

Arthur F. Batista

# **Utilizando Tropes em Modelos de Recomendação Híbridos**

Manaus

2016

Arthur F. Batista

# **Utilizando Tropes em Modelos de Recomendação Híbridos**

Dissertação apresentada ao Programa de Pós-graduação em Informática da Universidade Federal do Amazonas como requisito parcial para obtenção do grau de Mestre em Informática. Área de concentração: Banco de Dados e Recuperação da Informação.

PODER EXECUTIVO  
MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Orientador: Prof. Dr. Marco Cristo

Manaus

2016

## Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

B333u Batista, Arthur Félix  
Utilizando Tropes em Modelos de Recomendação Híbridos /  
Arthur Félix Batista. 2016  
56 f.: il. color; 31 cm.

Orientador: Marco Antônio Pinheiro de Cristo  
Dissertação (Mestrado em Informática) - Universidade Federal do Amazonas.

1. Sistemas de Recomendação. 2. Tropes. 3. Filmes. 4. Modelos Híbridos. 5. Slope One. I. Cristo, Marco Antônio Pinheiro de II. Universidade Federal do Amazonas III. Título



PODER EXECUTIVO  
MINISTÉRIO DA EDUCAÇÃO  
INSTITUTO DE COMPUTAÇÃO



PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

UFAM

## FOLHA DE APROVAÇÃO

"Utilizando Tropes em Modelos de Recomendação Híbridos"

ARTHUR FÉLIX BATISTA

Dissertação de Mestrado defendida e aprovada pela banca examinadora constituída pelos Professores:

Prof. Marco Antonio Pinheiro de Cristo - PRESIDENTE

Profa. Eulanda Miranda dos Santos - MEMBRO INTERNO

Prof. Thierson Couto Rosa - MEMBRO EXTERNO

Manaus, 16 de Dezembro de 2016

# Agradecimentos

Ao professor Dr. Marco Cristo, orientador desta dissertação, não só pelas longas horas de discussões políticas e filosóficas, mas, principalmente, pela paciência e empenho em ensinar-me os fundamentos da pesquisa científica.

Aos amigos do trabalho e academia, principalmente à Luiz Cavalcanti, Thomaz Silva e Diego Falcão, que tiveram participação direta neste trabalho.

À Ana Acris pelas suas revisões e, principalmente, por seu amor e companheirismo que estiveram sempre presentes comigo.

Aos professores Edleno Moura, Eulanda Santos, David Fernandes, João Cavalcanti, Altigran Soares e André Carvalho por repassarem o conhecimento durante as aulas.

Aos funcionários da UFAM por todo o suporte.

*“Sempre achei que um filme deve ser, antes de mais nada, uma experiência emocional. Deve fazer você rir, ou chorar, ou sentir medo. Mas deve também inspirar e provocar.*

*Fazer você refletir.”*

*William Friedkin*

# Sumário

<b>Lista de ilustrações</b> . . . . .	<b>vii</b>
<b>Lista de tabelas</b> . . . . .	<b>viii</b>
<b>1 Introdução</b> . . . . .	<b>3</b>
1.1 Motivação . . . . .	4
1.2 Hipótese, Problema e Questões de Pesquisa . . . . .	5
1.3 Objetivos . . . . .	6
1.3.1 Objetivos Específicos . . . . .	6
1.4 Metodologia . . . . .	6
1.5 Contribuição da Pesquisa . . . . .	7
1.6 Estrutura da Dissertação . . . . .	7
<b>2 Fundamentação Teórica</b> . . . . .	<b>9</b>
2.1 Tropes . . . . .	9
2.1.1 DBTropes . . . . .	11
2.2 Sistemas de Recomendação . . . . .	11
2.2.1 Filtragem Colaborativa . . . . .	12
2.2.1.1 Slope One . . . . .	12
2.2.1.2 Slope One Ponderado . . . . .	14
2.2.2 Filtragem Baseada em Conteúdo . . . . .	14
2.2.2.1 Distância Semântica . . . . .	14
2.2.2.2 Coeficiente de Jaccard . . . . .	17
2.2.2.3 Similaridade do Cosseno . . . . .	18
2.2.2.4 TF-IDF . . . . .	18
2.2.3 Slope One Ponderado Híbrido . . . . .	19
2.2.4 Limitações dos SRs . . . . .	20
2.2.5 Avaliação . . . . .	21
2.3 Trabalhos Relacionados . . . . .	23
2.3.1 Tropes . . . . .	23
2.3.2 Sistemas de Recomendação . . . . .	25
<b>3 Caracterização do TVTropes</b> . . . . .	<b>27</b>
3.1 Características do catálogo de tropes nos filmes . . . . .	27
3.1.1 Correlação entre tropes e metadados dos filmes . . . . .	28
3.2 Padrões de ocorrência de tropes entre filmes similares . . . . .	31
<b>4 Proposta de um SR Híbrido Baseado em Tropes</b> . . . . .	<b>34</b>
4.1 Arquitetura . . . . .	34
4.2 Construção de Base de Dados Integrada . . . . .	34
4.2.1 Crawler TVTropes . . . . .	35

4.3	Pré-processamento dos Dados . . . . .	35
4.3.1	Filtro de Tropes por Quantidade . . . . .	35
4.3.2	Apenas Tropes . . . . .	36
4.3.3	Combinação entre Tropes e Gêneros . . . . .	36
4.3.4	Filtro de Tropes por Nota . . . . .	36
4.4	Componente de Recomendação . . . . .	37
<b>5</b>	<b>Experimentos . . . . .</b>	<b>39</b>
5.1	Base de Dados . . . . .	39
5.2	Avaliação . . . . .	40
5.3	Protocolo Experimental . . . . .	40
5.4	Resultados . . . . .	41
5.4.1	RMSE . . . . .	41
5.4.2	Precisão, Revocação e F1 . . . . .	43
5.5	Discussão . . . . .	47
<b>6</b>	<b>Conclusão e Trabalhos Futuros . . . . .</b>	<b>51</b>
6.1	Trabalhos Futuros . . . . .	52
	<b>Referências . . . . .</b>	<b>54</b>



# Lista de ilustrações

Figura 1 – Exemplo de Tropes em Filmes. (A) “Et Tu, Brute ?”, (B) “Luke, Eu Sou Seu Pai”, (C) batalha da chuva, (D) o grande vilão. . . . .	9
Figura 2 – Matriz usuário-item . . . . .	12
Figura 3 – Representação básica do algoritmos Slope One. Fonte: (LEMIRE; MACLACHLAN, 2005) . . . . .	13
Figura 4 – Exemplo de um grafo Linked Data. Fonte: (PASSANT, 2010) . . . . .	15
Figura 5 – Abordagens para a combinação de FC e FBC. (A) FC e FBC são implementados separadamente e depois combinados para gerar a recomendação; (B) características da FBC são incorporadas na FC; (C) modelo unificado das filtragens colaborativa e baseada em conteúdo; e (D) características da FC são incorporadas na FBC. . . . .	19
Figura 6 – Histograma do número de filmes agrupados pela quantidade de tropes da coleção DBTropes. . . . .	28
Figura 7 – Média de tropes por gênero da coleção DBTropes $\cap$ IMDB. . . . .	29
Figura 8 – Distribuição dos gêneros dos filmes em um espaço bidimensional composto por tropes, reduzidos através da ferramenta t-SNE, da coleção DBTropes $\cap$ IMDB. . . . .	30
Figura 9 – Correlação da quantidade de tropes catalogados nos filmes com sua popularidade e preferência na coleção DBTropes $\cap$ IMDB. . . . .	31
Figura 10 – Arquitetura do sistema de recomendação híbrido. . . . .	34
Figura 11 – Distribuição da nota dos tropes derivada dos filmes. . . . .	37
Figura 12 – RMSE variando-se o valor de $\alpha$ e quantidade mínima de tropes com o filtro baseado nos tropes de forma pura. . . . .	42
Figura 13 – RMSE variando-se o valor de $\alpha$ e quantidade mínima de tropes com o filtro baseado na combinação de tropes e gêneros. . . . .	43
Figura 14 – RMSE variando-se o valor de $\alpha$ e quantidade mínima de tropes com o filtro baseado nos tropes mais bem avaliados pelos usuários. . . . .	45
Figura 15 – Precisão variando-se o valor de $\alpha$ e quantidade mínima de tropes com o filtro baseado nos tropes de forma pura. . . . .	46
Figura 16 – Precisão variando-se o valor de $\alpha$ e quantidade mínima de tropes com o filtro baseado na combinação de tropes e gêneros. . . . .	46
Figura 17 – Precisão variando-se o valor de $\alpha$ e quantidade mínima de tropes com o filtro baseado nos tropes mais bem avaliados pelos usuários. . . . .	47

# Lista de tabelas

Tabela 1	– Estatísticas sobre a quantidade de tropes catalogados nos filmes da coleção DBTropes. . . . .	27
Tabela 2	– Lista dos dez filmes com mais tropes catalogados na coleção DBTropes.	28
Tabela 3	– Lista com exemplos de filmes com menos tropes catalogados na coleção DBTropes. . . . .	28
Tabela 4	– Amostra de tropes catalogados para o filme “Coração Valente” e filmes similares . . . . .	32
Tabela 5	– Amostra de tropes catalogados para o filme “Toy Story” e filmes similares	32
Tabela 6	– Exemplo de Categorias de Tropes . . . . .	33
Tabela 7	– Estatísticas sobre a quantidade de tropes catalogados nos filmes da coleção <i>MovieLens 100k</i> $\cap$ <i>DBTropes</i> . . . . .	39
Tabela 8	– Estatísticas sobre a quantidade de notas por usuário da coleção <i>MovieLens 100k</i> $\cap$ <i>DBTropes</i> . . . . .	40
Tabela 9	– Melhores resultados de obtidos para a métrica RMSE em cada abordagem implementada. Os ganhos precedidos por “*” indicam que houve diferença estatística em comparação com o método de referência. As linhas em destaque denotam os maior ganhos entre todas as abordagens.	44
Tabela 10	– Melhores resultados de obtidos para a métrica de precisão e suas respectivas revocação e F1 em cada abordagem implementada. Os ganhos precedidos por “*” indicam que houve diferença estatística em comparação com o método de referência. As linhas em destaque denotam os maiores precisões entre todas as abordagens. . . . .	48

# Lista de abreviaturas e siglas

API	Application Programming Interface
FBC	Filtragem Baseada em Conteúdo
FC	Filtragem Colaborativa
IMDB	Internet Movie Database
LDS	Linked Data Semantic Distance
LDSd	Linked Data Semantic Distance Direct
LDSi	Linked Data Semantic Distance Indirect
RMSE	Root Mean Square Error
SOP	Slope One Ponderado
SOPH	Slope One Ponderado Híbrido
SR	Sistemas de Recomendação
t-SNE	t-Distributed Stochastic Neighbor Embedding
TF-IDF	Term Frequency - Inverse DocumentFrequency

# Resumo

Sistemas de Recomendação (SR) têm sido amplamente estudados nas últimas décadas. O crescimento da Internet e a consolidação da Web 2.0 contribuíram para o surgimento de diversos serviços, como redes sociais, blogs, plataformas colaborativas, entre outros, resultando no aumento significativo do volume de informação. Este cenário fomentou o desenvolvimento de pesquisas com o intuito de utilizar tal informação para mitigar algumas limitações dos SRs e melhorar sua qualidade. A recomendação de filmes tornou-se um dos tópicos mais abordados na literatura no domínio de SRs junto com comércio eletrônico. Essa popularidade foi alavancada pela indústria com o crescimento dos serviços de *streaming* (Amazon, Netflix, iTunes, Google Play). Neste contexto, diferentes fontes de informação, tipicamente disponibilizadas por usuários em sites da Web, vêm sendo exploradas para extração de características dos filmes. As fontes mais comuns proveem informações como gênero, direção, elenco, etc. Outras abordagens tentam extrair informações a respeito do conteúdo dos filmes, como estrutura da história, elementos da narrativa, personagens, de modo que se possa caracterizar a história em si. Tal conteúdo pode ser representado através de Tropes. Tropes são elementos que compõem uma história fictícia que pode ser contada através de filmes, livros, quadrinhos entre outros tipos de mídia. Neste trabalho, apresentamos um estudo sistemático sobre Tropes, investigando sua relevância para o contexto de uma história e como podem ser incorporados em Sistemas de Recomendação de Filmes. Os experimentos executados neste pesquisa sugerem que modelos híbridos baseados na combinação de tropes com os gêneros dos filmes, podem melhorar a precisão das previsões em quase 3% em comparação com o métodos tradicionais.

**Palavras-chaves:** sistemas de recomendação, tropes, filmes, filtragem colaborativa

# Abstract

Recommendation systems (SR) have been widely studied in recent decades. The growth of the Internet and the consolidation of Web 2.0 have contributed to the emergence of various services such as social networks, blogs, collaborative platforms, among others, resulting in increased volume of information. This scenario has fostered the development of new research on how to use such information to mitigate limitations of SRs and improve their quality. The recommendation of movies became one of the most discussed topics in the literature about SRs. The industry also contributed to its popularity with the growth of streaming services such as Amazon, Netflix, iTunes, and Google Play. In such scenario, different sources of information in Web have been exploited to extract features to describe movies. The most common approaches use features such as genre information, movie direction, cast, etc. Other approaches attempt to characterize the story itself by means of information about the content of movies, its story structure, elements of narrative and characters. Such content can be represented by Tropes. Tropes are the elements that make up a fictional story found in movies, books, comics and other contents. In this work, we present a systematic study of Tropes, investigating its relevance to the context of a story and how they can be incorporated in Movie Recommender Systems. The experiments performed in this research suggest that hybrid models based on the combination of tropes with the films genres can improve the precision of the predictions about 3% in comparison to traditional methods.

**Keywords:** recommendation systems, tropes, movies, collaborative filtering

# 1 Introdução

O crescimento da internet mudou a forma como nos comunicamos, como compramos algum produto e, até mesmo, como consumimos conteúdo de entretenimento. É cada vez mais frequente ver pessoas utilizando seus *smartphones* para jogos, ouvir músicas e assistir filmes. Essa nova realidade vem transformando a indústria de entretenimento, que aos poucos, vem substituindo a mídia física pela mídia digital implicando no surgimento de diversos serviços para a distribuição de conteúdo.

A consequência disso foi o aumento significativo no volume de informação que, às vezes, pode sobrecarregar o usuário, dificultando sua procura por itens relevantes. Neste contexto, os Sistemas de Recomendação (SR) desempenham um papel fundamental para auxiliar os usuários no processo de tomada de decisão.

Existem diferentes técnicas para a implementação de SRs. Entre os métodos mais tradicionais, estão aqueles baseados em Filtragem Colaborativa (FC), que tiveram rápida adoção pela sua eficácia. FCs baseiam-se apenas nas avaliações dos usuários para inferir suas similaridades com outros usuários do sistema, não levando em consideração as informações dos itens. Tais informações são exploradas pelas técnicas de Filtragem Baseada em Conteúdo (FBC) que analisam as características dos itens que interessam aos usuários e recomendam outros itens similares. Ambas estratégias possuem limitações, como novo usuário, novo item, super-especialização, etc, que podem ser mitigadas por abordagens de recomendação híbridas. Estas, por sua vez, combinam FC e FBC em um único modelo, reduzindo alguns dos problemas que essas técnicas possuem quando aplicadas individualmente.

No contexto de recomendação de filmes, alguns trabalhos têm utilizado as características como gênero, direção e elenco para representar os filmes. A ideia geral, relacionada ao uso destas evidências, é que elas poderiam ajudar a descrever a preferência dos usuários. Assim, um usuário poderia gostar de um certo filme porque ele é protagonizado por um certo ator (cujos filmes, em geral, estão entre os preferidos do usuário) ou poderia evitar um certo filme porque é de um gênero que o usuário não gosta.

Em um contexto geral, outra forma de descrever filmes é através dos *tropes* observados na narrativa. Tropes (STERLING, 2015) são ferramentas ou padrões narrativos, repetitivos, que escritores usam frequentemente. Representam arquétipos ou estereótipos que podem ser facilmente reconhecidos pelo público em geral, sendo possível encontrá-los em qualquer tipo de narrativa. São usados com diferentes objetivos. Por exemplo, podem ocorrer de forma inconsciente; podem ser parte do contexto/universo em que se dá a história, o que justifica a sua utilização; podem ser utilizados de forma a atender uma

expectativa natural do público para então, serem subvertidos de forma a causar surpresa, espanto, admiração; podem ser usados como recurso para fazer paródias. Exemplos de tropes culturalmente populares (PALACIOS, 2004) são “a mocinha em perigo”, “o cavaleiro de branco”, “o mordomo é sempre o assassino”, “a mulher vestida de preto”, “o final surpreendente/a reviravolta” e a “a grande revelação”.

## 1.1 Motivação

Considerando o potencial de tropes como elementos de descrição, é natural o interesse em avaliá-los no contexto de uma indústria tão importante como a de recomendação. De fato, SRs têm sido alvo de interesse em muitas pesquisas na comunidade científica e na indústria. Sites de e-commerce (Amazon<sup>1</sup>, eBay<sup>2</sup>, Submarino<sup>3</sup>) e provedores de conteúdo (Netflix<sup>4</sup>, Youtube<sup>5</sup>, Spotify<sup>6</sup>) têm investido bastante nessa área cujo retorno tem impacto direto em seus lucros. Uma prova disso foi a competição proposta pelo Netflix em 2007, chamada de Netflix Prize, que premiaria o trabalho que melhorasse a qualidade de suas recomendações (NETFLIX, 2009).

Apesar da sua possível utilidade no contexto de recomendação, até onde sabemos, tropes nunca foram estudados sistematicamente em SRs<sup>7</sup>. No caso de SRs de filmes, poderia se tirar proveito do fato de que tropes são comumente usados por usuários na descrição de suas preferências. Por exemplo, existem usuários que gostam de filmes com finais inesperados; outros preferem uma história de amor não correspondida; há aqueles que têm mais afinidade com um suspense psicológico em lugar de sustos mecânicos, etc.

O uso deste tipo de informação adicional poderia ser particularmente útil na resolução de limitações conhecidas dos SRs, em particular, os problemas de recomendação para usuário novo e recomendação de item novo. É complexa a recomendação nesses casos porque métodos baseados em FC necessitam de informação histórica relativa à avaliações de itens pelos usuários. Uma forma de lidar com este problema no contexto de filmes seria solicitar informação do usuário sobre suas preferências em termos de tropes e usar descrições dos filmes em termos de tropes. Estas informações adicionais podem ser úteis para melhorar a recomendação de filmes e, por conseguinte, aumentar as chances de um usuário associar-se a um item.

Uma razão para a falta de aplicação de tropes em SRs está relacionada à dificuldade

---

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

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

<sup>3</sup> <http://www.submarino.com.br>

<sup>4</sup> <http://www.netflix.com>

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

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

<sup>7</sup> Eles já foram, de fato, usados como evidência adicional, em um único trabalho (YANG; HU; QU, 2013). Contudo, seu impacto nunca foi analisado em particular.

em se compreender o conteúdo das obras de forma a catalogar os padrões narrativos de interesse. Para um processo automático de extração de tropes, a tecnologia necessária deveria envolver processamento de linguagem natural, classificação de padrões complexos e, em muitos casos, sofisticada capacidade de processamento de vídeo e áudio. Na prática, isto não é possível com o atual estado-da-arte e a alternativa viável seria a manutenção de um catálogo de tropes através da colaboração de seres humanos. De fato, este é outro fator motivante para este trabalho, a possibilidade de manutenção de conteúdo sofisticado via colaboração humana, através da Web 2.0.

O crescimento da Web 2.0 contribuiu para o surgimento de um novo tipo de usuário, que participa ativamente da produção e compartilhamento de conteúdo. Como ele, também emergiram várias plataformas colaborativas, como Wikipédia, Youtube, etc. Entre esse sites, está o TVTropes<sup>8</sup>. Nele, usuários descrevem tropes encontrados em obras diversas, o que inclui filmes, programas de TV, livros, revistas em quadrinhos, jogos de computador, séries, entre outros.

## 1.2 Hipótese, Problema e Questões de Pesquisa

Considerando o exposto previamente, neste trabalho, assumimos que a seguinte hipótese é válida e pode ser usada como base para a recomendação de filmes:

**Hipótese:** *Tropes são úteis para descrever filmes no contexto de sistemas de recomendação, uma vez que é comum usuários descreverem suas preferências em termos de tropes. Estes podem ser extraídos de forma viável de sites especializados da Web 2.0, mantidos por comunidades de usuários.*

Alguns desafios relacionados com esta hipótese incluem a possibilidade de que nem todos os tipos de tropes serem igualmente úteis e a possibilidade de baixa qualidade e cobertura de informações mantidas por comunidades de usuários na Web que atuam, geralmente, de forma amadora.

Assim, podemos agora descrever o problema a ser abordado nessa dissertação. Seja  $\mathcal{M} = \{m_1, m_2, \dots, m_j, \dots, m_{|\mathcal{M}|}\}$  um conjunto de filmes, onde  $m_j$  é um filme particular descrito por uma série de características  $f(m_j)$ . Seja  $\mathcal{U} = \{u_1, u_2, \dots, u_i, \dots, u_{|\mathcal{U}|}\}$  um conjunto de usuários. Seja  $R^{|\mathcal{M}| \times |\mathcal{U}|}$  uma matriz de notas dadas pelos usuários para os filmes. Seja  $r : R \times \mathcal{U} \times \mathcal{M} \rightarrow \mathbb{R}$  uma função que, dada a matriz de notas  $R$ , um usuário  $u_i$  e um filme  $m_j$ ,  $r$  estima a nota que o usuário  $u_i$  daria para o filme  $m_j$ . O problema abordado neste trabalho consiste em determinar uma função  $f(m_j)$  que incorpore tropes na descrição do filme  $m_j$ , de forma que o estimador  $r(R, u_i, m_j)$  retorne a nota mais próxima possível da que efetivamente seria dada pelo usuário  $u_i$  para o filme  $m_j$ .

---

<sup>8</sup> <http://tvtropes.org>



Considerando nossa hipótese e o problema que pretendemos resolver, as seguintes questões de pesquisa se tornam pertinentes:

- Como tropes podem ser usados em tarefas de recomendação? Em outras palavras, qual a melhor forma de definir  $f(m_j)$  e incorporá-la ao estimador  $r(R, u_i, m_j)$ ?
- Quão úteis podem ser tropes nestas tarefas, em termos de proporcionar recomendações de maior qualidade?
- Todos os diferentes tipos de tropes são igualmente apropriados para esta tarefa?

Baseados nestas questões, nas próximas seções, apresentamos nossos objetivos neste trabalho.

## 1.3 Objetivos

Avaliar como tropes podem ser utilizados como fonte de informação para a recomendação de filmes, determinando quão úteis eles podem ser.

### 1.3.1 Objetivos Específicos

- Investigar uma plataforma de tropes mantida na Web a fim de caracterizar o relacionamento entre tropes e filmes, determinando (i) padrões de similaridade entre filmes de acordo com os tropes usados, (ii) a relação entre tropes e a forma como filmes são avaliados e (iii) a relação entre tropes e outras características usadas para descrever filmes.
- Propor método de recomendação que, tirando proveito de tropes como fonte de evidência, seja capaz de fornecer recomendações melhores que aquelas que não utilizam tropes.
- Investigar o impacto de tropes para os resultados, verificando seus desempenhos relativos, de forma a determinar se certas classes de tropes têm mais impacto na recomendação que outras.

## 1.4 Metodologia

Inicialmente, pretendemos extrair tropes de um site da Web 2.0, o TVtropes<sup>9</sup>, e integrar suas informações com as de bases de referência usadas na literatura de recomen-

---

<sup>9</sup> <<http://tvtropes.org/>>

dação, como o Movielens<sup>10</sup>. A coleção de referência que pretendemos usar nos permite descrever filmes com uma variedade de atributos distintos, tais como o gênero do filme.

Vamos então investigar como tropes são caracterizados em relação aos filmes em que ocorrem. Com isso, pretendemos determinar a sua distribuição em relação aos interesses dos usuários, por exemplo, através das notas dadas aos filmes. Esta caracterização irá nos permitir ter uma ideia melhor de como explorar tropes como fonte de informação.

Entre as variáveis de interesse a serem investigadas, uma é a relevância do trope. Tropes que aparecem em filmes com características muito distintas, basicamente por serem populares, podem influenciar negativamente na tarefa de recomendação. Por outro lado, a sua ocorrência em filmes bem avaliados não implica necessariamente que eles sejam úteis em recomendação.

Quanto ao modelo de recomendação, pretendemos explorar recomendadores híbridos que combinem os métodos colaborativos e baseado em conteúdo, dada a sua popularidade na literatura. Em particular, utilizaremos a abordagem implementada por Yang, Hu e Qu (2013) que incorpora informações extraídas do TVTropes no algoritmo Slope One Ponderado.

Finalmente, para avaliar o método proposto e o impacto dos tropes, iremos utilizar métricas comumente usadas em SRs, ou seja, o erro de previsão de nota, a precisão e a cobertura do recomendador.

## 1.5 Contribuição da Pesquisa

Acreditamos que as principais contribuições deste trabalho para a área de sistemas de recomendação são:

- A condução de um estudo sobre um site de tropes da Web, o TVTropes, caracterizando as relações entre tropes, filmes e avaliação de filmes;
- A apresentação de estratégias para incorporar tropes em métodos de recomendação híbridos;
- Uma base de dados experimental contendo filmes com tropes cadastrados além das avaliações dos usuários.

## 1.6 Estrutura da Dissertação

O restante deste trabalho está estruturado da seguinte forma: no Capítulo 2 serão discutidos os fundamentos dos Tropes, Sistemas de Recomendação, bem como a revisão

---

<sup>10</sup> <<https://movielens.org/>>

---

da literatura que aborda ambos os temas. O Capítulo 3 apresenta o estudo feito sobre a base de dados do DBTropes a fim de caracterizar o relacionamento entre tropes e filmes. No Capítulo 4, discutiremos sobre a abordagem utilizada para incorporar tropes em SRs. Os experimentos e resultados obtidos serão abordados no Capítulo 5. Por fim, no Capítulo 6 apresentaremos as conclusões e os trabalhos futuros desta pesquisa, seguidos das referências.

## 2 Fundamentação Teórica

Neste capítulo será explorado com mais detalhes o conceito de tropes e discutidos os fundamentos e os tipos de SR.

### 2.1 Tropes

Tropes são definidos por Sterling (2015) como artifícios e convenções que autores acreditam estar presentes no inconsciente e na expectativa de sua audiência. Em outras palavras, tropes são elementos que compõem uma história, podendo ser encontrados em filmes, séries de TV, quadrinhos, literatura, etc. Uma das forma de registrar a ocorrência de tropes em obras fictícias é através do site TVTropes.org.

O TVTropes.org é uma plataforma colaborativa no formato de *wiki* criado com o objetivo de catalogar tropes. Existem milhares de tropes cadastrados. Cada um tem uma página contendo sua descrição, exemplos de itens onde o mesmo ocorre e uma explicação da sua relevância para o contexto de cada item. Os mesmos podem ser agrupados por categorias e sub-categorias. A seguir, citamos alguns exemplos de tropes catalogados em filmes e ilustrados na Figura 1.

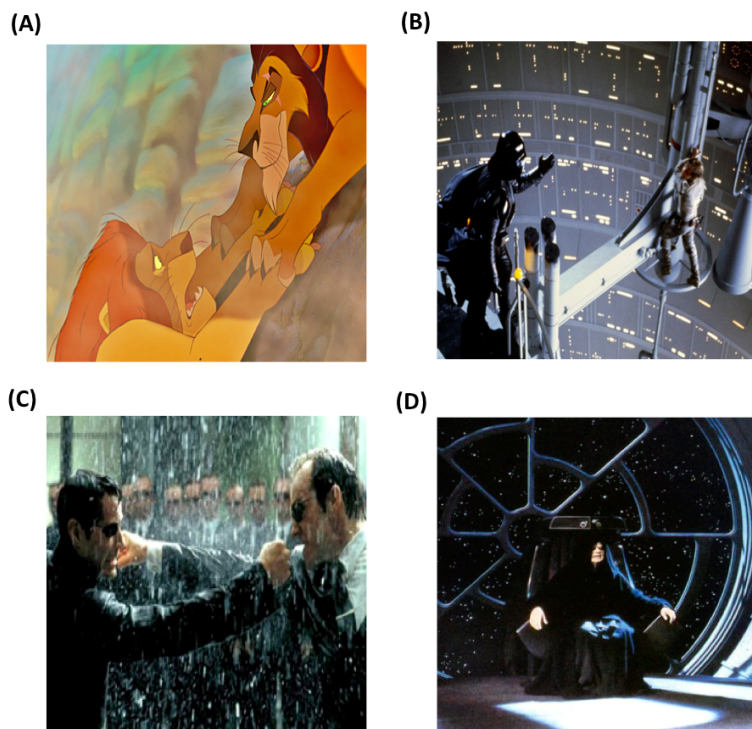


Figura 1 – Exemplo de Tropes em Filmes. (A) “Et Tu, Brute ?”, (B) “Luke, Eu Sou Seu Pai”, (C) batalha da chuva, (D) o grande vilão.

- **Et Tu, Brute ? (Figura 1a):** Frase em Latin que foi supostamente a última do ditador romano Julius Caesar para seu afilhado Marcus Brutus no momento do seu assassinato. Simboliza a traição vinda de uma pessoa inesperada. Exemplo:
  - Rei Leão: o rei Mufasa é assassinado pelo seu próprio irmão Scar.
  - Coração Valente: o protagonista Willian Wallace descobre que seu aliado Rober the Bruce faz parte da guarda real do seu inimigo, o rei Edward Longshanks.
  
- **Luke, Eu Sou Seu Pai (Figura 1b):** Momento em que é feita a revelação que dois personagens, sem relacionamento direto prévio, na verdade são pai (mãe) e filho. Exemplo:
  - Star Wars Episódio VI O Retorno do Jedi: esse filme dá nome ao trope e ocorre quando o vilão Darth Vader revela ser o pai do herói Luke Skywalker.
  - O Exterminador do Futuro: Kyle Reese, que é enviado ao passado para garantir que o exterminador não irá matar John Connor, irá se tornar o pai de John Connor na viagem no tempo.
  
- **Batalha na Chuva (Figura 1c):** Uma luta entre oponentes que ocorre em um ambiente durante a chuva aumentando a tensão deste evento. Exemplo:
  - Matrix Revolution: a luta final entre o personagem Neo e seu antagonista Agente Smith.
  - Blade Runner: a luta final entre o policial Deckard e o replicante Roy Batty. Neste filme, em particular, esta luta também contém outro trope, a *redenção do vilão na chuva*. Este ocorre quando o vilão tem um ato (normalmente, seu último) que o redime de todo o mal que causou. Neste caso, o replicante salva a vida do policial e morre em seguida, enquanto contempla a chuva.
  
- **O Grande Vilão (Figura 1d):** Responsável pelos maus eventos que acontecem em uma história. Antagonista ao herói. Pode ser um personagem ou uma situação onipresente, como um cometa vindo em direção à terra. Exemplo:
  - Trilogia Star Wars: o papel de vilão é desempenhado durante toda a trama pelo imperador Palpatine.
  - O mágico de Oz: a bruxa do Oeste é a vilã ao longo da história.

Tropes podem trazer metadados mais semânticos quando comparados com outras fontes. Por exemplo, enquanto a Wikipédia fornece informações sobre o elenco, direção ou o gênero de um filme, tropes podem informar se o filme possui um vilão, um anti-herói ou um fim surpreendente.

### 2.1.1 DBTropes

A informação no TVTropes.org pode ser acessada por meio de um serviço conhecido como DBTropes<sup>1</sup>. Este foi proposto por Kiesel e Grimnes (2010), como uma forma de automatizar o processo de extração de dados das páginas do site TVTropes que são fornecidos, originalmente, em HTML. O DBTropes oferece uma alternativa conveniente de parser para a tradução das páginas em uma estrutura de fácil acesso. Ele é implementado como uma camada de abstração baseada nos conceitos de *Linked Data*<sup>2</sup>. Logo, seu objetivo é fornecer as informações do TVTropes.org no formato RDF<sup>3</sup>. Em julho de 2016, o DBTropes continha 61.915 itens (filmes, livros etc), 27.254 atributos (tropes, links para outras páginas etc.) e 3.580.311 ligações entre itens e atributos.

## 2.2 Sistemas de Recomendação

A recomendação faz parte do dia-a-dia das pessoas. Estas confiam na opinião de terceiros para escolher livros, filmes, e músicas, entre outros itens, que atendam melhor as suas preferências e necessidades. Isso pode ser feito de diversas formas, como através de avaliações em guias de restaurantes ou interações boca-a-boca. Neste cenário, os SRs tornaram-se ferramentas fundamentais para auxiliar usuários nestes processos de escolha (RESNICK; VARIAN, 1997).

A tarefa de recomendação (ADOMAVICIUS; TUZHILIN, 2005) consiste, em geral, em estimar qual seria a avaliação que um usuário faria de um item novo, baseado intuitivamente em outros itens que ele avaliou, podendo também agregar outros tipos de informação na previsão. Dadas as avaliações de itens ainda não consumidos, estes podem ser recomendados de acordo com as estimativas recebidas.

Este problema pode ser formulado da seguinte forma: Seja  $\mathbf{U}$  um conjunto de usuários e  $\mathbf{I}$  sendo o conjunto de todos os possíveis itens que podem ser recomendados, como livros, filmes, restaurantes, etc. Considerando  $\mu$  uma função que mede quão útil é um item  $i$  para um determinado usuário  $u$ , tal que,  $\mu : \mathbf{U} \times \mathbf{I} \rightarrow \mathbf{R}$ , onde  $\mathbf{R}$  é um conjunto totalmente ordenado. Então, para cada usuário  $u \in \mathbf{U}$ , procura-se um item  $i' \in \mathbf{I}$  que maximiza a utilidade para o usuário. Isso pode ser definido formalmente através da equação 2.1:

$$\forall u \in \mathbf{U} \quad i'_u = \arg \max_{i \in \mathbf{I}} \mu(u, i). \quad (2.1)$$

<sup>1</sup> <http://skipforward.opendfi.de/wiki/DBTropes>

<sup>2</sup> Linked Data refere-se a um conjunto de melhores práticas para publicar e conectar dados estruturados na Web

<sup>3</sup> Resource Description Framework (RDF) é um arcabouço para a representação de informação na Web. O núcleo de sua estrutura é uma tripla constituída por um sujeito, um predicado e um objeto

Outra forma de representar o problema de recomendação, tipicamente encontrado na literatura, é através da matriz *usuário-item*, ilustrada na Figura 2. Nessa matriz, as linhas representam os usuários, enquanto as colunas representam os itens. Cada célula dessa matriz indica o relacionamento (preferência) de um usuário com um item. O objetivo dos SRs é prever os relacionamentos inexistentes (denotados por ?).

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
$u_1$	1	?	5	2	?
$u_2$	?	3	?	?	1
$u_3$	?	?	4	1	?
$u_4$	5	1	?	?	3
$u_5$	?	1	?	2	5

Figura 2 – Matriz usuário-item

A preferência do usuário pode ser expressada tanto explicitamente, através de avaliações, quanto indiretamente, utilizando uma notação binária indicando se o usuário acessou, visualizou ou comprou um item.

SRs são usualmente divididos em três categorias: filtragem colaborativa (FC), filtragem baseada em conteúdo (FBC) e filtragem híbrida, no qual agrega informações de ambas as técnicas (RESNICK; VARIAN, 1997).

## 2.2.1 Filtragem Colaborativa

A FC é a técnica mais bem sucedida e utilizada nos SRs. Baseia-se na seguinte hipótese: usuários que expressaram interesses similares no passado, compartilharão interesses semelhantes no futuro (SHI; LARSON; HANJALIC, 2014). Podemos usar como exemplo, um sistema de recomendação de músicas que identificou, através de notas dadas a itens, que os usuários  $a$  e  $b$  possuem gostos similares. Logo, o sistema recomendaria para  $b$  as músicas preferidas de  $a$  que  $b$  ainda não ouviu. A FC utiliza como fonte de informação a matriz *usuário-item*.

É possível encontrar na literatura uma variedade de métodos de filtragem colaborativa. Descreveremos a seguir os métodos que servirão como base para as abordagens híbridas utilizadas neste trabalho.

### 2.2.1.1 Slope One

O algoritmo *Slope One*, introduzido por Lemire e Maclachlan (2005), é uma das técnicas de filtragem colaborativa baseada em item mais populares na literatura por ter um conceito simples e de fácil implementação (ZACHARSKI, 2013). A intuição por trás deste algoritmo baseia-se no princípio de *diferencial de popularidade* entre os itens. Esse diferencial pode ser calculado subtraindo a média das avaliações de dois itens que será

usada para prever a avaliação de um item, dada a avaliação de outro. A Figura 3 ilustra melhor esse conceito, onde os usuários são representados por  $A$  e  $B$ , enquanto  $I$  e  $J$  representam os itens. Considerando que  $A$  avaliou os dois itens,  $I$  e  $J$ , dando as notas 1 e 1.5 respectivamente, enquanto  $B$  avaliou apenas o item  $I$ , observamos que a diferença de  $J$  para  $I$  é de 0.5 ( $1.5 - 1 = 0.5$ ). Logo, a nota prevista de  $B$  para  $J$  seria de 2.5.

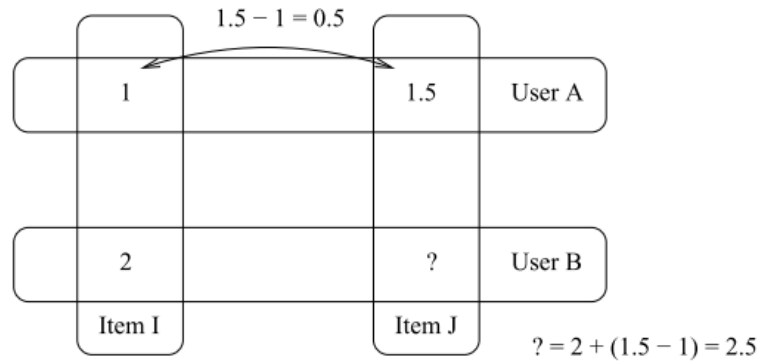


Figura 3 – Representação básica do algoritmo Slope One. Fonte: (LEMIRE; MACLA-CHLAN, 2005)

O Slope One pode ser dividido em duas partes. Primeiro, é necessário calcular o diferencial entre os itens, também chamado de *desvio*, formalmente definido como:

$$dev_{ij} = \frac{\sum_{u \in S_{ij}} (r_{ui} - r_{uj})}{|S_{ij}|} \quad (2.2)$$

onde  $U$  e  $S$  representam os conjuntos de usuários e itens respectivamente,  $r_{ui}$  e  $r_{uj}$  são as notas dadas pelo usuário  $u \in U$  para os itens  $i, j \in S$  e  $S_{ij}$  é o conjunto de usuários que avaliaram ambos itens  $i$  e  $j$ . Como resultado, temos uma matriz contendo os desvios entre os itens.

Uma vez calculada a matriz de desvios, a segunda parte do Slope One consiste em fazer a previsão da avaliação do usuário  $u$  para o item  $i$ , denotada como  $\hat{r}_{ui}$  e calculada da seguinte forma:

$$\hat{r}_{ui} = \frac{\sum_{j \in C_i} (dev_{ij} + r_{uj})}{|C_i|} \quad (2.3)$$

onde  $C_i$  corresponde a todos os itens avaliados por  $u$  que também foram avaliados em conjunto com  $i$  por outros usuários, tal que  $C_i = \{j | j \in R_u, j \neq i, |S_{ij}| > 0\}$ , onde  $R_u$  é a coleção de itens avaliados pelo usuário  $u$ .



### 2.2.1.2 Slope One Ponderado

Um dos problemas da equação 2.3 é que o número de vezes em que os itens  $i$  e  $j$  foram avaliados em conjunto não é levado em consideração. Imagine o seguinte cenário: um usuário  $U$  avaliou os itens  $I$  e  $J$  e queremos prever a nota que ele daria para o item  $L$ . Se 2 mil usuários avaliaram os itens  $I$  e  $L$  e apenas 20 avaliaram  $J$  e  $L$ , logo, o desvio do item  $I$  para o item  $L$  será mais robusto que o desvio de  $J$  para  $L$ . Portanto, o objetivo do algoritmo *Slope One Ponderado* (SOP) é dar mais peso aos desvios que foram calculados para os itens que tiveram maior número de avaliação. Esse método é formalmente definido como:

$$\hat{r}_{ui} = \frac{\sum_{j \in R_u} (dev_{ij} + r_{uj}) \times |S_{ij}|}{\sum_{j \in R_u} |S_{ij}|} \quad (2.4)$$

onde  $S_{ij}$  é a quantidade de vezes em que  $I$  e  $J$  foram avaliados em conjunto.

## 2.2.2 Filtragem Baseada em Conteúdo

A filtragem baseada em conteúdo (FBC) utiliza a similaridade entre os itens para gerar as recomendações (METEREN; SOMEREN, 2000). Em outras palavras, analisa a correlação de itens da preferência de um usuário com outros ainda não consumidos por ele. Podemos citar como exemplo um usuário que no passado avaliou positivamente um filme do gênero fantasia; logo, seriam recomendados para ele outros filmes de fantasia.

Esta abordagem ganhou bastante importância nos últimos anos graças a consolidação da Web 2.0. Existe uma tendência em incorporar nos SRs, conteúdos produzidos pelos usuários para melhorar a acurácia das predições. Estas informações adicionais podem ser extraídas de comentários, críticas, opiniões, assim como os relacionamentos nas redes sociais (BOBADILLA et al., 2013).

Para medir a similaridade entre os itens, podemos utilizar diferentes estratégias dentre as quais detalharemos a seguir as medidas implementadas neste trabalho.

### 2.2.2.1 Distância Semântica

O trabalho de Yang, Hu e Qu (2013), que serve como referência para esta pesquisa, empregou a abordagem *Linked Data Semantic Distance* (LDS), introduzida no trabalho de Passant (2010), para medir a semelhança entre itens através das informações provenientes da base de dados DBTropes.

O objetivo da medida LDS é calcular a distância entre dois recursos publicados como *Linked Data* e é comumente empregada no domínio da *Web Semântica*<sup>4</sup>. O resultado

<sup>4</sup> A W3C define a Web Semântica como uma extensão da Web tradicional, com intuito de fornecer um arcabouço que permita o compartilhamento e reuso dos dados em diferentes aplicações

dessa medida está normalizado no intervalo  $[0,1]$ , tal que quanto mais próximo de 0, mais semelhantes dois itens distintos são. Como nossa pesquisa não está diretamente relacionada a esse domínio (apesar de usarmos uma coleção concebida com tal conceito), não nos aprofundaremos neste assunto, daremos apenas uma visão geral de modo que possamos entender seus conceitos fundamentais.

Segundo Passant (2010), uma coleção segue os princípios de *Linked Data* quando é formada por um grafo  $G$  tal que  $G = (R, L, I)$  no qual  $R = \{r_1, r_2, \dots, r_n\}$  é um conjunto de recursos — identificados por sua URI —,  $L = \{l_1, l_2, \dots, l_n\}$  é um conjunto de *links tipados* — identificados por sua URI — e  $I = \{i_1, i_2, \dots, i_n\}$  representa as instâncias dos *links* entre os recursos.

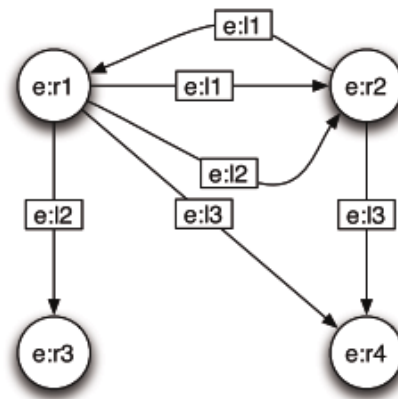


Figura 4 – Exemplo de um grafo Linked Data. Fonte: (PASSANT, 2010)

O grafo da figura 4 ilustra uma coleção baseada nas definições de *Linked Data*, no qual utiliza o domínio <http://exemplo.org/> (prefixo  $e$ ) como exemplo. Esse grafo pode ser representado como:

$$R = \{e:r1, e:r2, e:r3, e:r4\}$$

$$L = \{e:l1, e:l2, e:l3\}$$

$$I = \{\langle e:l1, e:r1, e:r2 \rangle, \langle e:l1, e:r2, e:r1 \rangle, \\ \langle e:l2, e:r1, e:r2 \rangle, \langle e:l2, e:r1, e:r3 \rangle, \\ \langle e:l3, e:r1, e:r4 \rangle, \langle e:l3, e:r2, e:r4 \rangle\}$$

Utilizando o grafo da figura 4 no contexto do site TVTropes,  $e:r1$  e  $e:r2$  podem representar, por exemplo, os filmes *Matrix Revolution* e *Blade Runner* respectivamente e  $e:r4$  o trope *Batalha na Chuva*. Neste caso, os filmes estão interligados diretamente, através do *link*  $e:l1$ , e indiretamente, através da conexão com o trope *Batalha na Chuva*.

Para calcular a distância semântica (LDS) entre dois recursos  $r_a$  e  $r_b$ , Passant (2010) apresenta diferentes mecanismos que utilizam os *links* diretos e indiretos (de entrada

e saída). Detalharemos a seguir, as duas formas implementadas por Yang, Hu e Qu (2013).

### Distância Direta

A distância direta é baseada nos *links* diretos — de entrada e saída — entre dois recursos. Para fundamentar essa medida, Passant (2010) apresenta a seguinte definição:

$C_d$  é a função que calcula a quantidade de links diretos distintos entre dois recursos em um grafo  $G$ .  $C_d(l_i, r_a, r_b)$  é igual a 1 se existir uma instância do tipo  $l_i$  do recurso  $r_a$  para o recurso  $r_b$  e é 0 se não existir. Logo,  $C_d$  pode ser usado para calcular o total de links diretos e distintos de  $r_a$  para  $r_b$  ( $C_d(n, r_a, r_b)$ ), bem como o total de links diretos e distintos de  $r_a$  para qualquer nó  $n$ , ( $C_d(n, r_a, n)$ ).

Assim, usando como exemplo o grafo da figura 4, a função  $C_d$  retornaria os seguintes valores para os recursos  $e:r1$  e  $e:r2$ :

$$C_d(e : l1, e : r1, e : r2) = 1$$

$$C_d(n, e : r1, e : r2) = 2$$

$$C_d(n, e : r2, e : r1) = 1$$

$$C_d(e : l1, e : r1, n) = 1$$

$$C_d(e : l2, e : r1, n) = 2$$

Baseado nessas definições, a medida de similaridade  $LDSD_d$  é então formalmente definida como:

$$LDSD_d(r_a, r_b) = \frac{1}{1 + C_d(n, r_a, r_b) + C_d(n, r_b, r_a)} \quad (2.5)$$

Por exemplo, a distância entre  $e:r1$  e  $e:r2$  no grafo da figura 4, é calculada como:

$$LDSD_d(e : r1, e : r2) = \frac{1}{1 + 2 + 1} = 0,25$$

### Distância Indireta

A outra versão do LDS D considera os *links* indiretos entre os itens que são extraídos através da conexão com outros recursos. Essa medida é conhecida como  $LDSD_i$ , porém, assim como na distância direta, Passant (2010) antecede o detalhamento desse método com a seguinte definição:

$C_{io}$  e  $C_{ii}$  são duas funções que calculam o número de links indiretos e distintos, de entrada e saída, entre recursos em um grafo  $G$ .  $C_{io}(l_i, r_a, r_b)$  é igual a 1 se

existir um recurso  $n$  que satisfaça tanto  $\langle l_i, r_a, n \rangle$  quanto  $\langle l_i, r_b, n \rangle$ , sendo 0 caso contrário.  $C_{ii}(l_i, r_a, r_b)$  é igual a 1 se existir um recurso  $n$  que satisfaça tanto  $\langle l_i, n, r_a \rangle$  quanto  $\langle l_i, n, r_b \rangle$ , sendo 0 caso contrário. Por extensão,  $C_{io}$  e  $C_{ii}$  podem ser usados para calcular o total de links indiretos e distintos entre  $r_a$  e  $r_b$  ( $C_{io}(n, r_a, r_b)$  e  $C_{ii}(n, r_a, r_b)$ ), de saída e entrada respectivamente, bem como o total de recursos  $n$  ligados indiretamente em  $r_a$  via  $l_i$  ( $C_{io}(l_i, r_a, n)$  e  $C_{ii}(l_i, r_a, n)$ ), de saída e entrada respectivamente

Desta forma, no grafo que serve como exemplo,  $C_{io}(e : l3, e : r1, e : r2) = 1$  (via links de saída para e:r4) e  $C_{ii}(e : l2, e : r2, e : r3) = 1$  (via links de entrada oriundos de e:r1).

Dito isto, a medida  $LDSD_i$  é formalmente definida pela seguinte equação:

$$LDSD_i(r_a, r_b) = \frac{1}{1 + C_{io}(n, r_a, r_b) + C_{ii}(n, r_a, r_b)} \quad (2.6)$$

Assim, a distância entre  $e:r1$  e  $e:r2$  no grafo da figura 4 é calculada como:

$$LDSD_i(e : r_1, e : r_2) = \frac{1}{1 + 1 + 0} = 0,5 \quad (2.7)$$

Nas próximas seções, detalharemos os métodos adotados em nossa pesquisa para substituir as estratégias baseadas na LDSD na tarefa de medir a similaridade entre os filmes. Nosso objetivo é empregar somente os tropes, eliminando outros tipos de relacionamento entre eles.

#### 2.2.2.2 Coeficiente de Jaccard

Uma estratégia bastante simples para medir a similaridade entre itens é considerá-los como documentos formados por vetores em um espaço  $n$ -dimensional, onde o número de dimensões corresponde aos termos distintos da coleção de itens e, em seguida, calcula-se a similaridade entre eles utilizando o Coeficiente de Jaccard (SHEPITSEN et al., 2008; U.; YAO; MAK, 2014). Nesta métrica, a correlação entre dois vetores  $x$  e  $y$  é medida da seguinte forma:

$$jac(x, y) = \frac{|x \cap y|}{|x \cup y|}. \quad (2.8)$$

Quanto mais próximo de 1, mais similares os vetores são. Deste modo, para usá-la como uma distância semântica, onde o comportamento é inverso, definimos a distância de Jaccard como:

$$dis_j(x, y) = 1 - jac(x, y). \quad (2.9)$$

### 2.2.2.3 Similaridade do Cosseno

Outra abordagem comumente referenciada na literatura para medir a similaridade entre vetores consiste em calcular o cosseno do ângulo formado por eles, formalmente definido como:

$$\cos(x, y) = \frac{x \cdot y}{\|x\| \times \|y\|}. \quad (2.10)$$

onde  $x$  e  $y$  são vetores binários indicando a presença ou ausência de determinado termo. Desta forma, a similaridade do cosseno é obtida através do produto dos vetores  $x$  e  $y$ , dividido pelo produto de suas normas  $\|x\|$  e  $\|y\|$ . Da mesma forma que fizemos para a distância de Jaccard, Equação 2.9, definimos a distância do cosseno como:

$$dis_c(x, y) = 1 - \cos(x, y). \quad (2.11)$$

### 2.2.2.4 TF-IDF

A abordagem anterior não leva em consideração a relevância relativa que cada termo pode ter de acordo com o documento em que é utilizado. Essa importância pode ser expressada através do método de ponderação TF-IDF (do inglês, Term Frequency - Inverse Document Frequency) que é bastante usado na área de recuperação da informação (RICCI et al., 2010).

No TF-IDF, termos que ocorrem com mais frequência em um documento (TF), mas raramente no restante da coleção (IDF), tendem a ser mais relevantes para este. Para nivelar pequenos documentos com outros mais extensos, é aplicada uma normalização. Formalmente, o TF-IDF de um termo  $t_k$  em um documento  $d_j$  é calculado através da seguinte função:

$$TFIDF(t_k, d_j) = TF(t_k, d_j) \times \log \frac{N}{n_k}, \quad (2.12)$$

tal que  $N$  denota o total de documentos da coleção e  $n_k$ , o número de documentos onde o termo  $k$  ocorreu. Para evitar que a relevância de um termo cresça proporcionalmente à sua frequência na função TF, uma das abordagens encontradas na literatura consiste em:

$$TF_{k,j} = \left\{ \begin{array}{ll} 1 + \log f_{k,j} & \text{se } f_{k,j} > 0 \\ 0 & \text{caso contrário} \end{array} \right\}, \quad (2.13)$$

onde  $f_{k,j}$  representa a frequência do termo  $k$  no documento  $j$ .

Uma vez calculados os vetores de cada documento, é possível medir a similaridade entre dois documentos utilizando a similaridade do cosseno. Ao longo deste texto, usaremos o termo TF-IDF para se referir à similaridade do cosseno usando ponderação TF-IDF.

### 2.2.3 Slope One Ponderado Híbrido

Segundo Adomavicius e Tuzhilin (2005 apud BOBADILLA et al., 2013), as técnicas híbridas têm contribuído bastante para a evolução dos SRs, cujo objetivo é melhorar a eficácia da recomendação combinando diferentes abordagens uma vez que extraem as vantagens de cada uma. Muitos trabalhos têm utilizado a combinação das filtragens colaborativa e baseada em conteúdo que pode ser classificada como na Figura 5.

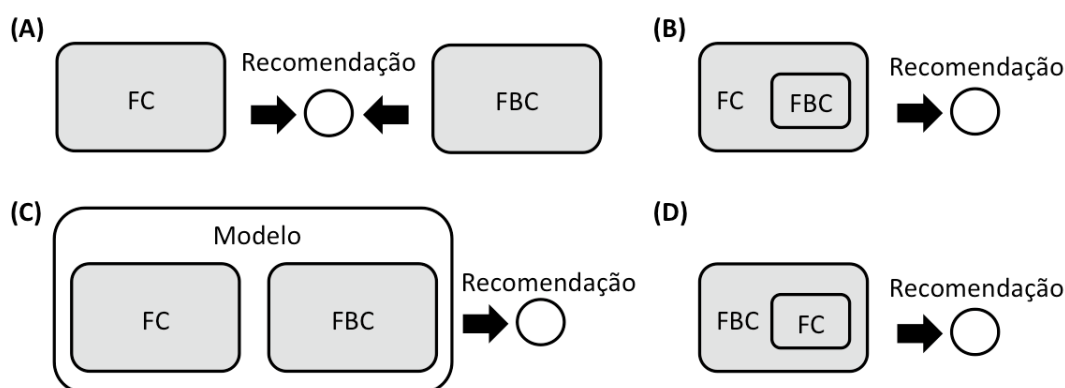


Figura 5 – Abordagens para a combinação de FC e FBC. (A) FC e FBC são implementados separadamente e depois combinados para gerar a recomendação; (B) características da FBC são incorporadas na FC; (C) modelo unificado das filtragens colaborativa e baseada em conteúdo; e (D) características da FC são incorporadas na FBC.

Nesta figura, observamos quatro tipos gerais de combinação. No primeiro, os métodos colaborativos e baseado em conteúdo são implementados separadamente e combinados posteriormente para gerar as recomendações, conforme ilustrado na Figura 5-A. No segundo, características de métodos baseados em conteúdo são incorporadas em abordagens colaborativas (Figura 5-B). O inverso é ocorre no método representados pela Figura 5-D, onde características de métodos colaborativos são incorporadas em abordagens baseadas em conteúdo. Finalmente, a Figura 5-C exemplifica um modelo unificado que incorpora ambas as filtragens colaborativas e baseadas em conteúdo. Neste trabalho, em particular, adotamos o método Slope One Ponderado Híbrido, baseado na última abordagem.

Desde seu surgimento, o algoritmo Slope One sofreu diferentes modificações na tentativa de incorporar outras características para melhorar a qualidade da recomendação (MENEZES et al., 2013; ZHAO; MA, 2013). O trabalho realizado por Yang, Hu e Qu (2013) apresentou diferentes abordagens para incorporar informações oriundas da Web

Semântica no SOP, em particular, do TVTropes. Dentre as estratégias implementadas, utilizamos em nossa pesquisa o método que modifica diretamente o cálculo do peso no SOP que agora leva em consideração a influência do relacionamento entre os itens. Denotamos esse método como *Slope One Ponderado Híbrido* (SOPH) que é formalmente definido como:

$$\hat{r}_{ui} = \frac{\sum_{j \in R_u} (dev_{i,j} + r_{u,j}) \times \frac{|S_{ij}|}{DS(i,j)}}{\sum_{j \in R_u} \frac{|S_{ij}|}{DS(i,j)}}, \quad (2.14)$$

onde  $DS(i, j)$  é distância semântica entre os itens  $i$  e  $j$  previamente descrita na seção 2.2.2.1. O objetivo dessa estratégia é que o peso  $|S_{ij}|$ , constituído da frequência que  $i$  e  $j$  são avaliados em conjunto, seja proporcional à força do relacionamento entre os itens, que é medido através da similaridade de seus atributos.

Para que seja possível observar a influência da distância semântica no SOP, este propõe a seguinte transformação linear, definida como segue:

$$\begin{aligned} \hat{r}_{ui} = (1 - \alpha) \times \frac{\sum_{j \in R_u} (dev_{ij} + r_{uj}) \times |S_{ij}|}{\sum_{j \in R_u} |S_{ij}|} \\ + \alpha \times \frac{\sum_{j \in R_u} (dev_{i,j} + r_{uj}) \times \frac{|S_{ij}|}{DS(i,j)}}{\sum_{j \in R_u} \frac{|S_{ij}|}{DS(i,j)}}. \end{aligned} \quad (2.15)$$

A primeira metade da equação 2.15 consiste do algoritmo Slope One Ponderado original, enquanto a segunda metade incorpora a distância semântica entre os itens. O parâmetro  $\alpha$  ajusta a proporção que cada parte terá sobre a recomendação final.

Em particular, como será observado mais adiante, em nossa proposta substituiremos a função  $DS(i, j)$ , que emprega a distância semântica, por medidas baseadas no coeficiente de Jaccard, na similaridade do cosseno e no esquema de ponderação TF-IDF além de usar somente os tropes para calcular a distância entre os itens.

## 2.2.4 Limitações dos SRs

Embora os SRs tenham sido utilizados com sucesso em diferentes aplicações, existem algumas limitações comumente referenciadas na literatura (BOBADILLA et al., 2013):

- **Novo Usuário:** Representa uma das maiores dificuldades enfrentadas pelos SRs. Ocorre quando o usuário não avaliou itens o suficiente para que o sistema possa entender suas preferências e encontrar usuários similares a ele, tornando difícil a tarefa de gerar recomendações personalizadas confiáveis. As estratégias frequentemente encontradas na literatura para reduzir o impacto deste problema baseiam-se em modelos

híbridos que incorporam outras fontes de informações como características dos itens, dados demográficos e informações sociais.

- **Novo Item:** Afeta diretamente métodos colaborativos uma vez que um novo item não possui avaliações o suficiente para que o sistema possa inferir sua similaridade para com outros itens, dificultando sua recomendação. Consequentemente, itens não recomendados não serão notados por grande parte dos usuários e dificilmente serão avaliados, gerando um ciclo vicioso de itens não recomendados. As abordagens baseadas em conteúdo não são afetadas por este problema uma vez que utilizam as propriedades dos itens, previamente cadastradas, para computar a similaridade entre os mesmos e gerar a recomendação.
- **Matrizes Esparsas:** Tipicamente, usuários avaliam apenas uma pequena quantidade de itens. A consequência disso se reflete em matrizes *usuário-item* esparsas. Esse problema prejudica métodos baseados em FC uma vez que estes precisam de um número de avaliações mínimo para inferir a similaridade de usuários e itens. Algumas abordagens utilizadas para mitigar esse problema baseiam-se nas estratégias adotadas no problema do *Novo Usuário* (ADOMAVICIUS; TUZHILIN, 2005). Outras abordagens encontrados na literatura para lidar com este cenário utilizam técnicas de redução de dimensionalidade baseadas na *Fatoração de Matrizes* ou a combinação dos métodos de Indexação Semântica Latente (*Latent Semantic Indexing* - LSI) e Decomposição de Valores Singulares (*Singular Value Decomposition* - SVD) (RICCI et al., 2010).
- **Super-especialização:** Refere-se ao fenômeno em que os usuários recebem apenas recomendações de itens similares aos que eles já consumiram. Por exemplo, para um usuário que assistiu um filme de romance dificilmente seriam recomendados outros filmes de outros gênero. Esse fenômeno é comumente observado em métodos de FBC, uma vez que eles empregam as características dos itens para medir a similaridade entre os mesmos e gerar as predições. Esse problema tornou-se bastante relevante nos últimos anos. Uma prova disso é que, atualmente, umas das formas de avaliar a qualidade de um SR consiste em medir o grau de novidade das suas recomendações.
- **Análise de Conteúdo Limitado:** FBCs são limitadas pelas características associadas aos itens. Esta caracterização de itens pode ser feita de forma manual ou automática. O desafio encontra-se na tarefa de extração de características que sejam relevantes para produzir recomendações confiáveis.

### 2.2.5 Avaliação

A avaliação da qualidade da recomendação sempre foi parte fundamental para testar e melhorar os SRs, contribuindo diretamente para o avanço da área, pois facilita a



comparação de diferentes soluções e linhas de pesquisa com resultados mais promissores (BOBADILLA et al., 2013).

Um ponto importante no processo de validação é a confiabilidade dos resultados. Deste modo, uma técnica comumente empregada é a validação cruzada de  $k$  partições, onde cada base de dados é dividida aleatoriamente em  $k$  partes, tal que, em cada rodada de experimentação, uma partição diferente é usada como conjunto de teste, enquanto o restante das instâncias é usado para treinar o modelo (HALL et al., 2009).

Adicionalmente, para garantir que todos os métodos de recomendação possuam diferenças estatisticamente significativas, podemos aplicar o teste de Wilcoxon pareado (SHANI, 2009). Este teste não paramétrico não faz suposições de distribuição sobre as diferenças. Neste contexto, as diferenças entre os métodos podem ser consideradas estatisticamente significativas quando os resultados alcançam um nível de confiança acima de 95%.

## RMSE

Para avaliar a acurácia dos SRs, umas das abordagens mais referenciadas na literatura são as métricas de erro de previsão (CREMONESI et al., 2008). Essas medidas são conhecidas pela facilidade na implementação e modelam os SRs como um problema de previsão numérica, ou seja, tentam prever as notas dos usuário aos itens. Dentre elas, podemos destacar a métrica Raiz do Erro Médio Quadrático (RMSE), que calcula diferença entre as previsões realizadas e os valores reais, penalizando os grandes erros. O RMSE é definida conforme a Equação 2.16.

$$RMSE_u = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{r}_{ui} - r_{ui})^2} \quad (2.16)$$

onde  $n$  é o número de instâncias de teste,  $\hat{r}_i$  é a nota prevista pelo método avaliado para a instância de teste  $i$  e  $r_i$  é a nota real da instância  $i$ . Note que o RMSE pode ser calculado tanto em relação aos usuários quanto em relação aos itens. Neste trabalho, em particular, vamos optar pelo cálculo em relação aos usuários. Assim, o RMSE é dado por:

$$RMSE = \frac{1}{|U|} \sum_u RMSE_u \quad (2.17)$$

## Precisão, Revocação, F1

Segundo Bobadilla et al. (2013), a confiança do usuário em um SR não depende diretamente da acurácia das previsões, mas sim o quão ele concorda com um conjunto reduzido de recomendações. Neste contexto, outra forma de avaliar o desempenho dos

SRs é através das métricas de *Precisão*, *Revocação* e *F1*, que são bastante conhecidas na área de recuperação da informação.

A precisão é definida como número de itens relevantes recuperados dividido pelo total de itens recuperados, enquanto a revocação é obtida pela razão entre o número de itens relevantes recuperados e o total de itens relevantes disponíveis. A métrica F1 é uma média harmônica derivada da combinação entre precisão e revocação.

Formalmente, sendo  $R_N(u)$  uma lista de recomendações com  $N$  itens,  $T(u)$  o conjunto de itens relevantes para um usuário e  $U$  o conjunto de usuários, temos:

$$Precisão = \frac{1}{|U|} \sum_{u \in U} \frac{|R_N(u) \cap T(u)|}{|R_N(u)|}, \quad (2.18)$$

$$Revocação = \frac{1}{|U|} \sum_{u \in U} \frac{|R_N(u) \cap T(u)|}{|T(u)|}, \quad (2.19)$$

$$F1 = \frac{1}{|U|} \sum_{u \in U} \frac{2 \times precisão(u) \times revocação(u)}{precisão(u) + revocação(u)}. \quad (2.20)$$

Perceba que assim como na métrica RMSE, estas medidas também serão calculadas por usuário.

Os itens são classificados como relevantes para um usuário se sua nota estiver acima de um limiar. Neste trabalho, foi utilizado um limiar específico para cada usuário baseado em sua nota média com objetivo de minimizar o viés particular de avaliação. Desta forma, um item é considerado relevante para um usuário, quando ele avalia o item como acima da sua nota média pessoal.

## 2.3 Trabalhos Relacionados

Esta sessão lista os trabalhos relacionados encontrados na literatura no domínio de SRs. Na Seção 2.3.1, abordaremos os autores que, de alguma maneira, utilizaram tropes como fonte de informação em suas pesquisas. Os trabalhos relacionados com SRs e, em particular, no contexto de recomendação de filmes, serão explorados na Seção 2.3.2.

### 2.3.1 Tropes

Tropes têm sido muito pouco usados como fonte de evidência em qualquer aplicação. Até onde sabemos, apenas três trabalhos usaram tropes. O primeiro usa a coleção de

tropes para detectar *spoilers*<sup>5</sup>; o segundo caracteriza personagens de um filme; e o último, os utiliza no contexto de recomendação.

O trabalho realizado por Boyd-Graber, Glasgow e Zajac (2013) propõe uma abordagem para detecção de *spoilers* em textos. Para isso, os autores coletaram informações do TVTropes.org para a criação de uma base de dados com textos rotulados como “spoiler” e “não spoiler”. A extração desses rótulos do TVTropes.org só é possível porque o site dispõe de uma funcionalidade que oculta partes de um texto onde ocorre spoiler. Logo, toda vez que o texto tem uma parte ocultada, ele é rotulado como spoiler. Isso é necessário porque uma característica de muitos tropes é a sua relação com características importantes da narrativa (ex: “o assassino é o mordomo”). O problema foi modelado como uma tarefa de classificação de textos no qual foi utilizado o algoritmo *Support Vector Machine* (SVM) para aprender a função objetivo. Um dos problemas encontrados pelos autores com relação ao TVTropes.org é que existem textos com spoiler que não são ocultados, gerando inconsistências na base de dados.

A pesquisa de Bamman, O’Connor e Smith (2013) apresenta modelos para a extração de classes latentes de personagens por meio de análise de textos. Através destes modelos é possível identificar os padrões de comportamento de um vilão, um herói e um agente secreto, entre outros tipos. Para isso, os autores utilizaram informações da Wikipédia e da Freebase<sup>6</sup>. O TVTropes foi utilizado para validar seu modelo, uma vez que este contém tropes que descrevem características dos personagens como “o executivo corrupto”, “o desastrado” e “o surfista”.

A utilização de Tropes em recomendação foi considerada por Yang, Hu e Qu (2013). Neste trabalho, é proposta uma abordagem híbrida para a recomendação de filmes baseada no algoritmo *Slope One Ponderado*. Para isso, os autores combinam dois tipos de informação: a matriz usuário-item proveniente do MovieLens (MOVIELENS, 2015) com informações coletadas do DBTropes, entre as quais, os próprios tropes. Nesse trabalho, a similaridade entre os filmes é calculada através da técnica *Linked Data Semantic Distance* (LDS) que computa a distância entre dois itens publicados como *Linked Data*. Mais especificamente, a similaridade de dois filmes é proporcional aos elementos que eles têm em comum, tais como elenco, gênero e tropes. A similaridade entre os itens é então usada como fator de ponderação no algoritmo *Slope One Ponderado*. Os resultados preliminares dos experimentos mostraram que a abordagem proposta apresentou melhor desempenho que o *Slope One* original.

<sup>5</sup> Revelação de um fato crítico a respeito do conteúdo de uma história. O termo vem do inglês, mais precisamente está relacionado ao verbo “To Spoil”, que significa estragar. Numa tradução livre, spoiler faz referência ao famoso termo “estraga-prazeres”.

<sup>6</sup> Freebase é uma base de conhecimento colaborativa, com objetivo de criar um repositório global que permita pessoas (e máquinas) acessarem informações comuns, efetivamente. O projeto foi desativado em 31 de agosto de 2016, porém uma cópia da base de dados pode ser encontrada em <https://developers.google.com/freebase/>.

Este último trabalho, até onde sabemos, foi o único a usar tropes como evidência para recomendação. Contudo, ao contrário da nossa pesquisa, não apresenta um estudo sistemático sobre tropes nem apresentou um modelo de recomendação diretamente focado em tropes. Nesse estudo, tropes foram usados apenas como evidência adicional, sem uma clara avaliação do seu desempenho particular.

### 2.3.2 Sistemas de Recomendação

SRs tornaram-se foco de muitos estudos uma vez que seus benefícios atraíram o interesse da academia e da indústria nas últimas décadas. Um importante fator que contribuiu para esta popularidade foi o aumento no volume de informação proporcionado pelo crescimento da Internet.

Trabalhos como os de Bobadilla et al. (2013) e Shi, Larson e Hanjalic (2014) apresentam uma boa descrição da literatura geral sobre SRs cobrindo inúmeros tópicos, como a evolução dos SRs, seus tipos, métodos, áreas de aplicação e formas de avaliação, entre outros. Destes trabalhos, fica clara a evolução dos métodos desde os sistemas baseados em conteúdo até os modernos métodos híbridos, passando pelos métodos colaborativos, dominantes em grande parte do tempo. Com a consolidação da Web 2.0, o cenário atual consiste em abordagens híbridas que incorporam diferentes fontes de informação para melhorar a qualidade de recomendação. Ambas as pesquisas destacam a importância das redes sociais neste contexto que adicionaram aos SRs, conceitos de amizade, seguidores e confiança. Outro aspecto que também vem sendo bastante explorado é o conteúdo gerado pelos usuários (tags, geotags, comentários, etc) por meio de plataformas colaborativas. Futuramente, a tendência dos SRs é a utilização das informações provenientes da *Internet das Coisas* como GPS, RFID, sensores cardíacos, entre outros dispositivos, possibilitando, por exemplo, recomendações contextualizadas com a localização do usuário.

Entre os vários itens alvos de SRs, um dos mais estudados são filmes (BOBADILLA et al., 2013). Nos próximos parágrafos, descrevemos alguns sistemas criados especificamente para a recomendação de filmes, que exploram diferentes fontes de informação na recomendação.

O sistema INTIMATE, desenvolvido por Mak, Koprinska e Poon (2003), recomenda filmes utilizando a técnica de categorização de texto extraídas das sinopses dos filmes coletados do *Internet Movie Database* (IMDB). Os resultados desta abordagem mostram um desempenho melhor em condições específicas quando comparado com modelos baseados em características, como gênero, elenco, direção e etc. Neste mesmo caminho, a ferramenta MOVIES2GO, proposta por Mukherjee et al. (2001) também explora a sinopse dos filmes, porém, sua solução principal gira em torno da preferência do usuário por determinados gêneros, atores ou diretores.

No trabalho de Lees-Miller et al. (2008), os autores fazem uso da Wikipédia para estimar a similaridade entre filmes e gerar previsões mais acuradas. Mais especificamente, a solução proposta explora o conteúdo e a estrutura de hiperlinks dos artigos da Wikipédia para gerar uma matriz contendo o grau de similaridade filme a filme. Essa matriz é utilizada posteriormente para prever as notas de um usuário combinando com os algoritmos *k-Nearest Neighbors* (k-NN) e um *Pseudo-SVD*. Infelizmente, essa abordagem não trouxe melhorias significativas na acurácia das previsões.

A abordagem utilizada por Szomszor et al. (2007) explora palavras-chave associadas a um filme, por meio de uma plataforma colaborativa de social *tagging*, para calcular a similaridade entre eles. A ideia por trás desta abordagem é sobre o uso de *tags* para a construção de perfis de usuários que melhor reflitam seu grau de interesse para diferentes tipos de filme. Neste trabalho foram utilizadas as bases de dados do IMDB e Netflix. O IMDB fornece informações sobre filmes e *tags*, enquanto o Netflix contém dados de avaliações dos usuários. A conclusão dos autores sobre o trabalho é que as *tags* podem ser utilizadas para a recomendação de filmes, embora terminem por sugerir a combinação com outros modelos baseados em conteúdo e colaborativo.

A pesquisa de Noia et al. (2012) destaca a importância do *Linked Data* para os SRs, incorporando as informações provenientes das bases de dados *DBPedia*, *Freebase* e *LinkedMDB*, em um método baseado em conteúdo para a recomendação de filmes. Utiliza o modelo vetorial para calcular a similaridade entre dois itens baseada nas propriedades dos filmes, como elenco, direção, gênero, categorias e etc. Sua abordagem recomenda filmes para um usuário, similares aos filmes já assistidos por ele.

O projeto *Entertainment Genome*<sup>7</sup> desenvolvido pelo SR Jinni tem o objetivo de mapear elementos (genes) do conteúdo de filmes, shows, entre outros tipos de vídeos. Os genes podem ser, por exemplo, características da trama (guerra dos sexos), estruturais (história não linear, história dentro história), *flags* (violência, nudez), entre outros. Além disso, inclui aspectos externos como premiações. O processo de mapeamento de genes é feito inicialmente por profissionais. Novos itens são indexados automaticamente por meio da análise das críticas feitas por profissionais/usuários e metadados, utilizando Aprendizado de Máquina e Processamento de Linguagem Natural. A recomendação gerada leva em consideração o tipo de conteúdo de interesse do usuário que é definido através das avaliações feitas em itens já consumidos por ele.

Como em nosso trabalho, vários dos métodos descritos utilizam diferentes evidências para descrever os filmes. Contudo, nenhum deles explora tropes.

---

<sup>7</sup> <http://www.jinni.com/info/entertainment-genome.html>

## 3 Caracterização do TVTropes

Neste capítulo, são descritos os estudos realizados no site TVTropes. O primeiro estudo apresenta uma investigação feita nesta plataforma com objetivo de identificar as razões que levam um filme a ser catalogado pela comunidade, enquanto o segundo explora preliminarmente, o relacionamento de tropes e filmes no contexto de recomendação.

### 3.1 Características do catálogo de tropes nos filmes

Um ponto importante nessa pesquisa para fundamentar nossas análises, consiste em investigar o padrão de ocorrência dos tropes nos filmes. Para realização desta atividade, utilizamos o DBTropes por este extrair e estruturar as informações do TVTropes. A coleção utilizada contém 9.002<sup>1</sup> filmes cadastrados que estão associados a 25.044 tropes.

Inicialmente, investigamos a quantidade de tropes por filme. Essa quantidade pode influenciar diretamente a qualidade das previsões, pois é utilizada para medir a similaridade entre os filmes. Em particular, a Tabela 1 apresenta estatísticas sobre a base de dados. A observação dos percentis e do valor máximo mostram claramente que se trata de uma distribuição de potência, o que é ilustrado também na Figura 6.

Quantidade de Tropes por Filme	Valor
Média	45.36
Desvio Padrão	73.28
Mínima	1
25%	7
50%	23
75%	51
Máxima	1204

Tabela 1 – Estatísticas sobre a quantidade de tropes catalogados nos filmes da coleção DBTropes.

A Figura 6 apresenta o histograma da distribuição dos tropes nos filmes. Observe a ocorrência do fenômeno de cauda longa, ou seja, a maioria dos filmes têm poucos tropes catalogados.

As Tabelas 2 e 3 listam amostras de filmes com maior e menor quantidade de tropes respectivamente. Nestas listas, podemos observar que alguns filmes não possuem ano preenchido, uma vez que tal informação não é obrigatória ao se catalogar um filme no TVTropes, podendo ser extraída de outras regiões do site. Séries de filmes, como “James

<sup>1</sup> De fato, a quantidade de filmes encontrados na coleção DBTropes foi maior, entretanto filtramos apenas aqueles que continham ao menos um trope catalogado.

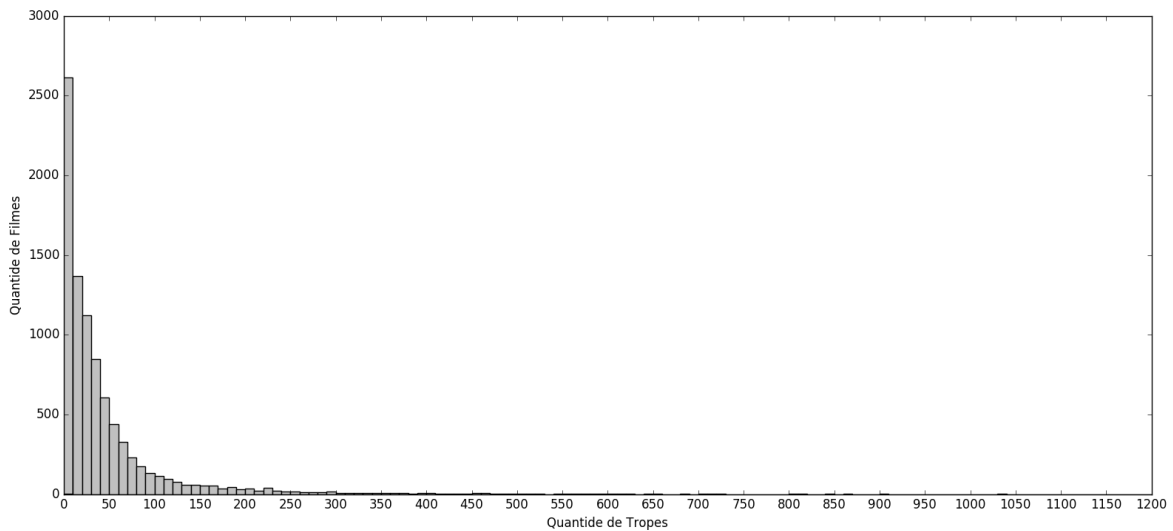


Figura 6 – Histograma do número de filmes agrupados pela quantidade de tropes da coleção DBTropes.

Bond” e a trilogia “The Lord Of The Rings”, são catalogados em uma única página, o que explica sua grande quantidade de tropes.

Filme	Tropes	Ano
James Bond	1204	-
The Lord Of The Rings	1036	-
Frozen	907	-
Aladdin	867	1992
The Dark Knight	849	2008
Wreck-It Ralph	814	2012
The Incredibles	811	2004
The Lion King	803	-
The Matrix	800	1999
Star Trek	789	1985

Tabela 2 – Lista dos dez filmes com mais tropes catalogados na coleção DBTropes.

Filme	Tropes	Ano
Falcon Rising	1	2014
Kronos	1	-
Choudenshi Bioman	1	1984
Blaze	1	1989
Kidou Keiji Jiban	1	1989
Hikari Sentai Maskman	1	1987
Choushinsei Flashman	1	1986
Choujinki Metalder	1	1987
Deathtrap	1	1982
Aquatic Wizards	1	-

Tabela 3 – Lista com exemplos de filmes com menos tropes catalogados na coleção DBTropes.

### 3.1.1 Correlação entre tropes e metadados dos filmes

Para uma caracterização mais detalhada, utilizamos a base de dados do *Internet Movie Database*<sup>2</sup> (IMDB) de onde é possível extrair diferentes informações sobre os filmes. Isso nos possibilitou uma investigação sobre o relacionamento dos tropes com o gênero, a popularidade e as preferências dos usuários em relação aos filmes.

O primeiro passo foi construir um subconjunto com a interseção entre as coleções do DBTropes e o IMDB. Para tanto, utilizamos uma função de equivalência de alfa-numéricos

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

simples para associar os filmes correspondentes nas duas bases de dados através do título. Além disso, utilizamos também o ano de lançamento (quando este esteve disponível) para aumentar a acurácia desta abordagem. Nesta filtragem, aproximadamente 83% dos filmes do DBTropes foram encontrados no IMDB.

A Figura 7 exibe um gráfico com a média de tropes por gênero. Nela, é possível observar que a média é maior em filmes de Animação, Aventura, Ação e Fantasia quando comparado com outros gêneros, especialmente Comédia e Drama, que possuem mais filmes cadastrados que tropes catalogados. Esse fenômeno pode ser interpretado como uma preferência da comunidade do TVTropes por filmes destes gêneros (Animação, Aventura, Ação e Fantasia) aliada à possibilidade de que eles sejam percebidos como mais formulaicos, sendo identificadas grandes variedades de tropes em sua construção.

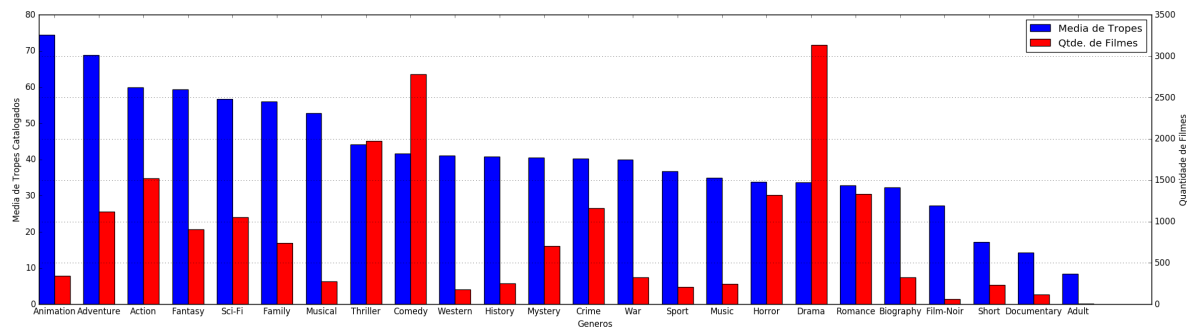


Figura 7 – Média de tropes por gênero da coleção DBTropes  $\cap$  IMDB.

Na Figura 8, é possível visualizar como os gêneros estão distribuídos no espaço de acordo com os seus tropes. A metodologia utilizada para construção deste espaço considerou os gêneros como vetores de tropes contendo a quantidade de vezes em que cada trope foi referenciado em um filme de determinado gênero. Como este vetor possui alta dimensionalidade, utilizamos a ferramenta t-SNE (MAATEN; HINTON, 2008) para a redução em um espaço bidimensional de modo a facilitar a visualização. Como resultado, podemos observar algumas aproximações de gêneros que apresentam intuitivamente, características similares. Por exemplo:

- Ficção-Científica, Aventura e Ação;
- Suspense, Crime, Mistério e Terror;
- Animação, Comédia, Família e Fantasia;
- Drama, Romance e Musical;
- História, Guerra e Biografias;

Note que este resultado sugere uma relação entre certos gêneros e os padrões de narrativa que eles empregam, o que é capturado pelos tropes. Esta relação sugere que



tropes podem ser explorados em recomendação nos mesmos contextos em que gêneros são, inclusive com a possibilidade de usar estas duas fontes de evidência em conjunto. Finalmente, é importante frisar que embora estes dados sugiram uma associação entre tropes e gêneros, há muitos tropes que são comuns a filmes independentes dos seus gêneros. Por exemplo, o trope “Oh, Crap!”, que indica o momento em que personagem percebe que algo ruim está prestes a acontecer<sup>3</sup>, ocorre em praticamente todos os gêneros.

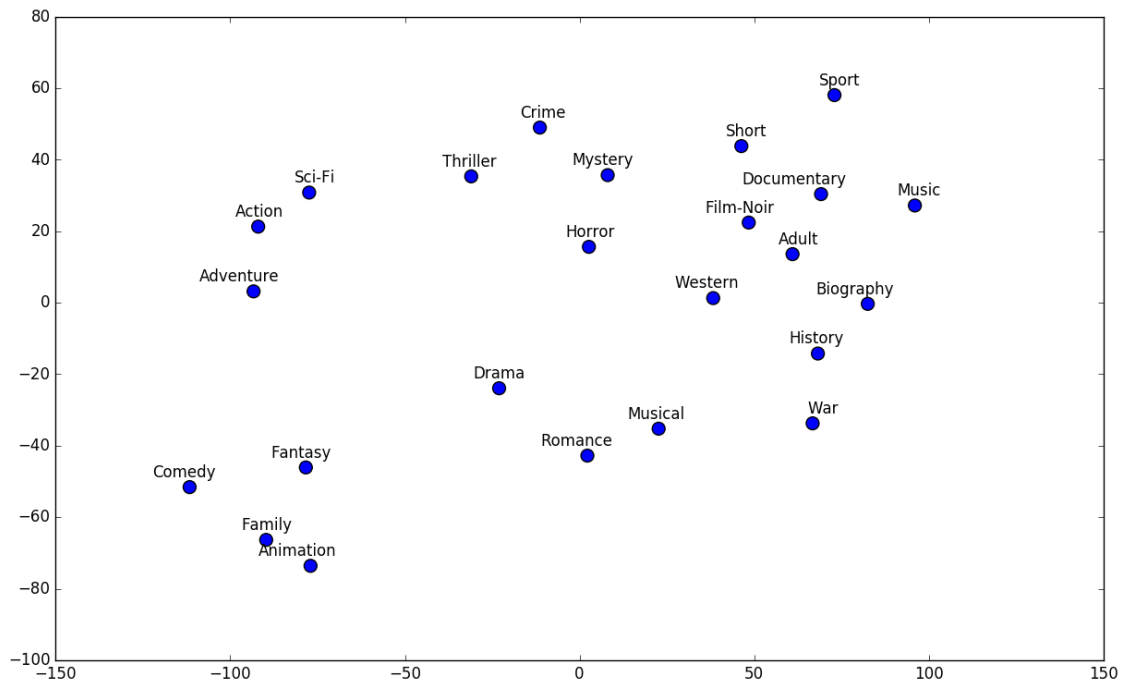


Figura 8 – Distribuição dos gêneros dos filmes em um espaço bidimensional composto por tropes, reduzidos através da ferramenta t-SNE, da coleção DBTropes  $\cap$  IMDB.

A Figura 9 apresenta a correlação entre a quantidade de tropes catalogados nos filmes e as popularidades e preferências destes filmes na coleção DBTropes x IMDB. Claramente, a popularidade dos filmes influencia a possibilidade de cadastro de tropes. Como observado na figura, os filmes que possuem mais votos no IMDB são os que apresentam maior quantidade de tropes em geral. Isso se explica pelo fato de que quanto mais popular um filme é, maior é a chance que membros da comunidade cataloguem tropes para ele. O mesmo fenômeno não é observado na preferência, uma vez que existe uma quantidade considerável de filmes com notas acima de 6, com baixo número de tropes.

<sup>3</sup> Mais especificamente, o personagem percebe algo ruim que é iminente ou que já aconteceu e já o deixou em sérios problemas. Neste momento, ele pode não dizer nada em particular que indique a situação para a audiência, mas esta fica clara na sua expressão facial.

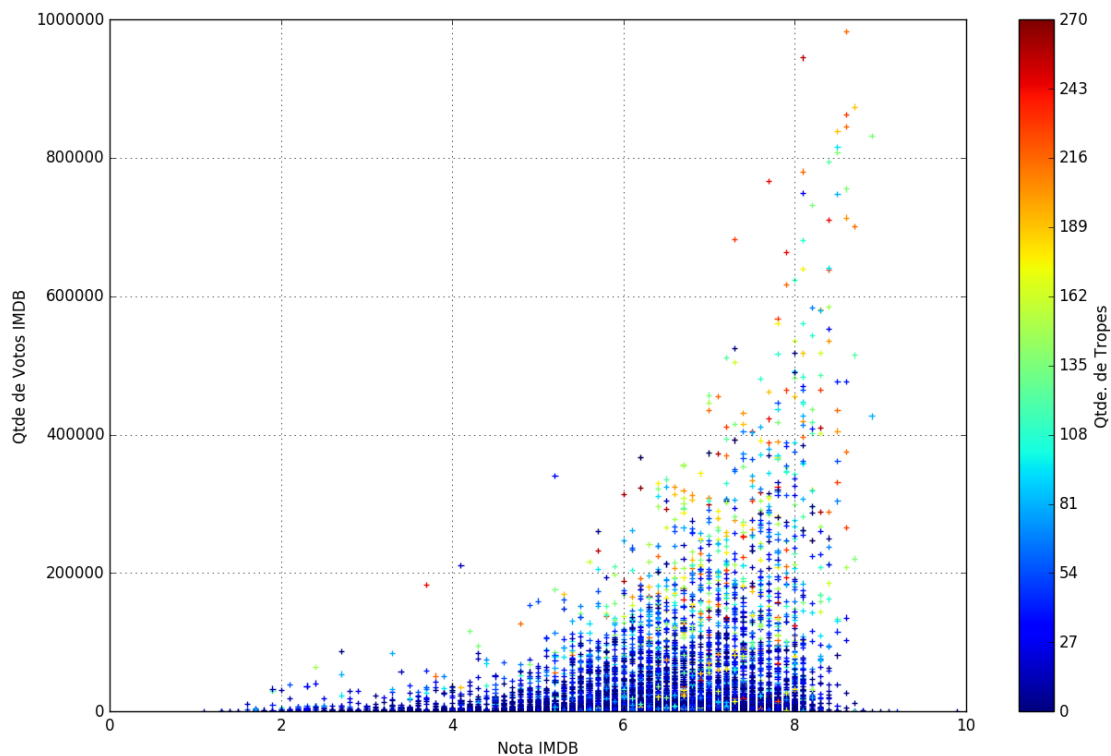


Figura 9 – Correlação da quantidade de tropes catalogados nos filmes com sua popularidade e preferência na coleção DBTropes  $\cap$  IMDB.

## 3.2 Padrões de ocorrência de tropes entre filmes similares

Para investigar o comportamento dos tropes no contexto de SR, fizemos um estudo inicial, de caráter qualitativo, com objetivo de analisar os padrões de ocorrência de tropes entre filmes com um certo grau de similaridade. Como não tínhamos a intenção de criar métodos para calcular essa similaridade, investigamos abordagens já existentes. Neste caso, selecionamos o site *Rotten Tomatoes*<sup>4</sup> que, através de uma *API*<sup>5</sup>, disponibiliza acesso ao seu conteúdo de críticas e avaliações de cinema. Uma das funcionalidades disponíveis consiste em listar filmes similares a um filme específico, ou seja, um filme-consulta. De posse dessa lista, extraímos todas as intersecções de tropes catalogados entre o filme-consulta e seus similares. As Tabelas 4 e 5 apresentam uma amostra do resultado obtido, considerando os filmes-consulta “Coração Valente” e “Toy Story”. Nestas tabelas são exibidos apenas os tropes catalogados para mais de três filmes da lista.

O primeiro ponto que observamos foi que a Tabela 5 (“Toy Story”) é maior e mais densa que a Tabela 4 (“Coração Valente”). Esse fenômeno pode ter diferentes causas como, por exemplo, os filmes mais similares a “Coração Valente” podem não ser tão similares quanto os relacionados a “Toy Story” do ponto de vista de tropes, ou possuem uma quantidade menor de tropes catalogados. Outro fator que pode ter influenciado pode

<sup>4</sup> <http://www.rottentomatoes.com>

<sup>5</sup> <http://developer.rottentomatoes.com>

Tropes	Coração Valente	Gladiador	O Patriota	300	Senhor dos Anéis	Robin Hood
Rousing Speech	x	x		x	x	x
The Determinator	x	x		x	x	
Rated M for Manly	x			x	x	
Badass Boast	x	x		x	x	
Historical Villain Upgrade	x	x	x	x		x
Historical Hero Upgrade	x	x	x	x		
Curb-Stomp Battle	x		x	x	x	

Tabela 4 – Amostra de tropes catalogados para o filme “Coração Valente” e filmes similares

Tropes	Toy Story	Vida de Inseto	Procurando o Nemo	Monstros S.A.	O Rei Leão	Up
Adult Fear	x		x	x	x	
Book Ends	x				x	x
Break Up Make Up Scenario	x	x	x	x	x	x
Chekhov’s Gun	x		x	x	x	x
Comically Missing the Point	x	x		x	x	
Darker and Edgier	x	x	x		x	
Deadpan Snarker	x		x	x	x	
Does This Remind You of Anything?	x	x			x	
Family Unfriendly Death	x	x			x	x
Foreshadowing	x		x	x	x	x
Funny Background Event	x		x		x	x
Getting Crap Past the Radar	x	x	x	x	x	x
Ironic Echo	x	x		x	x	x
Large Ham	x			x	x	
Meaningful Name	x	x	x		x	x
Oh Crap	x	x	x	x	x	x
Punctuated! For! Emphasis!	x	x		x	x	
Shout-Out	x	x	x	x	x	x
Stealth Pun	x		x	x		x
The Cameo	x	x	x			x
The Smurfette Principle	x		x	x	x	x
Those Two Guys	x		x	x	x	

Tabela 5 – Amostra de tropes catalogados para o filme “Toy Story” e filmes similares

estar relacionado ao gênero, uma vez que os filmes relacionados com “Toy Story” são todos animações, onde foi observado a maior média de tropes catalogados por filme (cf. seção 3.1.1).

Observamos também que o título de alguns tropes não traduz, em sua plenitude, seu valor semântico, tornando necessária a leitura de suas descrições para entender seu significado. Este é o caso, por exemplo, de “Oh, crap!” que constitui um título pouco intuitivo para o respectivo trope. Uma alternativa para mitigar este problema consiste em classificar os tropes conforme suas categorias, de forma a se caracterizar o contexto em que se aplicam. Esse tipo de informação pode ser utilizado, por exemplo, para explicar as recomendações geradas por algum método de previsão, que é uma das sub-áreas comumente estudadas no domínio de SR (BOBADILLA et al., 2013). A Tabela 6 apresenta a lista de categorias referentes aos tropes da Tabela 4 (“Coração Valente”).

<b>Tropes</b>	<b>Categorias</b>
Rousing Speech	Dialogue, Idealism, Speeches and Monologues
The Determinator	Bravery, Heroes, Heroic Spirit, Idealism, Obsession
Rated M for Manly	Action Adventure, Masculinity Tropes, Power Metal
Badass Boast	Word Power
Historical Villain Upgrade	Alternate History, Media Adaptation, Villains
Historical Hero Upgrade	Alternate History, Media Adaptation
Curb-Stomp Battle	Action Adventure, Combat, Fight Scene, Military and Warfare

Tabela 6 – Exemplo de Categorias de Tropes

Note que um problema em usar tropes para explicar recomendações é a natureza antecipatória que vários tropes têm em relação aos arcos narrativos. Por exemplo, tropes que indicam reviravolta, relacionamentos inesperados entre personagens ou repetição de cenas entre o início e o fim da história podem estragar a experiência do usuário ao antecipar elementos-chave da narrativa.

## 4 Proposta de um SR Híbrido Baseado em Tropes

Neste capítulo, apresentamos o sistema de recomendação híbrido que propomos para incorporar tropes em algoritmos de filtragem colaborativa. Inicialmente, é descrita a arquitetura geral do SR proposto. Em seguida, são detalhados os seus diversos componentes.

### 4.1 Arquitetura

O sistema de recomendação proposto é composto por três componentes principais: o agregador de bases de dados de entrada, o sub-sistema de pré-processamento e o recomendador. A integração entre estes componentes é ilustrada na Figura 10. Os detalhes dessa arquitetura são abordados nas seções seguintes.

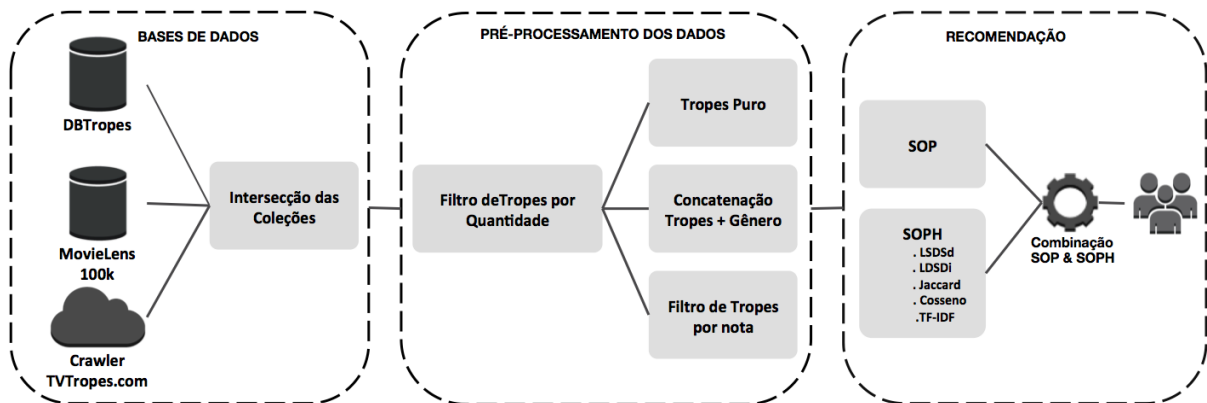


Figura 10 – Arquitetura do sistema de recomendação híbrido.

### 4.2 Construção de Base de Dados Integrada

Para a implementação do sistema de recomendação proposto é necessário uma base de dados que contenha dados referentes à preferência dos usuários, bem como relacionamento entre filmes e tropes. Por esse motivo, usamos a coleção MovieLens 100k<sup>1</sup> que contém as notas que os usuários deram para os filmes e é bastante popular em pesquisas no domínio de SRs. Também usamos a coleção DBTropes que provê informações sobre o catálogo de tropes.

<sup>1</sup> <http://grouplens.org/datasets/movielens/100k/>

Para integrar essas bases de dados, MovieLens e DBTropes, primeiro identificamos a interseção entre elas, de modo que possamos criar uma nova base de dados que contenha filmes avaliados com tropes catalogados. Para isso, foi utilizada a mesma estratégia adotada na Seção 3.1.1, onde os filmes foram comparados pelo título e ano de lançamento. Este processo resultou na base de dados usada como entrada para os algoritmos de filtragem colaborativa e baseado em conteúdo.

### 4.2.1 Crawler TVTropes

Um trope pode ocorrer mais de uma vez em uma mesma obra. Um exemplo disso é o trope *Bookends*<sup>2</sup> que possui duas ocorrências no filme *Matrix*, ou o trope “*Oh Crap!*” que é referenciado seis vezes em “*Star Wars: O império contra-ataca*”.

Apesar da coleção DBTropes fornecer dados sobre o catálogo dos tropes nos filmes, não é possível extrair a quantidade de vezes que o mesmo ocorre, pois tais ocorrências estão agrupadas em uma mesma estrutura. Desta forma, foi necessária a implementação de um *web crawler* para extrair este metadado do site TVTropes que é a fonte de informação utilizada para a construção do DBTropes.

A informação extraída nesta etapa é adicionada ao processo de construção da base de dados que será empregada nos experimentos para avaliar importância de um trope no contexto de um filme através do método TF-IDF.

## 4.3 Pré-processamento dos Dados

Após a construção da base de dados, o próximo passo consiste na aplicação de um pré-processamento nas informações coletadas com objetivo de explorar a influência dos tropes na recomendação de filmes sob diferentes aspectos.

### 4.3.1 Filtro de Tropes por Quantidade

Neste filtro são selecionados apenas os filmes que possuem uma quantidade mínima de tropes pré-definida. Por exemplo, se o filme  $F_1$  possui 7 tropes catalogados, enquanto  $F_2$  contém 10 e aplicarmos um filtro com quantidade mínima de tropes igual a 10 em cada filme, apenas os tropes de  $F_2$  serão adicionados à recomendação. Assim, as recomendações relacionadas ao filme  $F_1$  serão baseadas puramente na filtragem colaborativa.

A motivação por trás deste filtro consiste em investigar se filmes com poucos tropes pioram a qualidade da recomendação, dado que este número impacta diretamente os

<sup>2</sup> Repetição de uma cena no início e no fim da narrativa, frequentemente para mostrar como as coisas mudaram, no curso da história, ou demonstrar que nada mudou, nem mudará, de fato. É um dispositivo comumente usado para demonstrar o desenvolvimento (ou não) de um personagem.

métodos de similaridade. Este fenômeno pode ser exemplificado da seguinte forma: os filmes  $F_3$  e  $F_4$  possuem apenas o trope  $T_1$  catalogado. Neste caso, para o algoritmo baseado em distância de *Jaccard*, a similaridade entre eles será igual 100%. Se os usuários avaliam estes filmes como de preferências distintas, esta medida de similaridade pode influenciar negativamente a recomendação.

Como observado empiricamente, a consequência deste filtro será uma tendência do filmes terem suas recomendações baseadas somente na filtragem colaborativa, já que a grande maioria tem poucos tropes catalogados (cf. Seção 3.1).

### 4.3.2 Apenas Tropes

Essa é a estratégia mais simples para incorporar tropes na tarefa de recomendação implementada em nossa pesquisa. Nela, os tropes são utilizados em sua forma original e não é aplicado nenhum tipo de seleção além do filtro de quantidade executado na etapa anterior. Desta forma, quando a quantidade mínima de tropes for igual a 1, podemos considerar este cenário como sendo a maneira mais pura de empregar tropes.

Nosso objetivo é investigar se utilizar tropes sem nenhum tipo de pré-processamento já traz ganho para a qualidade das previsões.

### 4.3.3 Combinação entre Tropes e Gêneros

Tropes podem ser utilizados em filmes de diferentes gêneros, ou seja, podemos observar um mesmo trope sendo aplicado a um filme de terror ou uma animação. Caso uma pessoa prefira histórias de terror a infantis, uma recomendação baseada em tropes poderá sofrer influência negativa. Neste cenário, faz-se necessário avaliar os efeitos do relacionamento entre gêneros e tropes de um filme para recomendação.

A abordagem usada consiste na concatenação dos tropes de um filme com seu gênero, criando uma dependência entre eles. Essa solução pode ser ilustrada da seguinte forma: o trope  $T_1$ , utilizado em um filme de gênero  $G_1$ , será reescrito como  $G_1-T_1$ ; se outro filme de gênero  $G_2$  fizer referência para o trope  $T_1$ , ambos serão distintos já que o segundo será descrito com  $G_2-T_1$ .

A informação sobre os gêneros será extraída da coleção MovieLens. Nela, um filme pode estar relacionado a 1 ou mais gêneros. Um exemplo disso é caso do filme “O Poderoso Chefão” que é classificado com ação, crime e drama.

### 4.3.4 Filtro de Tropes por Nota

Por fim, o outro aspecto investigado em nosso estudo consiste em avaliar se a importância de um trope está relacionada à preferência do usuário. Essa importância é

derivada das notas dadas aos filmes na qual um trope é referenciado, sendo formalmente definida como:

$$\bar{r}_t = \frac{1}{|F_t|} \sum_{f \in F_t} \left( \frac{1}{|R_f|} \sum_{r \in R_f} r \right), \quad (4.1)$$

onde  $F_t$  representa o conjunto de filmes em que o trope  $t$  foi referenciado e  $R_f$  são as notas dadas ao filme  $f$ . Portanto, a nota de um trope, denotada por  $\bar{r}_t$ , é igual a média das notas médias dos filmes em que o trope foi catalogado.

Após o cálculo da nota de cada trope, um filtro pode ser aplicado para selecionar aqueles que atinjam determinado limiar. Neste trabalho em específico, foram selecionados apenas o tropes com nota acima da média. Como as notas da base de dados MovieLens se encontram em uma escala de 1 a 5, filtramos apenas aqueles com avaliação superior a 3.

A figura 11 exhibe a distribuição das notas dos tropes de acordo com a nossa abordagem, onde é possível observar que a grande maioria é superior a 3. Tal comportamento dá indícios que essa estratégia pode ser similar ao método que utiliza todos os tropes, já que poucos tropes serão podados na aplicação deste filtro.

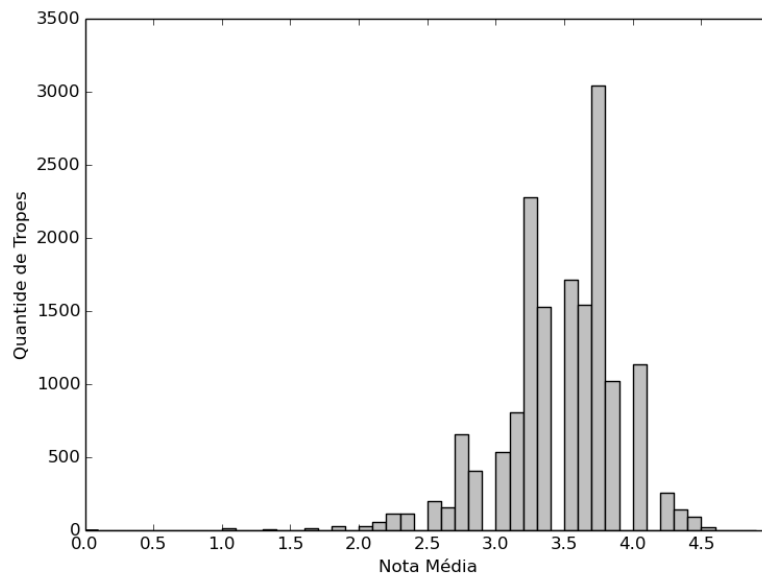


Figura 11 – Distribuição da nota dos tropes derivada dos filmes.

## 4.4 Componente de Recomendação

Nesta seção detalhamos o método de recomendação híbrida utilizado neste trabalho, o SOPH, proposto por Yang, Hu e Qu (2013). Este irá receber como entrada as avaliações



dos usuários, bem como os vetores de características de cada filme compostos pelos tropes pré-processados na etapa anterior.

O SOPH serve como referência para nossa pesquisa por ser, até onde sabemos, o único a utilizar tropes em sistemas de recomendação. Apesar disso, sua abordagem é mais voltada para a área de *Web Semântica*, pois utiliza as medidas de similaridade  $LDSD_d$  e  $LDSD_i$  que são frequentemente empregadas para mensurar a distância entre documentos *Web*.

Embora tenha obtido melhor resultado na pesquisa de Yang, Hu e Qu (2013), a medida  $LDSD_d$  não emprega tropes para calcular a similaridade entre os filmes, pois é baseada somente nas ligações diretas entre os filmes (ou seja, quando a descrição de um filme possui um link para a descrição de outro). Já a medida  $LDSD_i$ , baseia-se nas ligações indiretas dos filmes que podem ser representadas através de tropes. Note, contudo, que no trabalho de Yang, Hu e Qu (2013), esta métrica não considera apenas tropes, pois adiciona outros tipos de informação como relacionamento por gênero, elenco, diretor, etc.

Nossa proposta é baseada no método SOPH, porém usa apenas tropes que são modelados como vetores de características dos filmes, substituindo as estratégias de distância semântica utilizadas por Yang, Hu e Qu (2013) para calcular a similaridade entre dois itens. Em seguida, investigamos a similaridade entre os vetores implementando as métricas coeficiente de jaccard e similaridade do cosseno. Nesta última medida, em específico, testamos a versão binária onde cada posição de um vetor indica a presença ou ausência de um determinado trope; e a versão ponderada através do TF-IDF onde cada posição contém o grau de relevância de um trope para o filme. A versão ponderada será referenciada apenas como TF-IDF no restante desta pesquisa.

O coeficiente de jaccard e a similaridade do cosseno foram escolhidos por serem medidas de conceito simples e popularmente utilizadas na literatura para medir a similaridade entre vetores. Já o TF-IDF tem como objetivo explorar a importância dos tropes no contexto de um filme baseado nos conceitos de frequência e raridade, onde os tropes que ocorrem mais vezes em um filme, mas raramente no restante da coleção, tendem a ser mais relevantes para este. Desta forma esperamos, por exemplo, desenfaturar tropes usados indistintamente em filmes como, por exemplo, “Oh, crap!”. Nesta medida, em específico, serão usadas as informações coletadas pelo *web crawler* citado na Seção 4.2.1.

## 5 Experimentos

Neste capítulo são descritos os experimentos realizados neste trabalho para estimar a influência dos tropes na recomendação de filmes. Mais especificamente, testamos diferentes configurações entre os métodos híbridos e as medidas de similaridade especificadas na Seção 4.4, tendo como entrada as coleções geradas pelos filtros da Seção 4.3.

### 5.1 Base de Dados

Como citado na Seção 4.2, a coleção utilizada para realização dos experimentos foi criada através da interseção entre as coleções *DBTropes* e *MovieLens*.

A base de dados *MovieLens 100k* possui 943 usuários, 1.682 filmes e 100 mil avaliações em uma escala de 1 a 5. Fornece também o ano de lançamento e o gênero dos filmes que serão necessários para a construção da coleção utilizada nos experimentos. Além disso, é possível encontrar dados demográficos dos usuários como idade, ocupação, gênero, etc. Esta informação, porém, será ignorada por não fazer parte do escopo desta pesquisa (MOVIELENS, 2015). Já os detalhes da coleção *DBTropes* foram previamente abordados na Seção 3.1.

Semelhante à estratégia adotada na Seção 3.1.1, filtramos a interseção entre as bases de dados *DBTropes* e *MovieLens*. Esta filtragem resultou em uma coleção com 943 usuários, 845 filmes e 80.339 mil avaliações. Isso significa que apenas 50,24% dos filmes providos pelo *MovieLens* possuem um equivalente no *DBTropes*. Deste modo, a distribuição desta coleção com relação à quantidade de tropes catalogados por filme e à quantidade de itens avaliados por usuário pode ser observada nas Tabelas 7 e 8, respectivamente.

Quantidade de Tropes por Filme	Valor
Média	76.26
Desvio Padrão	93.53
Mínima	1
25%	19
50%	43
75%	95
Máxima	751

Tabela 7 – Estatísticas sobre a quantidade de tropes catalogados nos filmes da coleção *MovieLens 100k*  $\cap$  *DBTropes*.

Quantidade de Avaliação por Usuário	Valor
Média	85.20
Desvio Padrão	79.39
Mínima	10
25%	26
50%	52
75%	120
Máxima	475

Tabela 8 – Estatísticas sobre a quantidade de notas por usuário da coleção *MovieLens 100k*  $\cap$  *DBTropes*.

## 5.2 Avaliação

As métricas utilizadas para medir a qualidade das recomendações foram RMSE, Precisão, Revocação e F1. Estas foram primeiro estimadas individualmente para cada usuário e, então, reduzidas a um único valor referente a todos os usuários. Estas medidas foram previamente detalhadas na Seção 2.2.5.

Para avaliar a capacidade de generalização dos modelos, aplicamos o método de validação cruzada de *5-partições*<sup>1</sup>. Para cada rodada de experimentação, 80% da base de dados foi utilizada para o treinamento dos modelos e 20% para teste. Como as medidas foram calculadas por usuário, os resultados apresentados para cada métrica deste capítulo consistem da média de todos os usuários em todas as partições.

Todas as comparações foram verificadas quanto à sua significância estatística utilizando o teste de Wilcoxon. Com exceção dos casos onde é explicitamente informado o contrário, todas os resultados reportadas apresentam grau de confiança maior que 95%.

## 5.3 Protocolo Experimental

As diferentes configurações entre os pré-processamentos aplicados na base de dados (Tropes Puro, Tropes + Gênero e Filtro de Tropes por Nota) e as medidas de similaridade utilizadas no algoritmo SOPH (Jaccard, cosseno e TF-IDF), resultaram em uma gama de experimentos. Como nosso objetivo é analisar a influência dos tropes nos SRs, o que estamos avaliando, de fato, é o desempenho das abordagens híbridas que incorporam tropes comparadas ao desempenho das respectivas formas originais baseadas somente em filtragem colaborativa (SOP). Contudo, as abordagens de distância semântica  $LDSD_d$ ,  $LDSD_i$  implementadas no trabalho de Yang, Hu e Qu (2013), foram igualmente usadas como referência, por também terem usado informações provenientes do TVTropes.

<sup>1</sup> Foram utilizadas 5 partições por ser o padrão adotado na coleção MovieLens 100k, que fornece a base de dados já particionada com essa configuração, além da versão completa.

Adicionalmente, foi implementado um método baseado somente nos gêneros dos filmes para comparar com a estratégia “Tropes + Gênero”. Deste modo, podemos analisar o impacto específico do gênero nos resultados obtidos com esta técnica. Nesse caso, em particular, reportamos apenas os valores obtidos pelo cosseno, por terem apresentado melhor desempenho que as demais medidas implementadas, como Jaccard.

Os experimentos foram conduzidos utilizando a transformação linear que combina os algoritmos SOP e SOPH, descrita na seção 2.2.3. Nessa abordagem, é possível ajustar a proporção que cada algoritmo terá sobre a recomendação através do parâmetro  $\alpha$  que pode variar entre 0 a 1. O valor 0 significa que a recomendação é baseada puramente na filtragem colaborativa, enquanto o valor 1 representa uma recomendação totalmente híbrida. Investigamos o valor de  $\alpha$  crescendo a uma taxa de 0,05.

## 5.4 Resultados

Nesta seção são apresentados os resultados obtidos para os experimentos executados neste estudo. Estes resultados estão agrupados pelas métricas de avaliação. Logo, temos uma seção correspondendo à medida RMSE e outra, às medidas precisão, revocação e F1.

Antes de discutir sobre os resultados, é importante ressaltar que o método SOPH-LDSDd não utiliza tropes como evidência na recomendação. Reportamos seus resultados por este usar informações extraídas da coleção DBTropes e ter apresentado o melhor desempenho no trabalho de Yang, Hu e Qu (2013). Já o método SOPH-LDSDi é o que mais se aproxima das nossas estratégias, uma vez que também incorpora tropes. No entanto, utiliza tropes em conjunto com outras informações e nunca de forma isolada. Portanto, não é aplicado nenhum dos filtros implementados nessa pesquisa em ambas as abordagens variando apenas o valor do parâmetro  $\alpha$ .

Outro ponto a ser destacado é que com exceção do SOP, para todos outros modelos foi testado o valor de  $\alpha$  variando de 0,05 a 1. Isso explica razão da reta que representa o SOP ter apresentado um comportamento constante em todos gráficos dessa seção.

### 5.4.1 RMSE

A Figura 12 compara o desempenho dos métodos SOPH-Jaccard, SOPH-Cosseno e SOPH-TFIDF, baseadas puramente em tropes, tendo como referências os modelos SOP, SOPH-LDSDd e SOPH-LDSDi. Cada gráfico apresenta os resultados relacionados ao filtro de quantidade mínima de tropes variando entre 1, 30, 60 e 90, aplicado apenas nos métodos propostos nesta pesquisa.

Perceba que existe uma tendência dos métodos SOPH-LDSDd, SOPH-Jaccard, SOPH-Cosseno e SOPH-TFIDF, apresentarem menor erro de previsão quando a recomen-

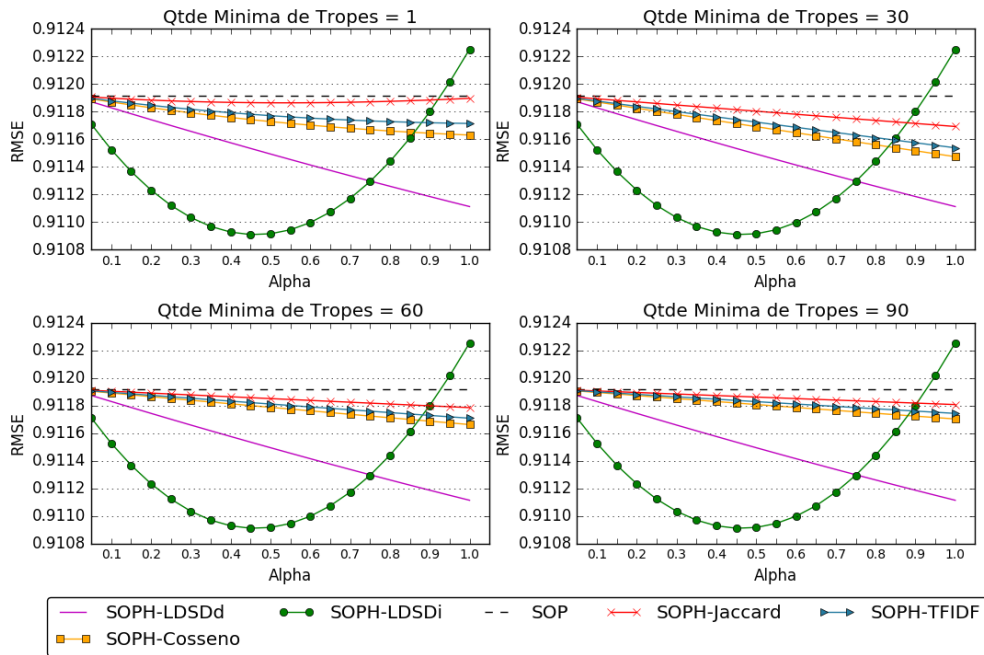


Figura 12 – RMSE variando-se o valor de  $\alpha$  e quantidade mínima de tropes com o filtro baseado nos tropes de forma pura.

dação é feita de forma totalmente híbrida ( $\alpha$  cada vez maior). No caso do SOPH-LSDi, o melhor resultado foi observado quando 55% da recomendação foi gerada baseada na filtragem colaborativa, tendendo a piorar quando a abordagem híbrida passa a ter maior influência. Embora o SOPH-LSDi use tropes, ele é muito pouco influenciado pela variação no número mínimo de tropes considerados (1, 30, 60 e 90). Entre os demais métodos, o que é mais impactado é o de Jaccard, especialmente quando comparamos mínimos de 1 e de 30 tropes.

A Figura 13 ilustra o comportamento da estratégia que utiliza como evidência a combinação de tropes com os gêneros dos filmes. Nestes gráficos destacamos o desempenho dos algoritmos SOPH-Cosseno e SOPH-TFIDF que apresentaram um ganho mais acentuado se comparados com a abordagem baseada somente em tropes. Apesar disso, os métodos apresentam tendências diferentes. No SOPH-Cosseno, o erro tem o crescimento mais intenso à medida que a previsão passa a ser dominada pela modelo híbrido, diferente do SOPH-TFIDF, onde o erro cresce em taxas menores. No fim, o SOPH-TFIDF é levemente melhor, o que indica que vale a pena penalizar tropes muito populares (que é o efeito do IDF sobre tropes como “Oh, crap!”, por exemplo).

A Tabela 9 sumariza os melhores resultados de cada experimento. Em todos os casos, a diferença para o desempenho do SOP foi estatisticamente significativa, o que é representado por um “\*”. Apesar disso, o ganho no desempenho da estratégia baseada somente em tropes é baixo, já que o maior valor corresponde apenas a 0,05%. Este ganho foi obtido pelos métodos SOPH-Cosseno e SOPH-TFIDF com a quantidade mínima de

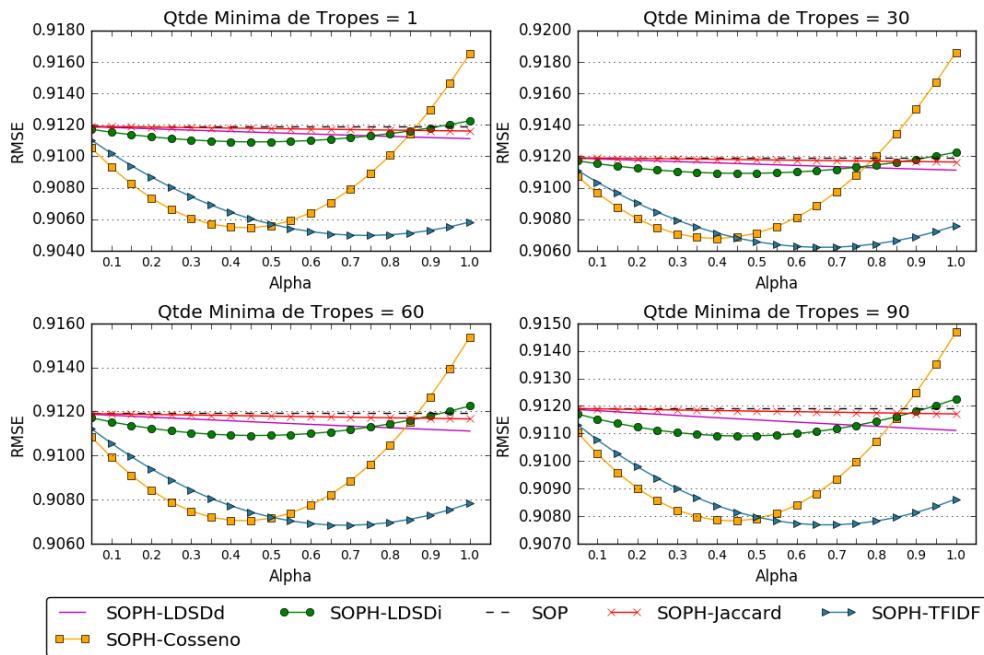


Figura 13 – RMSE variando-se o valor de  $\alpha$  e quantidade mínima de tropes com o filtro baseado na combinação de tropes e gêneros.

tropes igual a 30. Já entre os métodos de referência, o melhor resultado foi obtido pelo SOPH-LSDi que teve um ganho de 0,11% sobre o *baseline*.

Analisando a Tabela 9, percebe-se que o melhor resultado (menor RMSE) ocorre no SOPH-TFIDF quando a influência do modelo híbrido sobre a recomendação é de 75% ( $\alpha$ ) e a quantidade mínima de tropes igual a 1. Neste caso, o ganho chega a ser de 0,76%. Para investigar o quanto este ganho está dominado pelo gênero, é necessário observamos o desempenho da estratégia SOPH-Cos-Gênero, onde apenas o gênero é considerado. Neste caso, o ganho foi de 0,45% em comparação ao SOP, o que indica que a combinação entre tropes e gênero produz um ganho genuíno sobre o uso isolado dessas evidências.

A Figura 14 apresenta o erro ao variar o valor de  $\alpha$  e a quantidade mínima de tropes quando são considerados apenas os tropes mais bem avaliados pelos usuários, ou seja, quando é aplicado o filtro de tropes por nota. O desempenho deste filtro apresentou um comportamento semelhante à estratégia que utiliza todos os tropes isoladamente. Isso pode ser observado também na Tabela 9. Esse fato sugere que não há vantagem em usar todos os tropes ou filtrá-los de acordo com a sua nota, levando em consideração a metodologia adotada nesta pesquisa, para o conjunto de dados utilizado.

#### 5.4.2 Precisão, Revocação e F1

Os métodos implementados neste trabalho também foram avaliados com relação as métricas de precisão, revocação e F1, seguindo os mesmos procedimentos adotados na seção anterior. Apesar destas medidas serem correlacionadas, a precisão é de grande

Algoritmo	Pré-Processamento	Qtde. Tropes Min	Alfa	RMSE	Ganho (%)	
<b>SOP (Baseline)</b>	-		-	<b>0.9119</b>	-	
SOPH-LDSDd			1.0	0.9111	*0.09	
<b>SOPH-LDSDi</b>			<b>0.45</b>	<b>0.9109</b>	<b>*0.11</b>	
<b>SOPH-Cos-Gênero</b>			<b>0.2</b>	<b>0.9077</b>	<b>*0.46</b>	
SOPH-Jaccard	Tropes	1	0.55	0.9119	*0.0	
SOPH-Cosseno			1.0	0.9116	*0.03	
SOPH-TFIDF			1.0	0.9117	*0.02	
SOPH-Jaccard		30	1.0	0.9117	*0.02	
<b>SOPH-Cosseno</b>			<b>1.0</b>	<b>0.9115</b>	<b>*0.05</b>	
<b>SOPH-TFIDF</b>			<b>1.0</b>	<b>0.9115</b>	<b>*0.05</b>	
SOPH-Jaccard		60	1.0	0.9118	*0.01	
SOPH-Cosseno			1.0	0.9117	*0.02	
SOPH-TFIDF			1.0	0.9117	*0.02	
SOPH-Jaccard		90	1.0	0.9118	*0.01	
SOPH-Cosseno			1.0	0.9117	*0.02	
SOPH-TFIDF			1.0	0.9117	*0.02	
SOPH-Jaccard		Tropes por Gênero	1	1.0	0.9116	*0.03
SOPH-Cosseno				0.45	0.9055	*0.70
<b>SOPH-TFIDF</b>	<b>0.75</b>			<b>0.9050</b>	<b>*0.76</b>	
SOPH-Jaccard	30		1.0	0.9116	*0.03	
SOPH-Cosseno			0.4	0.9068	*0.56	
SOPH-TFIDF			0.65	0.9062	*0.63	
SOPH-Jaccard	60		1.0	0.9117	*0.02	
SOPH-Cosseno			0.45	0.9070	*0.54	
SOPH-TFIDF			0.7	0.9068	*0.56	
SOPH-Jaccard	90		1.0	0.9117	*0.02	
SOPH-Cosseno			0.45	0.9078	*0.45	
SOPH-TFIDF			0.7	0.9077	*0.46	
SOPH-Jaccard	Tropes por Nota		1	1.0	0.9117	*0.02
<b>SOPH-Cosseno</b>				<b>1.0</b>	<b>0.9114</b>	<b>*0.06</b>
SOPH-TFIDF		1.0		0.9115	*0.05	
SOPH-Jaccard		30	1.0	0.9117	*0.02	
SOPH-Cosseno			1.0	0.9115	*0.05	
SOPH-TFIDF			1.0	0.9115	*0.05	
SOPH-Jaccard		60	1.0	0.9118	*0.01	
SOPH-Cosseno			1.0	0.9117	*0.02	
SOPH-TFIDF			1.0	0.9117	*0.02	
SOPH-Jaccard		90	1.0	0.9118	*0.01	
SOPH-Cosseno			1.0	0.9117	*0.02	
SOPH-TFIDF			1.0	0.9117	*0.02	

Tabela 9 – Melhores resultados de obtidos para a métrica RMSE em cada abordagem implementada. Os ganhos precedidos por “\*” indicam que houve diferença estatística em comparação com o método de referência. As linhas em destaque denotam os maior ganhos entre todas as abordagens.

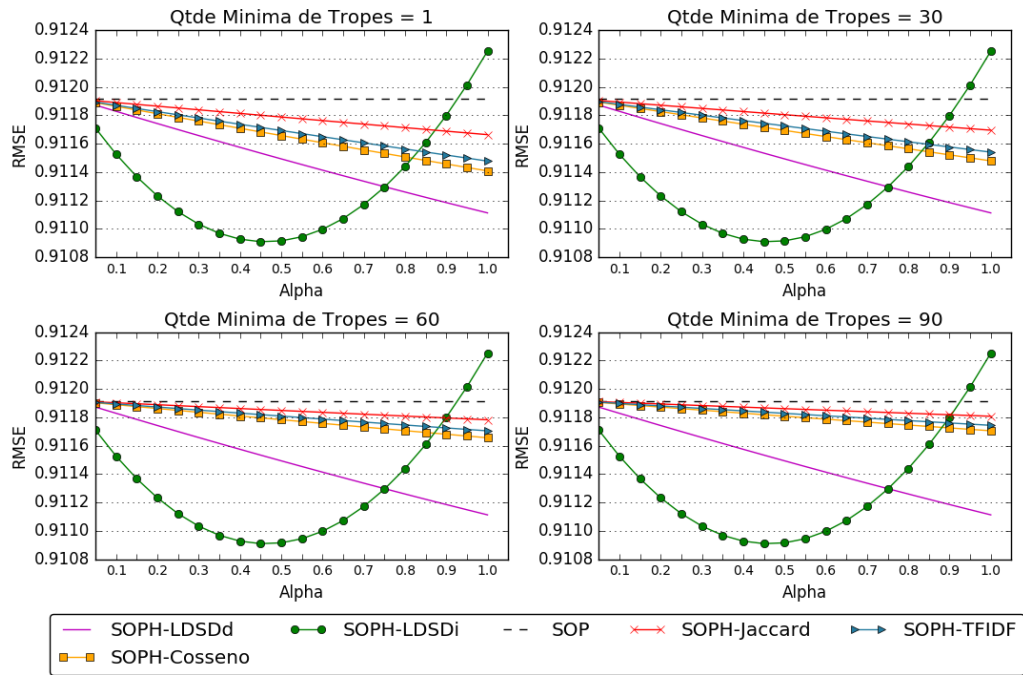


Figura 14 – RMSE variando-se o valor de  $\alpha$  e quantidade mínima de tropes com o filtro baseado nos tropes mais bem avaliados pelos usuários.

importância na área de SRs, uma vez que tem impacto direto sobre a confiança que o usuário deposita no sistema, que ele intuitivamente espera que produza poucos falsos positivos. Desta forma, é mais importante prover aos usuários uma lista pequena com itens de sua preferência (mesmo que não sejam fornecidos todos os itens relevantes), do que uma lista grande contendo tanto itens relevantes quanto irrelevantes (GEMMIS et al., 2008). Portanto, a análise dos resultados apresentados nesta seção foca mais na medida de precisão.

A Figura 15 apresenta a precisão ao se variar o valor de  $\alpha$  e a quantidade mínima de tropes, considerando o filtro baseado nos tropes de forma pura. Na figura, é possível observar que as estratégias puramente baseadas em tropes, bem como o método SOPH-LDSDd, apresentaram um comportamento semelhante ao *baseline*. O único método com comportamento distinto dos demais foi o SOPH-LDSDi onde é observado uma maior precisão para todos os valores de alfa.

A Figura 16 apresenta a precisão obtida com o filtro baseado na combinação de tropes e gêneros. Assim como o observado para a métrica RMSE, a combinação entre tropes e gêneros favorece os métodos SOPH-Cosseno e SOPH-TFIDF. Neste caso, em particular, foi observado o ganho mais expressivo de 2,93% obtido pelo método SOPH-TFIDF, utilizando a abordagem 100% híbrida ( $\alpha = 1$ ), com quantidade mínima de tropes igual a 30. Comparando com a abordagem que emprega somente os gêneros dos filmes, nossa estratégia também apresentou um melhor desempenho, dado que a precisão obtida pelo SOPH-Cos-Gênero não apresentou diferença estatística em comparação com o SOP,



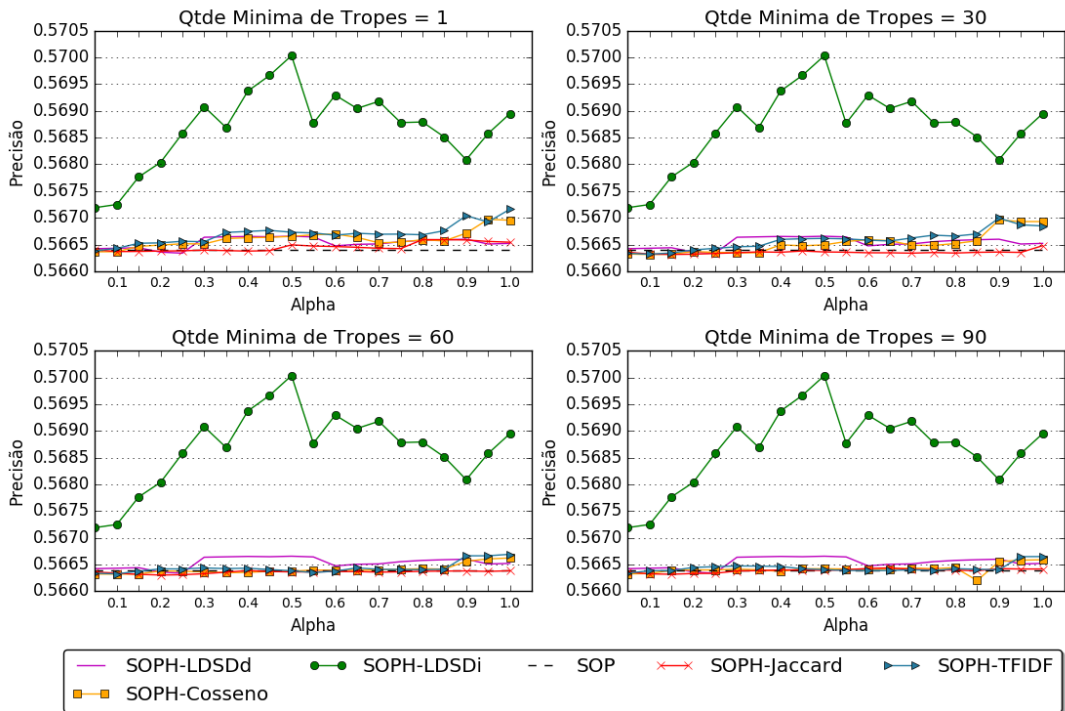


Figura 15 – Precisão variando-se o valor de  $\alpha$  e quantidade mínima de tropes com o filtro baseado nos tropes de forma pura.

apesar do ganho nominal ser de 2,03%. Esse fato reforça os benefícios da combinação dos tropes com os gêneros.

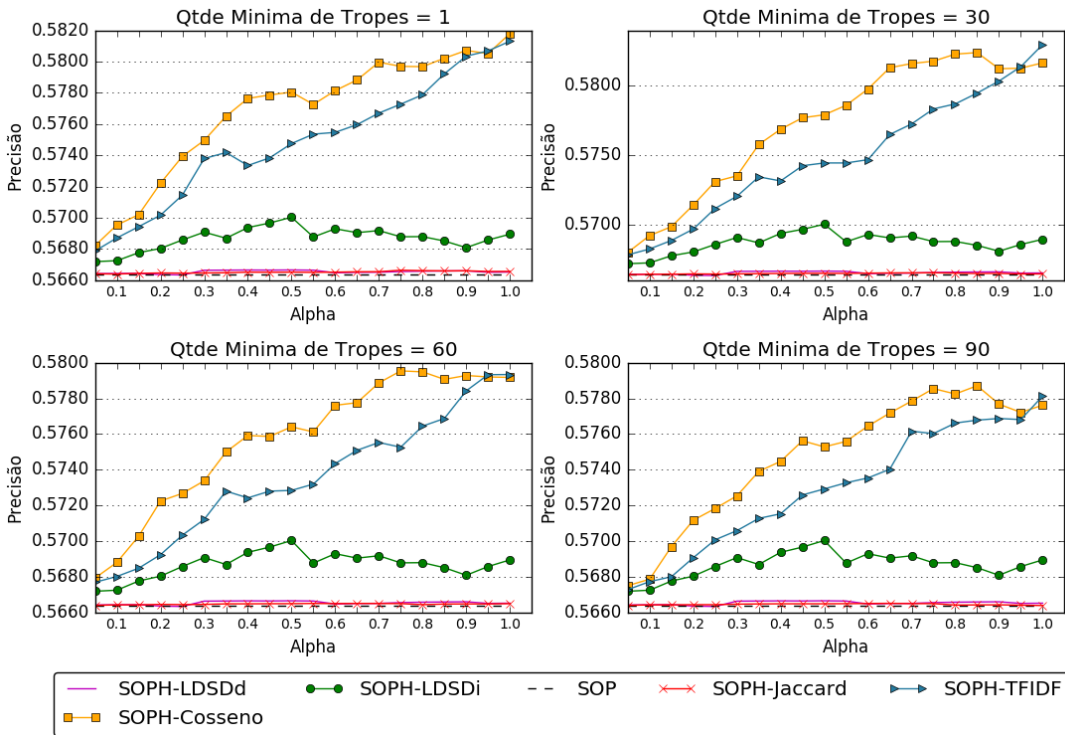


Figura 16 – Precisão variando-se o valor de  $\alpha$  e quantidade mínima de tropes com o filtro baseado na combinação de tropes e gêneros.

Por fim, a Figura 17 apresenta a precisão quando considerando o filtro de tropes por nota. Como já observado, o comportamento desse filtro com relação à precisão é bastante similar à estratégia que utiliza todos os tropes, onde também não houve diferença estatística para o SOP.

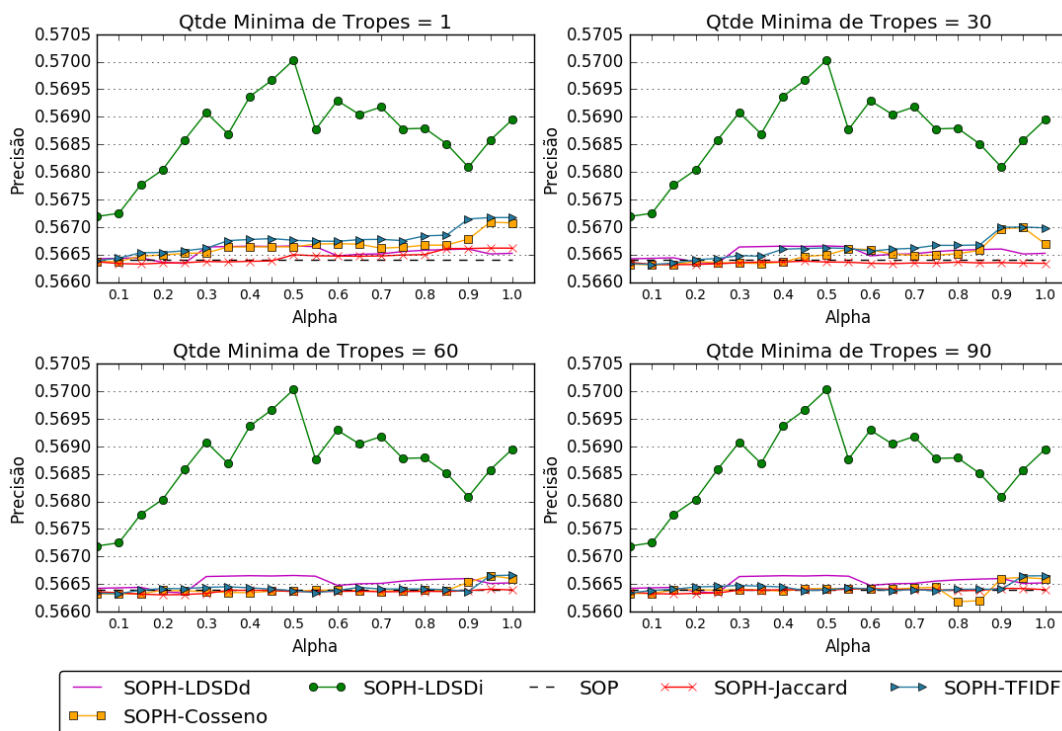


Figura 17 – Precisão variando-se o valor de  $\alpha$  e quantidade mínima de tropes com o filtro baseado nos tropes mais bem avaliados pelos usuários.

Os resultados com precisão, revocação e F1 são sumarizados na Tabela 10, onde os ganhos precedidos por “\*” indicam uma diferença estatística em comparação com SOP. Como mencionado no início desta seção, a precisão tem maior importância no campo de SRs, se comparada com a revocação e F1. A correlação entre essas métricas faz com que a priorização de uma implique no detrimento das demais. Desta forma, os melhores valores de precisão apresentados na Tabela 10 não correspondem necessariamente aos melhores valores de revocação e F1. Como observado nesta tabela, o ganho obtido pela estratégia baseada no método SOPH-TFIDF, a que apresentou a maior precisão, foi de 1,55% para a revocação e 2,4% em termos de F1. Apesar disso, o melhor desempenho para essas métricas foi obtido pela abordagem SOPH-Cosseno cujo o ganho da revocação e F1 foram de 1,81% e 2,95% respectivamente.

## 5.5 Discussão

Como observado, o uso de tropes isoladamente não apresentou uma vantagem expressiva quando adicionado no algoritmo SOPH, dado que na métrica RMSE o maior

Algoritmo	Filtro	Qtde. Tropes Min	Alfa	Prec.	Ganho (%)	Revoc.	Ganho (%)	F1	Ganho (%)
<b>SOP (Baseline)</b>	-		-	<b>0.5664</b>	-	<b>0.5907</b>	-	<b>0.5374</b>	-
SOPH-LDSDd			0.5	0.5667	0.05	0.5914	*0.11	0.5379	*0.09
<b>SOPH-LDSDi</b>			<b>0.5</b>	<b>0.5700</b>	<b>*0.64</b>	<b>0.5920</b>	<b>0.21</b>	<b>0.5390</b>	<b>*0.29</b>
<b>SOPH-Cos-Gênero</b>			<b>0.6</b>	<b>0.5779</b>	<b>2.03</b>	<b>0.6011</b>	<b>*1.76</b>	<b>0.5518</b>	<b>*2.67</b>
SOPH-Jaccard	Tropes	1	0.8	0.5666	0.04	0.5908	0.01	0.5376	0.03
SOPH-Cosseno			0.95	0.5670	0.11	0.5913	0.1	0.5380	0.11
SOPH-TFIDF			1.0	0.5672	0.14	0.5912	0.08	0.5379	0.09
SOPH-Jaccard		30	1.0	0.5665	0.02	0.5909	0.03	0.5376	0.03
SOPH-Cosseno			0.9	0.5670	0.11	0.5912	0.08	0.5379	0.09
SOPH-TFIDF			0.9	0.5670	0.11	0.5910	0.05	0.5378	0.07
SOPH-Jaccard		60	0.4	0.5664	0.0	0.5909	0.03	0.5375	0.01
SOPH-Cosseno			1.0	0.5666	0.04	0.5910	0.05	0.5376	0.03
SOPH-TFIDF			1.0	0.5667	0.05	0.5908	0.01	0.5375	0.01
SOPH-Jaccard		90	0.65	0.5664	0.0	0.5909	0.03	0.5376	0.03
SOPH-Cosseno			1.0	0.5666	0.04	0.5909	0.03	0.5375	0.01
SOPH-TFIDF			1.0	0.5666	0.04	0.5907	-0.01	0.5374	-0.01
SOPH-Jaccard	Tropes por Gênero	1	0.75	0.5666	*0.04	0.5912	*0.08	0.5378	*0.07
<b>SOPH-Cosseno</b>			<b>1.0</b>	<b>0.5818</b>	<b>*2.72</b>	<b>0.6014</b>	<b>*1.81</b>	<b>0.5533</b>	<b>*2.95</b>
SOPH-TFIDF			1.0	0.5813	*2.63	0.6008	*1.7	0.5507	*2.47
SOPH-Jaccard		30	0.75	0.5665	*0.02	0.5910	*0.05	0.5377	*0.05
SOPH-Cosseno			0.85	0.5824	*2.83	0.6012	*1.77	0.5520	*2.71
<b>SOPH-TFIDF</b>			<b>1.0</b>	<b>0.5830</b>	<b>*2.93</b>	<b>0.5999</b>	<b>*1.55</b>	<b>0.5503</b>	<b>*2.4</b>
SOPH-Jaccard		60	0.65	0.5665	0.02	0.5910	*0.05	0.5377	*0.05
SOPH-Cosseno			0.75	0.5795	*2.31	0.5971	*1.08	0.5471	*1.8
SOPH-TFIDF			1.0	0.5793	*2.28	0.5975	*1.15	0.5470	*1.78
SOPH-Jaccard		90	0.65	0.5665	*0.02	0.5910	*0.05	0.5377	*0.05
SOPH-Cosseno			0.85	0.5787	*2.17	0.5960	*0.89	0.5456	*1.52
SOPH-TFIDF			1.0	0.5781	*2.07	0.5962	*0.93	0.5452	*1.45
SOPH-Jaccard	Tropes por Nota	1	0.95	0.5666	0.04	0.5909	0.03	0.5377	0.05
SOPH-Cosseno			0.95	0.5671	0.12	0.5913	0.1	0.5380	0.11
SOPH-TFIDF			0.95	0.5672	0.14	0.5911	0.06	0.5378	0.07
SOPH-Jaccard		30	0.45	0.5664	0.0	0.5909	*0.03	0.5376	0.03
SOPH-Cosseno			0.95	0.5670	0.11	0.5911	0.06	0.5379	0.09
SOPH-TFIDF			0.95	0.5670	0.11	0.5909	0.03	0.5377	0.05
SOPH-Jaccard		60	0.95	0.5664	0.0	0.5909	0.03	0.5375	0.01
SOPH-Cosseno			0.95	0.5666	0.04	0.5910	0.05	0.5376	0.03
SOPH-TFIDF			1.0	0.5667	0.05	0.5908	0.01	0.5374	-0.01
SOPH-Jaccard		90	0.9	0.5664	0.0	0.5909	0.03	0.5375	0.01
SOPH-Cosseno			0.95	0.5666	0.04	0.5909	0.03	0.5375	0.01
SOPH-TFIDF			1.0	0.5666	0.04	0.5907	-0.01	0.5374	-0.01

Tabela 10 – Melhores resultados de obtidos para a métrica de precisão e suas respectivas revocação e F1 em cada abordagem implementada. Os ganhos precedidos por “\*” indicam que houve diferença estatística em comparação com o método de referência. As linhas em destaque denotam os maiores precisões entre todas as abordagens.

ganho foi de apenas 0,05%, enquanto na precisão não houve sequer diferença estatística em comparação com o SOP.

Dentre as técnicas híbridas implementadas, a estratégia que apresentou melhores resultados foi a combinação entre tropes e gêneros dos filmes. Neste caso, o ganho é mais acentuado quando a medida empregada para calcular a distância entre os filmes é a similaridade do cosseno. Em comparação com o modelo de referência, o melhor algoritmo foi o SOPH-TFID que apresentou ganho de 0,76% para o RMSE e 2,93% para precisão. Entretanto, a configuração utilizada para alcançar esse desempenho foi diferente para cada métrica. O melhor RMSE ocorreu quando utilizamos todos os tropes de todos os filmes (quantidade mínima de tropes igual a 1) e a recomendação foi baseada tanto na filtragens híbrida quanto na colaborativa. Já a melhor precisão foi observada quando podamos os tropes dos filmes que possuem menos de 30 tropes catalogados e a previsão foi dominada pelo modelo híbrido.

Com relação ao filtro de quantidade mínima de tropes, podemos inferir que um número ideal para a poda deve estar no intervalo entre 1 e 30 aproximadamente, visto que, fora deste intervalo, o desempenho começa a declinar. Nossa hipótese para explicar esse fenômeno é que previsão passa a ser dominada pela filtragem colaborativa, já que a grande maioria dos filmes tem poucos tropes catalogados, tornando a abordagem similar ao *baseline*.

Também analisamos a relevância dos tropes através do estudo da ponderação TF-IDF e da adoção de um filtro de tropes por nota. Em ambos os casos, o desempenho observado foi semelhante aos seus modelos de referência, indicando que tais estratégias não revelam uma distinção expressiva entre os tropes em termos de frequência nos filmes e na coleção em que ocorrem (o que é capturado pelo TF-IDF) quanto em termos da sua contribuição particular para as notas atribuídas aos filmes em que eles são observados (o que é capturado pelo filtro por nota). No caso do TF-IDF, esse resultado pode ter duas interpretações. A primeira é que os tropes não se repetem com tanta frequência em um mesmo filme, o que torna essa estratégia próxima da abordagem binária, já que as duas utilizam a similaridade do cosseno. A segunda hipótese é que a relevância não pode ser extraída utilizando os conceitos de raridade, ou seja, tropes comuns e raros possuem a mesma importância no contexto em que se aplicam no ponto de vista dos SRs. Quanto às notas, como já havia sido observado no capítulo anterior, tropes parecem se distribuir indistintamente por filmes com maiores e menores notas. Assim, a presença de um elemento particular de narrativa não implica que o filme seja bom ou ruim.

Por fim, analisando os resultados obtidos em nossos experimentos para os métodos SOPH-LDSDd e SOPH-LDSDi, observamos que existem divergências em comparação com os resultados apresentados no trabalho de Yang, Hu e Qu (2013). Investigando especificamente o método SOPH-LDSDi, o único que utiliza tropes, o ganho obtido em

---

sua pesquisa para a métrica RMSE foi de 0,83% com relação ao SOP, enquanto em nossa implementação foi de 0,11%. Essa diferença pode ter sido influenciada pela configuração das bases de dados de cada trabalho. Em nossos experimentos foi utilizada uma base de dados com 943 usuários e 80.339 avaliações. No caso da pesquisa de referência, foram usados apenas 442 usuários e 5.500 avaliações, o que torna nossos resultados mais robustos do ponto de vista estatístico.

## 6 Conclusão e Trabalhos Futuros

Neste trabalho foram propostas estratégias para melhorar a qualidade dos SRs de filmes através da aplicação de tropes em modelos híbridos. Nosso objetivo foi baseado na hipótese de que a preferência do usuário no contexto de filmes pode ser representada através de elementos que compõem a história, como personagens, padrões narrativos, ambientação, etc, indo além de evidências tradicionalmente usadas como elenco, direção e gênero ou até mesmo o comportamento histórico do usuário.

As informações sobre os tropes foram extraídas do site TVTropes através da coleção DBTropes. Essas coleções nos permitiram a condução de um estudo sistemático com intuito de caracterizar o relacionamento entre filmes e tropes, além de explorar aspectos que serviriam de evidências para o método de recomendação. Em seguida, buscamos na literatura um algoritmo que possibilitasse o uso de técnicas de FBC, o que resultou na seleção do método SOPH, que é inspirado em abordagens de FC e emprega a similaridade entre itens para melhorar a acurácia das previsões.

O próximo passo foi aplicar diferentes medidas de similaridade para calcular a distância entre os filmes do ponto de vista de tropes e adicionar tal informação no modelo de recomendação híbrido. Em nosso estudo, foram implementadas abordagens comumente utilizadas na literatura para calcular a similaridade entre vetores: coeficiente de Jaccard, similaridade do cosseno e TF-IDF. Este último, em específico, nos possibilitou analisar se a relevância dos tropes está relacionada à frequência com que são usados.

Com base no estudo realizado, podemos oferecer as seguintes respostas para as perguntas que motivaram o nosso trabalho:

- *Como tropes podem ser usados em tarefas de recomendação? Em outras palavras, qual a melhor forma de definir  $f(m_j)$  e incorporá-la ao estimador  $r(R, u_i, m_j)$ ?*

Os resultados obtidos em nossos experimentos indicam que a melhor estratégia para utilizar tropes em SRs é por combiná-los aos gêneros dos filmes.

- *Quão úteis podem ser tropes nestas tarefas, em termos de proporcionar recomendações de maior qualidade?*

O uso de tropes como evidência isolada resulta em ganhos estatisticamente significativos, porém pouco expressivos (inferiores a 1% de erro). Quando combinados com o gênero, tropes produzem ganhos melhores, porém ainda relativamente baixos. Eles parecem mais úteis para melhorar a precisão da recomendação (no sentido de reduzir o número de falso-positivos) que em reduzir o erro global de previsão (quando medido com o RMSE). É importante ressaltar, contudo, que a efetividade de tropes

como evidência pode ter sido muito prejudicada pela esparsidade das informações do TVTropes, que apresenta uma cobertura deficiente e muito desbalanceada das obras catalogadas.

- *Todos os diferentes tipos de tropes são igualmente apropriados para esta tarefa?*

Apesar do resultado com TF-IDF na métrica RMSE sugerir que tropes populares podem comprometer a qualidade da previsão, isto não foi claramente confirmado em outras avaliações. Quanto à associação de tropes com melhores avaliações de filmes, nenhum padrão claro foi observado. Diante disso, podemos concluir que a maioria dos tropes parecem ocorrer indistintamente entre filmes mais e menos preferidos. De forma geral, não fomos capazes de encontrar nenhum agrupamento particular de tropes que mostrasse de forma inequívoca que alguns tropes são mais ou menos apropriados para recomendação.

## 6.1 Trabalhos Futuros

Nosso estudo mostrou o potencial da utilização de tropes para e recomendação de filmes. Porém, faz-se necessária a exploração de outras características dos tropes a fim de tornar seu desempenho mais consistente. Deste modo, podemos destacar algumas estratégias que podem ser investigadas em trabalhos futuros.

O primeiro ponto é a aplicação da nossa abordagem em outras coleções que contenham notas de usuários a itens com objetivo de analisar se a solução mantém o mesmo comportamento. Além disso, outras bases de dados podem aumentar a quantidade de filmes a serem explorados, dado que apenas 50,24% dos filmes cadastrados no MovieInes foram utilizados em nossas experimentos devido a baixa intersecção com o DBTropes. Ainda nessa linha, um dos problemas enfrentados em nossa pesquisa foi a grande quantidade de filmes com poucos tropes catalogados na coleção DBTropes. Nesse caso, é necessário investigar formas para mitigar esse desbalanceamento, diferentes da nossa estratégia de podar os tropes dos filmes que não atingem determinado limiar de quantidade, pois não obteve o impacto esperado. Adicionalmente, poderíamos testar nossas abordagens em outros domínios como na recomendação de livros ou jogos eletrônicos, visto que este tipo de conteúdo também é referenciado pelo TVTropes.

Outro aspecto a ser explorado diz respeito a mecanismos para determinar a relevância dos tropes no contexto de um filme. Novamente, as abordagens aplicadas em nossa pesquisa não melhoraram expressivamente a qualidade da previsão. Uma estratégia alternativa a ser investigada seria expressar a importância dos tropes através de *spoilers*. Ou seja, tropes rotulados com esse metadado podem ser mais relevantes para determinado filme pois contém fatos críticos inerentes a história. Essa estratégia só é viável porque tal informação está presente no site TVTropes.

Também se faz necessária a utilização de outros cenários e métricas para analisar o uso de tropes sob diferentes óticas. Por exemplo, usar dados de teste que contenham novos usuários e itens para simular o problema de *coldstart* e analisar como os métodos que incorporam tropes se saem em comparação com outras abordagens. Podemos também usar métricas para avaliar não só a acurácia das recomendações, mas também a sua capacidade de surpreender o usuário (através da sugestão de novos itens), utilizando as medidas de *novidade* e *diversidade*. Nossa hipótese é que um algoritmo baseado em tropes pode apresentar melhor desempenho dado que daria ao usuário a oportunidade de conhecer filmes de diferentes gêneros ou com atores pouco conhecidos por ele, porém com elementos de narrativa do seu interesse.

Por fim, pretendemos investigar na literatura métodos híbridos mais eficazes que o SOPH, já que este é inspirado no algoritmo Slope One, que não constitui o estado-da-arte em recomendação, hoje dominada por complexos *ensembles* que incorporam estratégias de fatoração.



## Referências

- ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, v. 17, n. 6, p. 734–749, June 2005. ISSN 1041-4347. Citado 3 vezes nas páginas 11, 19 e 21.
- BAMMAN, D.; O’CONNOR, B.; SMITH, N. A. Learning latent personas of film characters. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Sofia, Bulgaria: Association for Computational Linguistics, 2013. p. 352–361. Disponível em: <<http://www.aclweb.org/anthology/P13-1035>>. Citado na página 24.
- BOBADILLA, J. et al. Recommender systems survey. *Know.-Based Syst.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 46, p. 109–132, jul. 2013. ISSN 0950-7051. Disponível em: <<http://dx.doi.org/10.1016/j.knosys.2013.03.012>>. Citado 6 vezes nas páginas 14, 19, 20, 22, 25 e 33.
- BOYD-GRABER, J.; GLASGOW, K.; ZAJAC, J. S. Spoiler alert: Machine learning approaches to detect social media posts with revelatory information. In: *ASIST 2013: The 76th Annual Meeting of the American Society for Information Science and Technology*. [S.l.: s.n.], 2013. Citado na página 24.
- CREMONESI, P. et al. An evaluation methodology for collaborative recommender systems. In: *Proceedings of the 2008 International Conference on Automated solutions for Cross Media Content and Multi-channel Distribution*. Washington, DC, USA: IEEE Computer Society, 2008. (AXMEDIS ’08), p. 224–231. ISBN 978-0-7695-3406-0. Disponível em: <<http://dx.doi.org/10.1109/AXMEDIS.2008.13>>. Citado na página 22.
- GEMMIS, M. de et al. Integrating tags in a semantic content-based recommender. In: *Proceedings of the 2008 ACM Conference on Recommender Systems*. New York, NY, USA: ACM, 2008. (RecSys ’08), p. 163–170. ISBN 978-1-60558-093-7. Disponível em: <<http://doi.acm.org/10.1145/1454008.1454036>>. Citado na página 45.
- HALL, M. et al. The weka data mining software: An update. *SIGKDD Explorations*, v. 11, n. 1, 2009. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Citado na página 22.
- KIESEL, M.; GRIMNES, G. A. Dbtropes: a linked data wrapper approach incorporating community feedback. In: VALKER, J.; CORCHO, O. (Ed.). *EKAW 2010 Demo e Poster Abstracts. International Conference on Knowledge Engineering and Knowledge Management (EKAW-10), 17th, October 11-15, Lisbon, Portugal*. o.A., 2010. Best Poster. Disponível em: <<http://dbtropes.org/>>. Citado na página 11.
- LEES-MILLER, J. et al. Does wikipedia information help netflix predictions? In: *Machine Learning and Applications, 2008. ICMLA ’08. Seventh International Conference on*. [S.l.: s.n.], 2008. p. 337–343. Citado na página 26.

- LEMIRE, D.; MACLACHLAN, A. Slope one predictors for online rating-based collaborative filtering. In: *Proceedings of SIAM Data Mining (SDM'05)*. [s.n.], 2005. Disponível em: <[http://www.daniel-lemire.com/fr/documents/publications/lemiremaclachlan\\_sdm05.pdf](http://www.daniel-lemire.com/fr/documents/publications/lemiremaclachlan_sdm05.pdf)>. Citado 3 vezes nas páginas vii, 12 e 13.
- MAATEN, L. van der; HINTON, G. E. Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research*, v. 9, p. 2579–2605, 2008. Citado na página 29.
- MAK, H.; KOPRINSKA, I.; POON, J. Intimate: a web-based movie recommender using text categorization. In: *Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on*. [S.l.: s.n.], 2003. p. 602–605. Citado na página 25.
- MENEZES, D. et al. Weighted slope one predictors revisited. In: *Proceedings of the 22Nd International Conference on World Wide Web*. New York, NY, USA: ACM, 2013. (WWW '13 Companion), p. 967–972. ISBN 978-1-4503-2038-2. Disponível em: <<http://doi.acm.org/10.1145/2487788.2488093>>. Citado na página 19.
- METEREN, R.; SOMEREN, M. Using content-based filtering for recommendation. In: *Proceedings of ECML 2000 Workshop: Matching Learning in Information Age*. [S.l.: s.n.], 2000. p. 47–56. Citado na página 14.
- MOVIELENS. *MovieLens*. 2015. Disponível em: <<http://grouplens.org/datasets/movielens/>>. Citado 2 vezes nas páginas 24 e 39.
- MUKHERJEE, R. et al. Movies2go: An online voting based movie recommender system. In: *Proceedings of the Fifth International Conference on Autonomous Agents*. New York, NY, USA: ACM, 2001. (AGENTS '01), p. 114–115. ISBN 1-58113-326-X. Disponível em: <<http://doi.acm.org/10.1145/375735.376018>>. Citado na página 25.
- NETFLIX. *Netflix Prize*. 2009. Disponível em: <<http://www.netflixprize.com/index>>. Citado na página 4.
- NOIA, T. D. et al. Linked open data to support content-based recommender systems. In: *Proceedings of the 8th International Conference on Semantic Systems*. New York, NY, USA: ACM, 2012. (I-SEMANTICS '12), p. 1–8. ISBN 978-1-4503-1112-0. Disponível em: <<http://doi.acm.org/10.1145/2362499.2362501>>. Citado na página 26.
- PALACIOS, G. *¿Qué es un Trope?* 2004. Disponível em: <<http://www.kissmybit.com/author/gpalacios/page/2/>>. Citado na página 4.
- PASSANT, A. Measuring semantic distance on linking data and using it for resources recommendations. In: *AAAI Spring Symposium: Linked Data Meets Artificial Intelligence*. AAAI, 2010. Disponível em: <<http://dblp.uni-trier.de/db/conf/aaais/aaais2010-7.html#Passant10>>. Citado 4 vezes nas páginas vii, 14, 15 e 16.
- RESNICK, P.; VARIAN, H. R. Recommender systems. *Communications of the ACM*, v. 40, n. 3, p. 56–58, 1997. Citado 2 vezes nas páginas 11 e 12.
- RICCI, F. et al. *Recommender Systems Handbook*. 1st. ed. New York, NY, USA: Springer-Verlag New York, Inc., 2010. ISBN 0387858199, 9780387858197. Citado 2 vezes nas páginas 18 e 21.

SHANI, A. G. G. *Evaluating Recommender Systems*. [S.l.], 2009. Disponível em: <<https://www.microsoft.com/en-us/research/publication/evaluating-recommender-systems/>>. Citado na página 22.

SHEPITSEN, A. et al. Personalized recommendation in social tagging systems using hierarchical clustering. In: *Proceedings of the 2008 ACM Conference on Recommender Systems*. New York, NY, USA: ACM, 2008. (RecSys '08), p. 259–266. ISBN 978-1-60558-093-7. Disponível em: <<http://doi.acm.org/10.1145/1454008.1454048>>. Citado na página 17.

SHI, Y.; LARSON, M.; HANJALIC, A. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 47, n. 1, p. 3:1–3:45, maio 2014. ISSN 0360-0300. Disponível em: <<http://doi.acm.org/10.1145/2556270>>. Citado 2 vezes nas páginas 12 e 25.

STERLING, B. *TVTropes.org*. 2015. Disponível em: <<https://tvtropes.org/>>. Citado 2 vezes nas páginas 3 e 9.

SZOMSZOR, M. et al. Folksonomies, the semantic web, and movie recommendation. In: *4th European Semantic Web Conference, Bridging the Gap between Semantic Web and Web 2.0*. [s.n.], 2007. Event Dates: 3-7th, June 2007. Disponível em: <<http://eprints.soton.ac.uk/264007/>>. Citado na página 26.

U., L. H.; YAO, K.; MAK, H. F. Pathsimplx: Revisiting pathsim in heterogeneous information networks. In: *Web-Age Information Management - 15th International Conference, WAIM 2014, Macau, China, June 16-18, 2014. Proceedings*. [s.n.], 2014. p. 38–42. Disponível em: <[http://dx.doi.org/10.1007/978-3-319-08010-9\\_6](http://dx.doi.org/10.1007/978-3-319-08010-9_6)>. Citado na página 17.

YANG, R.; HU, W.; QU, Y. Using semantic technology to improve recommender systems based on slope one. In: LI, J. et al. (Ed.). *Semantic Web and Web Science*. Springer New York, 2013, (Springer Proceedings in Complexity). p. 11–23. ISBN 978-1-4614-6879-0. Disponível em: <[http://dx.doi.org/10.1007/978-1-4614-6880-6\\_2](http://dx.doi.org/10.1007/978-1-4614-6880-6_2)>. Citado 11 vezes nas páginas 4, 7, 14, 16, 19, 24, 37, 38, 40, 41 e 49.

ZACHARSKI, R. A programmer's guide to data mining. In: . [s.n.], 2013. Disponível em: <<http://guidetodatamining.com/>>. Citado na página 12.

ZHAO, J.; MA, J. An improved slope one algorithm based on tag frequency. In: *Computer Science and Network Technology (ICCSNT), 2013 3rd International Conference on*. [S.l.: s.n.], 2013. p. 369–372. Citado na página 19.