

JOÃO HENRIQUE KERSUL FARIA

**SOCIAL TRACKS - SISTEMA DE
RECOMENDAÇÃO PARA MÚLTIPLOS
INDIVÍDUOS UTILIZANDO DADOS SOCIAIS**

São Paulo
2018

JOÃO HENRIQUE KERSUL FARIA

**SOCIAL TRACKS - SISTEMA DE
RECOMENDAÇÃO PARA MÚLTIPLOS
INDIVÍDUOS UTILIZANDO DADOS SOCIAIS**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para ob-
tenção do Título de Engenheiro Eletricista
com ênfase em Computação.

São Paulo
2018

JOÃO HENRIQUE KERSUL FARIA

**SOCIAL TRACKS - SISTEMA DE
RECOMENDAÇÃO PARA MÚLTIPLOS
INDIVÍDUOS UTILIZANDO DADOS SOCIAIS**

Trabalho apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Engenheiro Eletricista com ênfase em Computação.

Orientador:

Prof.^a Dra. Solange Nice Alves de Souza

São Paulo
2018

AGRADECIMENTOS

Queria agradecer a Profa.Solange, por todo o tempo dedicado e orientação no trabalho, sem ela esta pesquisa não seria possível.

A minha colega Alejandra, pelas discussões sobre sistemas de recomendação que ajudaram no desenvolvimento do projeto.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)(Processos: # 2014/04851-8 e # 2017/04658-1) pelo apoio.

RESUMO

O número de dados gerados nas interações dentro de uma rede social, ou nas interações dentro de recursos de uma plataforma (ex. cliques, acessos, compras), crescem exponencialmente com o passar do tempo. A popularização das redes e o aumento das interações permitem que esses dados possam ser analisados para prever os gostos e desejos de consumidores.

O uso de sistemas de recomendação para filtrar o conteúdo baseado nas características e gostos de um usuário já é largamente difundido e aplicado em diversas plataformas. No entanto, a aplicação de sistemas de recomendação para múltiplos indivíduos é um campo menos explorado. Para este trabalho, o projeto desenvolvido utiliza dados extraídos de redes sociais para recomendar listas de música para um grupo de indivíduos.

Ouvir música em grupo é algo que se faz com frequência, seja em um grupo de amigos, casais ou em festas. Os dados da rede social são utilizados para identificar o indivíduo mais influente no grupo. Além disso, para identificar as preferências, extrai-se características das músicas mais ouvidas pelos integrantes do grupo. Fatoração de matriz é empregada na predição dos interesses do grupo. Uma rede social foi criada para apoiar o experimento controlado, cujos resultados são apresentados.

Palavras-Chave - Sistemas de recomendação, redes sociais, técnicas de agregação, fatoração de matriz, recomendação para grupo.

ABSTRACT

The number of data generated through interactions within a social network, or interactions within a platform's resources (eg clicks, hits, purchases), grow exponentially over time. The popularization of social networks and the increase of interactions allow this data to be analyzed to predict the tastes and desires of consumers.

The use of recommendation systems to filter content based on the characteristics and tastes of a user is already widespread and applied across platforms. However, the application of recommendation systems to multiple individuals is a less explored field. For this project, data was gathered from social networks to recommend music playlists for a group of individuals.

Listening to music as a group is a common activity, be it with friends, couples or in parties. Social network data are used to identify the most influential individual in the group. In addition, to identify the preferences, the characteristics of the most heard songs by the members of the group are assembled. Matrix factorization is used to predict group interests. A social network was created to support the controlled experiment, where the results will be presented.

Keywords - Recommender systems, social networks, aggregation techniques, matrix factorization, group recommendation.

LISTA DE FIGURAS

1	Recomendações da Netflix	10
2	Exemplo simplificado dos modelos de fatores latentes	15
3	Fluxo da recomendação para o sistema proposto	24
4	Rede Social SocialTracks Mastodon	25
5	Tecnologias utilizadas para desenvolvimento da API de Gerenciamento	32
6	Etapas para Cálculo na API de Recomendação	33
7	Fluxo de Autenticação da Aplicação	34
8	Tela Aplicação Móvel - Estatísticas	35

LISTA DE TABELAS

1	Detalhes dos dados utilizados para geração das recomendações individuais .	25
2	Tabela da atividade musical para o Usuário A	38
3	Tabela da atividade musical para o Usuário B	39
4	Dados sociais obtidos do Usuário A	40
5	Dados sociais obtidos do Usuário B	41
6	Dados disponíveis para geração da recomendação individual	42
7	Resultado das recomendações individuais para o Usuário A	42
8	Resultado das recomendações individuais para o Usuário B	42
9	Resultado dos fatores de influência do Usuário A	43
10	Resultado dos fatores de influência do Usuário B	43
11	Recomendação final para o grupo formado pelos usuários A e B	44

LISTA DE SÍMBOLOS

EAP	Estrutura Analítica do Projeto
API	Application Programming Interface
ALS	Alternating Least Squares
SR	Sistema de Recomendação
SRG	Sistema de Recomendação para Grupos
TKI	Thomas-Kilmann Instrument

SUMÁRIO

1	Introdução	10
1.1	Objetivo	11
1.2	Metodologia	11
2	Aspectos Conceituais	13
2.1	Sistemas de Recomendação	13
2.2	Fatoração de Matriz	14
2.3	Técnicas de Agregação	16
2.4	Aspectos da Influência Social na Recomendação para Grupos	17
2.5	Métodos de Avaliação	18
3	Especificação de Requisitos do Sistema	20
3.1	Requisitos Funcionais para o SR proposto	20
3.2	Requisitos Não Funcionais	21
3.3	EAP	21
4	Sistema de Recomendação Proposto	23
4.1	Fluxo de Dados para a Recomendação	23
4.2	Ambiente Controlado	24
4.3	Dados para o Treinamento do SR	25
4.4	Identificação das Preferências Individuais	25
4.5	Modelando a Influência Social	26
4.6	Técnica de Agregação Empregada	29
5	Projeto e Implementação	30

5.1	API de Gerenciamento	30
5.1.1	Descrição do Funcionamento	30
5.1.2	Tecnologias Utilizadas	31
5.2	API de Recomendação	32
5.2.1	Descrição do Funcionamento	32
5.2.2	Tecnologias Utilizadas para a Etapa de Recomendação	33
5.3	Aplicação Móvel de Demonstração	33
5.3.1	Descrição do Funcionamento	34
5.3.2	Tecnologias Utilizadas	35
6	Testes e Resultados	37
6.1	Grupo de Teste	37
6.2	Resultados para as Recomendações Individuais	41
6.3	Resultados para os Fatores de Influência	43
6.4	Resultados Para a Recomendação do Grupo	43
7	Considerações Finais	45
7.1	Conclusões	45
7.2	Contribuições	45
7.3	Perspectivas de Continuidade	45
	Referências Bibliográficas	47

1 INTRODUÇÃO

Em aplicações atuais é cada vez mais comum o uso de sistemas de recomendação (SR) que sugerem produtos e/ou serviços de interesse aos seus usuários. Plataformas, *e-commerces* e redes sociais como Netflix, Spotify, Amazon e Facebook utilizam SR que sugerem ao usuário um conteúdo que se adéqua aos seus interesses pessoais.



Figura 1: Recomendações da Netflix

Fonte: netflix.com

Essas recomendações são geralmente baseadas em características individuais e nas aquisições já realizadas, ou nos interesses já apresentados. A proposta de SR é gerar recomendações únicas para cada usuário. A Figura 1 mostra recomendações feitas com base em filmes já assistidos pelo usuário. Os algoritmos utilizados avaliam os itens de interesse do usuário com base nas preferências identificadas e também com base em dados de outros usuários com interesse similares, mas ainda assim focando em uma recomendação personalizada.

Com a popularização das redes sociais e celulares conectados há maior geração e fluxo de dados diariamente (Quirino et al., 2015). Os dados provindos de redes sociais são ricos em informações sobre como os indivíduos se relacionam, formam grupos e compartilham os mesmos gostos. Vários trabalhos vem sendo publicados nos últimos anos usando dados de redes sociais para melhorar recomendações de produtos e serviços (Prando et al., 2017; Contratres et al., 2018; Gonzalez Camacho and Alves-Souza, 2018). Além disso, há várias

atividades que são feitas em conjuntos de 2 ou mais pessoas como viajar, jogar, assistir filmes, ouvir músicas. Assim, com a captação desses dados pode-se ampliar a atuação de SR, permitindo ir além de recomendações individuais, englobando grupos de pessoas.

É nesse contexto que se encaixam os Sistemas de Recomendação para Grupos (SRG), cujo objetivo é recomendar itens apropriados para um determinado grupo de indivíduos com interesses próprios e comuns aos demais no grupo. Esses interesses comuns são identificados e empregados pelo SRG para gerar as recomendações para o grupo.

Algumas aplicações para SRG, ou com múltiplos indivíduos, são: obter melhor local para viagem em um grupo de pessoas, TV que consegue adaptar sua programação de acordo com as pessoas que estão assistindo, músicas para serem tocadas em um carro com vários passageiros, dentre outras.

O uso de sistemas de recomendação para filtrar o conteúdo baseado nas características e gostos de um usuário já é largamente difundido e aplicado em diversas plataformas Contrates et al. (2018). No entanto, a aplicação de SR para múltiplos indivíduos é um campo menos explorado.

1.1 Objetivo

O objetivo deste trabalho é propor um sistema de recomendação que utiliza dados das redes sociais para identificar o grupo de amigos e a influência social entre os indivíduos do grupo e extrai de uma plataforma musical as preferências destes indivíduos para sugerir músicas que agradam a todos do grupo. Um experimento para validar a proposta faz parte do objetivo do trabalho. Para tal uma aplicação é desenvolvida para simular a interação com site de serviços, que identifica o grupo de amigos e o(s) mais influente(s) na rede social e as preferências musicais destes numa plataforma de músicas. Com base nestes dados a recomendação é apresentada ao grupo de usuários.

Optou-se por fazer um experimento controlado, criando uma rede social (*SocialTracks*) de forma a garantir que haja troca de informações na rede social e também informações musicais na plataforma musical para o grupo que receberá a recomendação.

1.2 Metodologia

O desenvolvimento deste trabalho pode ser dividido da seguinte forma:

- **Aspectos conceituais e análise de referências bibliográficas:** Análise das técnicas e algoritmos empregados no desenvolvimento de sistemas de recomendação e, mais especificamente, sistemas de recomendação para grupos de indivíduos.
- **Entendendo aspectos sociais:** Nesta parte é feita a análise das redes sociais, identificando as formas de interação que podem ser empregadas para determinar a influência entre as pessoas do grupo.
- **Entendendo plataformas musicais** Nesta parte plataformas de músicas são estudadas para entender o mecanismo e como extrair características das músicas ouvidas.
- **Pesquisa por *frameworks* para criação de redes sociais:** Nesta parte *frameworks* para a criação de redes sociais foram pesquisados e analisados para implementar o ambiente controlado.
- **Arquitetura e organização do sistema proposto:** Nesta etapa é definido como o sistema de recomendação proposto está organizado e se conecta com os dados dos usuários. Além de como a aplicação exemplo se comunicará com o sistema de recomendação.
- **Projeto e implementação:** Nesta etapa a implementação do sistema de recomendação proposto é efetivamente implementada.
- **Testes e resultados:** Nesta etapa são exibidos resultados do sistema de recomendação para grupos.

2 ASPECTOS CONCEITUAIS

2.1 Sistemas de Recomendação

Nas relações sociais é natural que pessoas recomendem livros, músicas e filmes uns à outros. Os SR surgiram para assistir e ampliar este processo natural de troca de conteúdo (Resnick and Varian, 1997), sendo geralmente utilizados quando há muitos itens a escolher e é inviável para o usuário ter ciência de todo o conteúdo disponível (Deng et al., 2014; Lalwani et al., 2015).

Os SR típicos são ferramentas de software, que agregam conteúdos que podem ser relevantes a um indivíduo. Para selecionar os itens adequados à um determinado usuário é criado um algoritmo de recomendação, capaz de entender os gostos daquele indivíduo e encontrar itens relacionados (Resnick and Varian, 1997; Al-Hassan et al., 2015).

Em relação as técnicas empregadas, pode-se distinguir os SR em 6 classes principais (Ricci et al., 2011):

- **Baseada em Conteúdo:** a recomendação é feita calculando a similaridade entre o conteúdo que o usuário já gostou anteriormente com novos conteúdos que ele ainda não conhece.
- **Filtragem Colaborativa:** é a técnica mais popular empregada em SR. Diferentemente da abordagem baseada em conteúdo, nesta o objetivo é calcular a similaridade de um usuário em relação a outros. Desta maneira, recomenda-se ao usuário ativo itens que outros usuários com gostos semelhantes gostaram no passado.
- **Demográfico:** recomenda-se itens baseado nos dados demográficos de um usuário. Geralmente são considerados dados como o gênero, a idade, o idioma e o país.
- **Baseado em Conhecimento:** dado um certo problema do usuário, o sistema estima o quanto um item resolve este problema. Um exemplo seria recomendar qual

a melhor câmera para um fotógrafo, baseado no orçamento disponível e o possível uso do produto.

- **Baseado em Comunidade:** utiliza os dados de um conjunto de amigos mais próximos, para aferir quais itens recomendar para um usuário. Nesta abordagem, entende-se que dentre o conjunto de amigos há um subconjunto para o qual a relação é mais forte e este subconjunto tem maior influência e/ou afinidade com o usuário que recebe a recomendação.
- **Sistemas de Recomendação Híbridos:** são combinações das abordagens anteriores. Geralmente são empregadas com o propósito de utilizar a vantagem de alguma técnica para suprir a desvantagem de outra.

2.2 Fatoração de Matriz

Para implementar a filtragem colaborativa, existem duas técnicas principais de atuação, os métodos de vizinhança e os modelos de fatores latentes (Koren et al., 2009; Desrosiers and Karypis, 2011). Os métodos de vizinhança são empregados para estimar avaliações para itens com base na relação entre itens ou entre usuários. Para itens, a relação é encontrada utilizando as avaliações feitas pelo usuário aos itens vizinhos ao não avaliado. (Desrosiers and Karypis, 2011). Similarmente, para o método de vizinhança baseado na relação entre usuários, a avaliação para um item para um determinado usuário é estimada com base nas avaliações feitas por outros usuários similares a ele ao item em questão.

Os modelos de fatores latentes operam caracterizando itens e usuários em fatores inferidos através das avaliações dos usuários. Em (Koren et al., 2009) esses fatores são comparáveis às categorias dos itens, como gênero de um filme, mas a diferença é que eles são inferidos computacionalmente, baseado unicamente nas avaliações dos usuários.

A figura 2 ilustra de maneira simplificada como os modelos de fatores latentes caracterizam itens. Para este exemplo consideram-se duas dimensões. O eixo X representa o gênero (feminino e masculino) e o eixo Y representa o realismo (sério e escapista). De acordo com suas características latentes, os filmes são distribuídos pelo plano. Assim, filmes mais voltados ao público masculino são plotados no lado positivo do eixo X, enquanto os mais voltados ao público feminino, no lado negativo. Por outro lado, filmes caracterizados como mais sérios aparecem no lado positivo do eixo Y e os mais escapistas (de maior ficção), do lado negativo. Da mesma forma, pessoas são plotadas no gráfico com base em suas avaliações prévias. Para este modelo específico, uma predição de um

filme para um usuário seria igual ao produto escalar da posição do filme com a posição do usuário nesse gráfico, ou seja, quanto mais próximos maior é o grau de predição desse item para esse usuário. Fatores latentes são, portanto, um mapa dos usuários e dos itens em um espaço de dimensão f (Koren et al., 2009).

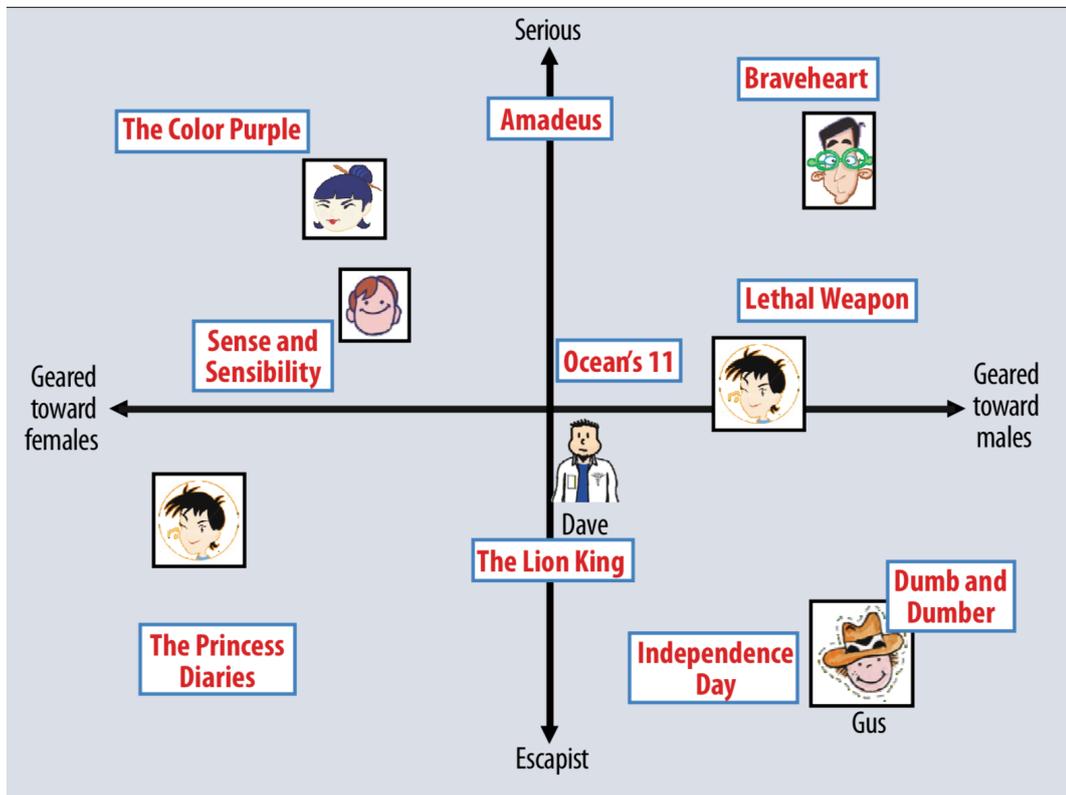


Figura 2: Exemplo simplificado dos modelos de fatores latentes

Fonte: (Koren et al., 2009)

Fatoração de matriz é uma abordagem de modelo de fator latente. (Koren et al., 2009) explica o funcionamento de um algoritmo básico de fatoração de matriz, no qual cada item i é associado à um vetor q_i de fatores latente estimados a partir das características dos itens. Cada usuário u é associado à um vetor p_u , que mede o quanto é o interesse do usuário u nos itens que possuem altos valores nos fatores correspondentes. O produto $q_i \cdot p_u$ corresponde à interação entre o usuário u e o item i . Ou seja, este produto corresponde à predição r_{ui} da avaliação do usuário para o item como se mostra na equação 2.1:

$$r_{ui} = q_i \cdot p_u \quad (2.1)$$

Em termos computacionais, o maior gasto de processamento ocorre no cálculo de q_i e p_u . Koren et al. (2009) apresenta dois modelos principais para realizar esse cálculo, o

método de Gradiente Descendente Estocástico (*Stochastic gradient descent*) e o método dos Mínimos Quadrados Alternados (*Alternating Least Squares* ou *ALS*).

Como evidenciado em Koren et al. (2009); Koren and Bell (2011) o ALS é um algoritmo muito eficiente. Primeiramente por permitir que os cálculos ocorram de forma paralelizada, uma vez que os parâmetros da matriz são calculados de maneira independente. Além disso, o ALS lida de forma mais eficiente em sistemas centrados em dados implícitos (Hu et al., 2008), como é o caso de aplicações musicais, onde nem sempre se tem um dado explícito da avaliação do usuário para uma determinada música, mas infere-se implicitamente através do número de vezes que ele a ouviu.

Portanto, usando fatoração de matriz é possível prever a avaliação de um usuário para um conjunto de itens, baseado na avaliação implícita ou explícita de outros usuários para este mesmo conjunto de itens (Koren and Bell, 2011). Tal informação apenas revela as recomendações individuais para um usuário, as recomendações para um grupo requerem a agregação dessas recomendações.

2.3 Técnicas de Agregação

Técnicas de agregação são utilizadas para combinar recomendações individuais são combinadas, gerando uma recomendação que pode satisfazer a um grupo de usuários.

As principais técnicas de agregação Ricci et al. (2011) são detalhadas a seguir.

- **Votação Plural:** utiliza a votação *first-past-the-post*, em que o item com mais votos é escolhido. Em seguida, retira-se o item com mais votos e repete-se a votação, gerando uma ordem de preferência.
- **Média:** obtém a média da avaliação para um dado item.
- **Multiplicativo:** multiplica-se a avaliação de cada indivíduo, obtendo um valor para a avaliação do grupo.
- **Contagem de Borda:** conta pontos para itens de acordo com o ranking deste item para cada indivíduo. O último item do ranking recebe 0 pontos, o seguinte, 1 ponto e assim por diante.
- **Regra de Copeland:** conta o quanto um item ganha de outros itens.
- **Votação por Aprovação:** conta o número de vezes que um item foi avaliado acima de um determinado limite.

- **Menor Insatisfação:** considera a menor avaliação como a avaliação do grupo.
- **Maior Satisfação:** considera a maior avaliação para a avaliação do grupo.
- **Média sem Insatisfação:** realiza a média das avaliações, desconsiderando itens com a avaliação menor que um valor limite.
- **Equidade:** itens são avaliados como se indivíduos estivessem escolhendo-os em turnos. Por exemplo, o primeiro item é o melhor avaliado do usuário A, o segundo é o melhor avaliado do usuário B e assim por diante.
- **Pessoa mais respeitada:** utiliza a avaliação do indivíduo mais respeitado.

Cada estratégia de agregação gera uma avaliação do grupo para um determinado item ou um conjunto de itens. Por exemplo, utilizando a técnica da "Média", se um item é avaliado por A como 8, por B como 9 e por C como 10, a avaliação do grupo composto por A, B e C é 9 (a média entre as três avaliações).

Nem sempre temos um número como resultado, podemos ter uma ordem de preferência, como é o caso da técnica "Votação Plural", em que ordenam-se os itens com a maior avaliação pela maioria do grupo. Por exemplo, em um grupo com 5 indivíduos, um item avaliado com 5 notas 10 fica na frente de um item avaliado com 4 notas 10 e uma nota 9.

2.4 Aspectos da Influência Social na Recomendação para Grupos

Quando se recomenda para um grupo, outros processos sociais precisam ser levados em conta que não precisariam no caso da recomendação individual. Por exemplo, caso uma recomendação possa ser de alguma forma embaraçosa para um dos integrantes do grupo, isso poderia afetar a percepção de todo o grupo (Masthoff and Gatt, 2006).

Masthoff and Gatt (2006) apresenta dois mecanismos da psicologia social que são essenciais para entender como a satisfação de um grupo pode ser obtida. O primeiro mecanismo é chamado de "Contágio Emocional", em que a satisfação de alguns usuários pode levar a satisfação de outros. O contrário também pode acontecer, ou seja, a insatisfação de alguns usuários podem levar a insatisfação de outros. O segundo mecanismo é chamado de "Conformidade", no qual a opinião de outros usuários pode influenciar a opinião de um usuário. A conformidade pode ser de dois tipos (Masthoff and Gatt, 2006):

influência normativa, na qual o indivíduo por querer fazer parte de um grupo, expressa uma opinião igual a do grupo, mesmo que interiormente não concorde. O outro tipo é a influência informacional, em que a opinião do indivíduo de fato muda, pois acredita que o grupo está correto.

Wang et al. (2014) comenta sobre o conceito de influência, em que um usuário pode ter maior poder de contágio e conformidade em um grupo. Descobrir o usuário com maior influência em um grupo passa a ser relevante, pois satisfazendo este usuário, aumenta-se as chances de satisfazer todo o grupo. Zhu and Huberman (2014) mostra como as escolhas de pessoas são afetadas por recomendações de outrem.

Ricci et al. (2011) combina as técnicas de agregação com os mecanismos de psicologia social aqui detalhados, culminando em três estratégias possíveis:

- Apoiar Fortemente o Mais Rabugento: esta estratégia recomenda o item que a pessoa menos satisfeita mais gosta.
- Suportar Fracamente o Mais Rabugento: esta estratégia seleciona itens que são razoavelmente apreciados pelo membro menos satisfeito (itens com avaliação 8/10 ou acima).
- Com Pesos: esta estratégia adiciona pesos aos usuários, dependendo da sua satisfação, assim, utiliza os pesos no momento de realizar a agregação.

As estratégias apresentadas só podem ser implementadas caso o SR possua um mecanismo de feedback, na qual a satisfação do usuário possa ser medida em tempo real.

2.5 Métodos de Avaliação

Uma vez definidos algoritmos, conjunto de dados e o processo de implementação do SR, é necessário criar processos claros para avaliar se o conteúdo oferecido fez sentido para o usuário.

Ricci et al. (2011) detalha 3 tipos de experimentos que podem ser conduzidos para avaliar a qualidade de um SR:

- Experimentos Offline: utiliza-se dados já obtidos de usuários, desta maneira é possível simular o usuário utilizando o sistema e interagindo com as recomendações. A vantagem deste tipo de experimento é que não requer a interação com usuários

reais. No entanto, é necessário que exista um conjunto de dados prévios para realizar a avaliação.

- Estudos com Usuários: pode-se convidar grupos de usuários para interagir com o SR. Durante essa interação é importante que todos os dados de ações no sistema sejam gravados para posterior análise. Um experimento tradicional para testar SR é pedir para diferentes grupos de usuários interagirem com conteúdos que lhes parecerem interessantes. Para cada grupo podem ser apresentados algoritmos distintos, podendo assim testar sua acurácia. A vantagem deste método é entender com o usuário interage com o sistema e como se comporta em relação às recomendações. O grande problema deste método é o custo necessário para realizar este experimento.
- Avaliação Online: é feita com usuários reais, utilizando o sistema de fato. Neste caso não há um questionário para o usuário, ele usará o sistema sem qualquer interferência de um mediador. Os dados para avaliar o SR são pelos próprios usuários ao interagirem com o sistema.

3 ESPECIFICAÇÃO DE REQUISITOS DO SISTEMA

Neste capítulo são especificados os requisitos para o SR proposto. Os requisitos são divididos em funcionais e não funcionais. Os funcionais são aqueles implementados por software. Os não funcionais devem ser atendidos pelo SR e para tal, um conjunto de fatores podem ser definidos/implementados ou seguidos para que determinado requisitos não funcional seja observado.

3.1 Requisitos Funcionais para o SR proposto

1. criação de rede social específica para o SR proposto - a criação de uma rede social é necessária para viabilizar o experimento controlado. Este é feito para garantir que os dados necessários para validar o SR proposto existam. Mais especificamente, um conjunto de usuários utilizam a rede social criada, interagem entre si, evidenciando grupos de usuários. Indivíduos pertencentes ao grupo ouvem músicas numa plataforma musical, da qual características das músicas podem ser recuperadas. Características dos grupos e de seus integrantes são extraídas da rede social e utilizadas para gerar a recomendação.
2. Rede social semelhante ao *Twitter* - entende-se que a similaridade com uma rede social largamente utilizada por indivíduos em geral facilita a interação e, consequentemente, o uso da rede social criada para o experimento controlado. Assim, além da interface similar a uma rede já conhecida, funcionalidades existentes nesta também devem estar presentes na rede social criada. Somente funcionalidades necessárias ao experimento controlado foram implementadas na rede criada.
 - (a) Criar grupos na rede social
 - (b) Adicionar amigos na rede social
 - (c) Ter um perfil e publicar textos pessoais

- (d) Curtir e comentar em publicações de outras pessoas
- 3. Aplicação móvel que receberá as recomendações das listas de música
 - (a) Conexão com a Rede social desenvolvida
 - (b) Conexão com o Spotify, para obter dados de música
 - (c) Criar grupos com os amigos da rede social
 - (d) Usar os grupos já criados na rede social
 - (e) Obter lista de músicas
 - (f) Exportar lista de músicas para o Spotify
 - (g) Ouvir lista de músicas no próprio aplicativo
- 4. API do Sistema de Recomendação
 - (a) Receber dados da Rede Social
 - (b) Receber dados do Spotify
 - (c) Receber dados de grupos da aplicação móvel
 - (d) Gerar recomendações utilizando os dados da rede social, Spotify e o grupo formado na aplicação móvel

3.2 Requisitos Não Funcionais

1. Tempo de resposta da recomendação menor que 5 segundos.
2. Garantir a privacidade dos usuários durante a recomendação

3.3 EAP

Para a especificação das tarefas e subdivisões do projeto, será utilizado a EAP (Estrutura Analítica do Projeto).

1. Documentação técnica
 - (a) Estudo sobre sistemas de recomendação
 - i. Obtenção de artigos sobre o assunto
 - (b) Estudo sobre técnicas de agregação e sistemas de recomendação para grupos

- i. Verificação da abrangência de pesquisas sobre os sistemas de recomendação para múltiplos indivíduos
 - ii. Obtenção de artigos com técnicas de agregação relevantes para o problema que será estudado
 - (c) Descrição das técnicas de agregação e uma análise sobre suas vantagens e desvantagens
 - (d) Descrição do funcionamento do sistema de recomendação
 - (e) Analisar dados que serão necessários da rede social para a agregação
 - (f) Descrição do funcionamento da rede social
2. Integração com Rede Social
- (a) Seleção de uma rede social aberta, onde a obtenção de dados seja possível
 - (b) Integração da rede social com a aplicação
 - (c) Obtendo dados da rede social
3. Criação da aplicação móvel
- (a) Criação da tela para conectar com rede social e Spotify
 - (b) Criação da tela para obter recomendações
 - (c) Criação de tela para deixar transparente ao usuário como a recomendação foi calculada
4. Prototipação do Sistema de Recomendação para Grupos
- (a) Obtenção de um conjunto de dados para treinamento do algoritmo
 - (b) Implementação de um algoritmo inicial, realizando análises de sua qualidade
5. Implantação do Sistema de Recomendação como um serviço
- (a) Criação de uma API, que receberá os dados da rede social para gerar as recomendações
 - i. Definição da linguagem da API e dos requisitos técnicos necessários
 - ii. Implementação da estrutura da API
 - iii. Implantação em um servidor para testes remotos
 - (b) Implantação do sistema de recomendação

4 SISTEMA DE RECOMENDAÇÃO PROPOSTO

Neste capítulo detalha-se o SR proposto, mostrando-se como acontece o fluxo de recomendação, as opções feitas para a implementação do ambiente controlado, os dados utilizados para o treinamento do SR, as técnicas empregadas. Além de detalhar a modelagem dos dados sociais.

4.1 Fluxo de Dados para a Recomendação

O projeto é constituído de duas partes principais que se relacionam diretamente: o SR proposto e um aplicativo para permitir a interação dos indivíduos com o SR. Este aplicativo recupera dados na rede social sobre a interação entre os indivíduos que integram o grupo, recupera de um provedor de atividade musical (*Spotify*) as músicas ouvidas por cada integrante do grupo, envia as músicas e suas características para o SR e recebe a lista de músicas recomendadas.

O fluxo da recomendação começa com um usuário se conectando ao aplicativo (aplicação móvel, ou seja acessível via um dispositivo móvel), usando suas credenciais da rede social. O usuário também informa o grupo com o qual ouvirá músicas (representando a atividade que farão em grupo). A Figura 3 esquematiza esse fluxo. A aplicação móvel se conecta ao mesmo tempo com a rede social e com o *Spotify*, recuperando dados da atividade musical do grupo de amigos, o que inclui o usuário. Usando a aplicação é ainda possível criar grupos baseado nos amigos fornecidos pela rede social, esta funcionalidade permite definir com qual o grupo de amigos se fará a atividade conjunta.

O SR utiliza os dados da rede social para aplicar as técnicas de agregação detalhadas na 2.3 para criar a recomendação para um grupo de pessoas, a qual é posteriormente mostrada também pelo próprio aplicativo.

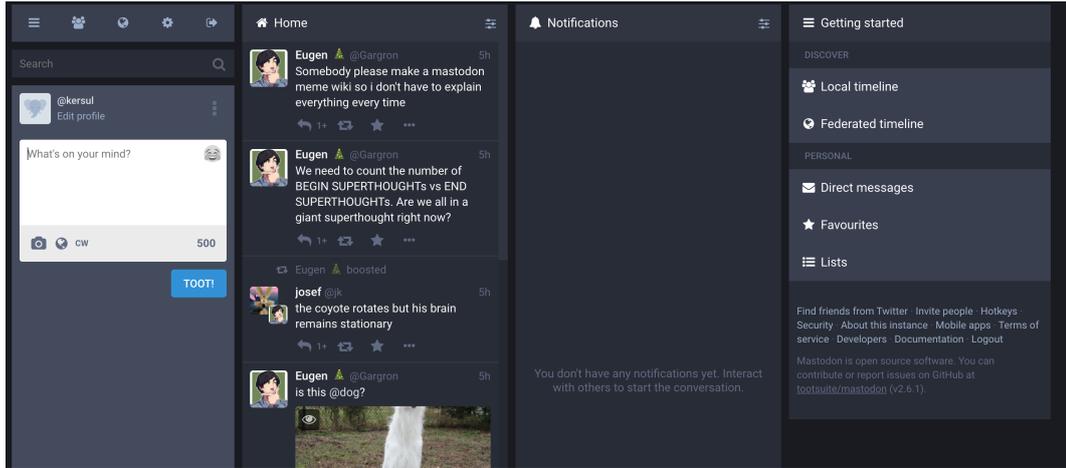


Figura 4: Rede Social SocialTracks Mastodon

Fonte: <https://socialtracks.masto.host/about>

4.3 Dados para o Treinamento do SR

Para treinar o algoritmo e realizar os testes necessários, foram utilizados dados da API pública do Last.fm. Este é um agregador online de dados musicais. Nessa API os dados de todos os usuários são abertos, diferentemente de outras APIs semelhantes, como a do Spotify, em que só é possível obter informações do usuário logado. A Tabela 1 exibe os tipos de dados obtidos.

Tipo de Dado	Detalhamento
Música	Nome, duração e artista
Usuário	Nome, país, idade e playcount (número total de vezes que o usuário escutou músicas)
Usuário-Música	Usuário, música e playcount (número de vezes que o usuário escutou determinada música)

Tabela 1: Detalhes dos dados utilizados para geração das recomendações individuais

4.4 Identificação das Preferências Individuais

A recomendação para o grupo é gerada a partir das preferências individuais. As técnicas de agregação detalhadas na 2.3 são utilizadas para esse fim. Para realizar a agregação das preferências dos usuários é necessário, primeiramente, ter conhecimento da afinidade dos usuários de um grupo para cada um dos itens disponíveis para a reco-

mendação. Assim, nesta etapa obtém-se os interesses individuais dos integrantes para posteriormente fazer a agregação desses interesses.

Os dados contantes da Tabela 1 são exemplos de dados de avaliações implícitas, pois não fornecem avaliações diretas dos usuários para músicas, mas sim o número de vezes que um usuário ouviu determinada música.

Para o SR proposto as recomendações individuais são definidas através da técnica de filtragem colaborativa, implementando a técnica de fatoração de matriz, mais especificamente, o ALS. Assim, emprega-se o ALS para inferir a afinidade de usuários para um conjunto definido de itens, utilizando avaliações implícitas de outros usuários.

4.5 Modelando a Influência Social

(Guo et al., 2016) detalham sobre a influência de um usuário perante outras pessoas, evidenciando como esta influência pode ser um fator chave para a recomendação. Para determinar a influência de um usuário em um grupo, os autores utilizam cinco fatores como base para os cálculos:

- *Expertise Factor*: conhecimento de um usuário sobre um dado tópico. Geralmente, a opinião de especialistas são mais adotadas que a de outros.
- *Susceptibility Factor*: considera o quanto um usuário é suscetível a opiniões e emoções de outros usuários.
- *Personality Factor*: definido pelo padrão de comportamento do indivíduo.
- *Intimacy Factor*: mensura quanto um grupo está conectado. Quanto mais os integrantes de um grupo são íntimos, maior a probabilidade de aceitarem as opiniões de outros.
- *Similarity Factor*: avalia grau de similaridade entre os integrantes do grupo. Obtido através de atividades e informações em comum.

Na abordagem proposta por (Guo et al., 2016), os fatores de influência são de difícil automatização. Por exemplo, o teste de identificação de tendências comportamentais (TKI) para definir o *Personality Factor*, é determinado realizando o teste TKI com cada integrante do grupo. Apesar de interessante, numa aplicação real, o tempo necessário

para fazer o teste pode ser um desestimulador para continuar no ambiente para receber a recomendação.

Assim, nesta monografia propõe-se um conjunto 4 novos fatores para determinar a preferência de grupos de usuários. Diferentemente de (Guo et al., 2016), para o cálculo destes fatores dados provindos de redes sociais são empregados com maior preponderância. Os fatores para definir a influência social propostos são:

- *Expert Factor*: conhecimento de um usuário em um dado tópico. Para música, o quanto determinado usuário ouve e conhece artistas.
- *Leader Factor*: avalia o quanto um usuário representa a figura de um líder para o grupo. É medido pela repercussão de suas postagens relativamente a dos demais integrantes do grupo.
- *Friendly Factor*: mensura quanto um grupo está conectado, calculado em pares de usuários. Pode ser calculado utilizando amigos em comum, conversas em comum e atividades dentro da rede social em comum.
- *Similarity Factor*: mensura quão similar é cada par de usuários de um grupo. Calculado observando as atividades em comum entre cada par de usuários dentro da rede social.

As equações 4.2, 4.4, 4.6 e 4.8 foram elaboradas para a determinação e normalização de cada um dos fatores de influência. Os cálculos são feitos construindo-se uma matriz para cada um dos fatores (equações 4.1, 4.3, 4.5 e 4.7). Nesta, a média dos valores das colunas de uma linha i da matriz corresponde ao fator para o usuário i . Para cada um dos cálculos, i e j se referem a dois usuários pertencentes à um grupo.

- *Expert Factor (Fexp)*

$$E_{ij} = \sum_{i,j=1,i \neq j}^{nu} \frac{\frac{nf_i}{nf_i+nf_j} + \frac{na_i}{na_i+na_j} + \frac{nm_i}{nm_i+nm_j}}{3} \quad (4.1)$$

Onde E_{ij} é a matriz que mede a *expertise* dos usuários, nu é número de usuários do grupo, nf é o número de seguidores no aplicativo musical do usuário, na é o número de artistas que o usuário segue e nm é o número de músicas que o usuário possui em sua biblioteca.

$$Fexp_i = \frac{\sum_{j=1}^{nu} E_{ij}}{nu} \quad (4.2)$$

Onde $Fexp_i$ é o *Expert Factor* para o usuário i , nu é número de usuários do grupo.

- *Leader Factor*

$$L_{ij} = \sum_{i,j=1,i \neq j}^{nu} \frac{\frac{tl_i}{tl_i+tl_j} + \frac{tm_i}{tm_i+tm_j}}{2} \quad (4.3)$$

Onde L_{ij} é a matriz para avaliar a repercussão das postagens dos usuários, nu é número de usuários do grupo, tl_i é o número total de curtidas que o usuário i recebeu, tm_i é o número total de menções que o usuário i recebeu. Análogo para tl_j e tm_j

$$Flead_i = \frac{\sum_{j=1}^{nu} L_{ij}}{nu} \quad (4.4)$$

Onde $Flead_i$ é o *Leader Factor* para o usuário i e nu é número de usuários do grupo.

- *Friendly Factor*

$$F_{ij} = \sum_{i,j=1,i \neq j}^{nu} \frac{\frac{m_{j \rightarrow i}}{tm_j} + \frac{c_{j \rightarrow i}}{tc_j}}{2} \quad (4.5)$$

Onde F_{ij} é a matriz que relaciona o total de menções e comentários sobre as postagens feitas por cada usuário do grupo, nu é número de usuários do grupo, $m_{j \rightarrow i}$ é o número de postagens que o usuário j mencionou o usuário i , $c_{j \rightarrow i}$ é o número de postagens que o usuário j comentou do usuário i .

$$Ffri_i = \frac{\sum_{j=1}^{nu} F_{ij}}{nu} \quad (4.6)$$

Onde $Fsoc_i$ é o *Friendly Factor* para o usuário i e nu é número de usuários do grupo.

- *Similarity Factor*:

$$S_{ij} = \sum_{i,j=1,i \neq j}^{nu} \frac{\frac{l_{j \rightarrow i}}{tl_j} + \frac{s_{j \rightarrow i}}{ts_j}}{2} \quad (4.7)$$

Onde S_{ij} é a matriz que relaciona as postagens de interesse e compartilhadas por cada usuário, nu é número de usuários do grupo, $l_{j \rightarrow i}$ é o número de postagens que o usuário j gostou do usuário i , $s_{j \rightarrow i}$ é o número de postagens que o usuário j compartilhou do usuário i .

$$Fsim_i = \frac{\sum_{j=1}^{nu} S_{ij}}{nu} \quad (4.8)$$

Onde $Fsim_i$ é o *Similarity Factor* para o usuário i e nu é número de usuários do grupo e S_{ij} .

Cada um desses fatores apresentados pelas equações X, Y, Z e E já estão normalizados para uma escala de 0 a 1, ou seja, um usuário com *Expert Factor* com valor 1 possui nota máxima nesse quesito. Estes fatores de influência serão considerados para o cálculo da avaliação do grupo na aplicação da técnica de agregação.

4.6 Técnica de Agregação Empregada

Ricci et al. (2011) detalham dois experimentos que foram conduzidos para avaliar qual técnica de agregação faz mais sentido para um grupo.

No experimento 1, os participantes receberam uma tabela com avaliações de um grupo de usuários e precisavam indicar qual seria a recomendação correta para este grupo. Percebeu-se que os participantes tentavam evitar a insatisfação, desconsiderando itens nos quais pessoas avaliaram muito mal. Verificou-se que os usuários usaram mais claramente as técnicas de "Menor Insatisfação", "Média" e "Média sem insatisfação".

No experimento 2, os participantes receberam sequência de itens obtidos por estratégias de agregação diversas, bem como as avaliações individuais para cada item. Possuindo as avaliações em mãos, os participantes decidiam qual sequência de itens seria melhor recebida pelo grupo. A técnica "Multiplicativa" foi a que obteve melhores resultados.

Portanto, pela simplicidade e os bons resultados nos experimentos de Ricci et al. (2011), no SR proposto neste projeto utiliza-se a técnica da "Média sem Insatisfação".

5 PROJETO E IMPLEMENTAÇÃO

Neste capítulo detalha-se sobre as opções feitas para a implementação do SR proposto.

A implementação do sistema de recomendação proposto foi dividida em três sistemas: API de Gerenciamento, API de Recomendação e Aplicação Móvel.

5.1 API de Gerenciamento

O passo anterior à recomendação é a obtenção dos dados sociais e da atividade musical do usuário. Estes dados precisam ser organizados e formatados. Após a recomendação ser realizada também é necessário a organização dos resultados para enviá-los de volta à aplicação móvel para ser apresentada ao usuário.

5.1.1 Descrição do Funcionamento

Dado as características do fluxo da recomendação (Figura 3), a parte do sistema, chamada de API de Gerenciamento, é responsável pela conexão entre a aplicação móvel e a parte que faz efetivamente a recomendação (API de Recomendação). A API de Gerenciamento possui as seguintes funções:

- **Conexão com serviços externos:** Estabelece a conexão com APIs do Spotify e do Mastodon, obtendo os dados necessários para identificação dos usuários, dados sociais e dados da atividade musical do usuário.
- **Autenticação dos usuários:** Gerencia se um usuário pode se logar no sistema, acessando seu cadastro na plataforma e seus dados.
- **Gerenciamento dos grupos:** Gerencia a criação e remoção de grupos dentro da plataforma. Armazena os usuários pertencentes a um grupo.

- **Compilação dos dados para o sistema de recomendação:** Organiza e compila os dados dos usuários e envia para o sistema de recomendação.
- **Gerencia as recomendações obtidas do sistema de recomendação:** Após uma recomendação ser recebida pelo sistema de recomendação, a API de Gerenciamento armazena esses dados e notifica a aplicação de demonstração que a recomendação está pronta.

Logo, a API de Gerenciamento trabalha como uma interface entre a aplicação que consumirá as recomendações e o SR que as fornecerá, gerenciando os grupos de usuários e compilando os dados.

5.1.2 Tecnologias Utilizadas

A API de Gerenciamento é implementada como um serviço web e deve fornecer à aplicação rápido tempo de resposta e alta disponibilidade. Este serviço não requer alto poder de processamento, pois os cálculos computacionais são relativamente simples, geralmente transmissão e transformação básica dos dados. A Figura 5 mostra as tecnologias empregadas no desenvolvimento da API de Gerenciamento. Para o desenvolvimento do serviço web foi utilizado o framework Express.js, um framework para criação de sistemas web para Node.js. O Express.js é um framework para criação de serviços web em Javascript. Para o banco de dados a aplicação exige flexibilidade no armazenamento dos dados, uma vez que serão um conjunto grande de dados providos de múltiplos sistemas externos, como Spotify e Mastodon. Banco de dados relacionais como MySQL e PostgreSQL possuem organização mais restrita, o que dificultaria o desenvolvimento da aplicação devido a natureza dos dados. Assim, optou-se por um banco de dados NoSQL orientado à documentos. O MongoDB foi escolhido devido a sua flexibilidade no armazenamento de dados, facilidade de uso e prévio conhecimento do autor. Este tipo de banco de dados permite que a informação seja gravada sem ser necessário um esquema pré-definido.



Figura 5: Tecnologias utilizadas para desenvolvimento da API de Gerenciamento

Fonte: expressjs.com, nodejs.org e mongodb.com

5.2 API de Recomendação

A aferição das avaliações dos usuários para um conjunto de itens é uma tarefa computacionalmente trabalhosa, como descrito em 4.5. É necessário a análise de um grande conjunto de dados, o que pode demorar muito tempo e exigir máquinas realmente potentes para realizar o cálculo. Por isso, o sistema foi arquitetado projetando-se a API de Recomendação totalmente separada dos processamentos básicos realizados pela API de Gerenciamento.

5.2.1 Descrição do Funcionamento

A API de Recomendação é o principal foco deste trabalho, é a parte do sistema em que os estudos realizados no capítulo 4 deste trabalho serão de fato aplicados.

Como apresentado na Figura 3, o cálculo da recomendação se inicia com o envio das informações da atividade musical e dos dados sociais do usuário. É de responsabilidade da API de Gerenciamento a aquisição, formatação e envio desses dados para a API de Recomendação.

Após o envio dos dados, o cálculo da recomendação é realizado em três etapas principais: cálculo da influência dos usuários do grupo, cálculo das recomendações individuais e a agregação das recomendações individuais considerando a influência (Figura 6).

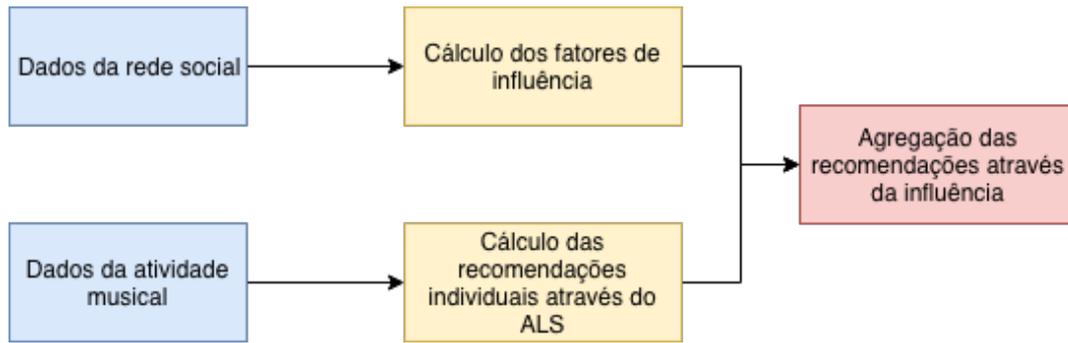


Figura 6: Etapas para Cálculo na API de Recomendação

A etapa mais dispendiosa em termos de processamento desse fluxo (Figura 6) é o cálculo das recomendações individuais através do ALS, pois o conjunto de dados a serem considerados é muito grande.

O cálculo dos fatores de influência (Figura 6), apesar de se tratar de um algoritmo dispendioso computacionalmente, é feito em pouco tempo, pois a aplicação possui uma limitação de grupo de até 5 indivíduos. Essa limitação se dá, justamente, para atender aos requisitos do sistema e atender o usuário em um tempo viável.

5.2.2 Tecnologias Utilizadas para a Etapa de Recomendação

Para realizar o cálculo das recomendações individuais foi utilizado a implementação do ALS provida pelo Apache Spark (spark.apache.org). Este software é voltado para análise e processamento de dados em larga escala, sendo ideal para ser empregado para bases de dados com milhares de músicas e milhares de usuários. O Apache Spark também permite processamento paralelizado e alta eficiência computacional.

Como o Apache Spark possui uma implementação para Python chamada de PySpark, foi conveniente para este projeto implementar a API de Recomendação utilizando a linguagem de programação Python. Em conjunto ao PySpark utilizou-se o framework Falcon (falconframework.org), para que a API de Recomendação ficasse disponível como um serviço web.

5.3 Aplicação Móvel de Demonstração

Para permitir demonstrar o funcionamento do sistema de recomendação foi criada uma aplicação para que o usuário interagisse com o SR. Usando esta aplicação o usuário fornece os dados de sua atividade musical e os dados de sua rede social, cria e gerencia

os grupos, recebe as recomendações para um grupo selecionado e pode ouvir a lista de canções recomendadas.

5.3.1 Descrição do Funcionamento

O fluxo inicial da aplicação (Figura 7) se dá pela conexão com o Spotify e com a rede social Mastodon, que são os serviços que fornecerão os dados necessários para a API de recomendação. A conexão com o Spotify cria um usuário no sistema e a conexão com a rede social vincula a conta do Mastodon à conta do usuário já criado.

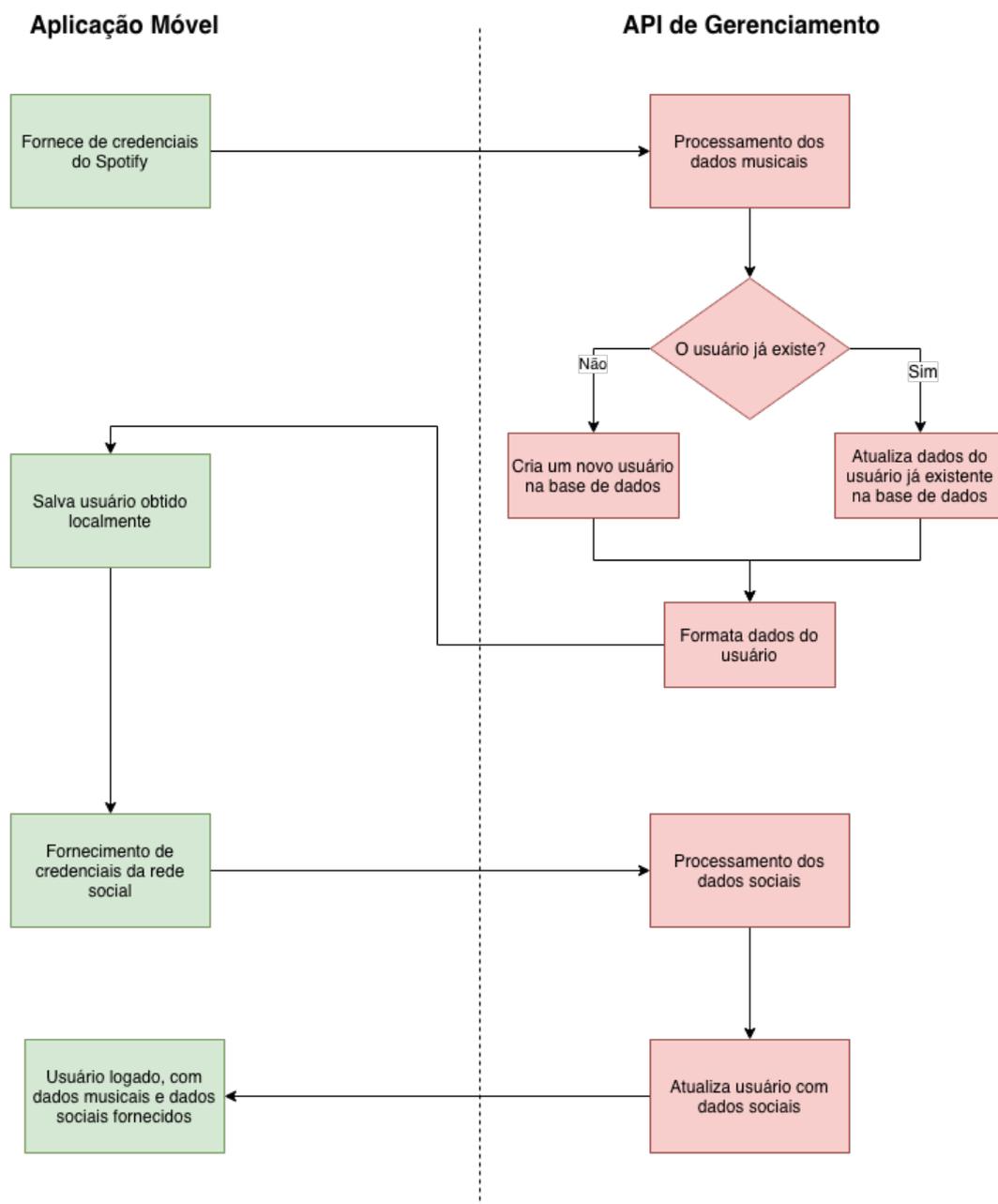


Figura 7: Fluxo de Autenticação da Aplicação

Após estar conectado com sua rede social, o usuário pode criar um grupo selecionando os amigos provindos da rede social. Os grupos criados ficam salvos com um nome, podendo ser utilizado outras vezes pelo usuário.

Para que fique transparente ao usuário a forma como as suas recomendações foram calculadas, há uma aba com estatísticas, onde é exibido os resultados do cálculos realizados para calcular os fatores de influência (seção 4.5) (Figura 8).

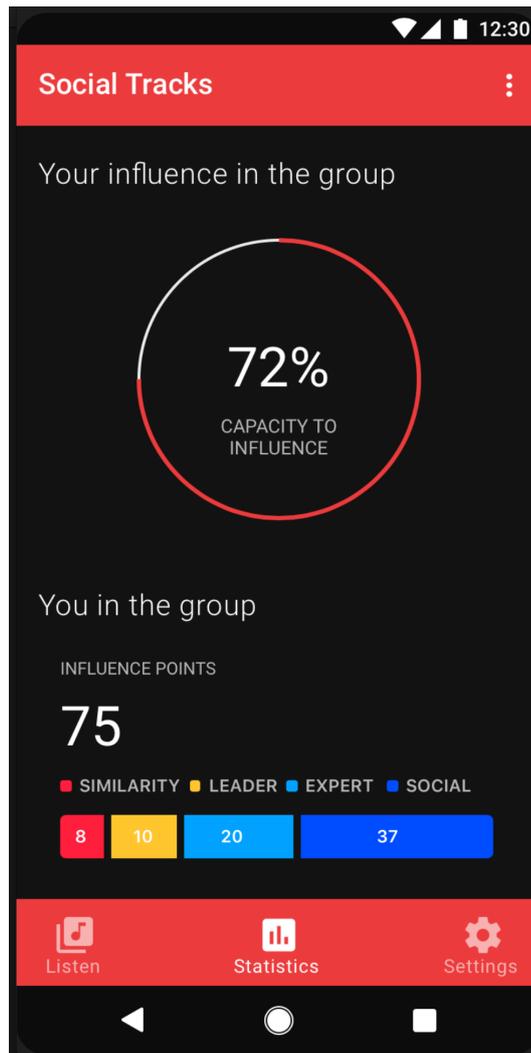


Figura 8: Tela Aplicação Móvel - Estatísticas

5.3.2 Tecnologias Utilizadas

Para que a recomendação para todo o grupo seja possível, todos os usuários pertencentes ao grupo precisam estar cadastrados no sistema. Portanto, é uma necessidade indispensável que a aplicação seja suportada em todos os dispositivos dos integrantes do grupo, ou seja, o sistema precisa suportar os principais sistemas operacionais móveis: iOS

e Android.

Atendendo o requisito multiplataforma, optou-se pelo framework React Native, desenvolvido pelo Facebook para criação de aplicações iOS e Android utilizando a linguagem JavaScript, possibilitando compartilhamento integral de código para desenvolvimento simultâneo para os dois sistemas operacionais.

6 TESTES E RESULTADOS

O sistema de recomendação desenvolvido neste trabalho passa por diversas etapas antes de fornecer uma recomendação para um grupo. Para este capítulo o foco será apresentar os resultados obtidos em cada uma dessas etapas, tomando como exemplo um grupo simples com dois indivíduos.

6.1 Grupo de Teste

Para os testes apresentados neste capítulo o grupo é composto por dois usuários, cujo os dados são reais e foram obtidos através do Spotify. Para garantia de privacidade eles serão chamados de *Usuário A* e *Usuário B*.

Para o Usuário A a tabela abaixo exhibe as 19 músicas obtidas com seus respectivos artistas e nomes, além de uma contagem de quantas vezes o Usuário A reproduziu cada uma das músicas.

Artista	Música	Contagem
A Banda Mais Bonita da Cidade	Boa Pessoa	33
Frank Zappa	Cosmik Debris	43
A Banda Mais Bonita da Cidade	Se Eu Corro	43
Cachorro Grande	Sinceramente	48
A Banda Mais Bonita da Cidade	Boa Pessoa	21
A Banda Mais Bonita da Cidade	Canção pra Não Voltar	23
Bread	Guitar Man	50
Rita Lee	Agora Só Falta Você	44
A Banda Mais Bonita da Cidade	Mercadoramama	26
ZZ Top	La Grange	71
A Banda Mais Bonita da Cidade	Potinhos	22
Black Sabbath	Children of the Sea	37
Cazuza	Ideologia	56
A Banda Mais Bonita da Cidade	Uma Atriz	23
Black Sabbath	War Pigs	63
Jimi Hendrix	May This Be Love	52
Little Richard	Tutti Frutti	62
Arnaldo Antunes	A Casa é Sua	46
A Banda Mais Bonita da Cidade	Oração	61

Tabela 2: Tabela da atividade musical para o Usuário A

O mesmo é apresentado abaixo para o Usuário B, mas neste caso foram obtidas 22 canções.

Artista	Música	Contagem
Eagles	Hotel California	80
P!nk	Just Give Me a Reason	74
R.E.M.	Losing My Religion	79
Of Monsters and Men	Six Weeks	49
Sia	Chandelier	57
Boston	More Than a Feeling	75
Of Monsters and Men	Little Talks	73
Of Monsters and Men	Your Bones	51
Aretha Franklin	I Say a Little Prayer	72
The Verve	Bitter Sweet Symphony	78
Avril Lavigne	When You're Gone	39
Christina Aguilera	Beautiful	70
Cher	Strong Enough	59
Arctic Monkeys	Do I Wanna Know?	81
Of Monsters and Men	Sloom	47
Of Monsters and Men	Mountain Sound	60
Bruno Mars	When I Was Your Man	80
Of Monsters and Men	Dirty Paws	67
Of Monsters and Men	Love Love Love	55
Of Monsters and Men	Lakehouse	48
U2	Beautiful Day	69
Of Monsters and Men	From Finner	50

Tabela 3: Tabela da atividade musical para o Usuário B

Em seguida é obtido os dados sociais para o cálculo da influência de cada um dos usuários. Para os Usuário A a Tabela 4 exibe os dados obtidos da rede social Mastodon e do Spotify.

Dado	Valor	Descrição
Total de Menções	7	Número total de menções realizadas pelo usuário
Total de Comentários	32	Número total de comentários realizadas pelo usuário
Total de Curtidas	32	Número total de curtidas realizadas pelo usuário
Total de Compartilhamentos	32	Número total de compartilhamentos realizadas pelo usuário
Menções de A em B	3	Menções que o Usuário A fez do Usuário B
Compartilhamentos de A em B	3	Compartilhamentos que o Usuário A fez de postagens do Usuário B
Comentários de A em B	3	Comentários que o Usuário A fez em postagens do Usuário B
Curtidas de A em B	3	Curtidas que o Usuário A fez em postagens do Usuário B
Número de Seguidores no Spotify	10	Número de seguidores que o usuário possui no Spotify
Número de Músicas no Spotify	1123	Número total de músicas na biblioteca do usuário no Spotify
Número de Artistas no Spotify	53	Número total de artistas que o usuário segue no Spotify

Tabela 4: Dados sociais obtidos do Usuário A

O análogo é obtido do Usuário B, exibido na Tabela 5.

Dado	Valor	Descrição
Total de Menções	14	Número total de menções realizadas pelo usuário
Total de Comentários	56	Número total de comentários realizadas pelo usuário
Total de Curtidas	121	Número total de curtidas realizadas pelo usuário
Total de Compartilhamentos	78	Número total de compartilhamentos realizadas pelo usuário
Menções de B em A	2	Menções que o Usuário B fez do Usuário A
Compartilhamentos de B em A	2	Compartilhamentos que o Usuário B fez de postagens do Usuário A
Comentários de B em A	6	Comentários que o Usuário B fez em postagens do Usuário A
Curtidas de B em A	12	Curtidas que o Usuário B fez em postagens do Usuário A
Número de Seguidores no Spotify	7	Número de seguidores que o usuário possui no Spotify
Número de Músicas no Spotify	400	Número total de músicas na biblioteca do usuário no Spotify
Número de Artistas no Spotify	87	Número total de artistas que o usuário segue no Spotify

Tabela 5: Dados sociais obtidos do Usuário B

6.2 Resultados para as Recomendações Individuais

Para obter as recomendações individuais o algoritmo utilizado é o Alternating Least Squares (ALS), já mencionado no capítulo 4. O ALS requer uma base de dados para criação de um modelo, para os testes a base de dados é constituída pelos dados exibidos na Tabela 6.

Dado	Quantidade	Descrição
Músicas	105.655	Canções únicas, com informações diversas como: nome, artista e duração.
Usuários	41.262	Usuários únicos do LastFM com dados como: total de músicas ouvidas, país e idade.
Usuário-Música	309.986	Relação de um usuário com uma música, ou seja, o número de vezes que um usuário ouviu uma determinada música.

Tabela 6: Dados disponíveis para geração da recomendação individual

Ao enviar a atividade musical dos usuários A e B para a API de Recomendação as recomendações individuais são calculadas. Na Tabela 7 consta as 5 músicas melhores avaliadas para o Usuário A e seu grau de predição. Uma nota importante é que o sistema não gerou apenas 5 recomendações individuais do usuário, o grau de predição é gerado para todas as músicas presentes do banco de dados exibidos na Tabela 6.

Artista	Música	Grau de Predição
Stairway to Heaven	Led Zeppelin	0.12158
Wish You Were Here	Pink Floyd	0.11658
Refrão De Bolero	Engenheiros Do Hawaii	0.11197
Something	The Beatles	0.11156
Como Nossos Pais	Elis Regina	0.1109

Tabela 7: Resultado das recomendações individuais para o Usuário A

O análogo é realizado para o Usuário B e as 5 músicas melhores avaliadas com seu grau de predição são exibidas na Tabela 8.

Artista	Música	Grau de Predição
Radioactive	Imagine Dragons	0.11811
Demons	Imagine Dragons	0.11716
Royals	Lorde	0.11613
Pumped Up Kicks	Foster the People	0.11497
The Scientist	Coldplay	0.11474

Tabela 8: Resultado das recomendações individuais para o Usuário B

6.3 Resultados para os Fatores de Influência

Para a determinação da influência são calculados os quatro fatores definidos na seção 4.5. Para o Usuário A a Tabela 9 exibe os resultados obtidos para cada um dos fatores.

Fator	Valor
Similarity	0.06240728968001695
Expert	0.5680557318134082
Friendly	0.125
Leadership	0.2871794871794872
Influência	0.2606606271682281

Tabela 9: Resultado dos fatores de influência do Usuário A

O mesmo é realizado para o Usuário B, os resultados são exibidos na Tabela 10.

Fator	Valor
Similarity	0.28735632183908044
Expert	0.43194426818659176
Friendly	0.3392857142857143
Leadership	0.37948717948717947
Influência	0.35951837094964145

Tabela 10: Resultado dos fatores de influência do Usuário B

Portanto, avaliando os resultados obtidos o Usuário B possui uma capacidade de influência maior que o Usuário A, isso fará com que as avaliações de B possuam maior peso que as avaliações de A.

6.4 Resultados Para a Recomendação do Grupo

A partir das recomendações individuais e os dados de influência é possível realizar a agregação das recomendações para a geração de uma recomendação única para o grupo. A recomendação final das 10 músicas com maior grau de predição exibida na Tabela 11.

Fator	Valor	Grau de Predição
Royals	Lorde	0.21873939211309348
Do I Wanna Know?	Arctic Monkeys	0.2177352314424269
R U Mine?	Arctic Monkeys	0.21576832295300877
Chandelier	Sia	0.20720682015666964
Shake It Out	Florence + the Machine	0.20411055501991582
Elephant Gun	Beirut	0.20077034262371685
Pumped Up Kicks	Foster the People	0.19968925884170713
Take Me to Church	Hozier	0.19934865326122678
Fluorescent Adolescent	Arctic Monkeys	0.19876480023500176
Wonderwall	Oasis	0.1986019536578552

Tabela 11: Recomendação final para o grupo formado pelos usuários A e B

7 CONSIDERAÇÕES FINAIS

7.1 Conclusões

Neste projeto foi possível realizar estudos sobre sistemas de recomendação para grupos e desenvolver um aplicação exemplo capaz de fornecer para um grupo de usuários um conjunto de músicas compatível com todos os integrantes do grupo.

Para este trabalho ainda é necessário um experimento com uma população estatisticamente relevante para avaliar o grau de satisfação obtido através da agregação por influência social desenvolvida neste projeto. No entanto, é válido ressaltar que o saldo do projeto é extremamente positivo, uma vez que foi possível aplicar conceitos avançados de engenharia de software de arquitetura de sistemas computacionais e de análise de algoritmos.

7.2 Contribuições

A principal contribuição deste trabalho é o desenvolvimento de um método alternativo para o cálculo da influência social e sua aplicação para agregação de recomendações individuais. Tal cálculo foi possível graças a recente emergência de dados provindos de redes sociais que cada vez mais conseguem reunir as relações entre pessoas de uma mesma comunidade.

7.3 Perspectivas de Continuidade

Em termos de funcionalidades, não foram considerados neste projeto a geração de recomendações a partir de um contexto, por exemplo, um grupo pode desejar obter uma recomendação para uma festa ou durante uma viagem. A consideração do gênero musical enriqueceria a recomendação e tornaria possível a análise por contexto.

Trabalhos futuros podem utilizar o sistema de recomendação desenvolvido nesse projeto para modelar outras maneiras de se calcular a influência social e antever seus impactos na satisfação do grupo. Trabalhos na área da Psicologia Social podem analisar os efeitos da conformidade e do contágio emocional na distribuição de conteúdos e como os indivíduos se comportam quando apresentados a uma recomendação em comum.

REFERÊNCIAS BIBLIOGRÁFICAS

- Al-Hassan, M., Lu, H., and Lu, J. (2015). A semantic enhanced hybrid recommendation approach: A case study of e-Government tourism service recommendation system. *Decis. Support Syst.*, 72:97–109.
- Contratres, F. G., Alves-Souza, S. N., Filgueira, L. V. L., and DeSouza, L. S. (2018). Sentiment analysis of social network data for cold-start relief in recommender systems. In *New Contributions in Information Systems and Technologies. Advances in Intelligent Systems and Computing*. in press.
- Deng, S., Huang, L., and Xu, G. (2014). Social network-based service recommendation with trust enhancement. *Expert Syst. Appl.*, 41:8075–8084.
- Desrosiers, C. and Karypis, G. (2011). *A Comprehensive Survey of Neighborhood-based Recommendation Methods*, chapter 4. Springer US.
- Gonzalez Camacho, L. A. and Alves-Souza, S. N. (2018). Social network data to alleviate cold-start in recommender system: A systematic review. *Information Processing & Management*, 54(4):529–544.
- Guo, J., Zhu, Y., Li, A., Wang, Q., and Han, W. (2016). A Social Influence Approach for Group User Modeling in Group Recommendation Systems. *IEEE Intelligent Systems*, 31(5):40–48.
- Hu, Y., Volinsky, C., and Koren, Y. (2008). Collaborative filtering for implicit feedback datasets. *Proceedings - IEEE International Conference on Data Mining, ICDM*.
- Koren, Y. and Bell, R. (2011). *Advances in Collaborative Filtering*, chapter 5. Springer US.
- Koren, Y., Bell, R., and Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37.
- Lalwani, D., Somayajulu, D. V. L. N., and Krishna, P. R. (2015). A community driven social recommendation system. *Proc. - 2015 IEEE Int. Conf. Big Data, IEEE Big Data 2015*, pages 821–826.

- Masthoff, J. and Gatt, A. (2006). In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. *User Modeling and User-Adapted Interaction*, 16(3-4):281–319.
- Prando, A., Contratres, F., Alves-Souza, S., and deSouza, L. (2017). Content-based recommender system using social networks for cold-start users. In *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. in press.
- Quirino, G. Z., Mals, N. P., Groterhorst, V. M., De Souza, S. N., and De Souza, L. S. (2015). Meneduca - Social school network to support the educational environment. In *Proceedings - 2015 41st Latin American Computing Conference, CLEI 2015*.
- Resnick, P. and Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3):56–58.
- Ricci, F., Rokach, L., Shapira, B., and Kantor, P. B. (2011). *Recommender Systems Handbook*, volume 53.
- Wang, J., Liu, Z., and Zhao, H. (2014). Group recommendation using topic identification in social networks. In *Proceedings - 2014 6th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2014*, volume 1, pages 355–358.
- Zhu, H. and Huberman, B. A. (2014). To Switch or Not To Switch: Understanding Social Influence in Online Choices. *American Behavioral Scientist*, 58(10):1329–1344.