



# Uma Revisão Acerca da Recomendação Personalizada de Conteúdo

Jaqueline Ferreira de Brito, *Mestrando em Sistemas de Informação, USP*,  
Luciano Antonio Digiampietri, *Professor no Bacharelado em Sistemas de Informação, USP*

**Resumo**—Os sistemas de recomendação de conteúdo surgiram para lidar com a sobrecarga de informações disponíveis diariamente na Web. A fim de estudar o cenário atual sobre recomendação personalizada de conteúdo, o presente artigo pretende identificar e analisar os métodos e as técnicas adotados. No processo de revisão sistemática, foram selecionados 22 artigos publicados nos últimos cinco anos nas principais bibliotecas digitais da área. Os resultados obtidos sugerem que há uma tendência na utilização de algumas estratégias para minimização das limitações e aumento da eficiência dos métodos tradicionais de recomendação.

**Palavras-chave**—Recomendação de Conteúdo, recomendação personalizada, sistemas de recomendação.

## A Study about Personalized Content Recommendation

**Abstract**—Recommender systems have been proposed as efficient strategies to deal with the information overload which is an Internet reality. In order to study personalized content recommendation, this paper aims at identifying and analyzing, systematically, methods and techniques adopted in the recommendation context. In the systematic review process, 22 papers have been accepted from the main digital databases in the area. The final results suggest a tendency in the use of strategies to reduce the limitations and to increase the efficiency of the recommendation methods.

**Index Terms**—Content Recommendation, Personalized Recommendation, Recommender Systems

### I. INTRODUÇÃO

**S**ISTEMAS de Recomendação são ferramentas de *software* e métodos que proveem sugestões de itens a serem utilizados por um usuário. As sugestões oferecidas apoiam seus usuários em vários processos de decisão, tais como: quais itens comprar, quais músicas ouvir ou quais notícias ler [1, 2]. Em geral, as seguintes técnicas de recomendação têm se sobressaído: **Baseado em conteúdo (Content-Based Filtering - CBF)**: Recomenda itens ao usuário, cujo conteúdo é similar ao conteúdo que o usuário tenha visto ou selecionado recentemente; **Colaborativo (Collaborative Filtering - CF)**: Faz recomendações baseadas nas avaliações dos itens realizadas por um grupo de usuários (vizinhos), cujos perfis de avaliações

são os mais similares ao do usuário alvo; **Híbrido (H)**: Combina ambos os métodos citados.

Além disso, a necessidade de personalizar o conteúdo está cada vez maior, pois a busca pela informação relevante e a minimização do tempo despendido tornaram-se fatores indispensáveis em mecanismos de busca e filtragem de informações [1]. Desse modo, esses sistemas permitem que o conteúdo tenha um significado semântico bem definido e que o usuário possa interagir melhor com a informação fornecida, dadas suas preferências.

Este artigo visa estudar o estado da arte dos sistemas de recomendação de conteúdo baseados no perfil do usuário. O restante deste artigo encontra-se estruturado da seguinte forma: a Seção II apresenta a metodologia utilizada, a Seção III apresenta os trabalhos analisados e nas Seções IV e V podem ser observadas, respectivamente, a discussão e a conclusão do trabalho.

### II. METODOLOGIA

**A** Fim de estudar o cenário atual sobre Recomendação de Conteúdo baseada no perfil do usuário final, primeiramente, realizou-se uma análise exploratória sobre o tema para familiarizar-se com os principais conceitos do objeto de estudo. A partir disso, o processo de revisão sistemática [3] contou com as seguintes etapas:

- 1) Planejamento: Etapa na qual define-se o problema de pesquisa e o protocolo da revisão. A presente revisão sistemática tem o propósito de esclarecer a seguinte questão: “Quais são as técnicas existentes para recomendação de conteúdo que são baseadas no perfil do usuário final?”.

Para cada base de dados selecionada, foram consideradas as seguintes *strings* de busca:

- *ACM Digital*: *Abstract*: (“content-based” AND “recommender system” AND (profile OR personalize))
- *IEEEExplore*: “Abstract”: “content-based” AND “recommender system.” AND (profile OR personalize)

O estudo contou com os seguintes critérios de seleção:

- Foram incluídos trabalhos publicados e disponíveis digitalmente nas bases de dados científicas da *ACM* e da *IEEE*;

- Foram incluídos trabalhos recentes (publicados a partir de 2008) que já possuam aprovação pela comunidade científica;
- Foram excluídos trabalhos que apresentam recomendação de conteúdo sem apresentar o método e a técnica utilizados;
- Foram excluídos trabalhos que não estão no domínio do conhecimento de recomendação de conteúdo baseada no perfil do usuário final.

2) Condução: Nesta etapa, 39 artigos atenderam os critérios de inclusão estabelecidos: 19 artigos da base ACM e 20 da base IEEE. Desse montante, alguns eram duplicados e outros foram rejeitados de acordo com os critérios de exclusão definidos acima [4]. A Figura 1, ilustra o *status* dos artigos selecionados.

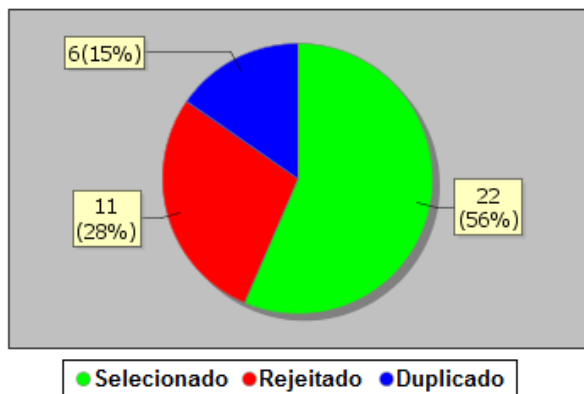


Fig. 1. Status dos artigos selecionados. Fonte: os próprio autores.

3) Extração: Etapa na qual os 22 artigos (5 artigos da base ACM e 17 da base IEEE) foram lidos na íntegra e suas principais características foram extraídas. A Tabela I apresenta uma síntese dos artigos com os campos de extração estabelecidos.

### III. RESULTADOS

DOS 22 artigos selecionados, a maioria (47,4%) foi publicada no ano de 2010 e a minoria (5,3%) no ano de 2012. A Figura 2 ilustra a distribuição dos artigos selecionados durante os últimos cinco anos.

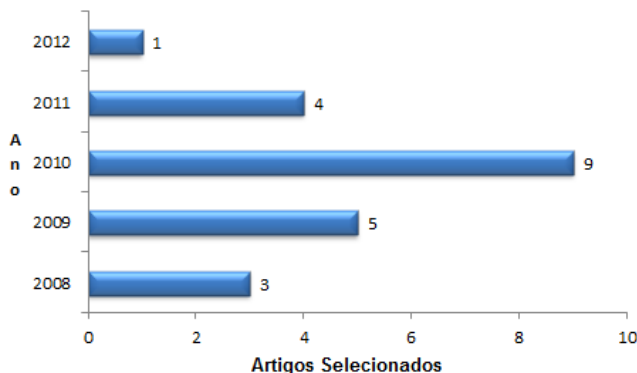


Fig. 2. Distribuição dos artigos selecionados no período de 2008-2012. Fonte: os próprio autores.

Além do perfil do usuário e das características dos itens, os trabalhos consideram outras informações para o processo de recomendação de conteúdo:

- **Informações de contexto:**

Com base nas variáveis de tempo e preferências do grupo, o modelo de Julashokri e outros [23] busca por produtos similares aos produtos comprados, computa o nível de interesse do usuário alvo e recomenda os N produtos mais similares. As medidas R (*Recency*) - período de tempo desde a última compra, F (*Frequency*) - número de compras ocorridas durante um período definido e M (*Monetary*) - a quantidade de dinheiro gasta pelo cliente durante o período definido, são consideradas pelo sistema.

Em [9], os problemas de *Sparsity Rating* (poucos itens avaliados pelos novos usuários) e *Without Contextual Information* (ausência de informações de lugar, tempo, dia, companhia) são tratados. *ModernizeMovie* baseia-se em multi critérios para criar as pseudo avaliações dos filmes e multi dimensões de contexto para compor o perfil do usuário. A avaliação do modelo Naive Bayes desenvolvido confirma que o sistema foi capaz de prover uma maior precisão de recomendação do que as técnicas puras de CBF e CF.

Alabastro e outros [16] lidam com métodos de recomendação CBF e CF com e sem informação de contexto em diferentes algoritmos. Um guia interativo foi desenvolvido para recomendar visitas personalizadas a um museu de arte. O melhor modelo obtido foi o colaborativo utilizando o algoritmo k-NN. O sistema de Iaquina e outros [17], também aplicado ao contexto de itinerário de museu, considera as informações de ambiente e comportamento do usuário para reorganizar os itens sugeridos e fazer recomendações serendipistas (surpreendentemente interessantes). Um perfil probabilístico dos interesses dos usuários é construído, a partir da classificação obtida dos itens.

Em [20], é apresentado um modelo de rede Bayesiana, para recomendação de notícias personalizadas para celulares que baseia-se em atividades, preferências, ações, localização, tempo, conteúdo de notícias, etc, para computar o interesse do usuário. Por fazer uso de tecnologias híbridas P2P, o sistema incorpora as informações obtidas e, automaticamente, oferece em tempo real as notícias consideradas relevantes. Os experimentos mostraram que o modelo é efetivo.

- **Folksonomies:**

Lops e outros [22] combinam *User Generated Content* (UGC) com análise semântica de conteúdo. Uma das maneiras de lidar com UGC é por meio de *folksonomy*, taxonomia gerada pelos usuários que colaborativamente categorizam recursos com

TABELA I  
SÍNTESE DOS ARTIGOS SELECIONADOS. FONTE: OS PRÓPRIO AUTORES.

Ref.	Método	Domínio	Modelo	Validação
[5]	H	anúncios	k-NN, k-means, k-means híbrido, ITCC, ITCC híbrido, CCAM, CCAM híbrido	MAE
[6]	CBF	anúncios	k-NN, CBF	feedback do usuário, Precision
[7]	CBF	artigos científicos	k-NN, CBF with memory-based and model-based attributes	feedback de usuários
[8]	CBF	<i>e-learning</i>	CBF	medida TF-IDF
[9]	H	filmes	Naïve Bayes, CF	MAE, F-measure, Coverage
[10]	CBF	filmes	Naïve Bayes, semantic CBF	5-fold cross validation
[11]	H	filmes	ontology CBF, Apriori	-
[12]	CF	filmes	K-medoid, cluster updating algorithm	MAE
[13]	H	filmes, programas de TV	k-NN, item-based CF, user-based CF, social CBF	estudantes
[14]	CBF	livros	trust-based recommendation, Temporal Difference learning	CDC, CDO, CDLOC, CBC - orientada e não orientada a aspecto
[15]	CBF	multimídia	smart multimedia	-
[16]	CBF, CF	museu de arte	Naïve Bayes, Árvore de decisão, k-NN, RNA, CBF with context, CF with context	Precision, Recall, F-Measure, Coverage, 10-fold cross validation, teste de aceitação
[17]	H	museu de arte	Naïve Bayes, algoritmo genético, heurística de serendipismo	usuários
[18]	H	notícias	<i>Framework</i> Bayesiano, CBF, CF	CTR, frequência de visitas diárias
[19]	H	notícias	semantic CBF, CF	estudantes, medida TF-IDF
[20]	H	notícias	Rede bayesiana	funcionários e estudantes
[21]	CBF	pesquisadores	OKAPI, KLD, PM, REL	Average recall, MRR
[22]	CBF	pintura	Naïve Bayes, semantic CBF	5-fold cross validation, Precision, Recall
[23]	H	produtos em geral	SOM, item-based CF	medida RFM
[24]	H	roupas	k-NN, item-based CF, CBF, user-based CF	informações de questionário e no site
[25]	H	<i>social bookmarking</i>	WebDCC algorithm, ontology CBF, Rocchio	Precision, Recall
[26]	H	web services	algoritmo least square, fuzzy	ANFIS, feedback de usuários

palavras-chaves escolhidas livremente (*tags*). O sistema estima a probabilidade de interesse do usuário com base nas características das pinturas. A adoção de *tags* aumentou a precisão da recomendação das obras de arte, em especial, para usuários não especialistas, cujos perfis ficaram em destaque.

Godoy e Amandi[25] propõem uma estratégia híbrida que tira proveito dos perfis dos usuários por meio de CBF e *tags* frequentemente utilizadas. Os experimentos mostraram que os perfis construídos superaram os métodos de recomendação baseados em popularidade de *tags*, como MPTU (*Most Popular Tags by User*) e MPTR (*Most Popular Tags by Resource*). No entanto, mais experimentos precisam ser feitos para a confirmação dos resultados preliminares.

Em [13], é apresentado um sistema híbrido baseado em nuvens de *tags* de usuários de filmes e programas de TV e em suas classificações para a construção de uma nuvem de *tags* para o usuário-alvo. O modelo indicou uma potencial eficácia de recomendação para os estudantes, porém ainda é necessário completar a fase de validação da solução.

- **Palavras-chaves e Ontologias (ou taxonomias):** Em [8], o sistema aprende o perfil do usuário de forma implícita, baseando-se nas palavras-chaves e seus relacionamentos criados automaticamente. A relação das palavras-chaves pode levar a um aumento da acurácia no cálculo da similaridade. Esse

método melhora a construção do perfil do usuário e aumenta a eficiência da aprendizagem com CBF para a recomendação de materiais.

Loh [6] investiga o uso de palavras-chaves (automaticamente extraídas do texto) e classes (extraídas de uma taxonomia ou ontologia) para representar os perfis dos usuários. A combinação de classes e palavras-chaves alcançaram melhores resultados do que quando foram utilizadas sozinhas, criando um perfil mais preciso. O desempenho da melhor técnica ficou abaixo de outras técnicas conhecidas, contudo, deve-se levar em consideração as diferentes condições experimentais que podem ter afetado o resultado final.

Em [7], adotou-se o modelo baseado na taxonomia da ACM para recomendar artigos científicos para usuários do *site* CiteSeerX<sup>1</sup>. O sistema foi capaz de prover recomendações corretas, sem precisar de muitos registros históricos de usuários ou *rankings* explícitos.

Cantador e outros [19] apresentam o sistema *News@hand*, que faz uso de tecnologia semântica para recomendar notícias. Os conteúdos de notícia e preferências de usuários são descritos em termos de conceitos que aparecem em um conjunto de ontologias de domínio. Neste artigo não está claro

<sup>1</sup><http://CiteSeerx.ist.psu.edu/>

se o processo de anotação é feito manualmente ou por meio de técnicas automáticas. A medida TF-IDF (*Term Frequency-Inverse Document Frequency*) [27] atribui pesos a todas as anotações criadas para os itens, refletindo a importância de cada uma delas.

Pan e outros [11] propõem o sistema multi-agentes OARS (*Ontology-Based Adaptive Personalized Recommender System*) que lida com algumas limitações dos sistemas de recomendação em geral: novo item, novo usuário, superespecialização, sugestão estática e esparsidade das classificações.

Moursi e outros [15] propõem um sistema descentralizado de multimídia (áudio, visual e audiovisual) que utiliza o *Smart Multimedia* (SM), um modelo que além de possuir a representação binária do arquivo de multimídia, contém o conhecimento sobre a semântica do conteúdo interno, junto com a capacidade de tomar decisão autônoma. Os *smart agents* são capazes de obter os perfis do usuário e do item, e trabalham colaborativamente para lidar com o problema da recomendação de acordo com as preferências do usuário.

- **Outros:**

Em [18], o *framework* Bayesiano desenvolvido utiliza informações dos logs de *clicks* dos usuários para prever as notícias de interesse do usuário, com base em suas atividades e nas notícias que são tendências demonstradas nas atividades de um grupo de usuários. O modelo melhorou a qualidade de recomendação e atraiu mais visitantes para o *website Google News*.

Em [24], um sistema é projetado para oferecer sugestões de roupas quando uma cliente VIP entra na loja com seu cartão (embutido com a tecnologia RFID) na bolsa. O modelo mostrou que o sistema é um mecanismo efetivo para produção de recomendação de alta qualidade aos clientes, porém possui ainda algumas limitações: necessita de uma base de dados enorme e o cálculo da distância entre todos os interesses do cliente é demorado.

Bedi e Agarwal [14] apresentam o AORS (*Aspect Oriented Recommender System*) que aprende as preferências do usuário de uma maneira modular melhor, baseada em multi-agentes. A validação do sistema contou com as medidas de: CDC (*Concern diffusions over components*) - que indica a necessidade de mais componentes, CDO (*Concern diffusions over operations*) - que indica a necessidade de mais operações (métodos/sugestões) para implementação do aprendizado baseado na abordagem orientada a agentes em vez da AOP (*Aspect Oriented Programming*), CDLOC (*Concern diffusions over LOC*) - indicativo de que a solução AOP é mais efetiva em termos de modularização

do aprendizado pelas linhas de código, CBC (*Coupling between components*) - para um sistema com e sem uso de aspecto. Os resultados apontam que o aprendizado das preferências com aspecto requer menos componentes, menos operações para representar o aprendizado e uma solução mais efetiva em termos de modularização do aprendizado, por meio da métrica de linhas de código.

Lops e outros [10] apresentam um sistema de recomendação de filmes em Inglês e Italiano. O sistema usa um modelo probabilístico para classificar um novo item baseado nas preferências anteriores. A partir dos experimentos realizados, os resultados foram muito semelhantes frente aos modelos bilíngues e monolíngues, porém o modelo que teve o aprendizado em inglês e recomendava itens também em inglês obteve um desempenho um pouco melhor. Uma das considerações importantes é que o erro na etapa de desambiguação introduzido nos modelos deve ser considerado.

Em [26], o sistema baseia-se em propriedades não-funcionais (tempo de resposta, reputação, custo, confiabilidade, etc) e nas preferências do usuário para recomendar serviços *web*. Utilizou-se ANFIS (*Adaptive Neuro Fuzzy Inference System*) para melhorar o conjunto de regras *fuzzy* estabelecido. O sistema possui uma característica dinâmica, sendo dependente do *feedback* de usuários anteriores de um determinado serviço *web*.

Gollapalli e outros [21] propõem um sistema de recomendação de pesquisadores com expertises similares. As técnicas exploradas foram: OKAPI - calcula a similaridade entre dois perfis, KLD (*KL Divergence*) - usada para quantificar a similaridade entre duas distribuições de probabilidade, PM (*Probabilistic Modeling*) - calcula a similaridade entre dois perfis usando probabilidade condicional, REL (*Trace-based Similarity*) - utiliza uma matriz de densidade para computar um *score* relevante entre dois perfis. As técnicas OKAPI e REL obtiveram melhores resultados com base nos valores de *average recall* e MRR (*Mean Reciprocal Rank*). Os resultados apontam um indicativo, porém ainda carecem de exploração de técnicas mais acuradas para representação dos perfis.

Em [5], o modelo *co-clustering augmented data matrix* (CCAM) proposto é capaz de lidar com a esparsidade dos dados. Após a comparação de desempenho entre modelos, constatou-se que a predição baseada em modelo (*k-means*, ITCC - *Information-Theoretic Co-Clustering* e CCAM) possui um desempenho muito melhor do que predição baseada em memória (*k-NN*). O *k-means* obteve um desempenho melhor na abordagem individual do que ITCC e CCAM, mas

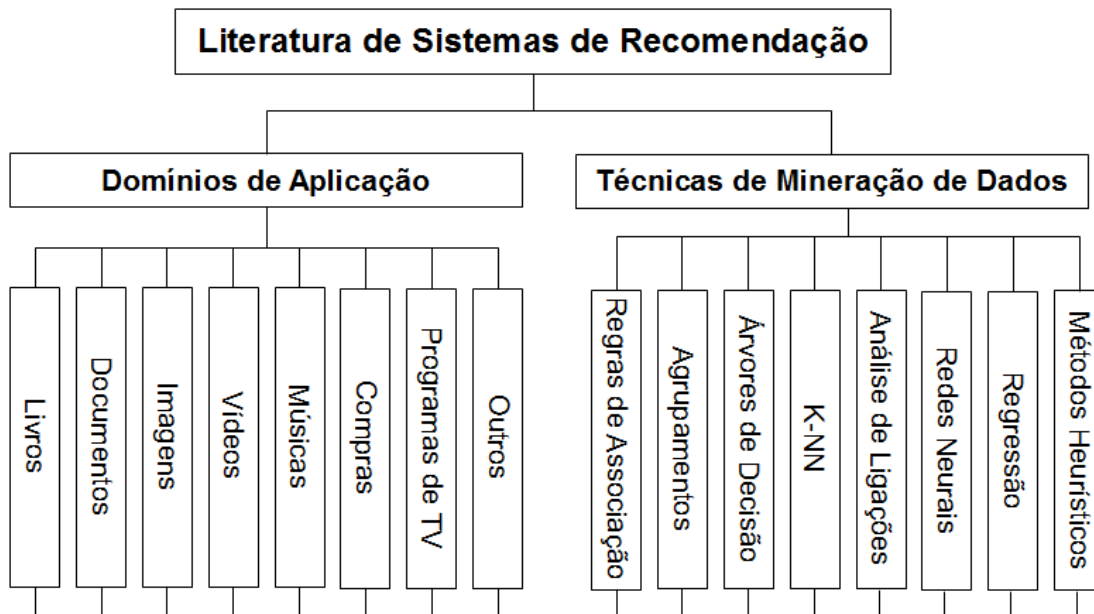


Fig. 3. Estruturação dos Sistemas de Recomendação. Fonte: adaptado de Deuk e outros [28].

o resultado foi próximo. O CCAM híbrido obteve um resultado muito melhor do que ITCC híbrido e o k-means híbrido. Isso pode ser justificado pelo fato do CCAM possuir um ganho de informação mais útil do que os demais, para minimização da perda de informação dos múltiplos dados relacionados.

Em [12], uma abordagem de *cluster* incremental para melhorar a escalabilidade do método de CF é apresentada. Como o processo de reconstrução de *clusters* é muito custoso, um algoritmo foi proposto para atualização dinâmica dos *clusters*. Comparando os valores de MAE (*Mean Absolute Error*) obtidos pelas situações dinâmicas e pela situação antes da clusterização inicial, a abordagem dinâmica obteve um resultado semelhante à situação inicial, porém mostrou-se mais vantajosa quando novos dados entraram no sistema.

#### IV. DISCUSSÃO

COMPARANDO a Tabela I com a Figura 3, percebe-se que os domínios de aplicação e as técnicas ou modelos de mineração de dados convergem. Os principais modelos aplicados nos trabalhos que trouxeram maior qualidade e precisão nos resultados são k-NN, Naïve Bayes e os algoritmos CF e CBF, frente aos variados campos de aplicação envolvidos.

Embora o desempenho do Naïve Bayes não seja tão bom quanto alguns outros métodos de aprendizagem estatística, tais como os classificadores k-NN e SVM (*Support Vector Machines*), foi visto que ele pode ter um desempenho surpreendentemente bom em tarefas de classificação nas quais a atualização da probabilidade computada não é importante. Outra vantagem da abordagem Naïve Bayes

é que ela é muito eficiente e fácil de ser implementada se comparada a outros métodos de aprendizagem [29].

O método de recomendação mais utilizado nos trabalhos analisados é o híbrido, pois reúne as vantagens dos métodos CBF e CF. Na Tabela II, as vantagens e limitações de cada método podem ser observadas [30].

O método CBF não precisa da avaliação do usuário sobre um item. Esse método compara todos os produtos disponíveis na base com o perfil do usuário. Entretanto, as características dos itens que merecem destaque não são levadas em consideração, já que o usuário não fornece notas nesse tipo de filtragem, assim, todas as características dos itens são tratadas de forma igualitária. O problema da superespecialização é sua outra limitação, pois as recomendações geradas são semelhantes às palavras-chave encontradas no perfil do usuário.

O método CF baseia-se nas avaliações dos usuários sobre os itens, independente da natureza do conteúdo que se quer recomendar. As recomendações de qualidade geradas por esse método dependem das avaliações (sejam elas positivas ou negativas) dadas pelos usuários. Além disso, esse método é capaz de fazer recomendações serendipistas. Porém, enquanto o produto não for avaliado, este nunca será recomendado e, se a base de dados for muito grande, é preciso que o usuário compre ou avalie um número considerável de produtos para que seja possível encontrar bons vizinhos com preferências semelhantes. Usuários que avaliam um número de itens inferior a 50 podem sofrer do problema do falso vizinho, pois, coincidentemente, avaliam itens iguais, mas seus perfis possuem poucos itens (não sendo possível generalizar a semelhança). Dessa forma, quanto mais produtos são avaliados igualmente, maior a chance dos perfis serem realmente semelhantes. Também, para formar grupos com gostos semelhantes, é necessário calcular a similaridade de um usuário-alvo com todos os

TABELA II  
VANTAGENS E LIMITAÇÕES DOS MÉTODOS DE RECOMENDAÇÃO. FONTE: OS PRÓPRIOS AUTORES.

Método	Vantagem	Limitação
CBF	não possui o problema do primeiro avaliador; todos os itens possuem chances de serem recomendados	não consideração de aspectos como qualidade do texto e renome do autor; superespecialização
CF	independência de conteúdo; geração de recomendações baseadas em preferências dos usuários; possibilidade de produzir recomendações inesperadas e de alta qualidade	o problema do primeiro avaliador; a dispersão da base de dados; ovelha negra (falso vizinho); custo de processamento
H	combina as vantagens e minimiza as limitações individuais de ambos métodos citados	o problema do <i>start up</i>

outros usuários, podendo ser um processo demorado e caro computacionalmente.

Como as desvantagens dos métodos citados são complementares, ou seja, as desvantagens de um são vantagens do outro, o método híbrido é capaz de fazer a junção de ambos os métodos, geralmente, com auxílio de outras técnicas para minimizar as limitações de cada um. A única desvantagem é o problema do *start up*, que refere-se ao tempo que o sistema leva para conseguir informações relevantes do usuário para gerar boas recomendações. Alguns autores sugerem recomendar uma lista de itens mais vendidos ou utilizar dados demográficos (idade, sexo, cidade, etc.) para gerar as recomendações iniciais.

Nesta revisão observou-se que o foco dos trabalhos revisados está mais na representação do perfil do usuário do que no próprio processo de recomendação, já que a melhoria da qualidade do primeiro reflete também na melhoria da qualidade e precisão do segundo. Os trabalhos sugerem que há uma tendência na utilização de informação de contexto, *folksonomies*, propriedade *serendipity*, palavras-chaves, ontologias (ou taxonomias) dentre outras estratégias, a fim de amenizar as limitações e aumentar a eficiência dos métodos aplicados. Além disso, a maioria dos trabalhos revisados apresenta estratégias de validação dos modelos construídos, destacando a importância dessa etapa no processo de análise dos sistemas.

## V. CONCLUSÃO

OS Sistemas de Recomendação têm chamado a atenção de profissionais e acadêmicos, pois são ferramentas que proveem o apoio aos usuários em variados contextos de tomada de decisão, diante do volume de informações disponibilizadas na *Web* a todo momento.

Dessa forma, identificou-se que o método de recomendação mais utilizado nos trabalhos é o híbrido, devido ao fato de reunir as vantagens dos métodos de aprendizado baseado em conteúdo e colaborativo. Além disso, os modelos mais frequentemente utilizados na etapa de aprendizagem do sistema foram k-NN, Naïve Bayes e os algoritmos CF e CBF. Em geral, esses modelos trouxeram um resultado de maior eficácia dentre os modelos aplicados.

A maioria dos trabalhos preocupou-se em validar suas soluções propostas, seja por meio de experimentos com estudantes participantes, técnicas de amostragem ou algum tipo de medida estatística específica. Observou-se também que o foco dos trabalhos revisados está mais na

representação do perfil do usuário do que no próprio processo de recomendação, refletindo numa melhoria global da qualidade do sistema.

Assim, os artigos apontam que além do perfil do usuário e das características dos itens, os trabalhos consideram outras informações no processo de recomendação de conteúdo: informação de contexto, *folksonomies*, propriedade serendipista, palavras-chaves, ontologias (ou taxonomias), etc. Isso sugere um indicativo de tendência, já que essas são estratégias para minimizar as limitações e, ao mesmo tempo, aumentar a eficiência da aprendizagem dos principais métodos de recomendação existentes.

## REFERÊNCIAS

- [1] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., *Recommender Systems Handbook*. Springer, 2011, ISBN: 978-0-387-85819-7.
- [2] S. C. Cazella, M. A. S. N. Nunes, and E. Reategui, “Jornada de atualização de informática,” in SBC, 10, ch. A Ciência do Palpite: Estado da Arte em Sistemas de Recomendação, pp. 161–216.
- [3] J. Biolchini, P. G. Mian, and A. C. C. Natali, “Systematic review in software engineering,” COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, Tech. Rep. RT-ES 679/05, 2005.
- [4] A. Zamboni, A. Thommazo, E. Hernandez, and S. Fabbri, *Start: uma ferramenta computacional de apoio à revisão sistemática*, 2012. [Online]. Available: <http://lapes.dc.ufscar.br/software/start-tool>.
- [5] M.-L. Wu, C. H. Chang, and R. Z. Liu, “Collaborative filtering with ccam,” in *Machine Learning and Applications and Workshops (ICMLA), 2011 10th International Conference on*, vol. 2, 2011, pp. 245–250. DOI: 10.1109/ICMLA.2011.47.
- [6] S. Loh, F. Lorenzi, G. Simões, L. K. Wives, and J. P. M. de Oliveira, “Comparing keywords and taxonomies in the representation of users profiles in a content-based recommender system,” in *Proceedings of the 2008 ACM symposium on Applied computing*, ser. SAC ’08, Fortaleza, Ceara, Brazil: ACM, 2008, pp. 2030–2034, ISBN: 978-1-59593-753-7. DOI: 10.1145/1363686.1364177. [Online]. Available: <http://doi.acm.org/10.1145/1363686.1364177>.
- [7] A. Kodakateri Pudhiyaveetil, S. Gauch, H. Luong, and J. Eno, “Conceptual recommender system for citeseerx,” in *Proceedings of the third ACM conference on Recommender systems*, ser. RecSys ’09, New

- York, New York, USA: ACM, 2009, pp. 241–244, ISBN: 978-1-60558-435-5. DOI: 10.1145/1639714.1639758. [Online]. Available: <http://doi.acm.org/10.1145/1639714.1639758>.
- [8] X. Wan, Q. Jamaliding, F. Anma, and T. Okamoto, “Applying keyword map based learner profile to a recommender system for group learning support,” in *Education Technology and Computer Science (ETCS), 2010 Second International Workshop on*, vol. 1, 2010, pp. 3–6. DOI: 10.1109/ETCS.2010.439.
- [9] N. Rattanajitbanjong and S. Maneeroj, “Multi criteria pseudo rating and multidimensional user profile for movie recommender system,” in *Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on*, 2009, pp. 596–601. DOI: 10.1109/ICCSIT.2009.5234808.
- [10] P. Lops, C. Musto, F. Narducci, M. de Gemmis, P. Basile, and G. Semeraro, “Learning semantic content-based profiles for cross-language recommendations,” in *Proceedings of the First Workshop on Personalised Multilingual Hypertext Retrieval*, ser. PMHR ’11, Eindhoven, Netherlands: ACM, 2011, pp. 26–33, ISBN: 978-1-4503-0897-7. DOI: 10.1145/2047403.2047409. [Online]. Available: <http://doi.acm.org/10.1145/2047403.2047409>.
- [11] P.-Y. Pan, C.-H. Wang, G.-J. Horng, and S.-T. Cheng, “The development of an ontology-based adaptive personalized recommender system,” in *Electronics and Information Engineering (ICEIE), 2010 International Conference On*, vol. 1, 2010, pp. V1–76 V1–80. DOI: 10.1109/ICEIE.2010.5559842.
- [12] P. Chakraborty, “A scalable collaborative filtering based recommender system using incremental clustering,” in *Advance Computing Conference, 2009. IACC 2009. IEEE International*, 2009, pp. 1526–1529. DOI: 10.1109/IADCC.2009.4809245.
- [13] A. Barragans-Martinez, M. Rey-Lopez, E. Costa Montenegro, F. Mikic-Fonte, J. Burguillo, and A. Peleteiro, “Exploiting social tagging in a web 2.0 recommender system,” *Internet Computing, IEEE*, vol. 14, no. 6, pp. 23–30, 2010, ISSN: 1089-7801. DOI: 10.1109/MIC.2010.104.
- [14] P. Bedi and S. Agarwal, “Preference learning in aspect-oriented recommender system,” in *Computational Intelligence and Communication Networks (CICN), 2011 International Conference on*, 2011, pp. 611–615. DOI: 10.1109/CICN.2011.132.
- [15] S. Moursi, M. Elsakhawy, and H. Ghenniwa, “Agent oriented media recommender system utilizing smart multimedia,” in *Computer Supported Cooperative Work in Design (CSCWD), 2011 15th International Conference on*, 2011, pp. 437–444. DOI: 10.1109/CSCWD.2011.5960110.
- [16] P. Alabastro, M. Ang, R. de Guzman, M. Muhi, and M. Suarez, “Mymuseum: integrating personalized recommendation and multimedia for enriched human-system interaction,” in *Digital Content, Multimedia Technology and its Applications (IDC), 2010 6th International Conference on*, 2010, pp. 421–426.
- [17] L. Iaquinta, M. de Gemmis, P. Lops, and G. Semeraro, “Recommendations toward serendipitous diversions,” in *Intelligent Systems Design and Applications, 2009. ISDA ’09. Ninth International Conference on*, 2009, pp. 1049–1054. DOI: 10.1109/ISDA.2009.215.
- [18] J. Liu, P. Dolan, and E. R. Pedersen, “Personalized news recommendation based on click behavior,” in *Proceedings of the 15th international conference on Intelligent user interfaces*, ser. IUI ’10, Hong Kong, China: ACM, 2010, pp. 31–40, ISBN: 978-1-60558-515-4. DOI: 10.1145/1719970.1719976. [Online]. Available: <http://doi.acm.org/10.1145/1719970.1719976>.
- [19] I. Cantador, A. Bellogin, and P. Castells, “Ontology-based personalised and context-aware recommendations of news items,” in *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT ’08. IEEE/WIC/ACM International Conference on*, vol. 1, 2008, pp. 562–565. DOI: 10.1109/WIAT.2008.204.
- [20] K. F. Yeung and Y. Yang, “A proactive personalized mobile news recommendation system,” in *Developments in E-systems Engineering (DESE), 2010*, 2010, pp. 207–212. DOI: 10.1109/DeSE.2010.40.
- [21] S. D. Gollapalli, P. Mitra, and C. L. Giles, “Similar researcher search in academic environments,” in *Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries*, ser. JCDL ’12, Washington, DC, USA: ACM, 2012, pp. 167–170, ISBN: 978-1-4503-1154-0. DOI: 10.1145/2232817.2232849. [Online]. Available: <http://doi.acm.org/10.1145/2232817.2232849>.
- [22] P. Lops, M. de Gemmis, G. Semeraro, P. Gissi, C. Musto, and F. Narducci, “Content-based filtering with tags: the first system,” in *Intelligent Systems Design and Applications, 2009. ISDA ’09. Ninth International Conference on*, 2009, pp. 255–260. DOI: 10.1109/ISDA.2009.84.
- [23] M. Julashokri, M. Fathian, and M. Gholamian, “Improving customer’s profile in recommender systems using time context and group preferences,” in *Computer Sciences and Convergence Information Technology (ICCIT), 2010 5th International Conference on*, 2010, pp. 125–129. DOI: 10.1109/ICCIT.2010.5711042.
- [24] Y. Liangxing and D. Aihua, “Hybrid product recommender system for apparel retailing customers,” in *Information Engineering (ICIE), 2010 WASE International Conference on*, vol. 1, 2010, pp. 356–360. DOI: 10.1109/ICIE.2010.91.
- [25] D. Godoy and A. Amandi, “Hybrid content and tag-based profiles for recommendation in collaborative tagging systems,” in *Latin American Web Conference, 2008. LA-WEB ’08.*, 2008, pp. 58–65. DOI: 10.1109/LA-WEB.2008.15.

- [26] S. Tiwari and S. Kaushik, “A non functional properties based web service recommender system,” in *Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on*, 2010, pp. 1–4. DOI: 10.1109/CISE.2010.5676953.
- [27] D. W. Oard, “The state of the art in text filtering,” *User Modeling and User-Adapted Interaction*, vol. 7, no. 3, pp. 141–178, Mar. 1997, ISSN: 0924-1868. DOI: 10.1023/A:1008287121180. [Online]. Available: <http://dx.doi.org/10.1023/A:1008287121180>.
- [28] D. H. Park, H. K. Kim, I. Y. Choi, and J. K. Kim, “A literature review and classification of recommender systems research,” *Expert Systems with Applications*, vol. 39, no. 11, pp. 10059–10072, 2012, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2012.02.038. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0957417412002825>.
- [29] P. Lops, M. de Gemmis, and G. Semeraro, “Content-based recommender systems: state of the art and trends,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., Springer, 2011, pp. 73–105, ISBN: 978-0-387-85819-7. [Online]. Available: <http://dblp.uni-trier.de/db/reference/rsh/rsh2011.html#LopsGS11>.
- [30] V. M. Filho, “E-recommender: sistema inteligente de recomendação para comércio eletrônico,” *Trabalho de Conclusão do Curso de Engenharia da Computação*, 2006.

**Jaqueline Ferreira de Brito** possui graduação em Sistemas de Informação pela Escola de Artes, Ciências e Humanidades da Universidade de São Paulo - EACH-USP (2010). Em 2012 ingressou no mestrado em Sistemas de Informação na EACH-USP. Tem experiência da área de Inteligência Artificial, com ênfase em Redes Neurais Artificiais e Reconhecimento de Padrões.



**Luciano Antonio Digiampietri (autor correspondente)** possui graduação em Ciência da Computação pela Universidade Estadual de Campinas (2002) e doutorado em Ciência da Computação pela Universidade Estadual de Campinas (2007). Desde abril de 2008 é professor pesquisador no Bacharelado em Sistemas de Informação na Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH-USP) e desde 2010 é docente permante no Programa de Mestrado em Sistemas de Informação da Universidade de São Paulo. Tem experiência na área de Ciência da Computação, com ênfase em Biologia Computacional, Bancos de Dados e Inteligência Artificial, atuando principalmente nos seguintes temas: workflows científicos, bioinformática, proveniência de dados, composição automática de serviços, rastreabilidade de experimentos e algoritmos.

E-mail: [digiampietri@usp.br](mailto:digiampietri@usp.br).