



AVALIAÇÃO DE TÉCNICAS DE FILTRAGEM COLABORATIVA
PARA SISTEMAS DE RECOMENDAÇÃO

PROPOSTA DE TRABALHO DE GRADUAÇÃO

Aluno: Flávio de Holanda Cavalcanti Júnior (fhcj2@cin.ufpe.br)
Orientador: Germano Crispim Vasconcelos (gcv@cin.ufpe.br)
Área: Inteligência Artificial

21 de Abril de 2014

1. Contexto

Com a explosão de informações da era digital, principalmente após o advento da Web, usuários encontram-se constantemente expostos a uma imensa variedade de informações, das quais apenas uma pequena fração têm a relevância adequada às suas reais necessidades. Estima-se que a quantidade de dados armazenada em bancos de dados ao redor do mundo é duplicada a cada 20 meses [8], tornando vital aprimorar-se técnicas de filtragem de grandes quantidades de conteúdo, almejando refinar tais dados e delegar ao usuário apenas informações (potencialmente) relevantes, ao encontro de seu perfil e/ou necessidade "individual". Outra necessidade desse tipo de serviço vem do interesse empresarial, que ao disponibilizar recomendações direcionadas ao usuário com uma maior probabilidade de aceitação aumentam também a sua receita, pelo maior número de aquisições de seus produtos. As empresas vêm se adequando ao conceito de que não basta fornecer um único serviço padrão a todos os usuários, mas sim múltiplos serviços, que satisfaçam as múltiplas necessidades de múltiplos usuários [1], a intenção está em personalizar serviços ofertados em função das necessidades/preferências de cada um de seus clientes/usuários, um princípio bem colocado em uma afirmação feita pelo CEO da Amazon, Jeff Bezos, que disse: *"If I have 2 million customers on the Web, I should have 2 million stores on the Web"* [1], traduzido como "Se eu tenho 2 milhões de clientes na Web, eu deveria ter 2 milhões de lojas na Web" (tradução nossa).

Um tipo de ferramenta que propõe resolver essa questão da sobrecarga de informação é o chamado Sistema de recomendação [3, 4]. São ferramentas capazes de providenciar recomendações de "itens" a usuários de uma forma personalizada, isto é, de sugerir os itens mais prováveis de satisfazerem interesses individuais do usuário de forma prioritária aos demais, ou seja, itens considerados interessantes são sugeridos ao usuário com maior destaque que itens "menos" interessantes. Um sistema de recomendação pode sugerir, por exemplo, aquisições de itens como livros, músicas, notícias, etc. para um usuário específico com base em suas preferências e restrições pessoais, características essas que, por sua vez, poderiam ser obtidas de forma explícita (através de feedback do usuário) ou implícita (monitorando a atividade do usuário durante a utilização do sistema). Um meio bastante comum de se encontrar aplicações de sistemas de recomendação são os serviços de *e-commerce*, como no portal online da Amazon.com, auxiliando o usuário no fornecimento de recomendações das (possíveis) melhores aquisições de produtos de seu interesse, em acordo com o seu perfil ou em decorrência do histórico de produtos já adquiridos pelo usuário; pode-se observar também a aplicação de técnicas de recomendação em serviços de outras empresas famosas como: Yahoo, Google, Netflix, eBay, etc. [1, 2, 6, 11]. Estima-se que 60% das visualizações do YouTube sejam de recomendações providas pelo serviço [2].

As técnicas utilizadas por sistemas de recomendação podem ser dissociadas em três vertentes: filtragem colaborativa, filtragem baseada em conteúdo, e filtragem por abordagem híbrida. Na filtragem colaborativa itens recomendados ao usuário são feitos levando em consideração os interesses de "usuários semelhantes" ao corrente; na filtragem baseada em conteúdo itens são sugeridos com base em "propriedades" ou "conteúdos" semelhantes aos de itens já anteriormente avaliados positivamente

pelo usuário; nos sistemas híbridos há uma mesclagem das duas classificações anteriores, tentando conciliar suas vantagens e mitigar desvantagens [1, 2, 6]. Analisaremos no decorrer deste documento algumas das técnicas de recomendação na vertente de filtragem colaborativa, comumente aplicadas em sistemas de e-commerce e, ao final, faremos uma experimentação com uma dessas técnicas sobre uma base de dados de teste.

2. Objetivo

Este trabalho busca analisar técnicas de recomendação utilizadas em sistemas de recomendação [3][4], com foco na abordagem por “filtragem colaborativa” [1][2][6], mais comumente encontrada em sistemas de *e-commerce*. Nesse intuito, serão pesquisadas técnicas de recomendação neste teor e então experimentada uma dessas técnicas sobre uma base de dados de teste.

3. Cronograma

Atividade	Mês												
	Abril			Maio			Junho			Julho			
Levantamento e Estudo do Estado da Arte	█	█	█	█	█								
Implementação da técnica de Recomendação				█	█	█	█	█					
Escrita do relatório					█	█	█	█	█	█			
Preparação da apresentação oral										█	█	█	

4. Lista de possíveis avaliadores

- Cleber Zanchettin (cz@cin.ufpe.br)
- George Darmiton da Cunha Cavalcanti (gdcc@cin.ufpe.br)

5. Referências

- [1] SCHAFER, J. Ben; KONSTAN, Joseph; RIEDL, John. *Recommender Systems in E-Commerce*.
- [2] DAVIDSON, James et al. *The YouTube Video Recommendation System*.
- [3] LESKOVEC, Jure; RAJARAMAN, Anand; ULLMAN, Jeff. *Mining of Massive Datasets*. 2ed. Disponível em: <http://www.mmds.org/#ver21>.
- [4] BEEL, Joeran et al. *Research Paper Recommender System Evaluation: A Quantitative Literature Survey*.
- [5] Ben-Gal I. *Bayesian Networks*. Disponível em: <http://www.eng.tau.ac.il/~bengal/BN.pdf>
- [6] LINDEN, Greg; SMITH, Brent; YORK, Jeremy. *Amazon.com Recommendations: Item-to-Item Collaborative Filtering*. Disponível em: <http://www.cin.ufpe.br/~idal/rs/Amazon-Recommendations.pdf>
- [7] PIATETSKY, Gregory; Interview with Simon Funk. KDNuggets.
- [8] WITTEN, Ian H.; FRANK, Eibe; HALL, Mark A.. *Data Mining: Practical Machine Learning Tools and Techniques*. 3rd Ed. 2011
- [9] CANDILLIER, Laurent; MEYER, Frank; BOULLÉ, Marc. *Comparing State-of-the-Art Collaborative Filtering Systems*. France Telecom R&D Lannion, France.
- [10] SU, Xiaoyuan; KHOSHOTAAR, Taghi M.. *A Survey of Collaborative Filtering Techniques*. Department of Computer Science and Engineering, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA
- [11] PORIYA, Anil et al. *Non-Personalized Recommender Systems and User-based Collaborative Recommender Systems*. International Journal of Applied Information Systems (IJ AIS). Foundation of Computer Science FCS, New York, USA. Volume 6 – No. 9, March 2014.

6. Assinaturas

Germano Crispim Vasconcelos
(Orientador)

Flávio de Holanda Cavalcanti Júnior
(Aluno)