



Rafael Jessen Werneck de Almeida Martins

**Recomendação de pessoas em redes sociais
com base em conexões entre usuários**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio.

Orientadora: Profa. Karin Koogan Breitman

Rio de Janeiro
Março de 2013



Rafael Jessen Werneck de Almeida Martins

**Recomendação de pessoas em redes sociais
com base em conexões entre usuários**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Profa. Karin Koogan Breitman

Orientadora

Departamento de Informática – PUC-Rio

Prof. Marco Antonio Casanova

Departamento de Informática – PUC-Rio

Prof. José Viterbo Filho

Universidade Federal Fluminense

Carolina Howard Felicíssimo

Schlumberger Serviços de Petróleo – Matriz

Prof. José Eugenio Leal

Coordenador Setorial do Centro

Técnico Científico – PUC-Rio

Rio de Janeiro, 5 de março de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e da orientadora.

Rafael Jessen Werneck de Almeida Martins

Graduou-se em Engenharia de Computação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) em 2009.

Ficha Catalográfica

Martins, Rafael Jessen Werneck de Almeida

Recomendação de pessoas em redes sociais com base em conexões entre usuários / Rafael Jessen Werneck de Almeida Martins ; orientadora: Karin Koogan Breitman. – 2013.

59 f. : il. (color.) ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2013.

Inclui bibliografia

1. Informática – Teses. 2. Engenharia de software. 3. Internet. 4. Site. 5. Rede social. 6. Sistema de recomendação. 7. Banco de dados. I. Breitman, Karin Koogan. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Aos meus pais, Manuel e Magali.

Agradecimentos

À Professora Noemi Rodriguez pelo incentivo à realização de meu mestrado.

À minha orientadora Professora Karin Koogan Breitman pelos ensinamentos, estímulo e parceria para a realização deste trabalho.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Ao Felipe Nascimento e toda equipe do Peladeiro pelo apoio e confiança.

À minha família pela educação e carinho.

Resumo

Martins, Rafael Jessen Werneck de Almeida; Breitman, Karin Koogan. **Recomendação de pessoas em redes sociais com base em conexões entre usuários**. Rio de Janeiro, 2013, 59 p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Os *sites* de redes sociais ganharam importância nos últimos anos. Neles, seus usuários podem se conectar com outros usuários para interagir entre si. Porém, geralmente o número de usuários cadastrados é muito grande e, por isso, encontrar outros usuários afins para se conectar não é fácil. Sistemas de recomendação são ferramentas de software que provem sugestões de diversos tipos de itens para usuários e podem ser aplicados para recomendar pessoas (outros usuários) em redes sociais. Sistemas que recomendam pessoas utilizam técnicas específicas e, devido às implicações sociais envolvidas nas relações pessoais, devem levar alguns fatores em consideração. A ausência de dados confiáveis torna a tarefa de gerar recomendações úteis mais difícil. Este trabalho discute o assunto e apresenta um sistema de recomendação de pessoas para *sites* de redes sociais com base em conexões entre usuários. Para testar o sistema apresentado, realizamos um experimento com o Peladeiro, um *site* real de uma rede social que conta com mais de 500 mil usuários, onde poucos dados estão disponíveis para serem utilizados.

Palavras-chave

Informática; Engenharia de Software; Internet; Site; Rede Social; Sistema de Recomendação; Banco de Dados.

Abstract

Martins, Rafael Jessen Werneck de Almeida; Breitman, Karin Koogan (Advisor). **People recommendation in social networks based in user connections**. Rio de Janeiro, 2013, 59 p. MSc. Dissertation – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Social networking websites have gained importance in recent years. In them, users can connect with other users to interact with. However, generally, the number of registered users is very large. Therefore, find other users with common interests is not easy. Recommender systems are software tools which generate suggestions for various types of items to users and can be applied to recommend people (other users) on social networks. Systems that recommend people use specific techniques and, due to the social implications involved in personal relationships, must take several factors into consideration. The lack of data available makes the task of generate good recommendations more difficult. This paper discusses the theme and presents a person recommendation system for social networking websites based in user connections. To test the system presented, we conducted an experiment with Peladeiro, a real website of a social network that has over 500 000 users, where few reliable data are available.

Keywords

Informatics; Software Engineering; Internet; Site; Social Network; Recommendation System; Database.

Sumário

1 Introdução	11
1.1. Motivação	11
1.2. Enunciado do Problema	12
1.3. Principais Contribuições	13
2 Sistemas de recomendação	14
2.1. Introdução	14
2.2. Categorias	15
2.3. Recomendação de pessoas	15
2.4. Trabalhos relacionados a sistemas de recomendação	16
2.4.1. MySpace	18
2.4.2. Beehive	18
2.5. Bancos de dados NoSQL e Neo4j	19
3 Sistema de recomendação proposto	21
3.1. Introdução	21
3.2. Arquitetura e implementação	25
3.3. Limitações	30
3.3.1. Recomendações para usuários sem conexões	30
3.3.2. Performance e escalabilidade	30
3.3.3. Identificação da recomendação	31
4 Testes e experimentos realizados	32
4.1. Implementação e banco de dados	32
4.2. Performance e Escalabilidade	39
5 Estudo de Caso: Peladeiro	41
5.1. Introdução	41
5.2. Sobre o <i>síte</i> Peladeiro	43
5.3. Inserção do sistema de recomendação no Peladeiro	43
5.4. Resultados das interações com os usuários	45
6 Conclusão	47
6.1. Resumo do trabalho	47
6.2. Principais contribuições	48
6.3. Trabalhos Futuros	48
7 Referências bibliográficas	50
8 Apêndices	53
8.1. Descrição dos casos de uso do sistema de recomendação	53
8.2. Descrição dos arquivos do sistema de recomendação	56
8.3. Consulta SQL da recomendação	57
8.4. Implementação Neo4j	59

Lista de figuras

Figura 1 - Exemplo de grafo de uma rede social	21
Figura 2 - Lista de recomendações de pessoas do Facebook	24
Figura 3 - Arquitetura: cliente e servidor	26
Figura 4 - Inserção do sistema de recomendação na página HTML	26
Figura 5 - Exemplo de Json retornado pelo servidor	26
Figura 6 - Lista de recomendados exibida para o usuário	27
Figura 7 - Casos de uso do sistema de recomendação	28
Figura 8 - Tabela utilizada pelo sistema de recomendação	28
Figura 9 - Feedback exibido para usuário	28
Figura 10 - Amigos em um grafo	33
Figura 11 - Diagrama UML das entidades do programa de testes	33
Figura 12 - Implementação 1 (PostgreSQL)	34
Figura 13 - Implementação 2 (consulta direta em PostgreSQL)	35
Figura 14 - Implementação 3 (repetição da primeira, em Neo4j)	35
Figura 15 - Implementação 4 (utilizando a API do Neo4j)	36
Figura 16 - Esquemas das tabelas do <i>site</i> Peladeiro	37
Figura 17 - Tempos de execução das implementações	38
Figura 18 - Tempos de execução para quantidades de usuários/amigos variadas	40
Figura 19 - Página inicial do usuário do <i>site</i> Peladeiro	41
Figura 20 - Exemplo de perfil com tom humorístico	42
Figura 21 - Recomendação de pessoas no Peladeiro	44
Figura 22 - Detalhe da lista de recomendação no Peladeiro	44
Figura 23 - Interações com o sistema de recomendação	45

Lista de tabelas

Tabela 1 - Tempos de execução das implementações	37
Tabela 2 - Tempos de execução para quantidades de usuários/amigos variadas	40