contributions are positive.
- Flexible. Any kind of contribution that produces interesting results.
- Moderately flexible. Contributions that respect my initial idea, but with flexibility for innovations.
- Inflexible. Contributions need to follow very close my initial idea.
- Other:

**7 - Regarding the personal musical taste, you prefer that the collaborators of your projects... \***

- have a very similar taste to mine.
- have similar taste, but with different influences.
- have a different musical taste, in order to contribute in an innovative way.
- have a complete diverse taste to mine, so the result will be unpredictable.
- Other:

**8 - Write about your experience composing online with other people (optional).**

**9 - Sugestions and observations about this research (optional)**

Never submit passwords through Google Forms.

Powered by  


This content is neither created nor endorsed by Google.  
[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

## ANEXO B <DEPOIMENTOS DOS ENTREVISTADOS>

A seguir estão transcritos todos os depoimentos livres dos entrevistados na pesquisa descrita no Capítulo 3.

**Entrevistado 4:** *“I have grown so much using Kompoz. I always recommend Kompoz to those I know who would be interested. The experience has also given me confidence in my work and also my performance. My co-collaborators have been so supportive and have also influenced me so much with their wide range of music experience.”*

**Entrevistado 6:** *“I joined online collaborating to find a new career, but have been very frustrated because of the lack of good vocalists willing to make a time commitment to an album project.”*

**Entrevistado 7:** *“I started on Kompoz because of a hand injury where I was no longer able to write songs with my guitar. I am a semi-pro musician and needed to keep writing. I was blown away with the talent I found on Kompoz. Once you get into the "in" crowd, you have a better chance of pulling off some major productions. My favorite thing about kompoz is that anyone can contribute. I would say 9 times out of ten, it is someone I don't know that adds something that takes the song in a innovative direction. I depend on people I don't know jumping in to create possibilities I never thought of.”*

**Entrevistado 11:** *“It goes in and out. Sometimes you get a really poor performance that detracts from the song and it's hard to tell people that it just doesn't work. Most often, it yields great results, though”*

**Entrevistado 18:** *“Very "click" oriented. Especially Kompoz, but applies to the others. People tend to stay generally within their groups. Also, there are always people who want to lead their clicks, then try to put the others down. It's a bit childish.”*

**Entrevistado 24:** *“I love collaborating online!! I have experience performing live but I've done way more recording since I've joined this site. It's exciting to be able to try out genres of music that I haven't really explored before and also to create with artists with other cultures & backgrounds. It's exhilarating although the intimacy of live jamming and music isn't present. I'm hoping that this will come in time... as I get to know players & their style (language) better.”*

**Entrevistado 25:** *“I have really enjoyed my collaborations on Kompoz. I love the international aspect. There are some very skilled players on the site.”*

**Entrevistado 29:** *“I have found it most satisfying to be able to collaborate and create music with others especially at Kompoz. I simply love creating music, and I have found that Kompoz is the greatest way to do that with other like minded musicians who share a common interest. To me I listen to songs from Kompoz more than the radio, because everything is original and fresh.”*

**Entrevistado 32:** *“As far as the collaboration websites for music go, it is a great idea. I've been doing this for many years and the one thing I have found is it becomes very cliky with most people, meaning projects now have only invites. I thought it would or had to go in that direction because even though the idea is a very good one when you have 20 people contributing to a project it gets overwhelming, so it was just a matter of time before people started to hook up with certain people.....hell I did it!”*

**Entrevistado 38:** *“Online compositions have been a very positive experience. The chance to work with really talented people from all over the world is amazing. One downfall of online compositions is people come and go, so not always easy to get people to contribute on songs.”*

**Entrevistado 39:** *“Online kompozng kicks ass and is a major part of music evolution in today's world. I don't think the general public or even a lot of musicians or labels realize this yet but this is the future of music and the breaking down of genres where one man or woman's ideas are smashed into another man or woman's ideas from every corner of the earth creating a truly global thing that is not driven by money or fake hollywood bullshit but just the true passion for music and creating something that is unique.*

*I've done a lot of collaborating with local musicians and it is hard to beat working in the same room as someone but sometimes working alone on someone else's idea can be great too, so you don't have to deal with egos or be led astray by other people's ideas of what you should be playing or writing ect.*

*For the most part my experiences collaborating online have been amazing.*

*Negative aspects might be the challenge of being able to express yourself or ideas through emails clearly, but I have also skyped with a few collaborators and that works great to help keep everyone on the same page.*

*I could write about this for hours but I got to fly. I hope this helps you in your research.”*

**Entrevistado 40:** *“online collaborating with other musicians has allowed me to continue being creative and passionate about my music when time & commitment to a 'band' is simply not possible.”*

**Entrevistado 45:** *“I think it is quite time consuming, particularly if you are trying to find the right projects to work with. Additionally getting other musicians to spot your project to collaborate with.*

*I've decided before to receive messages from Kompoz so I can listen to each project submitted.”*

**Entrevistado 46:** *“Generally very good. Sometimes dealing with someone who is not up to par is difficult.”*

**Entrevistado 48:** *“has been hard to find those who are dedicated and/or with time to contribute. Takes a very long time to get those that request to be a part of my session(s) to actually contribute, if they even do at all.*

*Also, there are few, if any, drummers who provide any tracks at all.”*

**Entrevistado 49:** *“I started on Indaba, now active only on Kompoz. Online collaborations is a great learning platform, both as a musician and producer, and also very satisfying on a personal level but it doesn't help to promote you as an artist. I have been on a handful of projects I consider excellent from an artistic point of view but nobody picks up on them. Without any chance of getting 'discovered' I don't think we will ever be able to recruit the top talent.”*

**Entrevistado 50:** *“Great fun and easily available from home when you have family and small kids and so on.”*

**Entrevistado 51:** *“Often find that I need to give more direction for my vision to be realized. But then you need a good relationship with collaborators so that they are not offended when you want different results from what they tried.”*

## ANEXO C <ARTIGO PUBLICADO>

### A Hybrid Approach to Recommend Partners in Collaborative Musical Environments

Guilherme G. Testa<sup>1</sup>, Marcelo S. Pimenta<sup>1</sup>, Damián Keller<sup>2</sup>

<sup>1</sup>Instituto de Informática – Federal University of Rio Grande do Sul (UFRGS) – Porto Alegre – RS – Brazil

<sup>2</sup>NAP – Federal University of Acre (UFAC) – Rio Branco – AC – Brazil

{ggtesta,mpimenta}@inf.ufrgs.br, dkeller@ccrma.stanford.edu

**Abstract.** *Collaboration is an important element in the human creative construction process, and takes special significance in music composition. Systems that support collaborative activities with creative focus have been growing largely lately. However, there is a lack of tools that assists users of these systems to find partners to collaborate with in their activities. In this context, this paper presents a hybrid approach – based on both content-based and collaborative filtering – to recommend partners in creative collaborative environments focused on their produced work. Its main characteristics and functionalities are demonstrated by its design on CODES, a platform for music prototyping for lay users.*

#### 1. Introduction

Collaboration is a key concept in the human creative process. Michael Farrel, in his extensive study about collaborative circles, mentions a passage of Henry James where he declares that “every man produces more when it is in company of people working in the same field of work” [Farrel 2003]. Farrel complements saying that collaborative circles affect significantly the creative work. Artists, writers, composers and politicians reported that cooperation is an indispensable part for the developing of their works. Several important studies of sociology, anthropology and psychology have demonstrated the importance of collective creativity to the intellectual and technological advancement of human society [Palispis 1996][Winthrop 1991][Tuomela 2000].

In general, collaborative processes to build products that are not generated by methodological steps require special support for collaboration. Creative activities tend to feed from mutual inspiration between its collaborators: ideas flow around the discussion of elements that composes the creative work in the pursuit of a satisfactory outcome for all (which is not always known *a priori*). This means that all parts must be tuned to the final results and having similar expectations. But at the same time, it is

desirable that every person could offer new and innovative ideas to contribute in a unique manner to the final result.

In this sense, conventional approaches to stimulate cooperation in the design of technical products, where the objectives and responsibilities are usually well defined, do not allow a systematic and opportunistic trading of ideas [Pimenta et al. 2011]. Therefore, they are not suitable for the dynamic and creative nature typically associated with collaboration among musicians. In this sense, a new concept of awareness is needed to define this new form of collaboration [Miletto et al. 2006]. Similarly, we believe that the recommendation of partners in non-technical byproducts should incorporate some main differences to optimize the suggestion of compatible partners.

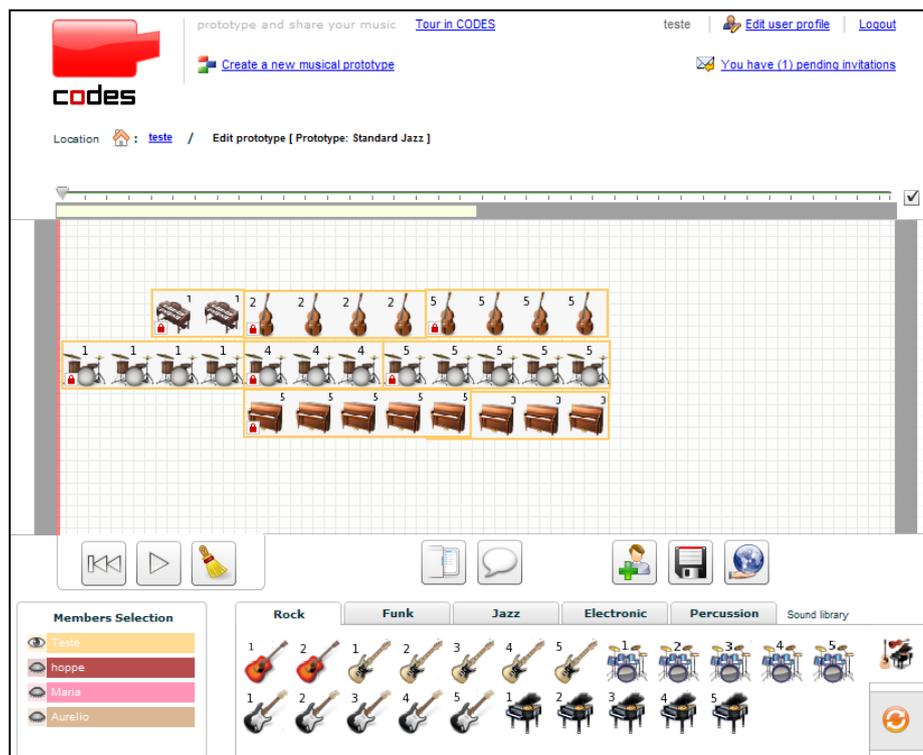
In this context, we present an approach for recommending people (partners) for collaborative musical environments. This proposal will be demonstrated from its implementation and usage in CODES, a collaborative music prototyping environment focused on people with no prior musical experience [Miletto 2010]. Our recommendation approach is hybrid, combining both content-based and collaborative filtering. Thus, characteristics of users are recovered to create musical profiles, so they can be compared and analyzed for partner recommendation.

This article is organized as follows: section 2 presents an overview of CODES. In section 3, we discuss some related works and studies. Section 4 explores the details of the recommendation system being proposed. Section 5 describes the user experience of the partner recommendation in CODES. Finally, section 6 presents some conclusions and future work.

## **2. Exploring CODES**

CODES is a Web-based environment designed to support Cooperative Music Prototyping (CMP), built with special focus on music novices. The main focus of CODES is to approximate the lay user to music creation, without a need of theoretical music concepts. CODES offers a high level music representation and user interface features to foster easy direct manipulation of icons representing sound patterns (predefined MP3 samples with 4 seconds of duration).

Built with appropriate support features, CODES users can create, edit, share and publish simple musical pieces - or Musical Prototypes (MPs) collaboratively to a restricted group or to the Web. These shared MPs can be repeatedly listened to, tested, and modified by the partners who cooperate on prototype refinement. Users can start a new MP just by choosing the name and dragging the musical instruments to the editing area. Edition in CODES includes intuitive actions to build a MP. It is possible to “drag-and-drop” sound patterns from the sound library, “move”, “organize”, “delete”, “expand” the duration, and “collapse” to listen to the final result. When sharing a musical prototype, the “owner” user can invite other users to listen and collaborate with his prototype, or may send explicit invitations via e-mail to non-members asking them for cooperation. When someone accepts such an invitation, the user becomes a prototype partner and can edit the MP like the owner does.



**Figure 1. Editing a MP in CODES**

The prototypical nature of CODES is designed and built to provide a novice-oriented perspective. All the interactions with the editing area are made as intuitive as possible, without any music theory terms. All prototype partners can discuss and change ideas about each step of the prototype refinement, in order to understand each other's decisions. It is possible to link arguments to every decision made. When someone considers that the resulting sounds are good, a "publication request" can be triggered and the group may discuss and deliberate about the publication of this musical prototype in the CODES home page. This activity is called musical prototype publishing. As an alternative to publishing their music, users can export their musical prototype, and share it at will.

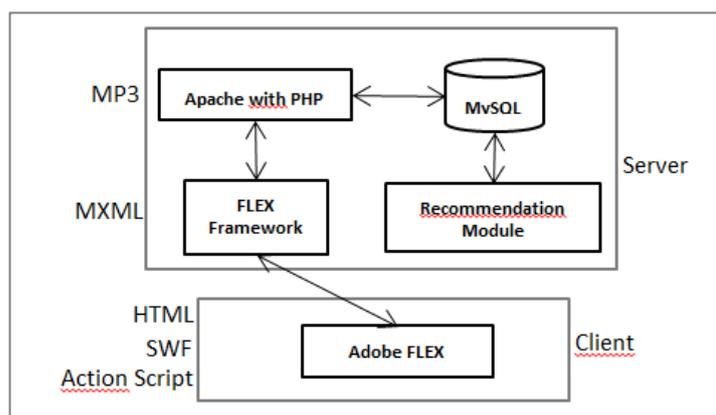
### 2.1. Awareness Mechanisms in CODES

Through CODES, ordinary users may have the opportunity to be the actors of their own musical experiences. Also, the system offers tools to allow a full cooperation among partners. This implies a focus not only on community management but also on experimenting and participating in specific design practices using a suitable interaction vocabulary. This process suggests the existence of noteworthy distinct kinds of cooperation activities. Awareness and conflict resolution are already considered critical issues in general Computer Supported Cooperative Work (CSCW). However, mechanisms existing in other systems need some adaptation to take into account the idiosyncrasies of the CMP context. The ultimate goal is to provide actual cooperation, social knowledge construction, argumentation and negotiation among the actors of the MP design activities. This type of cooperation is supported by a set of mechanisms borrowed from the Software Engineering and Human-Computer Interaction (HCI) areas and specially adapted for CODES [Miletto 2010].

## 2.2. Design and Development of CODES

The design and development of CODES adopted a user-centered and incremental approach, taking into consideration social aspects such as the characteristics of the users, contexts, purpose, minimal technology requirements, and the nature of its possible influence on the novice user.

CODES is based on the classical client-server architecture for Web applications. In the current version of CODES, special attention was given to aspects related to interaction flexibility and usability since one of the main goals is to implement an adequate support for manipulation of complex musical information, cooperative activities and group awareness, to provide an effective interaction of the users with each other and with the environment itself. Thus, in the client-side, CODES uses scripts embedded within standard HTML (Figure 2).



**Figure 2. CODES current architecture**

On the server side, CODES implements the Model-View-Controller (MVC) architecture. The “Model” part (Apache with PHP) connects the Web server with MySQL database, and represents all the information (the data) of the application; and the “Controller” part manages the communication of data and the business rules used to manipulate the data to and from the model. For this, CODES makes use of Adobe MXML (an XML-based language used to lay out user-interface components for Adobe Flex applications). The Adobe Flex as script language was chosen to allow actions like drag-and-drop, use of sliders, scalable windows, and other facilities to manipulate the sound samples provided for this technology, while the HTML 5 standard is still being on development [W3C 2013]. The sound files used in CODES are small MP3 files which can be quickly downloaded by the client-side ensuring a standard audio quality.

The recommendation module connects to the MySQL database to extract information about the users and to write the data structures of the recommendations. This communication does not interfere in the response of the server to the client.

## 3. Related Work

There are several studies that approximate partially to the work being proposed in this paper. CODES and its recommendation system draws from several research areas such as Recommendation Systems, Sound and Music Computing (SMC), Music Recommendation, Computer Supported Cooperative Work (CSCW) and so on.

To introduce the analysis of these areas of study, we will present, in the next subsections, important works classifying them into three main areas: people

recommendation, music recommendation and collaborative environments of music composition.

### **3.1. People Recommendation**

Although recommendation of general items is a problem already well researched both in academia as in business and commercial environments, it is not possible to assert the same about people recommendation. In CODES, our focus is to recommend other users (partners) to collaborate with in the process of music prototyping. Recommending individuals is a difficult task due to the subtle inherent and instinctive aspects of the human relations, especially in creative activities. Studies in this area are growing, mainly focusing on social networks.

In music, a person has greater empathy with people with a similar musical taste. This assertive actually can be propagated to several other areas of interest. When more things in common we have with another individual, more we tend to sympathize and relate better with him [Guy et al. 2010][Guy et al 2009]. The key point here is similarity. Briefly, the problem of recommending people is to construct a profile that reflects faithfully the main characteristics of a user, and then, compare those profiles to look for similarities.

Therefore, to make a good recommendation it is necessary an extra effort than just the application of the usual recommendation algorithms. These differences of recommending people instead of products became evident with the growth of virtual communities, and the direct application of the collaborative filtering technique in such systems demonstrated to be ineffective [Cai et al. 2010]. Individuals differ in ways that the standard algorithms are not adequate to present a good recommendation [Lopes et al. 2010]. It is necessary to adopt further measures to encompass all these details.

In this sense, CODES differs from the recommendation engines usually designed to recommend people because our focus is partnership. People recommendation is mainly driven by virtual communities centered on personal relationships (new friendships, for example), or for technical activities. In a creative environment like CODES, characteristics as empathy are not essentials for a good partnership. Otherwise, is very important that the collaborators have the same goals and are also able to work together efficiently.

### **3.2. Music Recommendation**

Several interesting academic studies about music recommendation can be found in the literature, using different focuses. In generic terms, music recommendation deal with the problem of recommending a list of songs that pleases a certain user. To achieve that, it is necessary to understand the elements that construct the musical taste of a user, including as many details as possible, to feed a personal musical profile of the user. In this sense, several information can be used to enrich this profile, like the explicit searches of the user, his reviews or feedback about songs listened, favorites songs and artists, his personal profile and others [Chen and Chen 2001][Shao et al. 2009][Bogdanov et al. 2011].

With all this information, another question reveals to be crucial: the automatic analysis of the music themselves. To categorize songs, artists and genres efficiently, systems extract semantic information from audio files and use it to group similar artists and songs. Algorithms that provide sound analysis require a considerable amount of

processing, but they are becoming viable with the advancements of hardware. Websites that implements these techniques are growing and becoming popular on the Internet and people are changing for listening downloaded music to streaming media as Grooveshark and Pandora

However, almost all of those websites are focused on simple playback of tracks and are not related to music creation or collaboration itself. Also, the focus of recommendation in CODES is not audio analysis, once the information about the genre and style are already explicit in the system.

### **3.3. Collaborative environments for music composition**

Collaborative environments for music composition are also growing with the increase of interactive community websites. Some interesting services are already available on the Internet today, and they have a significant number of users, which demonstrates the growing potential of this area.

Some of them are Indaba, Kompoz and SoundCloud. All these provide a platform for publication, storage and support collaboration in different ways. In Indaba, the user can also make use of an interesting tool to edit and save directly voice or instrument sounds using a microphone connected to the computer. Kompoz is focused mainly in collaboration itself, and it is possible to upload and publish recordings of musical material so that other users can use freely. Sound Cloud, differently, serves as a repository of sounds and songs.

CODES, unlike, focuses in users who have no previous knowledge of music theory and, in this sense, it differs from the examples cited. Additionally, none of these websites provides recommendations to bring users with similar musical tastes together for musical composition.

## **4. Recommendation in CODES**

There are two approaches widely used in recommender systems in general: content-based and collaborative filtering. The first uses representations of data that were accessed in the past to create user profiles. Based on these profiles, the system then recommends new items for a specific user that are possibly relevant, according to his activity history [Balabanović and Shoham 1997]. In collaborative filtering, similarities between the actions of different users are analyzed to predict a behavior. Users with similar profiles are grouped together to help the recommendation. Collaborative filtering is based on the premise that people who historically have similar activities, are more likely to behave similarly also in the future [Sarwar et al. 2001].

Looking to take advantage of the main benefits of both techniques of recommendation, some hybrid algorithms were created mixing characteristics of these two methods. In a hybrid approach, we use information about both the history of user activities as its similarity with other users to search for compatible profiles. The recommendation can also count with other relevant information to refine the similarity calculation [Kim 2008][Cai et al. 2009].

In CODES, as the user interact with the system, the database is being populated with relevant information that is used by the recommendation engine to produce a musical profile. CODES has a sound library with several predefined samples of instruments of different genres. A combination of these sounds patterns produces a MP.

The number of occurrences of different instruments and different genres are collected of the range of MPs created by the user to approximate the preferred genre for composition of a user. Then, with the musical profile created, similarity calculation takes place to achieve recommendation.

The musical profile of the user is represented by a matrix that is filled while the user composes MPs. These matrixes are then combined and compared with other parameters to mold the final profile. The initial genre matrix is populated as follows.

$$ncomp(p)_u = \frac{\sum_{i=0}^m nmp_{i,u} \times 0,9^i}{\sum_{i=0}^m 0,9^i} \quad (1)$$

Where in (1),  $ncomp(p)_u$  is the combined grade of the prototypes of the user  $p$  calculated separately for each genre  $u$ .  $nmp$  represents the value of the genre  $u$  for a specific music prototype, while  $i$  corresponds to the order of the MPs of the user, being 0 the most recent to the oldest  $m$ . Therefore, each MP will have a 10% decrease in the relative score of its immediate precedent. The sum of all the music prototypes of the user as their respective genres will result in a matrix of genres that takes into account a large number of compositions to quantify the musical preference of a user.

In order to avoid costly calculations that can compromise the scalability of the system, it is preferable to assign a fixed value to calculate the genre matrix for each user [Agrawal and Srikant 1994]. In this case, considering that the compositions themselves are the predominant factor to draw the music profile, this maximum value should not be too low. For this reason, it is initially assumed  $m = 20$ . The genre matrix is updated by the profile manager to each new MP edited or created by the user through the collaborative prototyping module.

#### 4.1. Evaluation of public items

The recommender system in CODES implements a well-known rating system to evaluate the opinion of the user to any item. This explicit information is an important way to get a user feedback about others' MPs. It helps to reflect interests and rejections that may go unnoticed by analyzing only the contributions of a single user.

CODES has a rating system of public MPs that follows a standard widely used in evaluations of items, with grades assigned from 1 to 5. The impact of an evaluation alters the influence of a single MP positively or negatively in the total of the public MPs calculated. Table 1 lists the relative influence of each grade.

**Table 1. Relative influence according different grades.**

Grade assigned	Opinion associated	Relative influence
1	Terrible	-0.25
2	Bad	-0.1
3	Neutral	0
4	Good	+0.1
5	Great	+0.25

For each user, it is constructed an evaluation matrix that stores information about the ratings given by the user to public MPs. Initially, the matrix is completed with equally values distributed to each genre. As the user evaluates MPs, the relative influence of each MP (according to Table 1) of the range of all MPs that the user gave feedback is calculated and the table updated. This matrix is constructed as follows:

$$\mathbf{natr}(p)_{i,u} = \mathbf{natr}(p)_{i-1,u} \times (\mathbf{1} + \mathbf{nmp}(p)_u \times \mathbf{feedback}) \quad (2)$$

In (2),  $\mathbf{natr}(p)_{i,u}$  corresponds to the grade given by the user  $p$  relative to the genre  $u$  to all public MPs that the user gave feedback until  $i$ . Initially,  $\mathbf{natr}(p)_{i,u}$  has a value equally distributed among all genres; and  $\mathbf{feedback}$  represents the value of the relative influence column on Table 1 according to the grade assigned.

#### 4.2. Clustering and similarity calculation

Clustering is an interesting way to reduce the universe of comparisons needed to build a recommendation. First, the information about the user activity history and the public evaluations of data items collected in the previous steps are combined. Then, the users are grouped according to a defined criterion and, finally, the similarity calculation is conducted, where users are actually chosen for the recommendation.

The genre and evaluation matrixes are unified using a weighted arithmetic mean (3), where each grade assigned to a public MP will have an effect of 5% of the final grade, up to a limit of 30%. I.e. the user's own MPs correspond to at least 70% of the final grade. This difference in weights seeks to give greater emphasis to the genres of MPs made by the user himself.

$$\mathbf{nfinal}(p)_u = \begin{cases} \frac{(20 - n) \times \mathbf{ncomp}(p)_u + n \times \mathbf{natr}(p)_u}{20} & \text{if } n < 6 \\ 0.7 \times \mathbf{ncomp}(p)_u + 0.3 \times \mathbf{natr}(p)_u & \text{if } n \geq 6 \end{cases} \quad (3)$$

In (3),  $\mathbf{nfinal}(p)_u$  represents the weighted final score for each of the genres  $u$  for the user  $p$ .  $\mathbf{ncomp}(p)_u$  refers to the relative grade of the genre  $u$  in the genre matrix of the user  $p$ . And finally,  $\mathbf{natr}(p)_u$  corresponds to the given grade of the genre  $u$  in the range of all MPs that the user gave feedback (as explained in section 4.1).

After set the unified genre matrix, the clustering part takes place. Each user is assigned to a specific group according to the majority of its genres of interest. This group separation is intended to reduce the universe of calculations needed for comparison, making the system faster and more scalable [Cataltepe and Altinel 2007]. In CODES, clustering is made according to each user's favorite genres, and these genres are those who contribute, jointly, with at least 80% of the total number of MPs in the unified genre matrix of the user.

The next step is the similarity calculation. One of the most used methods to calculate the proximity between users in recommend systems is the Cosine Similarity (or Vector Similarity), which presents an interesting measure for comparison and prediction [Vozalis and Margaritis 2003]. Note that the summations are calculated over a number of items for which both users  $u$  and  $p$  have expressed their opinions. The correlation then is calculated as follows:

$$dist_{u,p} = \sum_{i=1}^n \frac{r_{i,u} * r_{i,p}}{\sqrt{\sum_{i=1}^n (r_{i,u})^2} \sqrt{\sum_{i=1}^n (r_{i,p})^2}} \quad (4)$$

In (4),  $r_{i,u}$  is the rating of the user  $u$  for the corresponding item  $i$ , in the  $n$ -dimensional item space. If the user didn't rate the current item, the value is set with 0.

A similarity or neighborhood matrix,  $S$ , can now be generated, including the similarity values between all users of a given group. The entry at the  $x$ -th row and the  $y$ -th column corresponds to the similarity between two given users,  $u_x$  and  $u_y$ .

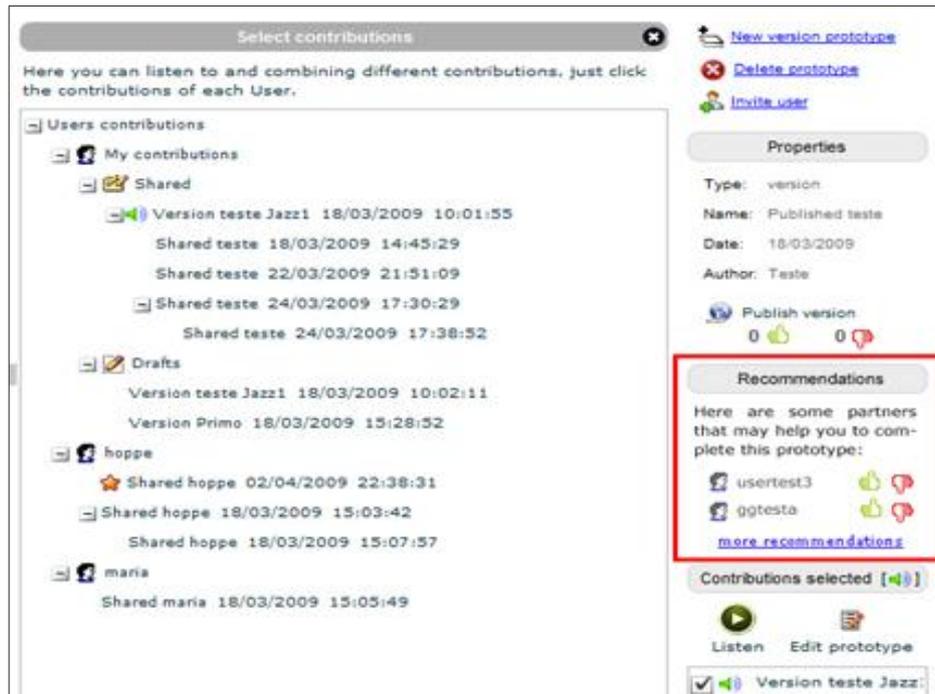
In possession of  $S$ , is possible to recommend the  $n$  closest users of the current user consulting the neighborhood matrix. The distance of the acceptable users for recommendation is dynamic, being evaluated and refined through the user feedback about the previous recommendations. As said in the introduction, in creative activities, sometimes people are not looking for a similar user to interact but, unlike, for someone that can contribute with different ideas. This creates a recommendation area, which is shaped with an increase or decrease of 10% of the similarity value for each evaluation of a collaborator recommended by the system.

The list of users of the recommendation area for a specific user  $u$  is stored in the user profile itself, and is updated regularly. This communication intends to keep the processing and time consumption of the similarity calculations transparent to the user, considering that the total amount of calculations for the whole recommendation procedure can be costly.

## 5. User Experience

As explained in the previous section, the process of recommendation is transparent to the user. A list of suggestions about potential partners is showed in three different situations: in the main page, in the summary page and in the music prototyping edition page.

In the first situation, the user invites a recommended partner to start a new MP. The user writes an invitation detailing his intentions, so the potential collaborator can analyze and decide to accept it or not. If the user denies a suggestion, it will vanish from the list of potential collaborators. Otherwise, if the user accepts and collaborates to prototype a MP, this user is allocated to a list of accepted recommendations. Both users then can leave a feedback about the mutual experience, and this feedback will be used to refine the next recommendation procedure.



**Figure 3. Editing a MP in CODES**

In the other two situations, the users will be invited to collaborate in a MP that is already being prototyped (by one or more users). In these cases, the MP is attached to the invitation, so the invited user can listen to the current MP that he is being invited to collaborate with, and can have a better decision.

## 6. Evaluation

A multi-criteria evaluation of the recommender system proposed in this article is in phase of development, which means that aren't preliminary results of it. The main focus of this evaluation is to get user feedback about the recommendations of partners in CODES, and to adjust eventual needs. The evaluation will be conducted in 5 steps, as follows:

1. Prototyping: a set of 20 different people without prior knowledge about music composition will be invited to create at least 5 music prototypes in CODES, respecting their own musical taste;
2. Evaluation: the set will be indicated to give feedback (rate) of at least 5 others' MPs;
3. Recommendation: with the data collected in stages 1 and 2, the recommendations will be followed by the participants;
4. Cooperation: the users will be guided to create at least 2 MPs with each of the top 2 user recommended;
5. Evaluation: a final questionnaire will be presented to the participants to evaluate the collaborative composition and the recommendation of the users.

This evaluation will be very important to measure the effectiveness of the recommender system presented, as its capacity to give good recommendations. Also it will be an important source to guide future revisions or changes in the recommendation approach.

## 7. Conclusions

In this paper we presented a hybrid approach to recommend partners in a musical collaborative environment. First, we introduce inherent factors that influence partnership work in collaborative creative activities, such as mu-sic. Then we proceeded with a brief presentation about CODES – a platform for music prototyping for lay users –, and an explanation about the recommender engine projected and being implemented in CODES.

Although the preliminary results of the system evaluation with real users have not being collected at this stage, we believe that a hybrid approach concerned and intended to include the subtle aspects of partnership in creative activities is the more suitable approach to make recommendations in collaborative musical environments.

## 8. References

- Agrawal, R. Srikant, R. (1994) “Fast Algorithms for Mining Association Rules in Large Databases”. VLDB '94 Proceedings of the 20th International Conference on Very Large Data Bases. Morgan Kaufmann Publishers Inc. San Francisco, CA.
- Balabanović, M. Shoham, Y. (1997) “Fab: content-based, collaborative recommendation”. Magazine of Communications of the ACM. CACM Homepage archive. Volume 40 Issue 3. ACM New York, NY. 1997.
- Bogdanov, D. Haro, M. Fuhrmann, F. Xambo, A. Gomez, E. Herrera, P. (2011) “A content-based system for music recommendation and visualization of user preferences working on semantic notions”. 9th International Workshop on Content-Based Multimedia Indexing (CBMI). 249-252.
- Cai, J. Francis, J. Gheysens, S. (2009) “Creating a Hybrid Music Recommendation System from Content and Social-Based Algorithms”. GSET '09, Governor's School of Engineering and Technology.
- Cai, X. Bain, M. Krzywicki, A. Wobcke, W. Sok, Y. K. Compton, P. Mahidadia, A. (2010) “Collaborative Filtering for People to People Recommendation in Social Networks”. AI 2010. Lecture Notes in Computer Science. Volume 6464/2011, 476-485.
- Cataltepe, Z. Altinel, B. (2007) “Music recommendation based on adaptive feature and user grouping”. 22nd International Symposium on Computer and Information Sciences, ISCIS,,07. 1-6.
- Chen, H. Chen, A. L. P. (2001) “A Music Recommendation System Based on Music Data Grouping and User Interests”. CIKM '01 Proceedings of the tenth international conference on Information and knowledge management. ACM New York, NY, USA.
- Farrel, M. P. (2003) “Collaborative Circles: Friendship Dynamics and Creative Work”. University Of Chicago Press (Nov. 2003). ISBN 978-0226238678. 328p.
- Guy, I. Jacovi, M. Perer, A. Ronen, I. Uziel, E. (2010) “Same places, same things, same people? Mining user similarity on social media”. Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work. 41-50. ACM New York, NY.

- Guy, I. Ronen, I. Wilcox, E. (2009) "Do you know?: recommending people to invite into your social network". IUI '09: Proceedings of the 14th international conference on Intelligent User Interfaces. ACM New York, NY.
- Kim, K. Lee, D. Yoon, T. Lee, J (2008). "A music recommendation system based on personal preference analysis". First International Conference on the Applications of Digital Information and Web Technologies, ICADIWT,,08. 102-106.
- Lopes, G. R. Moro, M. M. Wives, L. K. Palazzo, J. M. O. (2010) "Collaboration recommendation on academic social networks". ER'10 Proceedings of the 2010 international conference on Advances in conceptual modeling: applications and challenges. 190-199.
- Miletto, E. M. (2010) "CODES: an Interactive Novice-oriented Web-based Environment for Cooperative Musical Prototyping". Ph.D. Thesis. PPGC, Instituto de Informática, UFRGS.
- Miletto, E. M. Flores, L. V. Rutily, J. Pimenta, M. S. (2006) "CODES: supporting awareness in a web-based environment for collective music prototyping". In Proceedings of VII Brazilian symposium on Human factors in computing systems (IHC '06). ACM, New York, NY, USA, 50-55.
- Palispis, E. S. (1996) "Introduction to Sociology and Anthropology". Manila: Rex Printing Company, Inc. ISBN 971234875X. 313p.
- Pimenta, M. S. Miletto, E. M. Flores, L. V. Hoppe, A. (2011) "Cooperative mechanisms for networked music". Future Generation Computer Systems archive. Volume 27, Issue 1, January, 2011. 100-108. Amsterdam, The Netherlands.
- Sarwar, B. Karypis, G. Konstan, J. Riedl, J. (2001). "Item-based collaborative filtering recommendation algorithms". WWW'01 Proceedings of the 10th international conference on World Wide Web. ACM New York, NY, USA.
- Shao, B. Wang, D. Li, T. Ogihara, M. (2009). "Music Recommendation Based on Acoustic Features and User Access Patterns". IEEE Transactions on Audio, Speech, and Language Processing, Vol. 17, No. 8.
- Tuomela, R. (2000) "Cooperation: A Philosophical Study". Kluwer Academic Publishers, Netherlands. ISBN 0-7923-6201-2. 431p.
- Vozalis, E. Margaritis, K. G. (2003) "Analysis of Recommender Systems Algorithms". HERCMA'03 Hellenic European Research on Computer Mathematics & its Applications Conference. 732-745, Athens, Greece.
- W3C. (2013) "HTML 5". Available on <http://www.w3.org/html/>. 2013.
- Winthrop, R. H. (1991) "Dictionary of Concepts in Cultural Anthropology". Greenwood Press, New York. ISBN 0-313-24280-1.