



Escola Politécnica da Universidade de São Paulo
PMR2500 – PROJETO DE CONCLUSÃO DO CURSO I

Relatório de Revisão Bibliográfica

Nome

Antônio Guilherme Ferreira Viggiano

Fernando Fochi Silveira Araújo

Número USP

6846450

5894546

Orientador

Prof. Dr. Fábio Gagliardi Cozman

18 de março de 2014

Sumário

1	Objetivo do trabalho	2
2	Desafios científico-tecnológicos / Metodologia	2
3	Palavras-chave (Português-Inglês)	2
4	Procedimento de busca	3
5	Estado da arte do problema. Estado da arte das soluções	3
	Referências	5

1 Objetivo do trabalho

O objetivo do presente Trabalho de Conclusão de Curso é o desenvolvimento de um Sistema de Recomendação de produtos para lojas de comércio online, com foco na sugestão feita através de marketing via e-mail.

2 Desafios científico-tecnológicos / Metodologia

Um dos maiores desafios tecnológicos dos sistemas de recomendação é, atualmente, o da escalabilidade. O sistema de recomendação deverá ser flexível no sentido de poder operar igualmente bem tanto em conjuntos pequenos de itens e clientes quanto em grandes bases de dados, que podem chegar até centenas de milhões de usuários [1]. Isso significa que as recomendações devem ser suficientemente rápidas para poderem operar em tempo real e ainda assim proverem sugestões valiosas aos clientes.

O sistema de recomendação também deve prever quando enviar uma determinada recomendação, e não agir apenas mediante requisição do cliente [2]. É interessante, por exemplo, enviar recomendações de produtos com descontos a usuários que estão há algum tempo inativos no site, para que eles retornem a comprar. Da mesma forma, um sistema inteligente poderia sugerir produtos infantis a um usuário detectado como recém-casado.

Outro desafio científico ainda em estágio inicial de pesquisa é referente à diversidade das recomendações realizadas. Ao mesmo tempo que o sistema deve apresentar itens similares ao que o usuário está procurando, ele também deve sugerir itens que o usuário desconhecia ou nem sabia que poderiam interessá-lo.

Por fim, um desafio científico que este trabalho enfrentará é a execução de um sistema híbrido do ponto de vista de efemeridade e persistência [3], ao construir um modelo de recomendação que integre entre as preferências de curto e longo termo dos usuários. A análise dos dados de compras anteriores, bem como de dados demográficos [4], deverá portanto ser incorporada à análise de característica dos produtos, a fim de enriquecer a acurácia do sistema.

Esse tópico de pesquisa inclui diversos outros desafios científicos e tecnológicos, tais como a preservação da privacidade dos usuários, a criação de modelos de recomendação inter-domínios, o desenvolvimento de sistemas descentralizados operando em redes computacionais distribuídas, a otimização de sistemas para sequências de recomendações e ainda a otimização de sistemas para dispositivos e aplicativos móveis. Entretanto, esses desafios são menos relevantes ao nosso projeto porque não se aplicam diretamente à recomendação de produtos de e-commerces via e-mail, ou porque não se encaixam no formato dos dados que serão utilizados para análise.

3 Palavras-chave (Português-Inglês)

As palavras chave principais desse Trabalho de Conclusão de Curso são:

- Sistema de recomendação (recommender system)

- e-commerce (*idem*)
- Recomendação baseada em conteúdo (content-based recommendation)
- Filtragem colaborativa (collaborative filtering)
- Recomendação híbrida (hybrid recommendation)

4 Procedimento de busca

A busca por artigos científicos começou com a recomendação de livros didáticos pelo professor orientador ([2] e [5]). A partir da leitura preliminar dos capítulos introdutórios, observamos as palavras-chave ligadas à temática de Sistemas de Recomendação e iniciamos a pesquisa por artigos especializados.

No website Google Scholar buscamos as palavras-chave indicadas na seção e priorizamos os artigos mais citados, e principalmente os documentos de *surveys* e *reviews*.

Na base de dados IEEE Explorer também procuramos pelas palavras chave e selecionamos os artigos mais citados, mas filtramos por data a fim de obter apenas resultados recentes (artigos mais recentes que 2010).

Uma vez obtidos os documentos, avaliamos se o conteúdo de cada um deles era relevante ao escopo do trabalho. Descartamos artigos que se baseavam inteiramente em filtragem colaborativa baseada em usuários, tais como [6], pois nós não dispomos dessa quantidade de dados sobre os clientes. Em vez disso, priorizamos artigos que exploram técnicas de recomendação baseadas em itens ou mesmo técnicas híbridas, como [7], [8] e [9].

5 Estado da arte do problema. Estado da arte das soluções

As técnicas de recomendação baseada em conteúdo exploram os dados dos itens para calcular a sua relevância baseada no perfil do usuário. As pesquisas nesta área são realizadas em diversas frentes como a de Recuperação de Informação e Inteligência Artificial.

Os sistemas baseados em Recuperação de Informação partem do princípio que os usuários interessados em recomendações estão engajados em um processo de busca de informação. Em sistemas de Recuperação de Informação o usuário fornece uma lista de palavras-chave para uma busca específica, enquanto nos sistemas baseados em filtragem de conteúdo a informação base para a recomendação é o perfil do usuário. Os itens a serem recomendados podem possuir diversos atributos e formas de classificação, cada item pode ser descrito por uma pequena quantidade de atributos com valores conhecidos, porém, documentos, emails, websites ou reviews de usuários são compostos por textos sem estrutura definida e a abordagem por Recuperação de Informação é mais recomendada [10].

Na abordagem por Inteligência Artificial, a recomendação pode ser vista como um problema de aprendizado que explora os conhecimentos sobre o usuário. Na

sua forma mais simples, os perfis de usuários estão especificados por palavras-chave que refletem os interesses de longo prazo dos usuários. Muitas vezes é recomendado o aprendizado sobre o perfil do usuário com o seu uso contínuo em vez de forçar o usuário a responder diversas perguntas sobre seu perfil. Isto geralmente envolve problemas de aprendizado da máquina, onde o objetivo é aprender a categorizar novas informações baseado em informações previamente adquiridas e implícita ou explicitamente rotuladas como interessantes ou não pelo usuário.

Com estas informações em mão, métodos de aprendizado de máquina são capazes de gerar modelos preditivos que, com nova informação, recomendará um item que tem mais chances de ser do interesse do consumidor.

Referências

- [1] Leslie Tutol. *Amazon Launches ‘Login and Pay with Amazon’ for a Seamless Buying Experience*. Novembro de 2013. URL: <http://services.amazon.com/post/Tx2A98P3EKP6202/Amazon-Launches-Login-and-Pay-with-Amazon-for-a-Seamless-Buying-Experience>.
- [2] Pasquale Lops, Marco de Gemmis e Giovanni Semeraro. Em: *Recommender Systems Handbook*. Springer, 2011.
- [3] J Ben Schafer, Joseph Konstan e John Riedl. “Recommender systems in e-commerce”. Em: *Proceedings of the 1st ACM conference on Electronic commerce*. ACM. 1999, pp. 158–166.
- [4] Kangning Wei, Jinghua Huang e Shaohong Fu. “A survey of e-commerce recommender systems”. Em: *Service Systems and Service Management, 2007 International Conference on*. IEEE. 2007, pp. 1–5.
- [5] Toby Segaran. *Programming Collective Intelligence*. First. O’Reilly, 2007. ISBN: 9780596529321.
- [6] Pooyan Adibi e Behrouz Tork Ladani. “A collaborative filtering recommender system based on user’s time pattern activity”. Em: *Information and Knowledge Technology (IKT), 2013 5th Conference on*. IEEE. 2013, pp. 252–257.
- [7] Greg Linden, Brent Smith e Jeremy York. “Amazon.com recommendations: Item-to-item collaborative filtering”. Em: *Internet Computing, IEEE 7.1* (2003), pp. 76–80.
- [8] Badrul Sarwar et al. “Item-based collaborative filtering recommendation algorithms”. Em: *Proceedings of the 10th international conference on World Wide Web*. ACM. 2001, pp. 285–295.
- [9] Lei Ren, Junzhong Gu e Weiwei Xia. “An item-based collaborative filtering approach based on balanced rating prediction”. Em: *Multimedia Technology (ICMT), 2011 International Conference on*. IEEE. 2011, pp. 3405–3408.
- [10] J BEN SCHAFER, JOSEPH A KONSTAN e JOHN RIEDL. “E-Commerce Recommendation Applications”. Em: *Data Mining and Knowledge Discovery* 5 (2001), pp. 115–153.