

## 2 Conceitos

### 2.1 Redes

Segundo [Newman, 2003], uma rede é um conjunto de itens, chamados vértices ou nós, interconectados através de arestas. Sistemas com a forma de redes (ou grafos) emergem com frequência no mundo. Exemplos incluem a internet, a WWW, redes de empresas consumidoras e prestadoras de serviços, sistemas de logística, cadeias de presas e predadores, sistemas metabólicos, relacionamentos entre cientistas para co-autorias de artigos etc.

Uma revisão extensa das diversas classificações de redes e suas manifestações na natureza e nas criações humanas pode ser vista em [Albert and Barabasi, 2002].

### 2.2 Redes sociais

Uma rede social é caracterizada por um conjunto de pessoas com algum padrão de contato ou interações entre eles. [Wasserman et al., 1994, Scott, 2000] pode ser entendido também como qualquer rede cujos inter-relacionamentos derivem de padrões emergentes encontrados na sociedade humana.

Um serviço online para redes sociais tem como objetivo principal auxiliar o estabelecimento e registro de contatos entre usuários e grupos que compartilhem interesses e atividades. A maior parte desses serviços oferecem uma variada gama de formas de interação, armazenagem e compartilhamento de informações como bate-papo, troca de mensagens, email, video, conversa usando voz, troca de arquivos, blogging, grupos de discussão, rastros de atividades realizadas online e assim por diante.

#### 2.2.1 Exemplos

Podemos citar os padrões de amizade entre indivíduos, relações comerciais entre empresas, casamentos entre famílias, relações de co-autoria em artigos

científicos, padrões para troca de emails etc são exemplos de redes sociais extensamente estudadas no passado.

Como exemplos de serviços online focados em redes sociais temos o *Last.fm*, *Facebook*, *Orkut*, *Flickr*, *MySpace*.<sup>1</sup>

### 2.2.2

#### Caracterização e métricas

Em [Newman, 2003] são enumeradas diversas métricas e conceitos que podem ser usados para caracterizar redes sociais. Segundo [Miquel], algumas dessas métricas tem maior significado para o estudo de sistemas recomendadores que atuam sobre redes sociais. São elas:

**Tamanho** é o número de nós presentes na rede, sendo útil no cálculo de outras medidas. Ele nos dá uma ideia geral de como a rede social se apresenta. Por exemplo, para a tarefa de recomendação de funcionarios numa empresa com perfis e habilidades semelhantes ao usuário alvo, imagine uma empresa com dez trabalhadores. Nela provavelmente todos se conhecem e não haveria muito sentido no uso de sistemas que recomendam funcionários com interesses em comum. No entanto numa empresa com mais de mil trabalhadores tal recomendação faria sentido. E também, em redes sociais deste tamanho os funcionários seriam fragmentados em varios grupos e a proporção de elos entre as pessoas cairia e poderíamos esperar o surgimento de grupos particionados (componentes desconexos).

**Densidade** é a proporção entre todos os elos que poderiam existir e os que de fato existem interligando atores dessa rede. Uma densidade baixa indica alguma dificuldade inerente à rede para que os usuários estabeleçam elos entre sí. Uma densidade muito baixas para a rede social como um todo, para uma amostra do grafo de treinamento ou para a vizinhança de um usuário ou item alvo podem diminuir a taxa de confiança em relação às recomendações feitas quando o grafo de relacionamentos entre usuários é usado como entrada para o recomendador.

**Transitividade ou *Clustering*** seria a taxa da ocorrência de “triângulos” no grafo da rede. Em redes sociais, se o nó *A* está conectado a *B* e este a *C*, é esperado uma probabilidade maior que *A* também esteja conectado a *C* ao invés de um nó *D* aleatório em outra região do grafo. Segue definição matemática para o grau de *clustering* de uma rede:

$$\bar{C} = \frac{3 * \text{número de triângulos na rede}}{\text{número de triplas de vértices conectadas}} \quad (2-1)$$

<sup>1</sup><http://facebook.com>, <http://orkut.com> e <http://myspace.com>

Alternativamente,  $\bar{C}$  mede a taxa de triplas que possuem seu terceiro nó inserido no grafo para completar um triângulo. Assim como a medida de densidade, uma baixa transitividade pode implicar em resultados ruins para recomendadores baseados no grafo social, uma vez que a vizinhança de cada usuário seria relativamente menor, havendo também menor formação de clusters de usuários com interesses comuns (usuários com interesses comuns têm a tendência de se unir em comunidades).

**Distribuição de graus** também chamado de “centralidade de graus”, mede a frequência de ocorrência de nós na rede com determinados grau<sup>2</sup>. Definimos  $p_k$  como sendo a fração de vértices na rede que possuem grau  $k$ . Uma análise gráfica de  $p_k$  então pode ser feita com um histograma dos graus dos vértices. Uma maneira alternativa para apresentar dados de grau de uma rede social é fazer uma plotagem da função de distribuição cumulativa

$$P_k = \sum_{k'=k}^{\infty} p_{k'}, \quad (2-2)$$

que representa a probabilidade que o grau é maior ou igual a  $k$ . Heurísticas usadas na implementação de recomendadores podem utilizar a média de graus da vizinhança de um usuário ou o grau do próprio usuário como uma medida de sua autoridade ou relevância no momento em que os ratings dados por esse usuário são considerados para a seleção dos itens candidatos à recomendação.

**Participação em cliques** representa a tendência de uma rede em formar comunidades. Numa rede social, a existência de muitos nós com alta participação em *cliques*<sup>3</sup> indica a existência de comunidades pois os nós são altamente relacionados entre si. O impacto dessa métrica na implementação de recomendadores é semelhante ao das métricas de transitividade ou *clustering*.

### 2.2.3

#### Redes sociais semânticas

Conforme vimos, há atualmente um grande número de serviços online para compartilhamento de mídia social como mídia social entendemos posts em blogs, páginas wiki [Leuf and Cunningham, 2001], fotografias, bookmarks,

<sup>2</sup>número de arestas conectadas a um determinado vértice.

<sup>3</sup>subconjunto em uma rede em que os nós possuem todos os possíveis elos entre si, i.e. em termos de grafos, trata-se de um subgrafo completo maximal.

músicas, vídeos etc. No entanto, cada um desses serviços representa um silo isolado de conteúdo social e para cada contexto/serviço temos a replicação do grafo de relacionamentos entre usuários.<sup>4</sup>

Para combater essa redundância e falta de conectividade entre contribuições, identidades e relacionamentos, há então a necessidade de semânticas comuns (ontologias) que possam propiciar:

1. um modelo único e extensível para descrever conteúdos de mídia social;
2. interface e ferramentas únicas para busca, manipulação e apresentação desses dados<sup>5</sup> ao invés do uso de distintas APIs para cada rede social ou repositório de mídia social;
3. interligação dos dados, permitindo maior navegabilidade e possibilidades para descobrir conteúdo novo e determinar automaticamente relacionamentos implícitos entre conteúdos e contribuidores;
4. uma mudança de paradigma: de documentos para recursos, da WWW para a GGG<sup>6</sup>

Existem então iniciativas como o padrão FOAF (*Friend of a Friend*) [Brickley and Miller, 2007, Graves] e SIOC (*Semantically-Interlinked Online Communities*) [sio, 2008, Bojars et al., 2008, Breslin et al., 2005], tendo este segundo os seguintes objetivos em linhas gerais:

1. padronizar a forma com que o grafo social (interconexões entre pessoas e seus atributos) e o conteúdo colaborativo relacionado às pessoas é declarado;
2. descrever de forma completa o conteúdo e estrutura de sites sociais/orientados a comunidades;
3. viabilizar a integração da informação de comunidades online;<sup>7</sup>
4. possibilitar a navegação entre conteúdos online de mídia social de formas interessantes e inovativas.

<sup>4</sup>este problema está sendo também endereçado pela iniciativa <http://www.opensocial.org/>

<sup>5</sup>sendo SPARQL (<http://www.w3.org/TR/rdf-sparql-query/>) a sintaxe mais frequentemente usada para a busca e manipulação de dados no formato RDF [rdf, 2008].

<sup>6</sup>*Giant Global Graph*, termo cunhado por Tim Berners-Lee que pode ser descrito como a transição de conteúdo+ponteiros (que temos na WWW) para conteúdo+ponteiros+relacionamentos+descrições (visão da web semântica).

<sup>7</sup>não faz parte do escopo dessa dissertação investigar as implicações destas iniciativas em termos de privacidade e segurança para usuários finais.

O atendimento dos objetivos acima representam diversos benefícios ao processo de criação de sistemas recomendadores e serão discutidos na seção 3.1.1.

Como trabalhos relacionados ao emprego de ontologias no estudo de recomendadores, podemos citar o trabalho [Bouza et al., 2008] que propõe o uso de árvores de decisão baseadas em ontologias. Nele é empregado um avaliador e uma ontologia para generalizarem semanticamente as características do conteúdo dos itens, buscando assim melhoras na eficácia da árvore de decisão construída.

## 2.3

### Sistemas recomendadores

Existem diversas definições para sistemas recomendadores. Uma das primeiras na literatura foi apresentada por [Resnick and Varian, 1997]:

Sistemas clássicos de recomendação usam a opinião de uma comunidade de usuários para ajudar indivíduos dessa comunidade a identificar conteúdo de interesse próprio com maior facilidade.

Segundo [Herlocker et al., 2004], sob o ponto de vista do usuário final, as tarefas de recomendação podem ser classificadas entre as seguintes:

**Classificação dentro de um contexto.** Envolve a classificação automática dos itens entre determinadas categorias pré-definidas. Um exemplo de aplicação dessa tarefa consiste em realizar filtragens dentro de mensagens pertencentes a um fórum de discussão online, destacando as mensagens e discussões que seriam de potencial interesse para o usuário final.

**Encontrar itens bons.** Consiste em sugerir itens específicos para os usuários, apresentados como forma de ranking, podendo fornecer um valor expressando quanto cada item seria interessante ao usuário. Essa é a principal tarefa de recomendação e já foi tratada em uma vasta gama de pesquisas e sistemas comerciais.

**Encontrar todos os itens bons.** Pode ser considerado como um caso especial do problema anterior, onde o usuário esteja interessado em saber todos os itens com grau de interesse a cima de um limiar pré-definido.

**Recomendar seqüência.** Em alguns sistemas interativos pode ser interessante aos usuários não apenas um conjunto de itens e sim a ordem com que esses itens são recomendados. Por exemplo, os usuários de um site de música desejariam que uma playlist gerada para eles automaticamente

não só contenham músicas que eles gostem, mas que estejam ordenadas da maneira mais agradável possível. Outra aplicação seria a recomendação automática de cursos online na seqüência correta (primeiro o introdutório, depois um apanhado geral etc).

**Ajudar na navegação.** Freqüentemente recomendadores em sites web são avaliados em quão bem ajudam usuários a realizar decisões de compra, no entanto usuários navegam pelo site mesmo quando não têm intenção de compra, e nesses casos recomendadores ajudam a guiar a navegação ao sugerir itens de potencial interesse.

### 2.3.1

#### Classificação

Existem diversas taxonomias [Adomavicius and Tuzhilin, 2005, Balabanović and Shoham, 1997, Burke, 2002, Montaner et al., 2003, Schafer et al., 2001] para a classificação de sistemas recomendadores. Apresentamos a seguir um resumo das principais categorias descritas na literatura:

**Colaborativo:** Usuário vai ser recomendado itens que pessoas com **gostos similares** a ele preferiram no passado.

**Demográficos:** Usuário vai ser recomendado itens que de alguma forma são relevantes a pessoas da mesma **classe demográfica** que ele.

**Baseado em conteúdo:** Usuário vai ser recomendado **itens similares** aos que ele preferiu no passado.

**Baseados em utilidade ou conhecimento:** Usuário vai ser recomendado itens que melhor atendem suas **necessidades**, segundo alguma definição para *utilidade e atendimento*.

**Híbridos:** Combinam dois ou mais métodos acima.

Na literatura, uma extensa avaliação de algoritmos recomendadores é realizada em [Herlocker et al., 2004].

O trabalho em [Adomavicius et al., 2005] além de trazer uma classificação dos recomendadores baseado em conteúdo, colaborativos ou híbridos e realizar um apanhado geral de pesquisas anteriores, traz também uma abordagem multidimensional baseada no contexto em que dados são coletados e as recomendações são feitas. O método proposto além de suportar múltiplas dimensões, é hierárquico (em relação ao nível de detalhe das recomendações feitas) e pode trabalhar com *ratings* agregadas.

Em [Spertus et al., 2005] encontramos resultados de experimentos práticos com métricas de semelhança de usuários baseadas nos seus interesses e seu posicionamento na rede social.

Podemos dizer também que a forma como sistemas recomendadores são implementados, como é feita a interação com usuários finais e como são integrados aos sistemas aos quais servem de suporte tem evoluído significativamente nos últimos dez anos. A tabela 2.1 apresenta essa comparação.

2001: Agentes inteligentes	2008: Hubs sociais
Usuário interage com algoritmos	Usuário interage com outros usuários
Filtragem colaborativa em dados globais não diretamente acessíveis a todos	Ajuda construção de rede social: interação é um diálogo com o usuário e não uma transação
Foco em busca ou vendas	Dados de outros usuários expostos em tempo real

Tabela 2.1: Evolução do contexto onde sistemas recomendadores desempenham suas funções

Descrevemos nas seções seguintes as principais categorias de sistemas recomendadores em maiores detalhes.

### 2.3.2

#### Recomendadores colaborativos

Segundo [Burke, 2002], é provavelmente a abordagem mais familiar e madura para a construção de recomendadores. Recomendadores colaborativos agregam pontuações (*ratings*) dos itens, reconhecem padrões comuns para a escolha de ratings entre usuários e gera novas recomendações baseadas nessas comparações entre usuários. Tipicamente, o perfil de um usuário nesses sistemas consiste de um vetor de itens e seus ratings para cada item, continuamente incrementado à medida que o usuário interage com o sistema ao longo do tempo.

Empregam a técnica chamada **filtragem colaborativa** (FC) que foi desenvolvida e melhorada ao longo desta década, existindo atualmente uma grande variedade de algoritmos baseados em seus conceitos.

Em [Sandvig et al., 2008] encontramos um apanhado das diversas técnicas baseadas em modelos para a construção de algoritmos de FC, sendo elas: clusterização de usuários (*k-means* [Steinbach et al., 2000] e pLSA<sup>8</sup>), redução

<sup>8</sup>análise probabilística de semântica latente: técnica estatística para análise de co-ocorrência de dados com diversas aplicações, introduzida em 1999 por [Hofmann, 1999]. Este modelo identifica grupos de usuários com comportamento semelhante através de

de características (PCA [Gerbrands, 1981]), e regras de associação (*Apriori* [Agrawal and Srikant, 1994]).

### 2.3.3

#### Recomendadores demográficos

Funcionam através da categorização do usuário alvo da recomendação baseando-se em atributos pessoais (idade, sexo, localização geográfica etc), fazendo então as recomendações associadas à categoria do usuário. Os atributos pessoais relevantes são informados pelos usuários e em seguida classificados em categorias usando técnicas de *machine learning* por exemplo.

### 2.3.4

#### Recomendadores baseado em conteúdo

Nesses sistemas, os objetos de interesse (alvos das recomendações) são definidos por suas características, como por exemplo usando vetores que expressam a frequência de ocorrência de determinadas palavras-chave no corpo do texto. Um recomendador baseado em conteúdo então aprende um perfil dos interesses do usuário baseado nas características presentes nos itens com os quais o usuário interagiu ou para os quais apresentou *ratings*.

Um resumo contendo exemplos na literatura e maiores detalhes pode ser encontrado em [Pazzani and Billsus, 2007].

### 2.3.5

#### Recomendadores baseados em utilidade ou conhecimento

Tais recomendadores não constroem generalizações a longo prazo sobre seus usuários, mas sim baseiam suas recomendações numa avaliação de quão próximas as necessidades do usuário estão dos itens candidatos dentro do conjunto de opções disponível.

Recomendadores baseados em utilidade fazem sugestões calcadas no cálculo da utilidade de cada objeto para o usuário. O problema central passa a ser então criar e aprimorar essa função de utilidade para cada usuário.

Abordagens baseadas em conhecimento diferem das demais no sentido em que seus algoritmos possuem conhecimento funcional sobre como um determinado item atende às necessidades específicas de um usuário, podendo assim avaliar e decidir em função do grau de atendimento entre uma necessidade e uma possível recomendação.

atributos estatísticos latentes, permitindo que o comportamento dos grupos seja utilizado na recomendação.

### 2.3.6

#### Recomendadores híbridos

O uso de recomendadores híbridos permite evitar algumas limitações de abordagens baseadas em conteúdo ou colaborativa [Balabanović and Shoham, 1997, Basu et al., 1998, Claypool et al., 1999, Pazzani and Billsus, 2007, Ungar and Foster, 1998], aumentando o desempenho alcançado nas recomendações. As abordagens possíveis [Adomavicius and Tuzhilin, 2005] para se combinar recomendadores baseado em conteúdo e colaborativos são:

1. implementar métodos colaborativos e baseados em conteúdo separadamente, combinando suas predições;
2. incorporar algumas características baseadas em conteúdo numa abordagem colaborativa;
3. incorporar técnicas colaborativas numa abordagem baseada em conteúdo;
4. construir um modelo geral que incorpore tanto características baseadas em conteúdo quanto colaborativas.

De acordo com [Burke, 2002], podemos de maneira geral classificar as abordagens existentes para a construção de recomendadores híbridos em:

**Ponderada:** a pontuação de um item recomendado é calculada a partir dos resultados de todas as técnicas de recomendação presentes no sistema.

**Alternada:** o sistema usa algum critério para alternar entre as técnicas de recomendação presentes no sistema caso as recomendações feitas nas tentativas anteriores não alcancem um grau de confiança aceitável por exemplo. Tal critério pode envolver por exemplo características do usuário que solicitou a recomendação, estágio de treinamento do sistema etc.

**Misto:** os itens recomendados de forma independente pelas técnicas de recomendação presentes no sistema são apresentadas de forma combinada para o usuário.

**Combinação de características:** trata uma fonte de informação alternativa simplesmente como características adicionais de cada item avaliado, que são então consideradas por uma abordagem baseada em conteúdo em conjunto com os demais atributos originais do item. Isso permite ao sistema levar em conta dados colaborativos sem depender deles

exclusivamente, o que reduz a sensibilidade do algoritmo ao número de usuários que já opinaram sobre itens.

**Cascata:** uma técnica de recomendação é empregada primeiro para produzir uma ordenação grosseira dos itens candidatos e uma segunda técnica refina essas recomendações tendo como domínio de itens recomendáveis esse conjunto de itens e ordenação produzida no primeiro passo.

**Incrementação de características:** uma técnica é empregada para produzir uma *rating* ou classificação de um item e essa informação é então incorporada ao processamento do recomendador seguinte. Difere da estratégia em cascata pois naquele o resultado de um recomendador não é utilizado como entrada para o próximo (lá é usado apenas como um refinamento do espaço de busca para itens a serem recomendados). Também difere da estratégia de combinação de características no sentido em que naquele as informações de diferentes fontes (dados colaborativos por exemplo) são usadas como dados crús/atributos dos itens a serem avaliados, enquanto que nestes a saída de um recomendador (itens, *ratings*, grau de confiança etc) são usados como entrada direta do algoritmo.

**Meta:** quando o modelo gerado por um recomendador é usado como a entrada de outro. Difere dos anteriores pois aqui não são características dos itens ou resultados de recomendações que são reaproveitadas, e sim um modelo inteiro de aprendizagem/recomendação.

### 2.3.7

#### Recomendadores para redes sociais

Sistemas recomendadores online frequentemente auxiliam usuários em tarefas inseridas no contexto de redes sociais. Dada a crescente importância destas redes, observa-se na literatura o aparecimento de recomendadores que fazem o uso de informações contidas nessas redes, como a vizinhança de usuários, relacionamentos entre itens de conteúdo e grupos de usuários com interesses comuns etc.

[Seth and Zhang, 2008] propõe e avalia um sistema recomendador de conteúdo social baseado na modelagem bayesiana [Heckerman, 1995] dos relacionamentos entre usuário-usuário e usuário-item.

Em [Liu et al., 2006] o conteúdo textual de 100.000 perfis de usuários de uma rede social foram capturados e os principais termos presentes em cada perfil foram mapeados em uma ontologia abrangente de música, livros, filmes, comida etc e técnicas de aprendizagem de máquina foram aplicadas para inferir

o perfil de gostos dos usuários. Este trabalho discute também metodologias e técnicas reusáveis para mineração semântica de comunidades online.

O trabalho em [Zheng et al., 2007] realiza um estudo empírico comparativo de métodos para filtragem colaborativa usando dados explícitos da rede social de consumidores do site *Amazon.com*. Nele são examinadas duas maneiras para incorporar informações da rede social em filtragem colaborativa (FC): [1] usando a rede social para restringir o conjunto de recomendadores selecionados e [2] adicionando à abordagem anterior a informação de proximidade entre usuários na rede social para modificar o cálculo tradicional de FC. Os resultados deste trabalho mostram que a abordagem [1] implica em desempenho consideravelmente melhor e que [2] não parece trazer melhorias significativas.

[Passant and Raimond, 2008] detalha como redes sociais, contribuições de usuários e outros dados interconectados e publicados no escopo da iniciativa Linking Open Data (LODR<sup>9</sup>) podem ser combinadas em recomendadores híbridos.

### 2.3.8

#### Recomendadores baseados em conteúdo visual

Recomendadores híbridos que ajudam em tarefas orientadas a imagens podem ter seu desempenho melhorado através da combinação de técnicas como FC e recuperação de imagens baseado em conteúdo visual (CBIR – do inglês *content-based image retrieval*<sup>10</sup>).

O trabalho em [Kim et al., 2004] apresenta um recomendador de conteúdo visual para aplicativos web móveis. Trata-se de uma melhoria à abordagem clássica de FC onde a aprendizagem do recomendador é acelerada com o uso de dados adicionais de preferência dos usuários proveniente do módulo CBIR. Os resultados desse trabalho também mostram que é possível também contornar problemas de escaridade de *ratings*<sup>11</sup> disponíveis para imagens a serem recomendadas com o uso de técnicas de CBIR, pois as mesmas podem recomendar imagens sem nenhum rating de outros usuários, baseando-se apenas em características visuais.

Em [Kwon, 2003] é proposto um sistema recomendador multimídia sensível a contexto onde emprega-se uma metodologia baseada em agentes que inicialmente identificam atributos semânticos de imagens hipermídia com o intuito de derivar informações contextuais. A seguir o sistema recomenda serviços web, antecipando necessidades ou produtos que o usuário pode ter

<sup>9</sup><http://lodr.info/>

<sup>10</sup>um apanhado abrangente desta área pode ser encontrado em [Rui et al., 1999].

<sup>11</sup>também chamado de “problema-do-item-novo”.

interesse. Para interpretar as imagens e identificá-las dentro de uma ontologia são usadas técnicas de visão computacional.

## 2.4

### Frameworks para avaliação de sistemas recomendadores

Os conceitos teóricos, métricas para avaliação de recomendadores e metodologias básicas para tais estudos podem ser encontrados em [Herlocker et al., 2004] e [Herlocker et al., 1999].

A *framework* para avaliação proposta em [Fisher et al., 2000] é composta de três componentes: um motor de **predição**, um sistema de **avaliação** e um componente de **visualização**. O primeiro deles provê uma interface comum para a implementação de diferentes algoritmos de predição. O segundo deles provê uma metodologia para testes e métricas para analisar a precisão e desempenho em tempo de execução dos algoritmos de predição. Finalmente, o terceiro componente sugere como representações gráficas podem informar o desenvolvimento e análise destes algoritmos.

Um dos principais aspectos limitantes desse trabalho é a premissa de que os algoritmos de predição/recomendação devem trabalhar apenas em cima de dados de uso orientados a votos (*ratings*) de usuários sobre itens, o que parece ser restrito demais para diversas outras aplicações, como as recomendações baseadas em rede social e conteúdo propostas nessa tese. Além disso, o código publicado pelos autores desse trabalho encontra-se demasiadamente acoplado com o domínio de testes usado por eles: ratings de filmes.

[Herlocker et al., 1999] sugerem que as abordagens atuais para avaliação de sistemas recomendadores podem ser divididas em duas categorias: *off-line* ou *on-line*. Nesta o desempenho é avaliado com a ajuda das interações de usuários reais num sistema em execução enquanto que naquela as avaliações são realizadas usando apenas bases de dados existentes. Um estudo mais detalhado sobre técnicas para avaliação on-line pode ser encontrado em [Hayes et al., 2002].

## 2.5

### Filtragem colaborativa

Filtragem colaborativa é um método para realizar predições automaticamente sobre os interesses de um usuário, através da coleta dos hábitos de usuários semelhantes ao longo de suas interações com um sistema.

A premissa por trás da abordagem de filtragem colaborativa é que aqueles usuários cujos hábitos foram semelhantes no passado têm a tendência de concordar no futuro.

Por exemplo, um sistema de filtragem colaborativa para gostos musicais poderia fazer predições sobre qual música um usuário gostaria dada uma lista parcial dos gostos daquele usuário. Note que essas predições são específicas de um usuário, mas a informação a partir da qual as recomendações são derivadas advém de muitos usuários. Essa abordagem difere de uma mais simples onde é dada uma pontuação média não específica para cada item de interesse, por exemplo baseado no número de votos dados por todos os usuários.

Uma das principais características da filtragem colaborativa quando comparada com filtragem baseada em conteúdo é que o primeiro não tem nenhum conhecimento acerca do real conteúdo dos itens ou do que eles tratam. Isso significa que estes recomendadores confiam apenas em valores de preferência — tais como *ratings* submetidos por usuários — para a geração das recomendações. Isso também significa que estes recomendadores estão sujeitos a “efeitos de média” que causam os itens mais populares (para os quais se têm mais dados de *rating*) serem recomendados com maior frequência, levando a uma maior visibilidade e taxa de consumo pelos usuários, que se converte em mais *ratings*, fechando o ciclo vicioso.

Dois outros problemas são bem conhecidos e frequentemente associados a técnicas de filtragem colaborativa pura: o problema do Primeiro-Avaliador (“*First-Rater*”) e Inicialização-a-Frio (“*Cold-Start*”). O problema do Primeiro-Avaliador é causado por novos itens no sistema que obviamente ainda não receberam nenhum *rating* de usuários. O sistema é então incapaz de gerar interconexões semânticas envolvendo esses itens, que por isso nunca são recomendados. De forma similar, o problema de Inicialização-a-Frio é causado por novos usuários no sistema que ainda não submeteram nenhuma *rating*. Sem nenhuma informação acerca desses usuários, o sistema não consegue inferir suas preferências para gerar recomendações.

### 2.5.1 Funcionamento

O funcionamento de sistemas de filtragem colaborativa baseada em **usuários** normalmente compreende dois passos: [1] encontrar usuários que compartilham com o usuário alvo o mesmo padrão de ratings para os itens e [2] usar as ratings desses usuários com gostos parecidos ao do usuário alvo para calcular uma predição contendo itens, normalmente ordenadas segundo uma pontuação que reflita a probabilidade de acerto de cada item recomendado.

Alternativamente, filtragem colaborativa baseada em **itens**, conforme popularizado pela Amazon.com (“...*clientes que compraram X também compraram Y e Z...*”) e inicialmente proposta em [Vucetic and Obradovic, 2000]

funciona de forma centrada em itens, segundo os passos: [1] construir uma matriz item-item representando a semelhança entre pares de itens e [2] usar essa matriz e os dados (ratings) do usuário alvo para inferir seu gosto e formar predições.

Um exemplo da abordagem baseada em itens é a família de algoritmos denominados *Slope One*, introduzida em [Lemire and Maclachlan, 2005].

Essa abordagem pode ser combinada com métricas de semelhança entre itens baseada no conteúdo deles, usando tal métrica para inicializar os valores da matriz item-item. Demais possibilidades de combinação foram discutidos na seção 2.3.6.

Existem duas formas para realizar a coleta de dados necessários para realizar a filtragem colaborativa: **filtragem ativa** ou **passiva**. A primeira delas é a mais disseminada e consiste na elicitación com o usuário de uma pontuação ou *rating* para cada item no domínio de itens a serem recomendados. Uma desvantagem dessa abordagem é que a opinião de usuários pode conter vieses por ser subjetiva e nem todos usuários utilizam a mesma escala: há distorções na escala entre diferentes usuários e até mesmo para um dado usuário ao longo de suas sessões de interação com o sistema. Além disso, a filtragem ativa representa uma tarefa a mais a ser desempenhada pelos usuários, que dificilmente enxergam algum benefício imediato em colaborar com *ratings*, implicando numa menor quantidade de dados disponíveis.

Já a **filtragem passiva** pode ser realizada através de observações implícitas do comportamento dos usuários, ao invés do comportamento artificial imposto pela necessidade do usuário atribuir uma pontuação (*rating*) para cada item. Nesses sistemas, estatísticas de acessos e padrões de interação de todos os usuários são coletadas e usadas para alimentar o algoritmo recomendador. Demais ações dos usuários que poderiam agir como fontes de dados seriam a visualização, solicitação por mais detalhes ou compra de um item, a menção/referência a um site, o número de buscas realizadas usando determinadas palavras chaves etc.

## 2.6

### Métricas de desempenho de recomendadores

Avaliar o desempenho de sistemas recomendadores e seus algoritmos é inerentemente difícil por uma série de razões. Primeiro, diferentes algoritmos podem ser melhores ou piores quando empregados em diferentes conjuntos de dados. Segundo, os objetivos de uma avaliação diferem com a mesma frequência que as tarefas realizadas pelos recomendadores e seus contextos diferem entre si: há métricas focando na precisão das recomendações realizadas, no número

de recomendações incorretas, no potencial de emprego de um recomendador como um filtro e diversos outros aspectos que ultimamente representem o grau de satisfação dos usuários finais. Finalmente, há um desafio significativo em determinar qual a melhor combinação de métricas a ser usada numa avaliação comparativa de recomendadores.

Assim, existem diversas técnicas e métricas para a validação de recomendadores, que são discutidas em maiores detalhes em [Herlocker et al., 2004].

Destacamos a seguir as principais dimensões onde recomendadores podem ser avaliados.

**Relevância ou Precisão** mede quão bem os itens recomendados vão de encontro com o gosto do usuário e pode ser calculado segundo:

$$\text{precisão} = \frac{|\{\text{itens relevantes}\} \cap \{\text{itens recomendados}\}|}{|\{\text{itens recomendados}\}|} \quad (2-3)$$

Frequentemente essa medida vem acompanhada do número do rank de corte (“P@n”), indicando que a precisão alcançada considera que foram recomendados ao usuário os  $n$  itens com maior probabilidade de acerto da previsão.

**Novidade** indica quão bom o recomendador é em encontrar itens que o usuário desconhece. Um sistema pode produzir recomendações com alta precisão e cobertura e mesmo assim serem inúteis em termos práticos já que para um determinado contexto de recomendação o interesse do usuário é na verdade descobrir itens inéditos e não relacionados aos que ele já conhece.

**Alcance, Recall ou Cobertura** pode ser calculado da forma a seguir

$$\text{recall} = \frac{|\{\text{itens relevantes}\} \cap \{\text{itens recomendados}\}|}{|\{\text{itens relevantes}\}|} \quad (2-4)$$

e representa quão grande é o conhecimento que o recomendador possui da cauda longa<sup>12</sup> de itens que podem ser recomendados. [e Kartik Hosanagar, 2007] traz um estudo do impacto da cauda longa de itens para a construção de recomendadores. Sistemas com pouca cobertura podem ser de menor utilidade para usuários já que terão limites nas decisões

<sup>12</sup>[Jansen et al., 1998] é um dos primeiros trabalhos a constatar o efeito da cauda longa no contexto de *IR* (recuperação de informação) ao verificar que o histograma de frequência dos termos usados nas *queries* em um motor de busca não segue a forma tradicional de uma distribuição *Zipf* que normalmente representa a distribuição de palavras encontradas em textos escritos em linguagens naturais como o inglês. No início da curva do histograma dos termos usados em *queries* há uma queda abrupta e à medida em que se avança ao final dela nota-se discontinuidades e um **cauda longa** representando muitos termos com frequência unitária.

onde serão capazes de ajudar. Essa dimensão é então particularmente importante para a tarefa de encontrar “itens bons” (seção 2.3).

**Erro médio absoluto (MAE<sup>13</sup>)** mede o desvio absoluto médio entre o rating predito e o valor que realmente foi dado pelo usuário quando se faz uso de conjuntos de dados de treinamento

**Curva ROC<sup>14</sup>** provê métricas alternativas à precisão e *recall*. Tais curvas medem até que ponto um sistema para filtragem de informação é capaz de distinguir entre sinal (relevância) e ruído. Tais métricas teriam maior utilidade para análise de sistemas recomendadores que assistem usuários em tarefas de classificação.

**Velocidade de aprendizagem** pode ser um critério importante na comparação entre recomendadores já que muitos deles incorporam algoritmos de aprendizagem baseados em modelos estatísticos que dependem da disponibilidade de dados reais. Como resultado, seus desempenhos variam de acordo com a quantidade de informação disponível para aprendizado, tendo alguns deles uma resposta mais rápida do que outros no aumento do desempenho à medida em que esses dados são disponibilizados.

**Feedback de usuários online** pode por exemplo ser realizado através de experimentos numa interface web onde à medida em que as recomendações são apresentadas o sistema pergunta ao usuário para que ele opine (frequentemente numa escala numérica) o quão relevantes foram os resultados apresentados. Pode ser obtido indiretamente através da medição de clicks de usuários como indicativo da efetividade de um algoritmo de recomendação.

**Grau de confiança ou certeza** indica quão certo ou confiante um algoritmo de recomendação está a respeito da precisão de uma recomendação a ser feita. Ao ajudar usuários finais na tomada de decisões, pode ser vantajoso apresentar em conjunto com os itens recomendados o grau de confiança associado a cada predição. A capacidade de apresentar o grau de confiança junto com as recomendações e a forma como essa apresentação é feita pode distinguir um recomendador do outro. De maneira geral, essas possibilidades refletem a transparência do recomendador, que deve refletir quão bem este explica ao usuário as recomendações feitas.

**F<sub>β</sub>-measure** é calculada através da média harmônica generalizada entre a precisão e o *recall*, segundo equação abaixo. O parâmetro  $\beta$  reflete a im-

portância dada por um usuário à precisão em detrimento da abrangência das recomendações feitas e vice-versa.

$$F_{\beta} = \frac{(1 + \beta^2) \cdot (\text{precision} \cdot \text{recall})}{\beta^2 \cdot \text{precision} + \text{recall}} \quad (2-5)$$