

Estratégia de Pós-processamento Aplicada a um Sistema de Recomendação de Artigos para a Melhora da Diversidade

Ediana da Silva de Souza¹, Daniel Lichtnow², Isabela Gasparini¹

¹Universidade do Estado de Santa Catarina (UDESC) – Joinville – SC – Brasil

²Universidade Federal de Santa Maria (UFSM) – Santa Maria – RS – Brasil

edianadasilvadesouza@gmail.com

daniel.lichtnow@ufsm.br, isabela.gasparini@udesc.br

Abstract. *Recommender Systems seek to identify items that are relevant to their users. One of the recent topics of interest in the Recommender Systems area is the diversity issues. When considering aspects to increase diversity, we try to avoid prioritizing some of the user's preferences or recommending items that are very similar to each other. This work presents the implementation of a post-processing model in a Papers Recommendation System, aiming to promote the diversity of recommended papers. The system was used by 56 users from different institutions, with diversity and precision being evaluated by metrics. The post-processing model increased the diversity of recommendations while maintaining the precision.*

Resumo. *Sistemas de Recomendação buscam identificar itens relevantes para seus usuários. Um dos temas recentes na área de Sistemas de Recomendação são as questões sobre diversidade. Ao considerar aspectos para aumentar a diversidade tenta-se evitar a priorização de algumas das preferências do usuário ou ainda a recomendação de itens muito similares entre si. Este trabalho apresenta a implementação de um modelo de pós-processamento aplicado a um Sistema de Recomendação de artigos, visando a promoção da diversidade das recomendações. O sistema foi utilizado por 56 usuários de diferentes instituições, sendo a diversidade e a precisão avaliadas por métricas. O pós-processamento aumentou a diversidade mantendo a precisão.*

1. Introdução

Sistemas de Recomendação (SRs) fornecem sugestões de itens para seus usuários, sendo hoje incorporados à sites de comércio eletrônico, bibliotecas digitais e redes sociais. Para produzirem recomendações, os SRs consideram informações sobre as preferências dos usuários obtidas a partir da interação dos usuários com o sistema e, em alguns casos partir de outras fontes como as redes sociais [Yang et al. 2014]. Um problema reconhecido em SRs é a chamada superespecialização, que consiste no fato de que ao usuário podem vir a ser recomendados frequentemente muitos itens similares aos já recomendados. A mitigação deste problema, passa pela geração de recomendações de itens com maior diversidade. Este trabalho apresenta a criação de um modelo de pós-processamento aplicado a um SR de artigos que visa a promoção da diversidade das recomendações. Este modelo foi incorporado a um SR que usava uma abordagem Baseada em Conteúdo e

uma etapa de pós-processamento que visava equilibrar as preferências do usuário, onde era apontada a necessidade da aplicação de métricas de diversidade para análise das recomendações produzidas e ainda considerar outros atributos dos itens recomendados para promoção da diversidade [Cezar et al. 2021]. Estratégia similar pode ser aplicada a SRs que utilizam outras abordagens e/ou recomendam outros tipos de itens.

A Seção 2 traz uma revisão dos principais conceitos envolvidos no desenvolvimento da proposta e trabalhos relacionados; a Seção 3 apresenta a estratégia de pós-processamento proposta; a Seção 4 descreve os experimentos feitos com usuários. Por fim, a Seção 5 apresenta as considerações finais.

2. Conceitos e Trabalhos Relacionados

Um dos problemas em SRs, especialmente da abordagem Baseada em Conteúdo, é a chamada superespecialização, que se refere ao fato de que pode acontecer do SR recomendar sempre itens muito similares aqueles pelos quais o usuário demonstrou interesse. Assim, além do SR ser preciso, recomendando o que é do interesse de um usuário, um SR deve buscar atingir outros objetivos como novidade, serendipidade, cobertura e diversidade, que são discutidos em [Kaminskas and Bridge 2016]. O foco do trabalho é diversidade.

Em [Bradley and Smyth 2001] [Kunaver and Požrl 2017], diversidade é definida no contexto de SRs como o oposto de similaridade. [Barraza-Urbina et al. 2015] apresentam duas perspectivas que explicam porque a diversidade é importante para SRs, sendo elas: (1) lidar com a incerteza em torno do perfil do usuário e (2) a importância do tema para o conceito de novidade/serendipidade nos SRs. Ressalte-se que questões de diversidade estão ligadas ao entendimento de sociedade e, conseqüentemente, questões éticas relacionadas são afetadas, questões estas que não são abordadas no presente trabalho.

Oito trabalhos relacionados foram identificados, sendo que dois deles abordam a diversidade em SRs de artigos científicos - Tabela 1. O presente trabalho também foi incluído na tabela para que seja possível identificar a contribuição deste.

Domínio	Trabalho	Contexto	Abordagem	Estratégia	Avaliação	Métricas
2*Artigos Científicos	[Vellino 2010]	Artigos	Filtragem Colaborativa	Alg. Recomendação	Off-line	N/A
	[Küçüküktunç et al. 2013]	Artigos	Baseada em Conteúdo	Alg. Recomendação	On-line	Graph density, Expansion ratio
6*Diversos	[Ziegler et al. 2005]	Livros	Filtragem Colaborativa	Pós-processamento	Off-line e On-line	Intra-list Similarity (ILS)
	[Adomavicius and Kwon 2009]	Filmes	Filtragem Colaborativa	Pós-processamento	Off-line	diversity-in-top-N
	[Kunaver et al. 2014]	Filmes	Híbrida e Baseada em Conteúdo	Pós-processamento	Off-line	Intra-list Diversity (ILD)
	[Antikacioglu and Ravi 2017]	Filmes	Filtragem Colaborativa	Pós-processamento	Off-line	Diversidade agregada, Índice de Gini e Entropia
	[Lunardi et al. 2018]	Cidades Inteligentes	Híbrida	Alg. Recomendação e Pós-processamento	Off-line e On-line	N/A
	[Wang et al. 2020]	Filmes e Perfil de Usuário	Filmes e Perfil de Usuário	Filtragem Colaborativa	Pós-processamento	Off-line e On-line
Artigos Científicos	Proposto	Artigos	Baseada em Conteúdo	Pós-processamento	On-line	Intra-list Diversity (ILD)

Tabela 1. Tabela comparativa dos trabalhos relacionados

Existem problemas relacionados a definição de métricas para avaliar a diversidade em SRs, a comunidade ainda não adotou algumas delas como preferencial

[Kunaver and Požrl 2017]. No trabalho, a diversidade será medida utilizando a métrica de diversidade *Intra-list Diversity - ILD* por estar presente em mais trabalhos relacionados dentre aqueles identificados (ver Tabela 1). Serão usadas as definições apresentadas em [Kaya and Bridge 2019], sendo a ILD calculada usando a Equação 1 onde $d(a, b) = (1 - sim(a, b))/2$, e a similaridade entre itens é medida por uma função que compara os itens (vetores de termos que representam documentos são comparados entre si). No caso do presente trabalho, como em [Kaya and Bridge 2019] é usado o Cosseno. [Vargas 2015] apresenta distância como o complemento da similaridade, onde $|L|$ representa o número de itens recomendados a um usuário e $dist(i, j) = 1 - sim(i, j)$, sendo $sim(i, j)$ a similaridade entre o item i e o item j da lista.

$$ILD(L) = \frac{2}{|L|(|L| - 1)} \sum_{i \in L} \sum_{j \in L, j \neq i} dist(i, j) \quad (1)$$

Assim, a métrica ILD busca medir a diversidade dentro de uma lista de itens recomendados e para isso considera quão diferentes são os itens da lista entre si [Vargas 2015]. Quanto maior o valor obtido no cálculo do ILD, mais diversidade a lista apresenta.

3. Estratégia de Pós-processamento

A proposta de elaboração de um algoritmo de pós-processamento tem como base os trabalhos levantados e o ambiente ao qual o pós-processamento será aplicado - o Sistema de Recomendação de Artigos Científicos descrito em [Cezar et al. 2021]. No presente trabalho, a diversificação da lista de recomendação ocorre através de um algoritmo de pós-processamento, que efetua troca de itens da lista de recomendação levando em consideração dois atributos dos artigos:

1. **Ano de Publicação** dos artigos, onde busca-se diferenciar a lista de artigos recomendados de maneira a impedir um grande número de artigos apenas de um mesmo ano (é assumido que a discussão sobre um tema evolui ao longo dos anos);
2. **Termos/Palavras** presentes nos artigos. É verificado quais são as palavras que compõem o perfil do usuário e as ocorrências de cada palavra do perfil em cada artigo recomendado. Se houver artigos que compartilhem os mesmos termos do perfil, estes serão passíveis de substituição. Assim, busca-se evitar que sejam recomendados muitos artigos sobre o mesmo tema de interesse do usuário.

Esses atributos foram escolhidos por serem, no caso dos termos, os mais relevantes para determinar os interesses do usuário e também pela facilidade de realizar o processamento, uma vez que trabalhar com outros atributos como os nomes dos autores ou instituições, exigiria trabalhar com questões como deduplicação, homônimos, etc.

Inicialmente a recomendação é gerada usando um algoritmo tradicional de Filtragem Baseada em Conteúdo, já usado em [Cezar et al. 2021]. Sobre a lista de itens gerados é medida a diversidade intra-lista baseando-se em um dos dois critérios de diversificação (Ano ou Termos). Após obtido o valor de diversidade (usando a métrica ILD), é analisado se este se encontra em seu valor máximo (igual à 1) ou ainda é possível ser incrementado. Caso esteja em seu valor máximo, a lista é retornada sem ser modificada pelo pós-processamento de diversificação. Caso seja possível diversificá-la, uma busca na lista dos dez itens é feita para que sejam recuperados os artigos que tiveram mais de uma

ocorrência quando baseados em algum dos dois critérios diversificáveis (mesmo ano ou termos em comum considerando os termos presentes no perfil do usuário)

Para diversificar a lista, um novo item é recuperado no conjunto restante (artigos retornados da *query* do sistema mas que não se encontram entre os dez primeiros da lista de recomendação). Este artigo recuperado é inserido no índice do artigo analisado para substituição e verificado se, após essa mudança, a diversidade intra-lista é melhorada, calculando o ILD. O artigo a ser substituído é sempre aquele que está mais no final da lista dos recomendados. Caso o valor do ILD tenha aumentado com o pós-processamento, a nova lista é salva e o pós-processamento de diversificação é encerrado com essa substituição. Caso a diversidade não tenha melhorado, o próximo artigo do conjunto restante é analisado, permanecendo neste fluxo até que todos os artigos desse conjunto sejam testados. Se, ainda assim a diversidade não for melhorada, o próximo índice da lista de artigos muito similares é analisada, novamente tentando substituir com um novo artigo do conjunto restante retornado pelo sistema. A inspiração para utilizar o conjunto restante surgiu do estudo de [Kunaver et al. 2014].

4. Avaliação

Para avaliação da proposta de pós-processamento foi implementado um SR usando HTML, CSS, PHP e também o ElasticSearch [Gormley and Tong 2015], partindo o que havia sido implementado em [Cezar et al. 2021]. Para usar o sistema o usuário deve fazer um pequeno cadastro e para definir o seu perfil escolhe cinco de dez termos apresentados pelo sistema, indicando um grau de interesse por cada termo (estes termos de interesse foram identificados a partir do conteúdo dos documentos existentes na base, sendo que em [Cezar et al. 2021] a forma como estes termos foram obtidos é detalhada). Feito o cadastro, o usuário efetua o login e recebe uma lista com dez artigos recomendados. O objetivo do experimento com usuários reais é possibilitar calcular a *Precision@10* e realizar uma análise comparativa dessa métrica com os resultados apresentados em [Cezar et al. 2021]. O sistema foi usado durante 5 dias por 56 usuários de 12 instituições diferentes.

Para cada usuário que realizou o cadastro no SR, um critério de diversificação era atribuído. Têm-se, portanto, a geração dos dois grupos de avaliação do algoritmo de pós-processamento, sendo um deles baseado no critério de diversificação pelos anos dos artigos publicados, com 28 usuários, e outro pela distribuição dos termos do perfil do usuário nos estudos, com 28 usuários. Das 56 vezes que o algoritmo de recomendação foi executado, o pós-processamento de diversificação atuou em 54 vezes, o que equivale a 96,49%. O valor da diversidade das listas de recomendações geradas foram calculadas usando a métrica ILD.

A respeito dos *feedbacks* explícitos, isto é, aqueles nos quais os usuários deram um parecer sobre ter ou não gostado das recomendações, dos 56 usuários, 53 realizaram a avaliação dos artigos. Cabe ressaltar que para enviar a avaliação, o usuário não era obrigado a fornecer o parecer sobre todas as dez recomendações apresentadas (apenas uma). Com isso, tem-se o número de itens avaliados como gostei, não gostei e não avaliados dentre os 530 itens recomendados (10 para cada um dos 53 usuários). Os resultados são apresentados na Tabela 2, cabendo destacar que a *Precision@10*, considerando o total de itens avaliados como ‘Gostei’ (379 dos 530 itens possíveis) foi de 71,51%.

A Tabela 3 demonstra a comparação entre a precisão obtida no presente trabalho

Avaliação	Quantidade	Percentual
Gostei	379	71.51%
Não Gostei	137	25.85%
Não Avaliado	14	2.64%
Total	530	100%

Tabela 2. Análise da precisão obtida na proposta

quando comparada aos resultados dos trabalhos anteriores usados como *baseline*, apresentados em [Cezar et al. 2021], sendo que o Trabalho 1 consistia de um SR sem pós-processamento e o Trabalho 2 em um SR com pós-processamento, onde não era medida a diversidade da lista para realizar a troca de itens. Como o algoritmo de pós-processamento de diversificação rodou em quase todas as execuções (54 das 56 vezes), isso indica que a precisão não foi afetada pelo algoritmo de pós-processamento.

Algoritmo	Precision@10
Trabalho 1	67.50%
Trabalho 2	66.73%
Proposta	71.51%

Tabela 3. Precisão obtida no presente estudo comparada aos *baseline*

Ainda considerando apenas o item da lista de recomendação substituído pelo pós-processamento implementado neste trabalho, a precisão (a proporção de usuários que avaliaram positivamente o item) foi de 67,31%, que foi superior a precisão obtida pelo item substituído pelo algoritmo de [Cezar et al. 2021], que teve precisão de 55.17%.

5. Considerações Finais

Os resultados preliminares obtidos foram promissores, especialmente pelo fato da precisão não ter sido afetada pela maior diversificação. Em trabalhos futuros, alguns critérios de diversificação podem ser mais explorados, sendo combinados ou utilizados outros atributos presentes nos itens recomendados. Embora aplicada em um SR que utiliza abordagem Baseada em Conteúdo e recomenda artigos, a estratégia de pós-processamento proposta neste trabalho pode ser aplicada dentro de outros domínios onde os SRs estão presentes, como em redes sociais e comércio eletrônico, desde que seja possível que os itens recomendados possam ser comparados entre si por seus atributos. Agradecemos o apoio do CNPq 308395/2020-4, FAPESC N° 027/2020 Apoio a Infraestrutura para Grupos de Pesquisa da UDESC TO n° 2021TR795 e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

- Adomavicius, G. and Kwon, Y. (2009). Toward more diverse recommendations: Item re-ranking methods for recommender systems. In *Workshop on Information Technologies and Systems*, pages 79–84.
- Antikacioglu, A. and Ravi, R. (2017). Post processing recommender systems for diversity. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 707–716.

- Barraza-Urbina, A., Heitmann, B., Hayes, C., and Ramos, A. C. (2015). Xplodiv: Diversification approach for recommender systems. INSIGHT Centre for Data Analytics, National University of Ireland, Galway.
- Bradley, K. and Smyth, B. (2001). Improving recommendation diversity. In *Proceedings of the Twelfth Irish Conference on Artificial Intelligence and Cognitive Science, Maynooth, Ireland*, volume 85, pages 141–152.
- Cezar, N. L., de Borba, C., Gasparini, I., and Lichtnow, D. (2021). Applying a post-processing strategy to consider the multiple interests of users of a paper recommender system. In Araujo, R. D., Dorça, F. A., de Araujo, R. M., Siqueira, S. W. M., and Fontão, A. L., editors, *SBSI 2021: XVII Brazilian Symposium on Information Systems, Uberlândia, Brazil, June 7 - 10, 2021*, pages 49:1–49:7. ACM.
- Gormley, C. and Tong, Z. (2015). *Elasticsearch: the definitive guide: a distributed real-time search and analytics engine*. "O'Reilly Media, Inc."
- Kaminskas, M. and Bridge, D. (2016). Diversity, serendipity, novelty, and coverage: a survey and empirical analysis of beyond-accuracy objectives in recommender systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(1):1–42.
- Kaya, M. and Bridge, D. (2019). Subprofile-aware diversification of recommendations. *User Modeling and User-Adapted Interaction*, 29(3):661–700.
- Küçükünç, O., Saule, E., Kaya, K., and Çatalyürek, Ü. V. (2013). Result diversification in automatic citation recommendation. In *Proceedings of the iConference workshop on Computational scientometrics: theory and applications*, pages 1–4.
- Kunaver, M., Dobravec, S., Požrl, T., and Kosir, A. (2014). Increasing top-20 search results diversity through recommendation post-processing. In *UMAP Workshops*.
- Kunaver, M. and Požrl, T. (2017). Diversity in recommender systems – a survey. *Knowledge-Based Systems*, 123:154–162.
- Lunardi, G. M., Machado, G. M., and de Oliveira, J. P. M. (2018). Diversificação de recomendação em cidades inteligentes: Estudo e estrutura de abordagem. *Cadernos de Informática*, 10(1):28–44.
- Vargas, S. (2015). *Novelty and diversity evaluation and enhancement in recommender systems*. PhD thesis, PhD thesis, Universidad Autónoma de Madrid, Spain.
- Vellino, A. (2010). A comparison between usage-based and citation-based methods for recommending scholarly research articles. *Proceedings of the American Society for Information Science and Technology*, 47(1):1–2.
- Wang, Y., Zhang, X., Liu, Z., Dong, Z., Feng, X., Tang, R., and He, X. (2020). Personalized re-ranking for improving diversity in live recommender systems. *arXiv preprint arXiv:2004.06390*.
- Yang, X., Guo, Y., Liu, Y., and Steck, H. (2014). A survey of collaborative filtering based social recommender systems. *Computer communications*, 41:1–10.
- Ziegler, C.-N., McNee, S. M., Konstan, J. A., and Lausen, G. (2005). Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, pages 22–32.