

1

Introdução

1.1

Problemas a resolver e relevância

O volume de informações disponíveis em forma digital no ano de 2007 é de 281 exabytes¹, e a previsão é que em 2011 chegará a dez vezes esse valor. Temos atualmente (2008) 1,4 bilhão de pessoas (1/4 da população mundial) conectados na internet e em 2012 esse número chegará a 1,9 bilhão (1/3 do mundo).

Assim, técnicas de recomendação de itens para usuários podem se mostrar bastante úteis à medida em que o número de itens em uma categoria de dados (como música, filmes, livros, notícias, páginas web) se torna tão grande que um única pessoa não possa avaliar todas para poder selecionar as que são relevantes. Abordagens mais simples, como um sistema baseado em pontuações ou *ratings* médios fornecido por todos os usuários do sistema ignoram interesses específicos de um usuário e são particularmente pobres em contextos onde existe uma grande variação de interesse entre os usuários (como para a recomendação de músicas por exemplo).

Existem outros métodos para combater a explosão no volume de informação disponível, muitos deles oriundos do campo conhecido como recuperação de informação: indexação de termos, clusterização de dados, análise semântica latente [Deerwester et al., 1990], redes neurais, classificadores bayesianos [Freitag, 2000] etc, no entanto este trabalho focará no estudo de sistemas recomendadores, especialmente no contexto de redes sociais online.

Sistemas recomendadores, especialmente aqueles baseados em filtragem colaborativa existem há bastante tempo [Resnick and Varian, 1997]. Em sua essência, eles dependem de alguma relação entre os indivíduos para obter conhecimento adicional sobre os itens a serem recomendados, além das relações diretas (similaridade) entre esses itens.

Recentemente, observamos a emergência de sites de mídia social online

¹*The Diverse and Exploding Digital Universe* — IDC — março/2008

como *Flickr*², *Last.FM*³ e outros, que aglutinam pessoas a redor de um interesse comum — fotografias ou músicas nos exemplos citados. Tais serviços online naturalmente induzem uma variedade de problemas para recomendação, que naturalmente emergem da estrutura de sua rede social. Por exemplo no *Flickr*, recomendar um grupo para uma foto, ou uma pessoa como contato para outra, ou uma foto para uma pessoa etc...

Além dessas tarefas de recomendação, serviços de rede e mídia social oferecem também a oportunidade de fazer uso de conhecimento indireto disponível nas diversas relações entre as entidades dessa rede. Por exemplo, fazer uso de contatos de usuários ao recomendar uma pessoa para outra, ou suas fotos favoritas, tags usadas para descrever itens e diversas outras combinações.

Durante esse estudo, à medida em que fomos compreendendo essas tarefas, desenvolvemos uma *framework* que permite uma caracterização clara e sistemática das tarefas de recomendação no contexto de uma rede social online, baseado em ontologias abertas para a descrição desses conteúdos. Além disso, desenvolvemos nesse trabalho também uma *framework* em software que permite a instanciação de uma classe de recomendadores baseado na identificação e extensão dessas relações comuns entre entidades básicas presentes em diversos serviços de mídia e rede social.

1.2

Objetivos desse trabalho

Destacamos assim os seguintes objetivos para essa dissertação:

1. Adquirir conhecimentos sobre o assunto — recomendadores baseado em conteúdo e redes sociais — através de pesquisas na literatura;
2. Estudar a relação entre a construção de sistemas recomendadores e redes sociais online, principalmente tendo em vista as novas possibilidades proporcionadas por tecnologias da web semântica;
3. Investigar o emprego de recomendadores baseado em conteúdo para mídia social, buscando responder perguntas como: Quais são os objetivos de recomendações em redes sociais? Quais tipos de informações podem ser recuperadas nessas redes? Como os dados disponíveis em redes sociais podem ser utilizados em sistemas recomendadores? Quais métodos de recomendação podem ser aplicados em redes sociais? Quais os padrões de tarefas recorrentes na análise e construção de recomendadores para redes sociais?

²<http://flickr.com>

³<http://last.fm>

1.3

Organização da dissertação

Os capítulos restantes abordarão os seguintes assuntos:

Capítulo 2 Traz os conceitos teóricos utilizados nessa tese, realizando levantamento e comparativos sobre trabalhos nas áreas de redes sociais, recomendadores, recuperação de informações etc.

Capítulo 3 Descreve uma *framework*⁴ (ou arcabouço) para auxiliar na modelagem de redes sociais e identificação das tarefas de recomendação, incluindo também a construção e avaliação de sistemas de recomendação. Contará com a descrição das melhores práticas, entidades, processos e interfaces presentes na *framework*.

Capítulo 4 Apresenta um exemplo de instanciação desta *framework* para a rede social *Flickr*⁵, que é um sistema online para compartilhamento de fotos. Inclui também os resultados obtidos com o emprego de recomendadores para itens dessa rede.

Capítulo 5 Traz um resumo dos resultados obtidos e propostas para melhorias e trabalhos futuros.

Apêndices Apresentaremos o material auxiliar para o melhor entendimento do conteúdo principal.

⁴pode ser definida como uma abstração que une códigos comuns entre vários projetos de software provendo uma funcionalidade genérica. Um *framework* pode atingir uma funcionalidade específica, por configuração, durante a programação de uma aplicação.

⁵<http://www.flickr.com/>