

UNIVERSIDADE DO VALE DO RIO DOS SINOS – UNISINOS
MBA EM ADMINISTRAÇÃO DA TECNOLOGIA DA INFORMAÇÃO

MIGUEL BIRCK NEDEL

MODELO HÍBRIDO DE SISTEMAS DE RECOMENDAÇÃO
EM COMÉRCIO ELETRÔNICO

SÃO LEOPOLDO

2011

Miguel Birck Nedel

MODELO HÍBRIDO DE SISTEMAS DE RECOMENDAÇÃO
EM COMÉRCIO ELETRÔNICO

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do título de Especialista em Administração da Tecnologia da Informação, pelo MBA em Administração da Tecnologia da Informação, da Universidade do Vale do Rio dos Sinos.

Orientador: Prof. Dr. Silvio César Cazella

São Leopoldo

2011

RESUMO

O comércio eletrônico está consolidado no mercado brasileiro. Empresas que desenvolvem sistemas para suporte ao comércio eletrônico, como lojas virtuais, necessitam buscar um diferencial para agregar valor aos seus produtos. Os sistemas de recomendação são um importante meio de potencializar vendas e melhorar a experiência de compra dos usuários de comércio eletrônico. O objetivo deste trabalho é propor um modelo híbrido que utilize um conjunto de técnicas para utilização de sistemas de recomendação em comércio eletrônico. O modelo foi concebido conforme necessidade de uma empresa real. Foram identificadas as necessidades da empresa desenvolvedora através de um roteiro de entrevistas e, a partir daí, encontradas soluções na literatura que atendessem ao objetivo. Foi montado um modelo híbrido de utilização de sistemas de recomendação. Foram utilizadas as técnicas de recomendação baseada em conteúdo, filtragem colaborativa e recomendação baseada em comunidade para compor este modelo. A utilização de mais de uma técnica mostrou-se proveitosa, pois traz sinergia ao sistema de recomendação para obter melhores resultados.

Palavras-chave: Comércio eletrônico; sistemas de recomendação; filtragem híbrida.

SUMÁRIO

1 INTRODUÇÃO	6
1.1 SITUAÇÃO PROBLEMÁTICA E PERGUNTA DE PESQUISA	6
1.2 OBJETIVOS	7
1.2.1 Objetivo geral.....	7
1.2.2 Objetivos específicos.....	7
1.3 JUSTIFICATIVA.....	7
2 FUNDAMENTAÇÃO TEÓRICA	9
2.1 COMÉRCIO ELETRÔNICO.....	9
2.1 SISTEMAS DE RECOMENDAÇÃO	9
2.1.1 Sistemas de Recomendação Baseados em Conteúdo.....	10
2.1.1.1 Estrutura de um Sistema de Recomendação Baseado em Conteúdo.....	10
2.1.1.2 Limitações	12
2.1.2 Filtragem Colaborativa.....	12
2.1.2.1 Recomendações baseadas em vizinhança de usuários.....	13
2.1.2.2 Limitações	14
2.1.3 Recomendações Demográficas.....	14
2.1.4 Recomendações Baseadas em Conhecimento.....	15
2.1.5 Recomendações Baseadas em Comunidade	15
2.1.5.1 O Papel da Confiança nas Recomendações Baseadas em Comunidade.....	15
2.1.6 Sistemas Híbridos	16
3 MÉTODOS E PROCEDIMENTOS.....	18
3.1 DELINEAMENTO DA PESQUISA.....	18
3.2 DEFINIÇÃO DA UNIDADE DE ANÁLISE.....	18
3.3 TÉCNICAS DE COLETA DE DADOS	19
3.4 TÉCNICAS DE ANÁLISE DE DADOS	19
3.5 LIMITAÇÕES DO MÉTODO	19
4 APRESENTAÇÃO E ANÁLISE DOS DADOS.....	20
4.1 CADASTRO E IDENTIFICAÇÃO DE USUÁRIOS	21
4.2 RECOMENDAÇÕES BASEADAS EM CONTEÚDO	22
4.3 FILTRAGEM COLABORATIVA.....	23

4.4 RECOMENDAÇÕES BASEADAS EM COMUNIDADE.....	23
4.5 MODELO HÍBRIDO PROPOSTO.....	25
4.5.1 Sinergia do Modelo Proposto.....	26
5 CONSIDERAÇÕES FINAIS	27
REFERÊNCIAS	29
APÊNDICE A	32

1 INTRODUÇÃO

Este trabalho tratará da utilização sistemas de recomendação para suporte a lojas virtuais.

1.1 SITUAÇÃO PROBLEMÁTICA E PERGUNTA DE PESQUISA

Nos sistemas de comércio virtual atuais, os sistemas de recomendação são importantes ferramentas para promoção de vendas e auxílio a usuários. Sistemas de recomendação têm como objetivo auxiliar os usuários de um sistema a encontrar os itens que possam lhes ser mais úteis. O presente trabalho abordará a utilização de sistemas de recomendação em lojas virtuais desenvolvidas por uma empresa.

O Grupo Conectt é uma empresa de desenvolvimento de projetos para web com quinze anos de experiência, e constituiu-se como estudo de caso deste trabalho de pesquisa. Atualmente, a empresa está iniciando trabalhos na área de lojas virtuais. Tendo em vista que essa é uma área incipiente na empresa, é importante encontrar formas de agregar valor aos produtos desenvolvidos. Portanto, viabilizar a utilização de sistemas de recomendação nas lojas virtuais desenvolvidas pela empresa surge como uma oportunidade de expansão das capacidades do grupo.

Outro aspecto a ser considerado é que a empresa não conta com colaboradores que possuam conhecimentos e experiência na área de sistemas de recomendação. Dessa forma, existe a necessidade de agregar às capacidades da empresa conhecimento sobre esse tipo de tecnologia.

O foco deste trabalho está em estudar os sistemas de recomendação existentes na literatura e propor um meio de viabilizar a incorporação desse tipo de tecnologia nos sistemas desenvolvidos pela empresa.

Portanto, a pergunta de pesquisa é a seguinte: - como utilizar sistemas de recomendação nos sistemas de comércio eletrônico desenvolvidos pela empresa?

1.2 OBJETIVOS

1.2.1 Objetivo geral

Propor um modelo híbrido que utilize um conjunto de técnicas para utilização de sistemas de recomendação em comércio eletrônico.

1.2.2 Objetivos específicos

- a) Identificar as necessidades de sistemas de recomendação na empresa;
- b) Selecionar as práticas mais adequadas à empresa nas técnicas estudadas;
- c) Combinar as técnicas selecionadas da forma mais adequada à empresa.

1.3 JUSTIFICATIVA

O mercado de comércio eletrônico está em franca expansão. Segundo o E-Commerce Brasil (2011), as vendas virtuais brasileiras alcançaram um faturamento de R\$ 13,60 bilhões em 2010, com um crescimento de 30% em relação ao ano anterior. Dentro de um mercado tão atraente e competitivo, as empresas desenvolvedoras de lojas virtuais necessitam de um diferencial.

A utilização de sistemas de recomendação é capaz de prover uma diferenciação para uma loja virtual. Segundo Knidel (2010), os sistemas de recomendação ajudam a melhorar a experiência de compra do usuário, trazendo a ele a sensação de ter feito um bom negócio.

Ao final deste trabalho, a empresa em que o estudo será feito terá a capacidade de agregar valor aos produtos de software por ela desenvolvidos. Os sistemas de recomendação poderão ser aplicados a lojas virtuais futuramente desenvolvidas pela empresa, ou até mesmo adaptados para outros tipos de sistemas que a empresa desenvolva.

Esta pesquisa se apresenta interessante neste momento, pois a empresa está começando a desenvolver lojas virtuais. É importante aproveitar este estágio de desenvolvimento inicial para incorporar aos produtos desenvolvidos a ferramenta de sistemas de recomendação e, assim, agregar o máximo de valor aos produtos de *software* da empresa.

Conforme Adomavicius (2005), há uma constante evolução dos sistemas de recomendação. Mesmo assim, sempre há espaço para melhorias que deixem os métodos de

recomendação mais efetivos e aplicáveis. Dessa forma, se faz necessário um estudo organizado dos sistemas de recomendação existentes, a fim de encontrar aquele mais adequado à realidade da empresa em questão.

Este trabalho encontra-se estruturado em capítulos. No capítulo 2 é apresentada a fundamentação teórica, onde são demonstradas as definições de comércio eletrônico e sistemas de recomendação. O subcapítulo sobre sistemas de recomendação apresenta, ainda, a definição de cada uma das técnicas de recomendação estudadas (sistemas de recomendação baseados em conteúdo, filtragem colaborativa, recomendações demográficas, recomendações baseadas em conhecimento, recomendações baseadas em comunidade e sistemas híbridos). O capítulo 3 apresenta os métodos e procedimentos utilizados na condução da pesquisa realizada por este trabalho. O capítulo 4 constitui-se da apresentação e análise dos dados colhidos na pesquisa, onde são apresentadas as técnicas de recomendação selecionadas, além do modelo híbrido de sistemas de recomendação proposto. No capítulo 5 encontram-se as considerações finais do trabalho. O Anexo A contém o roteiro de entrevista aplicado durante a coleta de dados, bem como as respostas obtidas.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 COMÉRCIO ELETRÔNICO

Segundo Albertin (2002), o comércio eletrônico é a realização de toda a cadeia de valor dos processos de negócio em um ambiente eletrônico. Para tal, são aplicadas tecnologias de comunicação e informação para atender aos objetivos do negócio.

Cameron (1997) define o comércio eletrônico como qualquer negócio transacionado eletronicamente, podendo estas transações ocorrer entre dois parceiros de negócio ou entre um negócio e seus clientes.

2.1 SISTEMAS DE RECOMENDAÇÃO

Sistemas de recomendação são técnicas e ferramentas de software capazes de sugerir a um usuário itens que o mesmo gostaria de utilizar. Estes sistemas possuem a capacidade de aumentar o número de itens vendidos, a diversidade de itens vendidos, a satisfação dos usuários, a fidelidade dos usuários e o conhecimento sobre o que o usuário realmente deseja. Dessa forma, um sistema de recomendação eficiente deve satisfazer às necessidades do usuário e do provedor do sistema de forma equilibrada (RICCI, 2011).

Um sistema de recomendação baseia-se em resolver problemas de uma estrutura de classificação. Um usuário pode classificar os itens de um sistema de acordo com a utilidade que cada unidade em especial lhe traz. O sistema de recomendação deve ser capaz de prever qual classificação um usuário atribuiria a um item que ele nunca acessou no sistema. Dessa forma, o sistema é capaz de apresentar os itens de acordo com a sua utilidade para os usuários (ADOMAVICIUS, 2005).

Os sistemas de recomendação podem ser classificados em seis categorias segundo Francesco Ricci (2011): recomendações baseadas em conteúdo, filtragem colaborativa, recomendações demográficas, recomendações baseadas em conhecimento, recomendações baseadas em comunidade e recomendações híbridas.

2.1.1 Sistemas de Recomendação Baseados em Conteúdo

Os sistemas de recomendação baseados em conteúdo oferecem ao usuário itens similares àqueles que ele gostou no passado. Para determinar a similaridade entre itens distintos, são aplicadas métricas de similaridade. Essas métricas são determinadas pela comparação entre características e atributos compartilhados entre itens distintos, quanto mais atributos são compartilhados entre dois itens, mais similares eles serão considerados (BLANCO-FERNANDEZ, 2008).

Com métricas bem construídas, essa estratégia tem uma precisão muito grande na tarefa de determinar a similaridade entre itens. No entanto, corre-se o risco de trazer ao usuário sugestões superespecializadas, ou seja, apenas itens que o usuário já conheça. Assim, o grau de efetividade das sugestões feitas pelo sistema pode estar sendo pouco explorado (BLANCO-FERNANDEZ, 2008).

Essa abordagem analisa os itens avaliados pelo usuário no passado e constrói um perfil de interesses do usuário baseado nos atributos desses objetos. Os atributos desse perfil são, então, comparados aos atributos de qualquer outro item do sistema, determinando sua relevância para o usuário. O resultado da comparação, se positivo, indicará se um item deve ser apresentado ao usuário ou mesmo escondido dele, em caso negativo (MITCHELL, 1997).

2.1.1.1 Estrutura de um Sistema de Recomendação Baseado em Conteúdo

Conforme Ricci (2011), algumas técnicas são necessárias para criar uma representação dos itens e produzir um perfil para cada usuário a partir disso. A figura 1 apresenta um modelo de alto nível que descreve as etapas necessárias para esse processo.

A construção do perfil do usuário é feita a partir das avaliações que ele faz para cada item do sistema. A lista de itens avaliados é analisada pelo aprendiz de perfil e as características e atributos de cada item são computados para gerar o perfil do usuário. Quando um novo item é incluído no sistema, suas características são comparadas àquelas consolidadas no perfil do usuário. Se o resultado da comparação for positivo, o item é acrescentado a uma lista de recomendações a ser apresentada para o usuário. Esse aprendizado deve ser mantido em um ciclo, pois as preferências de um usuário variam ao longo do tempo. Também é interessante que o sistema possua alguma forma de o usuário reportar se os itens da lista de recomendações são satisfatórios, permitindo uma readaptação do perfil do usuário a partir desses dados (RICCI, 2011).

2.1.1.2 Limitações

A recomendação baseada em conteúdo está limitada à quantidade e qualidade das informações armazenadas para seus itens. O conhecimento do domínio onde os itens estão inseridos deve estar contemplado na descrição de cada item, para que a correlação entre eles seja efetiva. Ainda assim, o conteúdo informado nunca será suficiente para determinar todos os aspectos que influenciam a experiência do usuário (FELLBAUM, 1998).

Recomendações baseadas em conteúdo também possuem a tendência de ser superespecializadas. As recomendações sempre serão similares aos itens avaliados no passado. Assim, as recomendações fornecidas não possuem a capacidade de surpreender o usuário, pois o sistema não consegue capturar todo o espectro de preferências do mesmo (ABBASSI, 2009).

Outra limitação dessa abordagem é que não é possível fornecer recomendações para novos usuários. É necessária uma interação com o sistema para que as avaliações sejam feitas e o perfil do usuário seja montado. Ainda assim, a acurácia das recomendações pode ser baixa enquanto houver poucas avaliações (BASIRI, 2010).

2.1.2 Filtragem Colaborativa

A filtragem colaborativa é um algoritmo que gera recomendações para um usuário a partir de preferências de outros consumidores que sejam similares a ele. Para isso, o algoritmo

parte do pressuposto de que, se dois usuários avaliaram vários itens de forma semelhante, eles tendem a avaliar novos itens da mesma forma. Por exemplo, se um usuário u tem um histórico de avaliações positivas muito semelhantes ao de um usuário v , é esperado que um novo item i avaliado positivamente pelo usuário v seja útil para o usuário u . Nesse caso, o item i é um bom candidato a ser recomendado ao usuário u (KONSTAN, 1997).

Essa abordagem traz uma vantagem em relação à recomendação baseada em conteúdo, pois itens sem nenhuma semelhança entre si são recomendados ao usuário. Isso faz com que sejam criadas relações de semelhança entre os itens, mesmo sem levar em consideração o conteúdo dos mesmos. Além disso, são recomendados muito mais itens não esperados pelo usuário, aumentando o espectro de itens aproveitados (LINDEN, 2003).

2.1.2.1 Recomendações baseadas em vizinhança de usuários

Para utilização da filtragem colaborativa, é necessário conhecer o conceito de vizinhança de usuários. Esse conceito automatiza o princípio das recomendações boca a boca, onde um usuário valoriza as opiniões de pessoas semelhantes para formar sua própria opinião a respeito de um item (RICCI, 2011).

O sistema de recomendação deve ser capaz de construir, para cada usuário, uma lista de usuários vizinhos com interesses semelhantes. Essa lista será usada para gerar as recomendações. Além de descobrir quais são os usuários com interesses semelhantes, é preciso classificar a similaridade entre eles. Quanto maior for a similaridade entre um usuário u e um usuário v , maior deverá ser o peso das avaliações do usuário u no momento de gerar recomendações ao usuário v (SHEN, 2010).

O conceito de peso de similaridade determina o quanto um usuário é semelhante a outro. Esse conceito permite que sejam selecionados apenas vizinhos confiáveis nas predições e que seja dada maior ou menor importância a esses vizinhos nas predições. No entanto, se dois usuários avaliaram poucos itens de maneira muito semelhante, o peso da similaridade entre eles será alto. Todavia, se a quantidade de itens for muito pequena, talvez não seja suficiente para determinar que realmente haja similaridade entre eles (SHEN, 2010). Outra questão que deve ser considerada é a forma como os usuários classificam os itens. Por exemplo, em um sistema onde seja possível classificar um item em uma escala de 1 a 5, alguns usuários classificarão apenas alguns poucos itens que considerem muito especiais com o grau máximo cinco. Outros usuários poderão ter um padrão de comportamento diferente, e dar o grau máximo para a maioria dos itens que gostarem. Isso significa que dois usuários

podem atribuir diferentes valores de classificação para quantificar o mesmo nível de apreciação de um item. O sistema deve ser capaz de normalizar as classificações para que a comparação entre usuários seja mais assertiva (JIN, 2004).

2.1.2.2 Limitações

Segundo Adomavicius (2005), sistemas de filtragem colaborativa são suscetíveis ao problema do novo usuário, assim como os sistemas de recomendação baseados em conteúdo. Para que as recomendações sejam feitas corretamente, o sistema precisa aprender as preferências do usuário a partir das avaliações que ele faz. Enquanto o usuário não fizer avaliações, ele não poderá ser comparado a outros usuários para gerar recomendações.

Adomavicius (2005) lembra ainda que novos itens são adicionados regularmente ao sistema. Da mesma forma, um novo item adicionado só poderá ser recomendado a um usuário quando possuir avaliações anteriores de outros usuários. Sem essas avaliações, o sistema não poderá saber se o item pode ser indicado, já que, para indicar um item a um usuário, é necessário que outro usuário parecido com ele o tenha avaliado positivamente.

Outra dificuldade encontrada na filtragem colaborativa é a esparsialidade. Muitas vezes, a quantidade de avaliações disponível é insuficiente para atingir a qualidade de recomendações desejada. Itens com nenhuma ou poucas avaliações correm o risco de jamais serem recomendados. Da mesma forma, usuários que façam poucas avaliações ou que tenham preferências únicas em relação à massa de usuários do sistema podem não ser incluídos em nenhuma vizinhança de usuários e, portanto, não receber recomendações. A esparsialidade pode prejudicar ainda a qualidade das recomendações geradas. Se dois usuários avaliarem poucos itens no site, mas de forma semelhante, o sistema pode indicar uma falsa semelhança de perfil, trazendo recomendações incorretas no futuro (SCHEIN, 2002).

2.1.3 Recomendações Demográficas

Informações demográficas dos usuários podem ser utilizadas para determinar se um objeto deverá ser recomendado. Esse método assume que diferentes recomendações devem ser geradas para diferentes nichos demográficos, portanto, os itens devem ser categorizados de acordo com atributos demográficos do usuário para possam ser oferecidos. Os atributos demográficos considerados podem ser dados como localização, idade, sexo e língua (PAZZANI, 1999).

2.1.4 Recomendações Baseadas em Conhecimento

Sistemas baseados em conhecimento recomendam itens fundamentados em informações específicas sobre como determinadas características de um item atendem às necessidades e preferências de um usuário. Existem dois tipos de sistemas baseados em conhecimento – baseados em casos e baseados em limitações. No primeiro, uma função de similaridade estima o quanto as necessidades do usuário (descrição do problema) são compatíveis com as recomendações (soluções do problema). Já no segundo tipo, são exploradas regras específicas sobre como os requisitos de um usuário são atendidos pelas características dos itens (RICCI, 2011).

2.1.5 Recomendações Baseadas em Comunidade

Comunidades são grupos de indivíduos que podem compartilhar experiências e opiniões a respeito de itens e vendedores. Esses grupos são capazes de recomendar itens de maneira positiva ou negativa. Os sistemas de recomendação baseados em comunidade apresentam essas informações em um formato organizado, auxiliando na tomada de decisão de um usuário com base na comunidade a que ele faz parte (CHEN, 2009).

Essa técnica segue a máxima “diga-me quem são seus amigos e lhe direi quem és”. Existem evidências de que as pessoas tendem a confiar mais em recomendações de amigos do que de anônimos. Aliado a isso, o crescimento das redes sociais possibilita que mais dados dessas interações comunitárias sejam computáveis e utilizáveis em um sistema de recomendação (ARAZY, 2009).

Sinha (2001) apresenta em seu estudo que a percepção da qualidade das recomendações feitas por amigos é muito superior àquela feita por sistemas de recomendação tradicionais. Isso demonstra que agregar o círculo social do usuário ao sistema de recomendações pode aumentar a confiabilidade do sistema como um todo.

2.1.5.1 O Papel da Confiança nas Recomendações Baseadas em Comunidade

Goldbeck (2006), explica que dados de redes sociais tem um grande potencial para utilização em sistemas de recomendação, pois a utilização das redes é cada vez maior, os dados estão disponíveis publicamente e, principalmente, são facilmente legíveis por

aplicações. Essa integração pode ser feita a partir de uma leitura inteligente de características específicas das conexões sociais dos usuários. Essas características incluem o tipo de relação (amigo, colega de trabalho, parente, etc.) e a intensidade da relação (amigo, melhor amigo, conhecido, etc.).

Existe uma correlação entre confiança e similaridade de usuários. As pessoas costumam desenvolver conexões em redes sociais com pessoas que possuem preferências similares às suas. O grau de desenvolvimento dessas relações denota a confiança mantida entre os usuários. No entanto, a confiança social depende de fatores que não são facilmente modeláveis em um sistema computacional, como experiências do passado, opiniões sobre atos da outra pessoa, fatores psicológicos e outras motivações. Por isso, para utilização em sistemas de recomendação, o conceito de confiança deve ser focado e simplificado. Confiança em uma pessoa passa a ser o comprometimento com uma ação baseado na crença de que as futuras ações da pessoa levarão a um bom resultado. Por exemplo, se um usuário *a* assiste um filme (comprometimento com uma ação) recomendado pelo usuário *b*, podemos dizer que o usuário *a* confia no usuário *b* no contexto de filmes (o usuário *a* acreditou que não seria uma perda de tempo assistir o filme indicado) (GOLDBECK, 2006).

2.1.6 Sistemas Híbridos

Os sistemas de recomendação híbridos são aqueles que combinam duas ou mais das técnicas apresentadas anteriormente para alcançar alguma sinergia entre elas. Os sistemas de filtragem colaborativa, por exemplo, apresentam problemas para dar recomendações a usuários novos, que ainda não têm muitas avaliações feitas para utilizar a técnica colaborativa. Para esses, poderia ser utilizada a técnica baseada em conhecimento até que o perfil do usuário seja mais conhecido. Dessa forma, um sistema de recomendação híbrido conjuga diferentes técnicas com o objetivo de produzir recomendações de melhor qualidade (BURKE, 2007).

Burke (2007) indica sete técnicas diferentes para combinar sistemas de recomendação:

1. **Ponderada:** a pontuação obtida a partir de diferentes tipos de sistema é combinada numericamente;
2. **Comutada:** o sistema escolhe uma dos resultados obtidos a partir de diferentes tipos de recomendação e o aplica;

3. **Mista:** as recomendações obtidas a partir de sistemas diferentes são apresentadas em conjunto;
4. **Combinação de Características:** características derivadas de fontes de conhecimento distintas são combinadas para produzir um algoritmo de recomendação único;
5. **Cascata:** cada técnica de recomendação possui um nível de prioridade e as de prioridade mais baixa desempatam as recomendações de pontuação semelhante obtidas nos níveis mais altos;
6. **Aumento de Características:** As saídas obtidas por uma técnica de recomendação são utilizadas como entrada para outra técnica de recomendação;
7. **Meta-nível:** um modelo é gerado por uma técnica de recomendação e serve então de entrada para a técnica seguinte.

3 MÉTODOS E PROCEDIMENTOS

Este trabalho é classificado como uma pesquisa de caráter qualitativo. Conforme Demo (2004), esse tipo de pesquisa parte da formulação de problemas a ser investigados e que possam ser reformulados durante o processo. Sua principal característica é a coleta de materiais em diversas fontes oriundas do ambiente natural, através de contato direto e prolongado entre o pesquisador e os atores sociais. Assim, é possível fundamentar exercícios de descrição para que os fenômenos investigados sejam compreendidos segundo a perspectiva dos participantes da situação em estudo.

Primeiramente, será feito um estudo das técnicas de sistemas de recomendação existentes. Após, será feita uma entrevista com um colaborador capacitado da empresa para determinar quais dessas técnicas deverão ser utilizadas pela empresa. Por fim, será aprofundado o estudo das técnicas selecionadas com a finalidade de propor um modelo de aplicação de sistemas de recomendação na empresa.

3.1 DELINEAMENTO DA PESQUISA

O método de pesquisa será um estudo de caso, com o conceito de pesquisa participante, conforme proposto por Yin (1994). Segundo ele, o estudo de caso não é apenas uma observação dos eventos, pois o pesquisador também participa da ação e deve fazer a análise dos resultados obtidos.

3.2 DEFINIÇÃO DA UNIDADE DE ANÁLISE

O estudo de caso foi realizado no Grupo Conectt, sediado em Porto Alegre, Rio Grande do Sul. Esta empresa está em processo de aquisição de conhecimento em sistemas de comércio eletrônico e a utilização de sistemas de recomendação é uma ótima oportunidade para esse objetivo.

A incorporação de novas funcionalidades que tragam diferencial competitivo é considerada importante na empresa. Desta forma, as técnicas de sistemas de recomendação selecionadas para o modelo final devem possibilitar o atendimento desse objetivo.

3.3 TÉCNICAS DE COLETA DE DADOS

Um roteiro de entrevista semiestruturado foi elaborado para determinar quais são os objetivos da empresa na utilização de sistemas de recomendação. Sua finalidade é identificar, diretamente com os colaboradores da empresa, quais as necessidades de agregação de valor aos produtos da empresa. A partir das respostas obtidas, foi possível definir os parâmetros desejados pela empresa e buscar, na literatura, as técnicas de sistemas de recomendação que atendam a essas expectativas.

3.4 TÉCNICAS DE ANÁLISE DE DADOS

A análise foi feita através do processo de análise de conteúdo, que visa ao exame de dados qualitativos. Trata-se de um conjunto de técnicas que permite a análise das comunicações a fim de obter, de maneira objetiva, a descrição do conteúdo das mensagens. Primeiramente, é preciso definir quais os objetivos da coleta de dados e, após, delimita-se o material a ser analisado e, então, é feita a análise (BARDIN, 1977).

3.5 LIMITAÇÕES DO MÉTODO

Esse estudo não poderá ser generalizado para outras empresas, haja vista que será feito de acordo com a realidade específica nela encontrada. Não há garantia de que o resultado da aplicação do conhecimento aqui adquirido será o mesmo para qualquer empresa.

4 APRESENTAÇÃO E ANÁLISE DOS DADOS

O roteiro de entrevistas (apêndice A) foi aplicado com o objetivo de determinar quais as técnicas de sistemas de recomendação que seriam úteis para a empresa. A entrevista foi realizada junto a um analista de sistemas que é colaborador da empresa. Esse colaborador foi o arquiteto da primeira solução de comércio virtual desenvolvida na empresa e é considerado pela empresa como o único especialista em comércio eletrônico do quadro de funcionários.

Logo foi diagnosticado que o conhecimento formal da empresa sobre sistemas de recomendação era limitado. Devido a esse fato, uma breve explicação sobre as técnicas existentes foi fornecida ao colaborador antes da entrevista. Dessa forma, o colaborador teve mais condições de estabelecer critérios de comparação e adequar as respostas às necessidades reais da empresa.

A entrevista evidenciou que existe uma preocupação da empresa em alcançar competitividade no sistema de recomendação desenvolvido, tendo a capacidade de atender aos padrões mínimos do mercado atual. Também foi evidenciada a necessidade de atingir um diferencial competitivo, distinguindo o produto desenvolvido pela empresa de alguma maneira.

Outro aspecto importante que a entrevista apontou, foi o de que não devem ser empreendidos esforços, em um primeiro momento, em técnicas que não tenham um retorno tangível. Isso implica que o modelo proposto deve ser o mais genérico possível, para que possa ser aplicado e adaptado facilmente em qualquer tipo de sistema de comércio virtual.

As técnicas de recomendações baseadas em conhecimento e recomendações demográficas foram consideradas inadequadas devido aos motivos expostos acima. São técnicas de uso muito específico e sua utilização nos tipos de comércio virtual desenvolvidos pela empresa não se mostrou atrativa.

As técnicas selecionadas para a pesquisa foram recomendações baseadas em conteúdo, filtragem colaborativa e recomendações baseadas em comunidade. As duas primeiras foram selecionadas para ser utilizadas com o objetivo de atender a um padrão de mercado, para que o produto desenvolvido pela empresa seja comparável ao que existe

atualmente. Já as recomendações baseadas em comunidade foram selecionadas para ser trabalhadas como um diferencial do sistema desenvolvido, agregando ao produto um valor diferenciado em relação ao mercado.

A conclusão final da análise é que o modelo de sistema de recomendação será híbrido, pois agregará mais de uma técnica de sistema de recomendação.

Abaixo é explicado como cada uma das técnicas selecionadas deverá ser utilizada no contexto de desenvolvimento de sistemas de comércio virtual da empresa. Por fim, é elucidado como será feito o encadeamento entre as técnicas, possibilitando a proposta de um modelo de sistema de recomendação híbrido. Como a empresa não possui um sistema próprio, mas sim desenvolve de acordo com a demanda do cliente, é esclarecido o que o sistema deve possuir para que as recomendações possam ser geradas.

4.1 CADASTRO E IDENTIFICAÇÃO DE USUÁRIOS

Para que o sistema de recomendação seja utilizado de maneira plena, é muito importante que sejam identificados seus usuários. Para isso, deve ser utilizado o próprio cadastro de usuários feito para a efetivação de compras. Dessa forma, enquanto o usuário estiver conectado ao site, é possível registrar suas ações e empregá-las como fonte de dados para o sistema de recomendação. Nesse cadastro será registrado o perfil do usuário, que será alimentado e utilizado pelas técnicas de recomendação explicadas abaixo.

O perfil do usuário será montado de acordo com o uso que ele faz do sistema. No entanto, essa característica faz com que só possam ser feitas recomendações a usuários recorrentes. Para amenizar esse problema, pode ser criada uma forma de o usuário informar suas principais preferências, com o intuito de montar um perfil inicial e utilizá-lo nas primeiras recomendações. O preenchimento do cadastro deve ser apenas sugerido, pois nem sempre haverá disposição, por parte do usuário, em completar esse tipo de informação.

A forma mais segura e assertiva de registrar dados sobre o comportamento do usuário é através do cadastro acima. No entanto, existe uma gama de usuários que podem visitar o site sem estarem registrados e não faria sentido restringir a utilização de uma loja virtual apenas a usuários registrados. Para isso, é preciso que sejam registradas as visitas anônimas da loja em um cadastro de usuários anônimos. Assim, será possível montar o perfil de usuários anônimos e utilizá-lo nas comparações da filtragem colaborativa, por exemplo.

Para identificar um usuário anônimo, pode ser gravado um *cookie* na máquina do usuário e, no futuro, utilizá-lo para detectar seu retorno à loja.

Outra motivação do registro de usuários anônimos é o fato de que, enquanto o usuário não se cadastra, suas ações podem ser registradas. Usuários podem fazer uma série de navegações na loja antes de tomarem uma decisão de compra. Os dados dessa navegação ficariam perdidos se não houvesse uma identificação de usuários anônimos. No momento que houver seu cadastro, o usuário anônimo passará automaticamente a ser tratado como um usuário identificado e seu perfil já possuirá dados.

4.2 RECOMENDAÇÕES BASEADAS EM CONTEÚDO

Para que a recomendação baseada em conteúdo seja utilizada, o sistema deverá possuir um cadastro de itens bem especificado. Devem ser definidos quais atributos dos itens serão utilizados para determinar a similaridade entre eles. Dentre os atributos utilizados, devem ser criados critérios para estabelecer a comparação entre os atributos de dois itens distintos. Cada atributo deve ter um peso, que indicará o quão importante ele é no momento de gerar uma recomendação, para que a pessoa tome uma decisão. A determinação de quais atributos terão mais peso dependerá da área de negócio para a qual estiver sendo desenvolvido o sistema de comércio virtual. Por exemplo, em uma livraria, o atributo “gênero” possivelmente será mais importante que o atributo “número de páginas” e, portanto, deve possuir um peso superior.

Cada item do sistema também poderá ser avaliado pelos usuários do site. Para facilitar a computação dos dados, o critério de avaliação deve preferencialmente ser em escala (de 1 a 5, por exemplo). Todos os itens do site devem possuir o mesmo critério de avaliação para que possam ser comparados da mesma maneira.

Também devem ser criados critérios de avaliação implícita que possam denotar preferências dos usuários. Podem ser registrados o tempo de navegação em um item, tempo de navegação em uma categoria de produtos, compartilhamento de um item em redes sociais, entre outros. Outro fator a ser registrado é se um usuário efetivamente compra um item, pois essa é a indicação mais clara de preferência sobre um item do site. Os atributos do item comprado deverão ser agregados ao perfil do usuário para futuras comparações.

Após a identificação e criação dos atributos importantes do item, são gerados os meios para o usuário avaliar esses itens e, a partir daí, deve ser montado o perfil do usuário. Ele será montado a partir dos atributos de cada item e as avaliações realizadas pelo usuário.

4.3 FILTRAGEM COLABORATIVA

Para utilização da filtragem colaborativa, deverão ser utilizadas as avaliações feitas pelos usuários sobre cada item do sistema. As avaliações de cada usuário devem ser comparadas com as dos demais. Havendo similaridade entre as avaliações, os usuários poderão ser considerados vizinhos, ou seja, possuem similaridade de preferências. Um usuário poderá ter, portanto, uma lista de vizinhos. Todos os itens bem avaliados pela lista de vizinhos do usuário poderão lhe ser recomendados, assim como os itens mal avaliados pelos vizinhos de um usuário poderão ser escondidos.

Cada avaliação realizada entrará em um banco de avaliações não processadas. Um processo periódico percorrerá esse banco e montará uma lista de vizinhos para cada usuário. Essa lista deve ser ordenada de acordo com o grau de similaridade entre os usuários. A partir disso, será possível varrer a lista de produtos avaliados pela vizinhança de usuários, montando a lista de produtos a recomendar e a lista de produtos a esconder. Os produtos avaliados pelos usuários mais similares deverão ser os primeiros a ser recomendados (quando a avaliação for positiva) ou escondidos (quando a avaliação for negativa).

A utilização de um processo periódico para gerar a lista de recomendações melhora o desempenho do sistema, pois seria muito custoso redefinir a vizinhança de usuários a cada avaliação realizada. Assim, pode ser gerada uma nova lista de recomendações a cada avaliação feita. Também é importante destacar que, quando um usuário faz novas avaliações, toda sua vizinhança poderá ser modificada, incluindo novos vizinhos ou aumentando a similaridade entre vizinhos. Isso significa que não apenas a lista de produtos do usuário avaliador terá de ser refeita, mas a também a de toda a sua vizinhança.

4.4 RECOMENDAÇÕES BASEADAS EM COMUNIDADE

As redes sociais serão utilizadas para potencializar a coleta de informações a respeito de cada usuário. Para que esses dados possam ser utilizados, o usuário deverá ativamente permitir que seu cadastro seja vinculado a uma rede social da qual ele participe. O sistema deverá estar preparado para vincular os perfis de usuários e coletar dados de redes

sociais. A utilização destes dados servirá para aumentar as características das duas técnicas anteriores.

A escolha de quais redes sociais serão utilizadas dependerá do momento em que o sistema for construído e do nicho de mercado do detentor da loja virtual. Deve ser dada preferência às redes sociais mais utilizadas no momento e também àquelas que possuem mais usuários integrantes do alvo de mercado proposto pela loja virtual. Também deve ser avaliado o quanto é possível obter dados da rede social, ou seja, se é viável obter os dados automaticamente a partir de um sistema.

Quando o usuário permite a vinculação de seu cadastro a uma rede social, pode ser obtida sua rede de relacionamentos. A partir da obtenção dessa rede, deve-se avaliar o quão forte é a relação entre o usuário e seus contatos na rede, que pode ser medido a partir do número de interações que o usuário possui. Se algum dos contatos do usuário também estiver cadastrado na loja virtual, poderá ser estabelecida uma relação de similaridade entre eles. Esses usuários passarão a ser considerados vizinhos, da mesma forma que na filtragem colaborativa. Com isso, todo histórico desse par de usuários poderá ser utilizado para gerar recomendações.

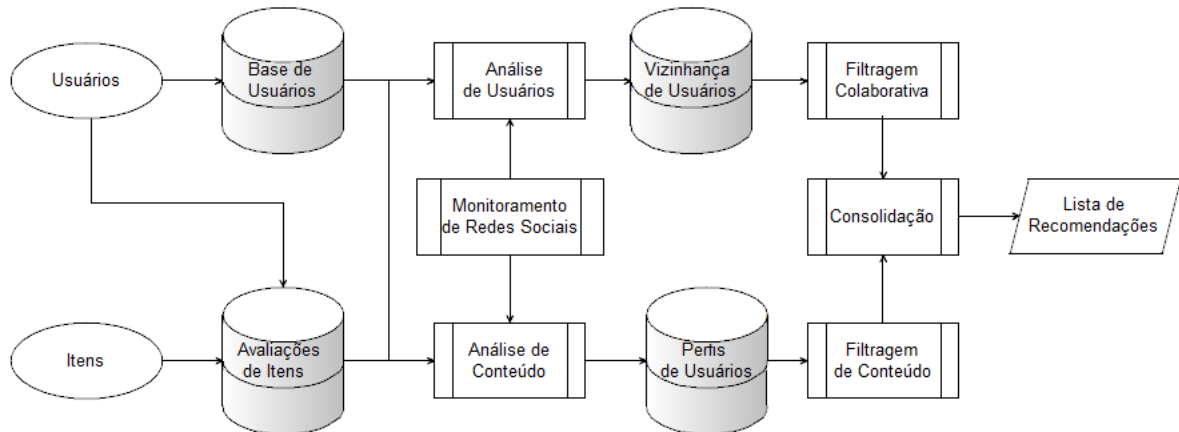
A partir do momento em que o usuário tem seu cadastro vinculado a uma rede social, suas atividades na rede poderão ser monitoradas. Aquilo que o usuário escreve na rede deve ser monitorado e, se houver relação com atributos de itens da loja virtual, pode ser utilizado para atualizar o perfil do usuário para o sistema de recomendação. Se a loja for de filmes, por exemplo, quando o usuário escrever em uma rede social que gosta de um filme de ação, pode-se aumentar o valor do atributo filmes de ação no perfil do usuário, passando a recomendar também itens desse gênero.

Da mesma forma, interações futuras entre dois usuários na rede social podem ser utilizadas para aumentar o perfil de similaridade entre eles. No momento do cadastro, dois usuários podem ter sido identificados como pouco similares, mas isso não significa que continuarão assim ao longo do tempo. Por isso, é importante monitorar as interações entre os usuários à medida que elas acontecem, pois acrescenta dinamismo à atualização de perfis dos usuários do sistema.

4.5 MODELO HÍBRIDO PROPOSTO

A figura 2 apresenta como as técnicas selecionadas a partir da entrevista realizada serão utilizadas em conjunto para produzir recomendações.

Figura 2 – Modelo Híbrido Proposto



Fonte: elaborado pelo autor

As principais fontes de informação do modelo são os usuários (com seus dados cadastrais e avaliações) e os itens do sistema (com seus atributos). Eles são utilizados, respectivamente, no algoritmo de filtragem colaborativa e na filtragem de conteúdo. O monitoramento das atividades dos usuários em redes sociais serve como meio auxiliar de obtenção de informações úteis aos algoritmos de recomendação.

Em uma etapa anterior à recomendação por filtragem colaborativa, é preciso determinar a vizinhança de usuários. Isso é feito no processo chamado de Análise de Usuários. Já para montar as recomendações baseadas em conteúdo, é preciso montar os perfis de usuários, que são feitos no processo Análise de Conteúdo, que avalia as preferências do usuário, de acordo com as avaliações que já foram realizadas no passado.

Ambos os processos têm, como entrada, os dados obtidos no monitoramento de redes sociais. Dessa forma, está se empregando a técnica de aumento de características descrita por Burke (2007). Esses dados servem para determinar o grau de similaridade

Assim, são geradas duas listas de recomendações, uma pela filtragem colaborativa e outra pela filtragem de conteúdo. É preciso consolidar essas informações em uma lista única, o que é feito no processo de consolidação. Esse processo emprega a técnica mista descrita por Burke (2007), onde os resultados obtidos a partir de fontes diferentes são apresentados em conjunto.

4.5.1 Sinergia do Modelo Proposto

A utilização de um modelo híbrido traz vantagens, pois permite uma sinergia entre as técnicas utilizadas, reforçando características ou mesmo corrigindo deficiências entre elas. Assim, a qualidade das recomendações é aumentada, aumentando a gama de benefícios obtidos.

A técnica de recomendação baseada em conteúdo possui tendência a produzir recomendações superespecializadas. Isto é corrigido através da utilização da filtragem colaborativa, que reconhece preferências inesperadas através da identificação de usuários similares.

Outra fragilidade da recomendação baseada em conteúdo é que esta depende da interação do usuário com o site, para que sejam coletadas as avaliações de cada item disponível. Se o número e frequência das interações forem abaixo do esperado, podem ser causados problemas como perfil incompleto ou desatualizado. A monitoração das atividades dos usuários em redes sociais amenizará esse problema, pois permitirá o incremento e atualização de perfis sem a necessidade de intervenção dos usuários.

Tanto a recomendação baseada em conteúdo quanto a filtragem colaborativa possuem o problema de não conseguir produzir recomendações para novos usuários, pois dependem da interação dos usuários com o sistema. Mais uma vez, a monitoração das redes sociais auxiliará na resolução desse problema, pois será possível coletar dados tanto sobre o perfil do usuário, como sobre a sua lista de vizinhos. Isso permite um aumento na gama de dados disponíveis para avaliação pelo sistema de recomendação, tanto no início do processo quanto ao longo do uso do sistema.

5 CONSIDERAÇÕES FINAIS

A internet é um esplêndido meio de obtenção de informações. Ao mesmo tempo, a grande quantidade de dados disponível torna custosa a tarefa de encontrar a informação efetivamente desejada. Essa realidade também se aplica a sistemas de comércio eletrônico, onde o número de produtos disponíveis, por vezes, é muito grande. Por isso, é importante aplicar alguma inteligência à tarefa de recomendar itens a um usuário. Os sistemas de recomendação são a ferramenta ideal para esse fim.

O presente trabalho estudou as técnicas de recomendação disponíveis para elaborar um modelo que possibilite a utilização de sistemas de recomendação nos sistemas de comércio eletrônico desenvolvidos pela empresa alvo da pesquisa. Foram estudadas as técnicas de recomendação baseada em conteúdo, filtragem colaborativa, recomendação baseada em conhecimento, recomendações demográficas e recomendações baseadas em comunidade.

A aplicação da entrevista junto à empresa evidenciou suas necessidades em sistemas de recomendação. Isso permitiu a seleção das técnicas que atenderiam a suas necessidades, permitindo criar uma base para o desenvolvimento de sistemas no futuro.

Como as necessidades da empresa se mostraram amplas, foi necessário utilizar mais de uma técnica de recomendação no modelo proposto. Isso se mostrou proveitoso, pois cada técnica possui vantagens e desvantagens e a utilização de um modelo híbrido permite o equilíbrio entre elas, produzindo melhores recomendações.

O modelo proposto utiliza três técnicas de recomendação, mas também não restringe que, no futuro, sejam retiradas algumas delas ou mesmo adicionadas outras. Isto é importante na realidade da empresa estudada, pois os sistemas desenvolvidos por ela são feitos sob medida para cada cliente. A realidade encontrada em cada cliente pode ser diferente e é importante estar preparado para adaptar-se às suas necessidades específicas.

Outro ponto positivo do trabalho é a utilização de dados de redes sociais no modelo. Esta característica torna as recomendações mais agregadas à realidade atual de utilização da internet. Também é importante manter um ponto de atenção em relação a esse

aspecto, pois as mudanças tecnológicas nas redes sociais são constantes e o modelo deve estar preparado para alterações quando necessário.

O estudo demonstrou ser de grande validade, tanto para o pesquisador, quanto para a empresa selecionada. O pesquisador pôde aplicar conhecimentos adquiridos no curso e a empresa pôde identificar necessidades em relação a sistemas de recomendação e obter uma proposta de solução. O fato de o modelo ter sido elaborado com base nas necessidades de apenas uma empresa não permite a generalização do mesmo. Assim, seriam interessantes pesquisas futuras visando à aplicação do modelo em outras empresas.

O trabalho possui a restrição de ser apenas um modelo teórico, sem trazer dados quanto à sua aplicação real. Sugere-se que, no futuro, o modelo seja submetido à avaliação da empresa e, após sua aprovação, seja desenvolvido um protótipo. Assim, será possível testar as recomendações produzidas e ajustar o modelo para obter o melhor desempenho. Após essa etapa, deve ser testada sua implementação em uma aplicação real, colhendo os resultados para avaliar o desempenho da aplicação do modelo final.

REFERÊNCIAS

ABBASSI, Zeinab; AMER-YAHIA, Sihem; LAKSHMANAN, Laks; VASSILVITSKI, Sergei; YU, Cong. **Getting Recommender Systems to Think Outside the Box**. 2009. Disponível em: <research.yahoo.com/files/recsys09-otb.pdf>. Acesso em 20/07/2011 às 16:05.

ADOMAVICIUS, Gediminas; TUZHILIN, Alexander. Toward the Next Generation of Recommender Systems: a survey of the state-of-the-art and possible extensions. **IEEE Transactions on Knowledge and Data Engineering**, Washington, vol. 17, nº 6, p. 734-749, Junho/2005.

ALBERTIN, Alberto Luiz. **Comércio Eletrônico: modelo, aspectos e contribuições de sua aplicação**. 4. ed. São Paulo: Atlas, 2002.

ARAZY, O.; KUMAR, N.; SHAPIRA, B. Improving social recommender systems. **IT Professional**, vol. 11, nº 4, p. 38–44, 2009.

BARDIN, Laurence. **Análise de Conteúdo**. 1. ed. Lisboa: Edições 70, 1977.

BASIRI, J.; SHAKERY, A.; MOSHIRIU, B.; HAYAT, M. Z. Alleviating the cold-start problem of recommender systems using a new hybrid approach. **2010 5th International Symposium on Telecommunications (IST)**, 2010.

BERRY, Michal J. A.; LINOFF, Gordon S. **Data Mining Techniques: for marketing, sales and customer relationship management**. 2. ed. Indianapolis: Wiley Publishing, 2004.

BILLSUS, D.; PAZZANI, M.J. A Hybrid User Model for News Story Classification. In: **Proceedings of the Seventh International Conference on User Modeling**. Banff, Canada, 1999.

BLANCO-FERNANDEZ Y., PAZOS-ARIAS J. J., G.S.A., RAMOS-CABRER, M., LOPEZ-NORES, M. Providing Entertainment by Content-based Filtering and Semantic Reasoning in Intelligent Recommender Systems. **IEEE Transactions on Consumer Electronics**, Washington, vol. 54, nº 2, p. 727–735, Maio/2008.

BURKE, R. **Hybrid web recommender systems**. In: *The AdaptiveWeb*, p. 377–408. Springer. Berlin: Heidelberg, 2007.

CAMERON, D. **Electronic Commerce: the new business platform of the Internet**. Charleston: Computer Technology Research Corp., 1997.

CHEN, Pei-Yu, CHOU, Yen-Chun, KAUFFMANN, Robert J. Community-Based Recommender Systems: analyzing business model from a system operator's perspective. **Proceedings of the 42nd Hawaii International Conference on System Sciences**, Hawaii, p. 1-10, 2009.

DEMO, Pedro. **Pesquisa participante: saber pensar e intervir juntos**. Brasília: Líber Livro, 2004.

E-COMMERCE BRASIL. **Vendas Comércio Eletrônico – Brasil**. 2011. Disponível em: <<http://www.e-commerce.org.br/stats.php>>. Acesso em 20/06/2011 às 13:24.

FELLBAUM, C. **WordNet: An Electronic Lexical Database**. MIT Press, 1998.

GOLDBECK, Jennifer. **Generating Predictive Movie Recommendations from Trust in Social Networks**. 2006. Disponível em: <<http://trust.mindswap.org/papers/iTrust06.pdf>>. Acesso em 15/07/2011 às 21:43.

JIN, Rong; Si, Luo. A Study of Methods for Normalizing User Ratings in Collaborative Filtering. **Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval**, 2004.

KNIDEL, Helder. **O Uso de Sistemas de Recomendação e a Melhoria na Experiência de Compra**. 2010. Disponível em: <<http://www.e-commercebrasil.com.br/artigos/o-uso-de-sistemas-de-recomendacao-e-a-melhoria-na-experiencia-de-compra/>>. Acesso em 20/06/2011 às 13:38.

KONSTAN, J.A., MILLER, B.N., MALTZ, D., HERLOCKER, J.L., GORDON, L.R., RIEDL, J. GroupLens: applying collaborative filtering to usenet news. **Communications of the ACM**, vol. 40, n° 3, p. 77–87, 1997.

LINDEN, G., SMITH, B., YORK, J. Amazon.com recommendations: item-to-item collaborative filtering. **IEEE Internet Computing**, Washington, vol. 7, n° 1, p. 76–80, 2003.

MITCHELL, T. **Machine Learning**. McGraw-Hill, New York, 1997.

PAZZANI, M.J. A framework for collaborative, content-based and demographic filtering. **Artificial Intelligence Review**, vol. 13, p. 393–408, 1999.

RICCI, Francesco; ROKACH, Lior; SHAPIRA, Bracha; KANTOR, PaulB. **Recommender Systems Handbook**. 1. Ed. Nova York: Springer, 2011.

SCHEIN, A.I., POPESCU, A., UNGAR, L.H., PENNOC, D.M: **Methods and metrics for cold-start recommendations**. In: SIGIR '02: Proc. of the 25th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, p. 253–260. ACM, Nova York, 2011.

SHEN, Lei; ZHOU, Yiming. A New User Similarity Measure for Collaborative Filter Algorithm. **Second International Conference on Computer Modeling and Simulation**, 2010.

SINHA, Rashmi; SWEARINGEN, Kirsten. **Comparing Recommendations Made By Online Systems and Friends**. In: Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries, 2001.

YIN, Robert K., Case Study Research: design and methods. 2. ed. Sage, 1994.

APÊNDICE A

ROTEIRO DE ENTREVISTA

1. O que a empresa conhece sobre sistemas de recomendação?

A empresa conhece alguns padrões do mercado. Por exemplo, sistemas que mostram que pessoas que compraram o item x também compraram o item y. Também conhecemos recomendações que mostram itens semelhantes aos que estão sendo vistos pelo usuário em determinado momento. Além de recomendações baseadas no perfil do usuário (ofertas especiais para você). No entanto, a empresa nunca desenvolveu sistemas de recomendação.

2. O que é considerado mais importante: um sistema que seja capaz de recomendar itens imediatamente a usuários novos do site ou que aprenda as preferências do usuário ao longo do tempo?

Depende do tipo de produto que será vendido. Se o site possuir usuários recorrentes, é muito importante que ocorra um aprendizado com o usuário, para que seja possível refinar as buscas. Já se o site possuir mais usuários eventuais, o esforço do sistema deve estar em captar a atenção do usuário de imediato, oferecendo itens úteis imediatamente. No entanto, essa segunda maneira demanda um esforço para identificar o tipo de usuário com a precisão necessária para gerar boas recomendações.

3. O sistema deve possuir a capacidade de recomendar produtos de acordo com a similaridade entre eles?

Sim, pois consideramos o desenvolvimento desse tipo de sistema como relativamente simples. Acreditamos que não poderia ser oferecido um sistema de recomendação sem esse tipo de funcionalidade, devido à perda de competitividade no mercado.

4. O sistema deve possuir a capacidade de recomendar produtos a partir da similaridade entre usuários?

Sim, pois assim como na resposta anterior, acreditamos que um sistema de recomendação sem essa funcionalidade não seria competitivo.

5. É importante que informações demográficas dos usuários (localização, idade, sexo, língua) sejam consideradas no sistema de recomendação?

Não, porque a obtenção desses dados pode ser muito custosa e, mais ainda, a criação de uma heurística complexa sobre esse tipo de dado não traria resultados aproveitáveis.

6. Haveria espaço para incorporar ao sistema regras específicas sobre como os requisitos de um usuário são atendidos pelos produtos oferecidos?

Não, pois é muito difícil estipular esse tipo de regra de uma forma genérica. Haveria um custo muito alto para implantar essa funcionalidade e um retorno baixo em recomendações úteis ao usuário.

7. A integração com redes sociais, aproveitando as relações nelas estabelecidas, deve ser trabalhada pelo sistema de recomendação?

Sim, pois hoje esse tipo de estratégia é fundamental para fazer com que o cliente conheça o site e venha a utilizá-lo mais de uma vez. Há uma facilidade para obter dados de redes sociais e a aderência das pessoas a esse tipo de ferramenta é muito alta. Consideramos esse um dos melhores recursos para fazer com que o cliente volte.

8. O sistema deve possuir mais de um tipo possível de recomendação ao usuário?

Acreditamos que sim. Recomendações baseadas em conteúdo e filtragem colaborativa devem fazer parte do sistema por uma questão de competitividade. Também acreditamos que recomendações baseadas em comunidade podem ser um diferencial, tornando o sistema mais aderente às tendências atuais de comércio virtual.

9. Dadas as cinco técnicas de sistemas de recomendação abaixo, ordene-as da mais importante à menos importante:

- (1) Sistemas de recomendação baseados em conteúdo
- (2) Filtragem colaborativa
- (4) Recomendações Demográficas
- (5) Recomendações Baseadas em Conhecimento
- (3) Recomendações Baseadas em Comunidade