



COPPE/UFRJ

UMA TAXONOMIA E PROPOSTA DE SISTEMA DE RECOMENDAÇÃO
PARA REDES DE VAREJO DE MODA

Pedro Peralva Figueiredo

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Produção, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Produção.

Orientador: Marcos do Couto Bezerra Cavalcanti

RIO DE JANEIRO, RJ - BRASIL

AGOSTO DE 2009

UMA TAXONOMIA E PROPOSTA DE SISTEMA DE RECOMENDAÇÃO
PARA REDES DE VAREJO DE MODA

Pedro Peralva Figueiredo

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE PRODUÇÃO.

Aprovada por:

Prof. Marcos do Couto Bezerra Cavalcanti, Ph.D.

Prof. Elton Fernandes, Ph.D.

Prof^a. Raquel Borba Balceiro, Ph.D.

RIO DE JANEIRO, RJ - BRASIL

AGOSTO DE 2009

Figueiredo, Pedro Peralva

Uma Taxonomia e Proposta de Sistema de Recomendação para Redes de Varejo de Moda/ Pedro Peralva Figueiredo. – Rio de Janeiro: UFRJ/COPPE, 2009.

XII, 77 p.: il.; 29,7 cm.

Orientador: Marcos do Couto Bezerra Cavalcanti

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de Engenharia de Produção, 2009.

Referencias Bibliográficas: p. 74-77.

1. Sistemas de Recomendação. I. Cavalcanti, Marcos do Couto Bezerra. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Produção. III. Título.

Dedico aos que abriram o caminho para que eu pudesse seguir.

AGRADECIMENTOS

Ao Criador por permitir que minha jornada tenha sido mais do que proveitosa e agradável;

Também aos amigos do plano espiritual que sempre estiveram presentes e ajudaram com recomendações sobre coisas que eu não conseguia enxergar;

Aos meus pais Cristina e Reginaldo pelo apoio incondicional e irrestrito;

Ao Alvaro pelo coração, paciência, apoio, e amor;

À Mariazinha e ao Joaquim pelo carinho e amor;

Ao Marcos Cavalcanti pela paciência infundável e por me mostrar que sempre há luz no fim do túnel;

À Raquel pela garra e disciplina que sempre me inspiraram;

Ao professor Elton por ter me lembrado que as forças e oportunidades andam juntas;

À Paula e ao Damião pelo bom-senso e alegria no trabalho do Crie;

À Tatiana e à Patrícia pela inspiração, ao abrirem meus olhos para um mercado completamente novo; e

À equipe da secretaria do Programa de Engenharia de Produção pela atenção e apoio nas minhas demandas.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

UMA TAXONOMIA E PROPOSTA DE SISTEMA DE RECOMENDAÇÃO
PARA REDES DE VAREJO DE MODA

Pedro Peralva Figueiredo

Agosto de 2009

Orientador: Marcos do Couto Bezerra Cavalcanti

Programa: Engenharia de Produção

O estudo de sistemas de recomendação é relativamente novo, razão inicial para a motivação desta dissertação. Ao estudar melhor o mercado de varejo de moda no Brasil, percebe-se o vasto espaço para a aplicação dessas soluções. Basta correlacionar as peças e os clientes para que isso fique ainda mais evidente. Em contrapartida, o estudo faz perceber a falta de aplicações desse tipo de sistema em ambientes *off-line*. Este trabalho parte da revisão teórica do tema em questão e propõe a aplicação de soluções de recomendação nas redes de varejo de moda, como forma de melhor aproveitar os investimentos em *marketing* e aperfeiçoar a reposição dos estoques das lojas.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

A RECOMMENDER SYSTEM TAXONOMY AND PROPOSAL
FOR THE FASHION RETAIL INDUSTRY

Pedro Peralva Figueiredo

August 2009

Advisor: Marcos do Couto Bezerra Cavalcanti

Department: Production Engineering

The study of recommender systems is relatively new, which provided the early reason for this dissertation. While studying the fashion retail industry in Brazil, it became clear that there is room for the application of such solutions. The correlation of the articles and the customers makes it even more apparent. On the other hand such solutions are not commonly used on off-line environments. This work goes from the basic revision of concepts and proposes the use of recommender systems in the fashion retail networks obtaining better profits from marketing investments and optimizing the stock reposition in stores.

SUMÁRIO

1.	INTRODUÇÃO.....	1
1.1	Motivação.....	3
1.2	Objetivos	5
1.3	Justificativa e relevância do trabalho.....	5
1.4	Delimitação da dissertação.....	5
1.5	Organização do texto	6
1.6	O que foi visto neste capítulo	7
2.	SISTEMAS DE RECOMENDAÇÃO.....	8
2.1	Apresentação.....	8
2.1.1	A origem dos sistemas de recomendação: o boca a boca.....	9
2.1.2	Tapestry: usando filtragem colaborativa	10
2.1.3	Da diferença entre recuperação e filtragem de informações.....	11
2.2	Taxonomia de sistemas de recomendação	12
2.3	Exemplos de aplicações de sistemas de recomendação.....	46
2.3.1	O laboratório GroupLens	46
2.3.2	Aplicações de comércio eletrônico.....	46
2.3.3	Aplicações voltadas ao entretenimento	47
2.3.4	Alguns produtos e serviços de recomendação	47
2.4	Outros aspectos associados aos sistemas de recomendação	47
2.4.1	Volatilidade das informações.....	47
2.4.2	Mudanças nas preferências dos usuários	48
2.4.3	Recomendações imprevisíveis	48
2.4.4	Peças que ainda não foram vendidas não podem ser comparadas.....	48
2.5	O que foi visto neste capítulo	49
3.	METODOLOGIA DE PESQUISA.....	50
3.1	Procedimentos metodológicos e caracterização da pesquisa.....	50
3.2	Delineando o problema	51
3.3	Etapas da pesquisa	52

3.3.1	Justificando o sistema de recomendação para redes de varejo de moda...	53
3.3.2	Definindo a abordagem do sistema face à taxonomia adotada.....	53
3.3.3	Detalhando a operação da solução proposta.....	58
3.4	O que foi visto neste capítulo	59
4.	DESENVOLVIMENTO DO PROTÓTIPO	60
4.1	Visão geral do mercado de varejo de moda	60
4.2	As funções identificadas no protótipo.....	61
4.2.1	Compra as peças.....	63
4.2.2	Registra a venda no sistema de retaguarda	63
4.2.3	Exporta os dados do sistema de retaguarda	63
4.2.4	Recebe os dados que foram exportados.....	63
4.2.5	Realiza a correlação entre peças	64
4.2.6	Disponibiliza os resultados.....	68
4.2.7	Recebe os resultados das correlações	69
4.2.8	Desenha as promoções que vai realizar.....	70
4.3	O protótipo	70
4.4	Dos resultados	71
4.5	O que foi visto neste capítulo	71
5.	CONTRIBUIÇÕES E PERSPECTIVAS FUTURAS.....	72
5.1	Contribuições.....	72
5.2	Perspectivas futuras	72
6.	REFERÊNCIAS BIBLIOGRÁFICAS	74
6.1	Bibliográficas e documentais	74
6.2	Sítios visitados	76

ÍNDICE DE FIGURAS

Figura 1 - Origem dos conceitos de filtragem.	12
Figura 2 - Principais entidades dos sistemas de recomendação.....	13
Figura 3 - Taxonomia dos sistemas de recomendação.	14
Figura 4 – Classificação da justificativa do sistema.	14
Figura 5 - Consulta às anotações do livro “O Pequeno Príncipe” no sítio Submarino...	15
Figura 6 - Sugestões do tipo "veja também" no sítio da Richads.....	16
Figura 7 - Interface do sítio Youtube que oferece uma seqüência de itens.	18
Figura 8 - Exemplo de tela da solução BeeTV.....	19
Figura 9 - Exemplo de itens oferecidos por ordenação.....	20
Figura 10 - Exemplo de como a Amazon oferece produtos.	21
Figura 11 – Classificação da abordagem do sistema.	22
Figura 12 – Classificação do perfil do usuário.	23
Figura 13 - Exemplo de dados coletados de forma implícita.	26
Figura 14 – Classificação das características dos itens.....	28
Figura 15 - Exemplo de apresentação de itens por meio de taxonomia.	29
Figura 16 - Exemplo de apresentação de itens usando avaliações.	30
Figura 17 - Exemplo de itens selecionados com base em atributos.	31
Figura 18 – Classificação da entrega.	32
Figura 19 - Exemplo de recomendação efêmera.	33
Figura 20 - Exemplo de lista de recomendação.	39
Figura 21 - Tela de recomendações para o usuário no sítio da Amazon.	40
Figura 22 - Exemplo de avaliação de um item.....	42
Figura 23 – Classificação da operação do sistema.	43
Figura 24 - Diagrama de como o Loomia funciona (original em inglês).....	44
Figura 25 - Exemplo de filtro de e-mail baseado em regras.....	45
Figura 26 - Passos da pesquisa.	52
Figura 27 – Relação entre conceitos das redes de varejo de moda.....	61
Figura 28 - Diagrama funcional do macro-processo do protótipo.	62
Figura 29 - Quantidades de peças de camisetas básicas e de manga comprida.	65

Figura 30 - Tela de consulta a uma peça.	68
Figura 31 – Tela do relatório de recomendações.....	69

ÍNDICE DE TABELAS

Tabela 1 – Justificativas do sistema de recomendação.....	53
Tabela 2 – Classificação do perfil do usuário no protótipo.....	54
Tabela 3 – Classificação das características dos itens no protótipo.	55
Tabela 4 – Classificação das características do protótipo em relação à entrega.....	56
Tabela 5 – Classificação das características do protótipo em relação à operação.	58
Tabela 6 – Quantidade de peças compradas pelos consumidores.	64
Tabela 7 – Valores utilizados no cálculo do coeficiente de Pearson para dois artigos selecionados. Fonte: elaborado pelo autor.....	67
Tabela 8 – Resultados obtidos usando-se o coeficiente de Pearson.	68
Tabela 9 – Estrutura do registro de transação.	70

1. INTRODUÇÃO

“The way to maximize freedom is to maximize choice.”¹ - Barry Schwartz, 2005 (informação verbal).

Quando os primeiros homens começaram a fazer o registro de suas atividades - por exemplo, sobre qual seria a melhor forma de caçar ou plantar -, observou-se a passagem de um conhecimento tácito para a sua forma explícita.

Pode-se também dizer que, no caso acima, o autor recomendou uma forma de caçar ou plantar para a próxima pessoa que visse aquele registro. Assim, ele transmitiu a informação de que “um local tem peixes em abundância”, ou “boas condições de plantio”.

Essencialmente falando, o que ele fez foi uma recomendação. Não foi uma recomendação solicitada por outrem, nem tampouco foi uma recomendação personalizada, mas foi essencialmente uma recomendação.

<p>Recomendação s.f. 1) ato ou efeito de recomendar. 2) Derivação: por metonímia. Aquilo que adverte; conselho, advertência, aviso Ex.: esquecia com frequência as <i>recomendações</i> da mãe. 3) qualidade de recomendável. Ex.: a grande clientela é sua melhor <i>recomendação</i>.</p>
--

Quadro 1 – Uma definição de recomendação.
Fonte: dicionário Houaiss da Língua Portuguesa.

No mundo real, a recomendação de produtos, serviços ou qualquer tipo de item se constitui em uma prática muito antiga, podendo ser feita por meio do que Shardanand e Maes (1995) chamaram de “recomendação boca-a-boca” (ou *word of mouth* no original).

A criação do alfabeto, a invenção da prensa, e a popularização da Internet permitiram grandes saltos na comunicação, a partir do momento em que as informações se tornaram, em cada um desses eventos, disponíveis a um público cada vez maior.

¹ A forma de maximizar a liberdade é maximizar as opções.

No entanto, o progresso das tecnologias de comunicação e informação (alfabeto, prensa e Internet), nunca foi tão grande quanto o observado nos últimos anos. Esse fato levou a um fenômeno conhecido como sobrecarga de informação (ou *information overload* em inglês), onde se tem acesso a um número de informações muito maior do que se pode consumir.

As primeiras soluções para resolver o problema de sobrecarga de informação foram os sistemas de recuperação e indexação de dados. Nestes sistemas, é preciso que se saiba exatamente qual dado se deseja recuperar. É preciso que se saiba *como* perguntar, pois cada sistema de recuperação de informação possui sua própria sintaxe.

O que permite perguntar: em um cenário onde a oferta de dados cresce de forma exponencial, como se pode obter algo quando não se sabe exatamente o que é que está se buscando?

Foi nesse ambiente de excesso de informação que surgiram os primeiros estudos sobre sistemas de recomendação personalizada.

Cabe dizer que a principal diferença entre uma recomendação pura e simples e uma recomendação personalizada está no fato de que, no primeiro caso, o emissor recomenda algo para vários receptores indiscriminadamente. Não importa se os receptores possuem recursos ou interesses diferentes, a recomendação é geral.

O *marketing* de massa trabalha essencialmente dessa forma. Um anunciante recomenda que todos comprem o seu produto. Quanto mais pessoas receberem a mensagem, maior a chance de sucesso. É o mesmo princípio aplicado aos meios de comunicação de massa, onde o que importa é o número de receptores da mensagem, e não as particularidades deles.

No caso da recomendação personalizada, o diferencial está no fato de que ela será feita somente àqueles que, por alguma razão (que pode advir da aplicação de um

algoritmo ou da execução de um método), tenham interesse na informação. É como se o anunciante, em vez de se preocupar em alcançar o maior número de pessoas, voltasse sua atenção e recursos para o maior número de pessoas com potencial interesse em seu produto.

Segundo Montaner (2003), as pesquisas na área de inteligência artificial têm empregado vários recursos na solução do problema de busca e recomendação da correta informação. Tais recursos oferecem uma extensão aos motores de busca, utilizando como exemplo: motores personalizáveis; sistemas de recomendação; agentes de software; e a mineração de dados na Internet.

1.1 Motivação

A motivação para essa dissertação adveio do fato de que grande parte dos autores que escreveram e estudaram os sistemas de recomendação se basearam em cenários *on-line*.

Tal abordagem é absolutamente compreensível, uma vez que a coleta de informações pela Internet confere diversas vantagens, tais quais:

- Permite alcançar um grande número de pessoas de forma rápida;
- Cada pessoa pode fornecer os dados de entrada na hora e local que lhes forem mais convenientes; e
- Os dados seguem o padrão estabelecido pela interface da ferramenta de coleta, podendo já ser tabulados em sua origem.

No entanto, como o foco deste trabalho são as redes de varejo de moda, observou-se que a maioria da entrada dos dados é realizada por outro meio, uma vez que nem todas as redes possuem um canal de venda baseado na Internet.

O meio mais freqüente de entrada de dados usados nessas redes são os terminais de ponto de venda, ou popularmente conhecido como POS, *point of sale* em inglês. Por meio dos dados coletados nesses terminais, os lojistas são capazes de saber quais produtos estão sendo vendidos e em quais lojas, além de uma série de dados derivados.

Em termos práticos, calcular as recomendações a partir dos dados coletados nos pontos de venda das redes representa uma grande vantagem, uma vez que se trabalha com transações já realizadas.

No mundo real não se mede o fato do consumidor “ter visto” uma peça, nem tampouco o consumidor pode “marcar a peça como favorita” da mesma forma como ocorre na Internet.

Assim, as informações coletadas no mundo real são muito mais fiéis ao real padrão de compra dos clientes, uma vez que não consideram os diversos fatores que representam as intenções de compra.

Esse fato, por si só, é a grande motivação para este trabalho. Aplicar a mecânica dos sistemas de recomendação usando dados reais em um ambiente fora da Internet, não somente é algo pioneiro, como de fundamental importância para que os lojistas conheçam melhor seus padrões de venda e possam programar seus esforços de *marketing* de maneira mais inteligente.

Em vez de criarem promoções dando descontos em toda a sua grade, ou mesmo em parte da grade, os lojistas poderão criar promoções personalizadas, predizendo inclusive, as chances de venda de cada ação para cada consumidor em particular.

1.2 Objetivos

Pode-se afirmar que o principal objetivo deste trabalho é mostrar que o campo para aplicação de sistemas de recomendação personalizada não se restringe a sistemas *on-line*, mas que também a recomendação pode, e deve ser cada vez mais, aplicada a processos de negócio *off-line*.

Outro objetivo é mostrar um possível caminho para a construção de sistemas de recomendação baseados em bases de dados transacionais do varejo de moda *off-line*. Além disso, pretende-se ressaltar os principais problemas que podem surgir, e indicar algumas soluções para o fato de trabalhar somente com dados transacionais.

Finalmente, essa dissertação se propõe a discutir, por meio de um protótipo, a criação e as vantagens dos sistemas de recomendação baseados em dados reais de venda, utilizando por base o procedimento metodológico proposto no capítulo 3.

1.3 Justificativa e relevância do trabalho

Dois fatores foram preponderantes na escolha em se trabalhar com o varejo de moda *off-line*: o primeiro é o fato da maioria dos sistemas de recomendação ser voltada à recomendação de livros, filmes, músicas e outros artigos comuns em lojas virtuais.

O segundo, é que poucos são os sistemas de recomendação voltados para ambientes fora da Internet. O desafio de demonstrar os ganhos diretos e indiretos ao se aplicar os modelos de recomendações a tais sistemas é, no mínimo, excitante.

1.4 Delimitação da dissertação

O texto define o que são sistemas de recomendação, desde sua origem histórica e os seus principais conceitos correlatos por meio de uma revisão da bibliografia. Em

seguida, explica a metodologia de trabalho adotada para a criação do protótipo do sistema a ser aplicado no varejo de moda.

É importante ressaltar que aspectos como escalabilidade, desempenho, bem como outros aspectos da arquitetura de *hardware* e *software* não são vistos como relevantes à pesquisa, uma vez que seu foco é demonstrar as vantagens do uso de sistemas de recomendação fora do ambiente da Internet.

1.5 Organização do texto

O texto foi organizado de maneira relativamente simples. Após esta introdução, o segundo capítulo trata das origens e principais conceitos associados aos sistemas de recomendação. Em seguida apresenta-se a taxonomia adotada nesse trabalho para descrever o sistema de recomendação. E por fim, outros aspectos associados à filtragem colaborativa são vistos em um nível maior de detalhe.

O terceiro capítulo emprega uma abordagem indutiva para descrever a metodologia de pesquisa utilizada, quais foram os passos seguidos e de que forma a taxonomia explicada no capítulo anterior foi aplicada ao protótipo.

Já o quarto capítulo, de caráter eminentemente dedutivo, trata da construção do protótipo, começando por uma explicação sobre o mercado de varejo de moda; passando pela especificação das funções de entrada e saída desenhadas; continuando no comentário da interface construída; e finalizando com os experimentos e resultados obtidos.

Finalmente, o último capítulo encerra esse trabalho revelando as conclusões e resultados obtidos em cada etapa do trabalho, e sugerindo um caminho para a realização de pesquisas no futuro.

1.6 O que foi visto neste capítulo

Neste capítulo foram expostos os aspectos introdutórios deste trabalho, explicando: de que forma a motivação guiou o autor; quais são os objetivos a serem alcançados; qual é a sua justificativa e sua relevância para a comunidade acadêmica; qual é a delimitação do texto; e como ele está organizado.

2. SISTEMAS DE RECOMENDAÇÃO

Tandis que les siècles s'écoulent, la masse des ouvrages s'accroît sans cesse, & l'on prévoit un moment où il serait presque aussi difficile de s'instruire dans une bibliothèque, que dans l'univers, & presque aussi court de chercher une vérité subsistante dans la nature, qu'égarée dans une multitude immense de volumes; il faudrait alors se livrer, par nécessité, à un travail qu'on aurait négligé d'entreprendre, parce qu'on n'en aurait pas senti le besoin. – Dennis Diderot, Encyclopédie (1755)²

2.1 Apresentação

A literatura mostra que os sistemas de recomendação tiveram origem nos sistemas de filtragem que buscavam resolver o problema do excesso de informação presente nos ambientes de mensagens eletrônicas (MALONE et al., 1987). Ainda de acordo com Malone, é preciso estabelecer o conceito de filtragem no âmbito desses sistemas:

Even though the term has a literal connotation of leaving things out, we use it here in a more general sense that includes selecting things from a larger set of possibilities. [...] this positive kind of filtering (i.e. selection) may be much more important in information sharing systems than the negative kind of filtering (i.e. removal).³ (MALONE et al, 1987, p.3)

Assim, o autor contrapõe duas formas clássicas de filtragem: a positiva e a negativa. Na primeira, selecionam-se os itens mais relevantes a partir de regras pré-estabelecidas; enquanto que na segunda os itens menos relevantes são retirados do espaço amostral.

A partir dessa abordagem, são sugeridas três formas de filtragem: a filtragem cognitiva; a filtragem social; e a filtragem econômica.

² Na medida em que os séculos passam, a massa de obras cresce sem cessar e pode-se prever um momento no qual será quase tão difícil se instruir em uma biblioteca quanto no universo; e quase tão curto (sic) buscar uma verdade subsistente na natureza, que afastada em uma plêiade de volumes. Seria necessário então dispor, por necessidade, de um trabalho que teríamos negligenciado empreender, porque não teríamos sentido a necessidade.

³ Ainda que o termo tenha uma conotação literal no sentido de deixar as coisas de fora, nós o usamos aqui em um sentido mais geral que inclui a seleção de coisas a partir de um conjunto maior de possibilidades. [...] esse tipo de filtragem positivo (i.e. seleção), pode ser muito mais importante em sistemas de compartilhamento de informação do que o tipo negativo (i.e. remoção).

A filtragem cognitiva se dá pela correspondência entre características do item pesquisado (por exemplo: palavras-chave, formato, e data de criação), e os parâmetros da pesquisa.

A filtragem social complementa a cognitiva incluindo como parâmetros de pesquisa dados sobre as pessoas e comunidades que trabalharam com o item (seja atribuindo uma nota ao item, ou comentando-o ou enviando-o para outrem, por exemplo).

E, finalmente, a filtragem econômica considera as relações de custo-benefício implícitas ou explícitas como, por exemplo, o tamanho da descrição de um item (que consumiria tempo para ser lido); ou o número de pessoas para os quais um item foi enviado (o que demonstraria pouca preocupação com a personalização, reduzindo em tese o valor do item).

2.1.1 A origem dos sistemas de recomendação: o boca a boca

O problema da sobrecarga de informações não é necessariamente recente, já tendo sido comentado por Diderot em sua Enciclopédia (DIDEROT, 1755).

O fato de haver um grande número de fontes de informação disponível, ou o fato de não se saber avaliar qual a melhor informação para resolver um dado problema, pode fazer com que as pessoas sintam-se desamparadas no momento em que devem tomar uma decisão.

Antes da popularização das tecnologias de informação e comunicação (TICs), recorria-se à opinião de pessoas melhor informadas para cada assunto. Assim, era comum ter uma pessoa que pudesse recomendar um mecânico, outra para recomendar um pediatra, e assim em diante de acordo com a necessidade. Eram os chamados *experts*.

Esse tipo de prática foi identificado por Maes e Shardanand como recomendação boca-a-boca, ou *word-of-mouth* (WoM) do original em inglês (SHARDANAND e MAES, 1995). Ainda que tenha sido a primeira forma de recomendação usada em uma escala maior do que entre duas pessoas, cabe esclarecer que o WoM continua a evoluir apoiado, sobretudo, nas TICs.

Uma definição para o WoM é fornecida pela agência PQ Media, que o define como uma estratégia alternativa de *marketing*, apoiada em pesquisa e tecnologia a fim de incentivar os consumidores a dialogar sobre produtos e serviços. Ele é entregue por meio de diversas táticas tanto *on-line* quanto *off-line*, freqüentemente auxiliadas por parceiros influentes, comunidades e diversos defensores da marca.

Segundo uma pesquisa divulgada em agosto de 2009, ainda pela PQ Media, os gastos com o *marketing* baseado no WoM subiram 14,2% nos Estados Unidos alcançando 1,54 bilhões de dólares em 2008, apesar da maior recessão econômica dos últimos 70 anos. A pesquisa prevê também um ritmo de crescimento na ordem de 10,2% para o ano de 2009, colocando-o como um dos segmentos de propaganda e *marketing* que cresce mais rapidamente. A título de comparação, estima-se que a economia americana, assim como o setor de propaganda e *marketing*, devam se retrair em 2009 pela primeira vez desde a recessão que seguiu a quebra da bolsa de 1929.

Entre o WoM original (sem o apoio das TICs), e a primeira aplicação efetiva de um sistema de recomendação, muito tempo se passou. Até que surgiu o Tapestry como será visto a seguir.

2.1.2 Tapestry: usando filtragem colaborativa

Pode-se dizer que o Tapestry é o exemplo da primeira aplicação de sistema de recomendação baseado em filtragem colaborativa. Ele foi criado para ser um sistema de correio eletrônico no qual as pessoas não precisassem se preocupar em criar filtros,

ou organizar as mensagens em pastas, para poder localizar facilmente aquelas que lhes interessavam (GOLDBERG, NICHOLS, OKI, 1992, RESNICK, VARIAN 1997).

Ele parte do pressuposto que as opiniões e comentários da comunidade de usuários do sistema de correio eletrônico são suficientes para auxiliar a ferramenta na atividade de recomendar as mensagens com maior possibilidade de atender às demandas dos usuários.

Cabe notar que a filtragem colaborativa é uma das formas de gerar recomendações. Este e outros aspectos específicos envolvendo sistemas de recomendações serão tratados adiante.

É válido fazer uma pausa para analisar que recuperar uma informação de uma base de dados, e filtrar um subconjunto de informações a partir dessa mesma base de dados são duas coisas bastante distintas.

2.1.3 Da diferença entre recuperação e filtragem de informações

Os termos recuperação e filtragem de informações podem gerar confusão uma vez que possuem muitas semelhanças (BELVIN e CROFT, 1992)

De início, cabe dizer que a recuperação de informações atende às pessoas que sabem exatamente o que procuram, e é baseada em tecnologias de indexação de conteúdo tais como construção de taxonomias, *thesaurus* e ontologias. Seu principal objetivo pode ser transcrito segundo Robertson (apud Belvin, 1992) como aquele que tem “a funcionalidade de conduzir o usuário para aqueles documentos que melhor irão habilitá-lo a satisfazer as suas necessidades de informação”.

As ferramentas de busca disponíveis na Internet constituem um dos melhores exemplos de sistemas de recuperação de informação. Tudo o que o usuário precisa é

conhecer algumas palavras que descrevam o que está buscando, e informá-las ao sistema.

Já um sistema de filtragem de informação remove os dados redundantes ou indesejáveis de um conjunto de dados antes de entregá-los ao usuário. E o faz comparando o perfil do usuário a algumas características de referência, que podem ser atributos do item (abordagem baseada em conteúdo), ou dados dos usuários (filtragem colaborativa).

Os sistemas de recomendação são sistemas de filtragem que buscam oferecer itens que possam interessar aos usuários. Eles geralmente são baseados na filtragem colaborativa, ou na combinação desta com a filtragem baseada em conteúdo.

A figura abaixo recupera o que foi visto até o momento sobre as origens dos sistemas de recomendação, e estabelece a base para seguir adiante.

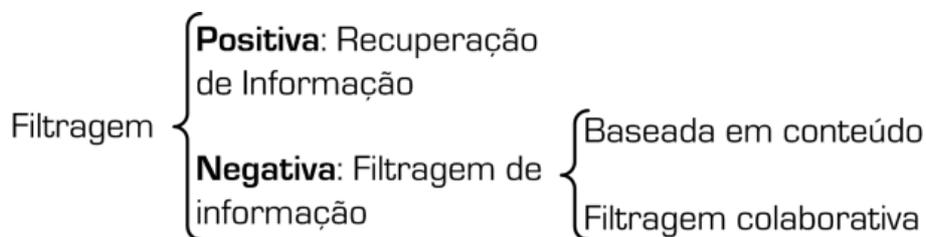


Figura 1 - Origem dos conceitos de filtragem.
Fonte: elaborado pelo autor.

2.2 Taxonomia de sistemas de recomendação

Entende-se que os sistemas de recomendação podem ser divididos em três entidades principais (TORRES, 2004a). A figura a seguir ilustra estas entidades (itens disponíveis, sistema de recomendação e usuário), e como elas estão relacionadas.

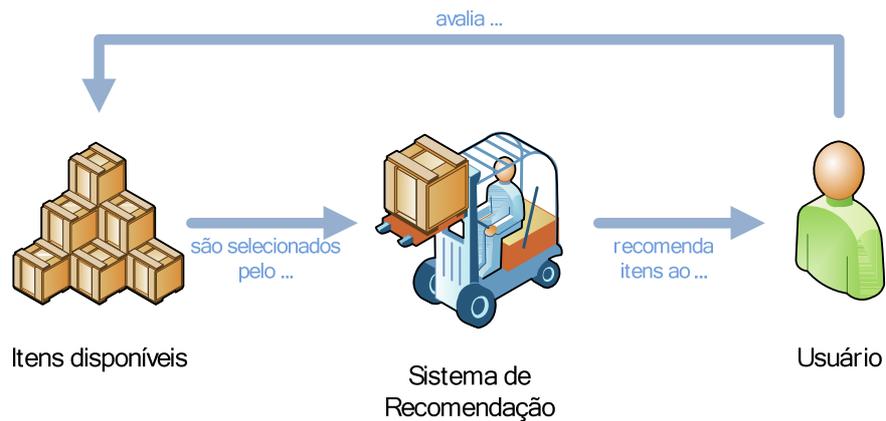


Figura 2 - Principais entidades dos sistemas de recomendação.
Fonte: elaborado pelo autor.

Uma descrição sucinta destas entidades pode ser dada por:

- Os **itens disponíveis** são aqueles sobre os quais as avaliações são feitas pelos usuários, e dentre os quais o sistema de recomendação seleciona os que irá recomendar aos usuários;
- O **sistema de recomendação** *per se*, é o sistema que serve de elo entre os usuários e os itens; e
- O **usuário** é aquele quem avalia os itens e recebe as recomendações.

O trabalho de Manouselis e Costopoulou (MANOUSELIS e COSTOPOULU, 2007), unificou diversas taxonomias já disponíveis e as agrupou nas seguintes categorias: justificativa, abordagem e operação. Essa parte do texto trata da explicação da taxonomia proposta pelos autores supracitados e adotada por tratar-se de uma revisão mais abrangente das taxonomias propostas em trabalhos anteriores.

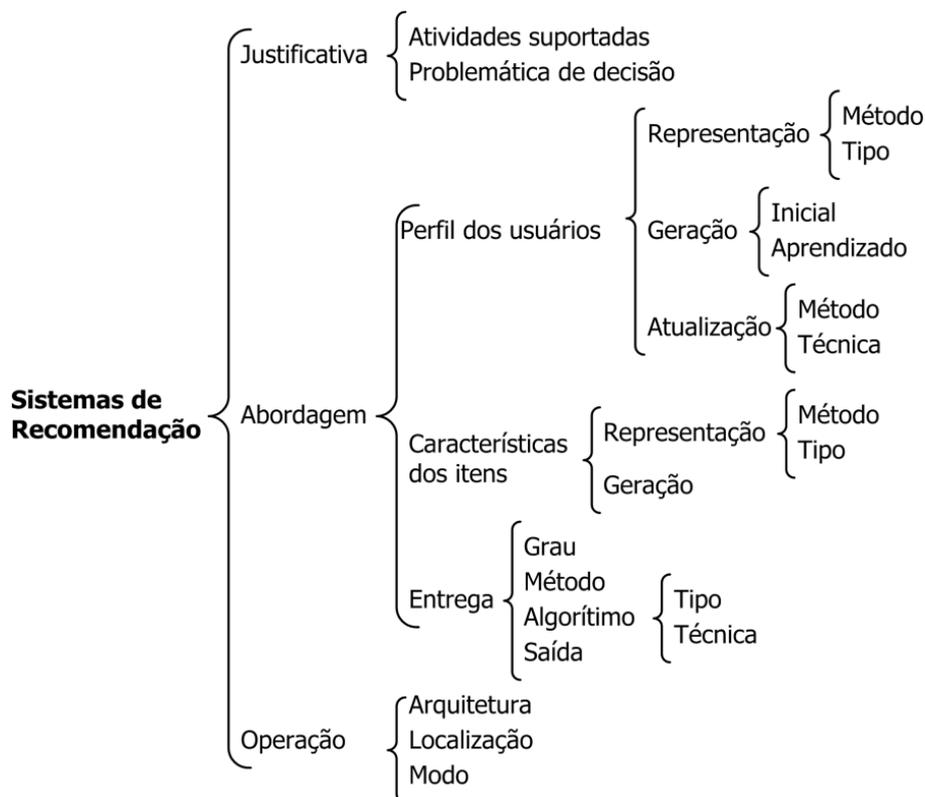


Figura 3 - Taxonomia dos sistemas de recomendação.
 Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

A) Justificativa

No que concerne a justificativa para o sistema de recomendação, o modelo de taxonomia identifica as categorias das atividades suportadas e da problemática de decisão. A figura a seguir mostra a estrutura hierárquica da taxonomia proposta:

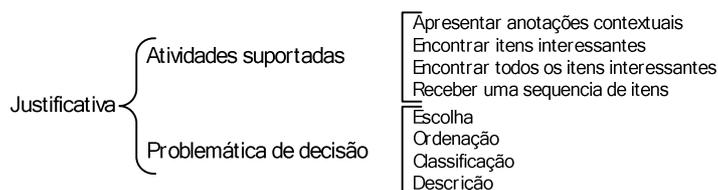


Figura 4 – Classificação da justificativa do sistema.
 Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

- i) As **atividades suportadas** refletem os tipos de resultados que as pessoas podem obter dos sistemas de recomendação. Como principais exemplos têm-se, o que a taxonomia adotada identifica como:

a) Anotações contextuais

Quando o sistema de recomendação é uma funcionalidade adicional que faz parte de outro sistema e possui como objetivo ajudar ou informar melhor ao usuário final sobre um determinado item. Um exemplo dessa aplicação pode ser encontrado nos quadros de comentários disponíveis para os usuários se expressarem sobre os itens em um sítio web.

A figura 4 ilustra de que forma uma anotação contextual pode ser realizada em um sítio de comércio eletrônico. No exemplo a seguir, o usuário tem acesso às opiniões já registradas por outros consumidores e é convidado a emitir uma opinião sobre determinado produto do catálogo (item).



Figura 5 - Consulta às anotações do livro “O Pequeno Príncipe” no sítio Submarino.

Fonte: sítio <http://www.submarino.com.br>, acessado em junho de 2009.

b) Encontrar itens interessantes

Nesse exemplo o sistema sugere itens específicos ao usuário. É o caso clássico de sistemas de recomendação, onde para um determinado item são dadas sugestões de outros itens que também possam agradar ao usuário final. Um exemplo são os quadros “veja também” disponíveis em sítios de comércio eletrônico.

O exemplo a seguir (figura 6) ilustra três sugestões de produtos que combinam com a camisa que está sendo visualizada pelo comprador no sítio de comércio eletrônico da Richards.

The screenshot shows the Richards website interface. At the top, there is a navigation bar with links for 'CONHEÇA A RICHARDS', 'COLEÇÃO', 'LOJAS', 'COMPRAS ON LINE', and 'SACOLA' (with a shopping cart icon and 'R\$ 0,00'). The main content area is divided into three columns. The left column is a navigation menu with categories like 'MASCULINO', 'FEMININO', and 'OFF MASCULINO'. The middle column features a large image of a 'CAMISA BUSINESS' (code 54303) with options for color (VERDE) and size (01, 02, 03, 04). Below the main image are smaller thumbnails and buttons for 'AMPLIAR' and 'VOLTAR A LISTAGEM'. The right column, titled 'SUGESTÕES', displays three recommended items: 'CALÇA MILÃO' (R\$ 240,00), 'BOTA RYAN' (R\$ 289,00), and 'JAQUETA MÔNACO' (R\$ 490,00). A red arrow points to the 'SUGESTÕES' header.

Figura 6 - Sugestões do tipo "veja também" no sítio da Richards.
Fonte: sítio <http://www.richards.com.br>, acessado em junho de 2009.

Na literatura, essa aplicação é conhecida como algoritmos de recomendação *top-n*⁴ posto que, dentro de um universo de diversos itens, o sistema seleciona os mais relevantes e os oferece ao usuário.

c) Encontrar todos os itens interessantes

Em alguns casos não convém selecionar somente os primeiros n itens mais próximos ao item pesquisado. Isso é comum com bases de dados médicos ou legais, em que todas as possibilidades conhecidas devem ser oferecidas ao usuário final, juntamente com o coeficiente de similaridade, para que seja dele a decisão sobre quais itens são relevantes para o caso dele.

d) Receber uma seqüência de itens

Esse tipo de recomendação é bastante encontrado em situações onde o usuário precisa receber os itens sugeridos de forma seqüencial, como ocorre, por exemplo, com serviços que oferecem conteúdo audiovisual.

Um exemplo bastante comum pode ser encontrado no sítio de vídeos Youtube (ver figura 7 abaixo), que exibe os vídeos relacionados de forma seqüencial (dois a dois em um total de seis vídeos) logo após a exibição do vídeo corrente.

⁴ Recomendação *top-n* é um termo comum na área de sistemas para indicar que os n primeiros registros serão trabalhados ou retornados pela função.



Figura 7 - Interface do sítio Youtube que oferece uma seqüência de itens.
Fonte: sítio <http://www.youtube.com>, acessado em junho de 2009.

- ii) No que diz respeito à recomendação, o objeto de decisão é um item i que pertence a um conjunto de todos os itens prováveis I .

A fim de expressar a justificativa por trás da decisão Roy faz referência ao conceito de **problemática da decisão** (ROY, 1996). Os quatro tipos mais comuns e que podem ser considerados válidos no contexto de recomendação, são:

a) Escolha

Quando o usuário do sistema tiver que escolher uma das opções fornecidas pelo sistema de recomendação para que o seu trabalho possa continuar a ser realizado.

Um exemplo desse tipo de decisão pode ser encontrado nos produtos da empresa italiana BeeTV (o sítio está disponível no endereço <http://www.bee.tv>). O que eles oferecem é um sistema para recomendar programas de TV em função das informações fornecidas pelo usuário. Para que ele possa assistir a um programa, ele deverá

selecionar um dos itens que foram oferecidos em uma tela, como por exemplo, a da figura abaixo:

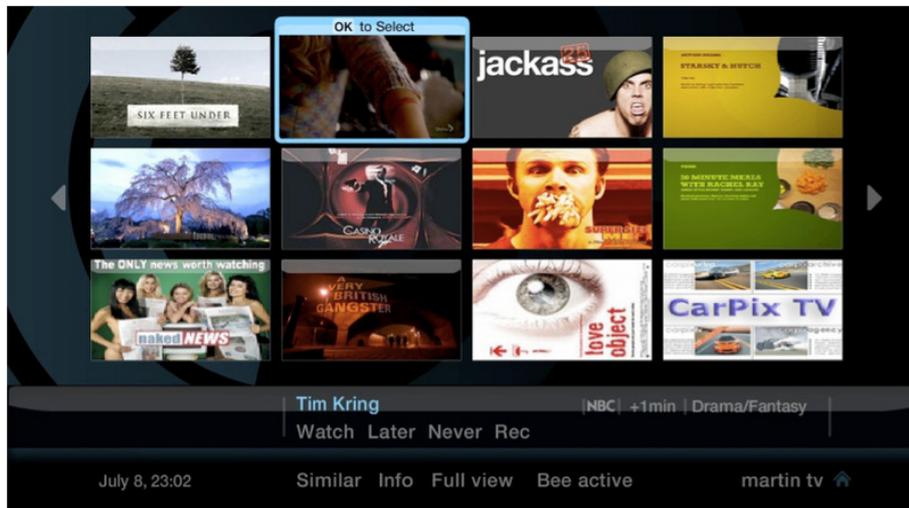


Figura 8 - Exemplo de tela da solução BeeTV.

Fonte: sítio <http://www.wired.com/gadgetlab/2008/09/beetv-gets-in-y/>, acessado em junho 2009.

b) Ordenação

É caso onde o sistema de recomendação ordena os resultados seguindo critérios pré-estabelecidos (por exemplo: data de validade, grau de relevância, confiabilidade do dado, ou uma categorização já definida pelo sistema).

O usuário pode ser compelido a decidir pelo item em uma lista ordenada, por exemplo, quando ele obtém a lista dos produtos mais vendidos em um sítio de comércio eletrônico, como ilustrado na figura a seguir:

Vitrine **Mais vendidos** Menor Preço

- 

Cj. Panelas Inox c/ Tampa de Vidro 07 pcs
 De: R\$ 199,00
 Por: R\$ 139,00
 ou 4 x de 34,75

FRETE GRÁTIS
- 

Cj Panelas Inox c/Tampa de Vidro 05 Pcs - BK0495
 De: R\$ 149,00
 Por: R\$ 99,00
 ou 3 x de 33,00

FRETE GRÁTIS
- 

Conjunto de Panelas Inox com Tampa de Vidro Inox 5 Pcs
 De: R\$ 159,00
 Por: R\$ 109,00
 ou 3 x de 36,33

FRETE GRÁTIS

Figura 9 - Exemplo de itens oferecidos por ordenação.
 Fonte: sítio <http://www.shoptime.com.br>, acessado em junho de 2009.

c) Classificação

Aqui os itens são apresentados do mais relevante ao menos relevante. Embora muito parecido com a ordenação supracitada, a principal diferença é que no caso do *ranking* o resultado é classificado de acordo com a sua relevância para o item em questão e não por outro critério estipulado.

A relevância pode ser obtida de várias formas, sendo que normalmente cada serviço, determina a forma que utilizará. Neste trabalho, usou-se o coeficiente de Pearson, como será visto mais adiante. No caso da Amazon, por exemplo, a fórmula da relevância constitui um segredo de negócio.

A Amazon é exemplo de um sítio de comércio que emprega de forma eficaz a classificação dos resultados ao oferecer itens para seus usuários. A figura 10, por exemplo, permite visualizar quais foram os livros que as pessoas compraram após comprar determinado livro. Com

isso, estes livros são recomendados de acordo com a sua relevância em relação ao livro que está sendo consultado.

Customers Who Bought This Item Also Bought



Figura 10 - Exemplo de como a Amazon oferece produtos.
Fonte: sítio <http://www.amazon.com>, acessado em junho de 2009.

d) Descrição

Nesse caso os itens são descritos em função de suas características como, por exemplo: posição no *ranking*, categorias associadas, data em que foram avaliados, ou número de avaliações.

Novamente a Amazon faz uso de maneira exemplar dessa abordagem, fornecendo um grande número de informações ao comprador, de forma que ele sinta-se confortável ao comprar um item em seu sítio.

B) Abordagem

Em termos de abordagem, os sistemas de recomendação podem ser divididos em três categorias que serão examinadas nesta seção. A figura abaixo ilustra os critérios para classificação segundo a abordagem:

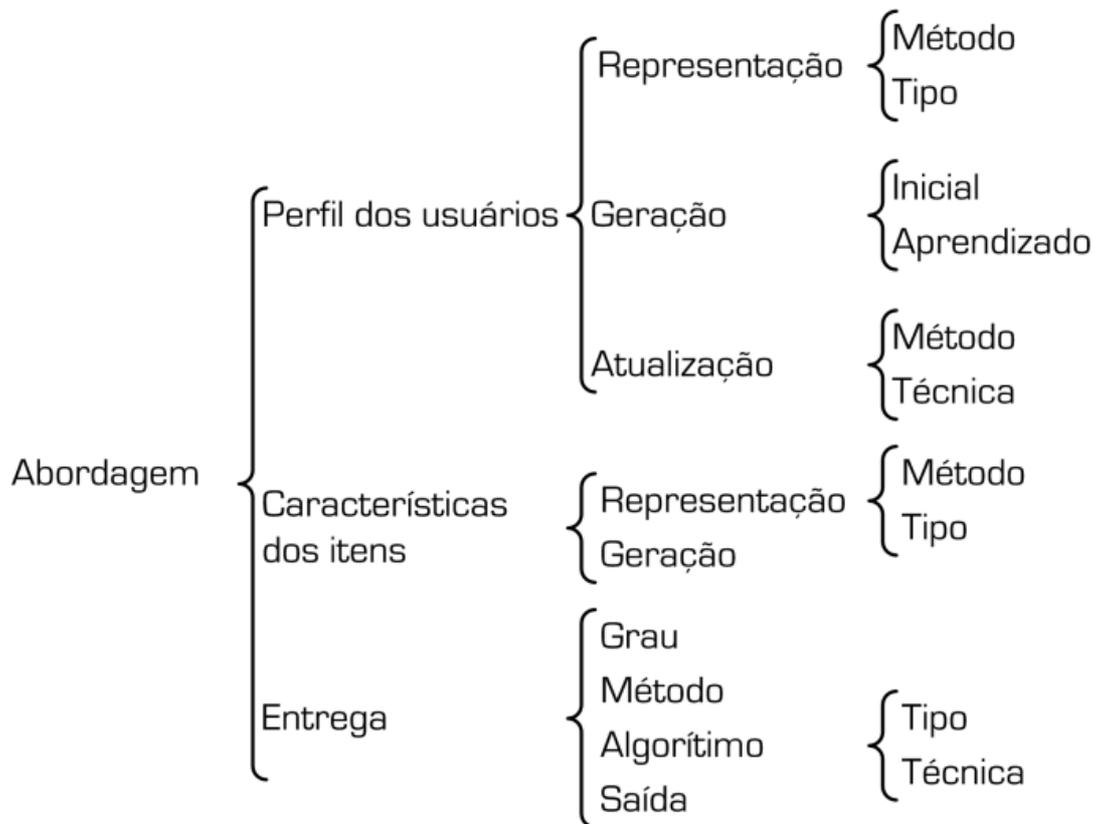


Figura 11 – Classificação da abordagem do sistema.
 Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

i) Perfil do usuário

Este primeiro critério se refere às características gerais que o usuário do sistema possui.

Por exemplo, ao preencher um cadastro, ele (o usuário) pode ser perguntado sobre sua faixa etária, ou sobre o seu gosto musical. Essas informações farão parte do seu perfil e poderão ser levadas em consideração no processo de geração das recomendações.

É importante notar que dependendo da função projetada para o sistema, o perfil do usuário poderá ser inexistente ou não ser levado em consideração, pelo mecanismo de recomendação. Entende-se que o fato de haver o perfil do usuário, pretende-se alcançar um objetivo determinado, seja ou não, recomendar alguma coisa.

A figura a seguir ilustra quais são as classificações relativas ao perfil dos usuários.



Figura 12 – Classificação do perfil do usuário.

Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

Neste trabalho, onde somente as características das peças vendidas serão consideradas, os dados do perfil do usuário são irrelevantes. No entanto, é importante descrever a forma como a taxonomia adotada trabalha tais dados.

a) Representação

Reflete a forma como o perfil do usuário pode ser representado considerando o método e o tipo de representação que se deseja realizar.

(i) Para abordar a representação do perfil do usuário, identificam-se os seguintes **métodos**:

- 1) características demográficas;
- 2) matrizes de avaliações dos usuários;
- 3) modelos baseados em classificadores;
- 4) modelos baseados em valores;
- 5) modelos baseados em vetores;
- 6) modelos de desagregação;
- 7) modelos de otimização;
- 8) modelos históricos;
- 9) redes associativas;
- 10) redes semânticas; e
- 11) relações de exclusão (ou *outranking*).

(ii) Já no tocante ao tipo de representação do perfil do usuário, Manouselis e Costopoulou (2007) sugerem a adoção de um dos seguintes **tipos**:

- 1) baseadas em *funcionalidades*;
- 2) lógica difusa (*fuzzy*);
- 3) mensuráveis; e
- 4) probabilísticos.

b) Geração

Dependendo dos objetivos, o perfil do usuário precisará ser alimentado com seus dados. Esse processo pode ser dividido em duas fases:

(i) Uma **fase inicial** onde a criação do perfil do usuário é realizada por meio do preenchimento de seus dados. Identificam-se as quatro formas seguintes:

- 1) Os perfis começam vazios e vão sendo gradativamente preenchidos pelos usuários na medida em que utilizam o sistema.
 - 2) Os perfis são integralmente fornecidos pelo usuário logo ao se cadastrar no sistema.
 - 3) A base de perfis é preenchida de acordo com algum estereótipo com o qual o usuário se identifica. Esse procedimento é bastante comum em serviços de recomendação de música onde o usuário classifica os gêneros musicais que mais gosta e se identifica como um fã do gênero *a* ou *b*.
 - 4) Os perfis são preenchidos por meio das respostas dadas pelo usuário a um conjunto pré-estabelecido de perguntas formuladas para montar o seu perfil. É uma forma mais sutil e menos invasiva, no entanto também é menos precisa.
- (ii) A geração do perfil do usuário também pode ser realizada pela **aprendizagem**, que pode ser feita a partir de diversas técnicas usando os dados que já foram coletados. Dentre as técnicas, destacam-se as enumeradas a seguir:
- 1) aprendizagem de máquina (*machine learning*);
 - 2) segmentação;
 - 3) técnicas de classificação (árvores de decisão, baseadas na resolução de casos (*case-based reasoning*), redes neurais, e redes Bayesianas); e
 - 4) técnicas de redução de dimensão.

c) Atualização

Uma vez a base de dados tenha sido alimentada com o perfil do usuário, é feita a escolha quanto ao método e a técnica para manter esses dados em dia.

(i) O **método** de atualização pode ser classificado como:

1) Implícito

As informações são coletadas sem que o usuário precise interagir com o sistema, por exemplo, por meio da análise dos registros de navegação, ou do histórico de compras, dentre outras formas.

A figura abaixo ilustra a forma adotada pelo o sítio da Amazon para fornecer ao usuário o resultado dos dados coletados de forma implícita.

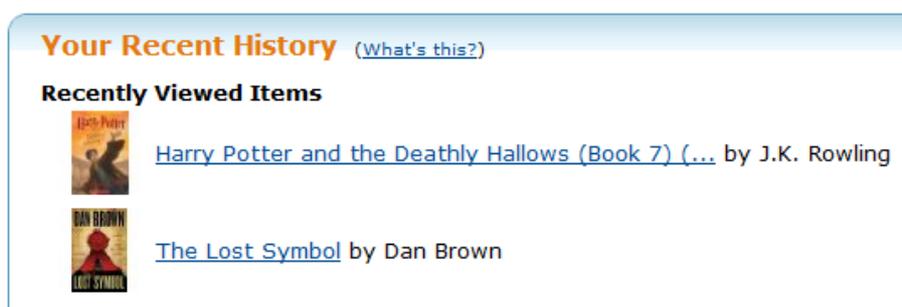


Figura 13 - Exemplo de dados coletados de forma implícita.
Fonte: sítio <http://www.amazon.com>, acessado em junho de 2009.

2) Explícito

Já no modo explícito o usuário indica formalmente os itens que gosta ou não, atualizando a base de dados sempre que achar necessário. Ele pode fazê-lo respondendo a um

questionário ou declarando as suas preferências em um formulário, mas sempre de forma explícita.

3) Híbrido

Finalmente no modelo híbrido ocorre a combinação dos dois métodos supracitados mediante as necessidades específicas de cada caso.

(ii) No tocante à **técnica** para manter os dados do perfil do usuário atualizados, o modelo propõe cinco opções:

- 1) Na primeira o usuário atualiza os seus dados manualmente, ou seja, parte dele a iniciativa de informar o que mudou em seu perfil, mantendo-o atualizado.
- 2) O sistema também pode acrescentar novos dados ao perfil do usuário a partir das interações que ele teve com o sistema. Por exemplo, pode-se coletar o dia da semana ou o horário em que o usuário usa o sistema com mais frequência e armazená-los de modo automático.
- 3) Outra técnica emprega o descarte gradual de informações antigas. No passado usuário poderá ter dito que gostava bastante de determinada empresa ou serviço que não exista mais, devendo ser descartada.
- 4) Técnicas de seleção natural, onde um curador determina quais dados devam permanecer e quais devam ser descartados também podem ser usadas.

- 5) Por fim, o estabelecimento de um período de tempo no qual alguns dados serão levados em consideração é uma das técnicas para manter o perfil atualizado.

ii) Características dos itens

De forma análoga ao perfil do usuário, as características dos itens representam os atributos dos itens pertencentes ao sistema de recomendação. A figura abaixo ilustra a forma como essas características são classificadas de acordo com a taxonomia adotada.

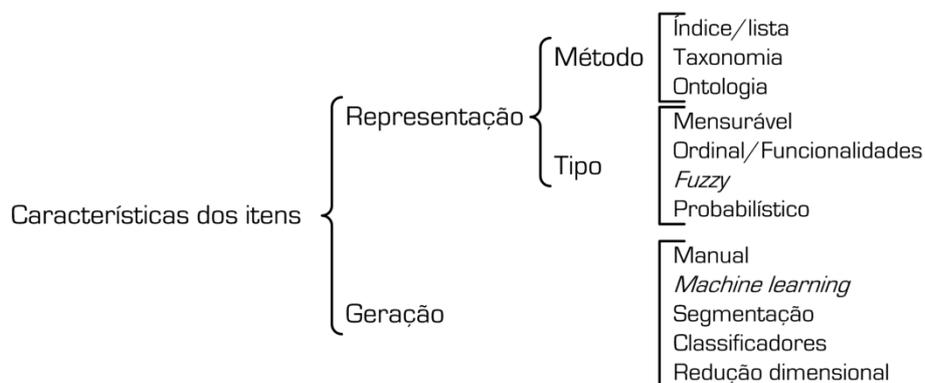


Figura 14 – Classificação das características dos itens.

Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

a) Representação

A representação indica de que forma essas características serão apresentadas ao usuário do sistema. Para efeitos de classificação, subdivide-se a representação em um método e um tipo. O método descreve como as características dos itens serão apresentadas, enquanto que o tipo determina:

- (i) Os itens que fazem parte do sistema de recomendação costumam ser apresentados de três formas (métodos) aos usuários.

- 1) A primeira delas é sob a forma de uma **lista de itens**, onde os itens são apresentados sem uma hierarquia ou classificação.
- 2) A segunda forma de apresentar os itens é por meio de uma **taxonomia**, na qual os itens são distribuídos hierarquicamente segundo conceitos pré-estabelecidos. Este método de representação é bastante comum em lojas virtuais que costumam separar seus artigos por categorias.



Figura 15 - Exemplo de apresentação de itens por meio de taxonomia.
 Fonte: sítio <http://www.pontofrio.com.br>, acessado em junho de 2009.

- 3) Por último têm-se as **ontologias** nas quais os itens recebem atributos, também pré-determinados, que servirão a descrevê-los e distingui-los dos demais itens que compõem o acervo. Tom Gruber (2008) define ontologia como sendo “*a set of representational primitives with which to model a domain of knowledge or discourse*”⁵.
- (ii) Já no tocante ao **tipo** de representação a literatura sugere a adoção dos seguintes tipos:
- 1) Baseados em métricas (i.e. **mensuráveis**), como a avaliação (*ratings*) por parte dos usuários e de curadores.

⁵ Um conjunto de primitivas representativas com as quais se modela um domínio de conhecimento ou discurso.

Neste caso, os usuários avaliam os itens atribuindo-lhes uma nota, e esses itens são apresentados de acordo com as notas que foram atribuídas.

Um exemplo de utilização deste tipo de representação pode ser observado na figura a seguir. Nela observa-se que foram selecionadas as cinco notícias melhor avaliadas.



Figura 16 - Exemplo de apresentação de itens usando avaliações.
Fonte: sítio <http://www.oglobo.com.br>, acessado em julho de 2009.

- 2) Baseados em características e/ou atributos (i.e. **ordinal ou funcional**), pré-definidos para representar o item em questão como, por exemplo, a data em que foi incorporado ao acervo, ou a classe a qual pertence.

A figura abaixo é um exemplo onde os editores (i.e., os curadores) selecionaram 50 discos com preço de até US\$ 5,00. Nela, pode-se concluir que a seleção foi feita com base no preço e no fato do disco ter sido marcado pelo editor.

50/\$5: 50 Albums for \$5 Each

This month, our editors picked 50 great albums to share with you for just \$5 each. Check back next month for a new set of featured \$5 albums.

50/\$5: 50 Albums for \$5 Each for July 2009



[Night In The Ruts](#)
by Aerosmith
Price: \$5.00



[Invaders Must Die](#)
by The Prodigy
Price: \$5.00



[Conversation With A Devil \[Explicit\]](#)
by Andre Nickatina
aka Dre Dog
Price: \$5.00



[Ray Sings, Basie Swings](#)
by Ray Charles
Price: \$5.00



[Innovative Life](#)
by Arabian Prince
Price: \$5.00

Figura 17 - Exemplo de itens selecionados com base em atributos.
Fonte: sítio <http://www.amazon.com>, acessado em junho de 2009.

- 3) Baseados em lógica difusa (*fuzzy*) que atende à necessidade de representar os itens com atributos não quantificáveis, indicando, por exemplo, se um vídeo é bom, mais ou menos, ou ruim.
- 4) E por fim, a representação baseada em **probabilidades**, utiliza modelos probabilísticos para representar os itens do acervo.

b) Geração

- (i) A criação do acervo de itens é freqüentemente feita de modo manual, onde os itens devem ser incluídos um a um na base de dados.

Entende-se que caso os itens sejam migrados de uma base de dados pré-existente para a base de dados do sistema de recomendação, esse processo poderá ser automatizado. No entanto, ainda assim, será necessária a inclusão dos itens na base de dados de origem.

Existem outras formas descritas na taxonomia: aprendizagem de máquina (*machine learning*), segmentação, classificadores e

redução dimensional. No entanto, elas não serão tratadas neste trabalho haja vista o fato de que o acervo de peças do varejo ser essencialmente fornecido pelo lojista.

O sistema não tem, em nenhum cenário, autonomia para criar ou aprender quais são as peças, sem que elas tenham sido manualmente informadas pelo lojista.

iii) Entrega

Esta categoria trata da forma como o sistema apresenta as recomendações aos usuários. A entrega pode ser classificada quanto ao grau, ao método, algoritmo empregado, e formato de saída.

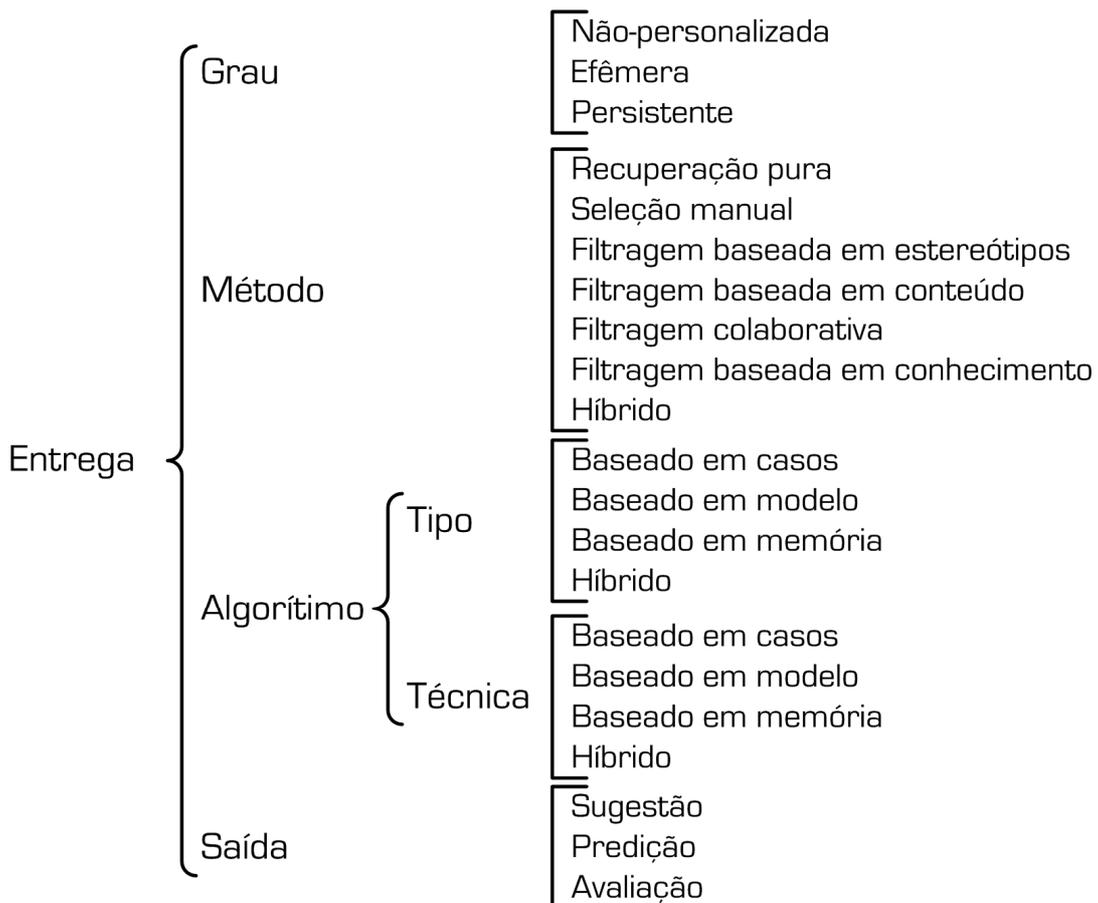


Figura 18 – Classificação da entrega.

Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

a) Grau

- (i) A recomendação pode ser não-personalizada quando os mesmos itens forem recomendados para todos os usuários sem distinção. Isso pode acontecer, por exemplo, quando o curador desejar divulgar determinado item independentemente do interesse que o usuário tenha sobre ele.

A figura 17 (p.31) mostrou os discos que foram selecionados pelo curador. Nesse exemplo, a entrega foi feita sem considerar o interesse do usuário nos itens. Logo, ela não foi personalizada.

- (ii) A recomendação também pode ser considerada como efêmera quando realizada tomando por base somente os dados de um curto espaço de tempo. Normalmente esse espaço é o tempo de uma sessão de interação.

A figura abaixo exemplifica uma recomendação efêmera, feita quando o usuário não havia se identificado e válida somente no contexto da sessão em questão.

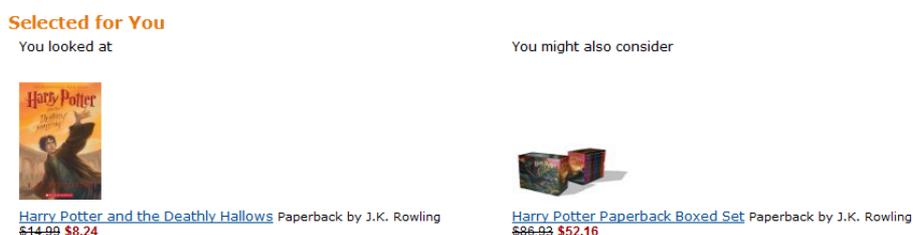


Figura 19 - Exemplo de recomendação efêmera.
Fonte: sítio www.amazon.com, acessado em junho de 2009.

- (iii) Já as recomendações persistentes, consideram praticamente todas as interações do usuário na hora de recomendar um novo

item. Em função disso, essas recomendações tendem a ter mais qualidade, pois compreendem um número maior de ações.

A Amazon é a empresa que fornece bons exemplos de recomendações persistentes. Ao acessar seu sítio de forma identificada, ou seja, após ter realizado o *login*, são oferecidas recomendações que consideram todo o histórico do usuário.

Isso permite que o usuário receba uma lista rica de produtos, mas também pode mascarar outros problemas, que serão vistos adiante no item 2.3.

b) Método

- (i) O primeiro e um dos mais simples métodos para gerar recomendações é chamado de **recuperação pura** de dados (*raw retrieval*). Ele representa os resultados sem nenhuma forma de personalização envolvida, ou seja, sem considerar os dados do perfil do usuário na hora de gerar o resultado.

Tal método é bastante comum quando se deseja recomendar, por exemplo, os 10 itens mais vendidos, mais comentados, ou mais vistos pelos usuários.

A utilização de recuperação pura e simples pode ser bastante útil para resolver o problema de itens novos, que ainda não tenham sido avaliados pelos usuários e que conseqüentemente não tiveram nenhuma correlação realizada.

- (ii) Outra forma de gerar recomendação é permitindo que um curador realize a **seleção manual** dos itens que devem ser recomendados.

Essa forma é bastante comum quando o objetivo é fazer com que o usuário conheça os itens que foram selecionados por um *expert*. Nessa situação, a reputação do curador é associado ao item, o que por sua vez justifica a sua seleção.

- (iii) Existe também uma forma de gerar recomendações tomando por base **dados demográficos** (como, por exemplo, a faixa etária, o sexo, ou domicílio) ou então estereótipos dos usuários. Deste modo o sistema pode, por exemplo, oferecer um mesmo conjunto de itens a usuários que compartilhem a mesma faixa etária ou qualquer combinação de dados demográficos.
- (iv) Há também a **filtragem baseada em conteúdo**, onde o sistema utiliza características dos itens e as características dos perfis dos usuários e faz a associação entre eles gerando as recomendações.

Um exemplo desse método é associar músicas sertanejas aos usuários que tenham declarado, em seus perfis, gostar deste estilo musical.

É importante destacar que a principal característica desse tipo de filtragem é considerar somente a relação do usuário com os itens do acervo, descartando as opiniões dos outros usuários.

- (v) Já a **filtragem colaborativa** é diferente da anterior por levar em consideração o que outras pessoas, com gostos semelhantes ou não, também gostaram no passado.

Assim, ela considera o histórico de interações dos outros usuários a fim de localizar pessoas com gostos similares e melhorar a qualidade dos itens que serão recomendados ao usuário atual.

Motta (1999) também identifica a *filtragem baseada em utilidade* na qual o usuário recebe as recomendações em função da utilidade.

Na filtragem baseada **em conhecimento** o processo sugere itens por meio de inferências lógicas sobre as preferências dos usuários. Uma representação do conhecimento, como por exemplo uma regras, a respeito de como um item atende a uma necessidade de um usuário se faz necessária.

- (vi) Por fim existem as **abordagens híbridas**, que Renato Torres (TORRES, 2004b) descreve como aquelas que “[...] se aproveitam das vantagens da filtragem colaborativa e da filtragem baseada em conteúdo, de forma a neutralizar suas desvantagens”.

c) Algoritmo

(i) Tipo

Encontra-se na literatura cinco tipos de algoritmo para os sistemas de recomendação.

- 1) Ele pode ser **baseado em memória**, quando todos os dados disponíveis (sobretudo dados históricos) das interações do usuário com o sistema são considerados para a geração das recomendações. Assim, os dados do perfil do usuário, as avaliações que ele fez para os itens, os itens que ele disse gostar, enfim tudo o que se sabe a respeito do usuário e da relação dele com os itens da base, será considerado para que o sistema crie uma recomendação.

- 2) Já nos algoritmos **baseados em modelos**, o sistema estabelece um modelo probabilístico a partir das avaliações feitas pelos usuários. Essa é uma abordagem probabilística e é normalmente usado quando se deseja saber o quanto um usuário gostará de determinado item em função das avaliações que ele deu para outros itens.
- 3) Normalmente usam-se algoritmos de aprendizagem automática (*machine learning algorithms*) como: redes Bayesianas, que trabalham com a criação de modelos probabilísticos; segmentação, que agrupam usuários com gostos semelhantes e trabalham o problema de filtragem sob a ótica de classificação; e as abordagens baseadas em regras, que aplica um conjunto de regras para descobrir associações entre os itens.
- 4) Há também os algoritmos baseados em **modelos heurísticos** onde as correlações entre os itens que os usuários gostaram têm por base o fato de que as pessoas que gostaram dos mesmos itens no passado, tenderão a concordar no futuro.
- 5) Nos algoritmos **baseados em instâncias**, as correlações já foram estabelecidas de antemão e o sistema verifica se a opção feita pelo usuário foi a mesma que a instância sugeriu. Normalmente, os cenários de instância são gerados por especialistas, mas também podem ser alimentados pelas avaliações dos usuários.
- 6) E finalmente, os **algoritmos híbridos** que trabalham a combinação dos tipos já descritos adaptando-os de acordo

com a necessidade e características do problema a ser resolvido.

(ii) Técnica

Cabe citar que no tocante à técnica, os algoritmos podem ser classificados em:

- 1) **Baseados em atributos** – quando usam os atributos dos itens.
- 2) **Correlações item-a-item** – quando usam as avaliações que os usuários fizeram para os itens.
- 3) **Correlações usuário-a-usuário** – quando comparam usuários em busca de padrões semelhantes de avaliação.
- 4) **Híbridos** – quando usam uma combinação de duas ou três das classificações acima.

d) Saída

Existem diversas formas de entregar as recomendações aos usuários.

(i) Sugestões

A forma mais comum de entrega se dá por meio de sugestões de itens diferentes ao usuário em função do item que ele estiver examinando.

Schafer et al. (2001) descreve algumas formas clássicas de como os sítios de comércio eletrônico, por exemplo, oferecem suas recomendações aos usuários.

1) Listas de recomendação

É a forma mais simples de oferecer itens recomendados aos usuários. Consiste essencialmente de uma lista de artigos que atendam a critérios como “produtos mais vendidos”, ou “temas para o Natal”, por exemplo.

A figura a seguir é um exemplo de uma lista de recomendação:



Figura 20 - Exemplo de lista de recomendação.
Fonte: sítio <http://www.amazon.com>, acessado em julho de 2009.

É importante citar que neste caso, pode ou não haver personalização, de acordo com a forma de entrega adotada (ver p.32).

No exemplo da figura acima, pode-se inferir que os itens foram selecionados por um curador com base em critérios que atenderam às diretrizes da empresa. Assim, o curador selecionou os sete livros mais significativos de julho.

No entanto, também poderiam ter sido usados outros critérios tais quais: livros mais vendidos, livros mais comentados, por exemplo.

2) Recomendações para você

O sistema utiliza os dados do perfil do usuário para gerar as recomendações personalizadas.

A figura abaixo é um exemplo dessa utilização:

These recommendations are based on [items you own](#) and more.

view: [All](#) | [New Releases](#) | [Coming Soon](#) More results

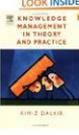
- 1.**  **Knowledge Management in Theory and Practice**
by Kimiz Dalkir (Jun 16, 2005)
Average Customer Review: ★★★★★ (4)
In Stock
List Price: \$67.96
Price: \$46.36
51 used & new from \$30.99
[Add to Cart](#) [Add to Wish List](#)
[I own it](#) [Not interested](#)
Recommended because you purchased **Knowledge Management Toolkit, The** and more ([Fix this](#))
- 2.**  **The Knowledge-Creating Company (Harvard Business Review Classics)**
by Ikujiro Nonaka (Dec 8, 2008)
Average Customer Review: ★★★★★ (15)
In Stock
Price: \$8.95
37 used & new from \$4.53
[Add to Cart](#) [Add to Wish List](#)
[I own it](#) [Not interested](#)
Recommended because you purchased **Working Knowledge** and more ([Fix this](#))
- 3.**  **Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation**
by Georg von Krogh (Jun 1, 2000)
Average Customer Review: ★★★★★ (9)
In Stock
List Price: \$46.00
Price: \$32.40
48 used & new from \$12.90
[Add to Cart](#) [Add to Wish List](#)
[I own it](#) [Not interested](#)
Recommended because you purchased **Working Knowledge** and more ([Fix this](#))

Figura 21 - Tela de recomendações para o usuário no sítio da Amazon.
Fonte: sítio <http://www.amazon.com>, acessado em junho de 2009.

Nela observa-se claramente que, em função de compras que foram realizadas anteriormente, o sistema recomendou novos livros que o usuário ainda não possui.

Cabe destacar a presença de algumas funções auxiliares oferecidas pela Amazon, como:

- *fix this* (conserte isso) que permite ao usuário dizer se o item que foi tomado por base para gerar a recomendação foi comprado para ser dado como presente, ou que ele não deve ser levado em conta na geração de recomendações;
- *I own it* (eu possuo isso) para que ele indique que já possui o item em questão e que ele não deverá ser recomendado numa próxima vez; ou
- *not interested* (não estou interessado) que permite ao usuário dizer que não tem interesse no produto que foi recomendado.

3) “Quem se interessou por x também se interessou por y ”

Esse tipo de saída é um dos exemplos mais claros de recomendação. Embora não seja personalizada, esse uso é capaz de oferecer ao usuário itens que ele desconhecia por levar em conta o que outros usuários comentaram, viram ou compraram.

4) Examinar itens semelhantes

Finalmente também existe a possibilidade de recomendar itens que possuam características em comum com o item em questão.

Por exemplo, se o usuário estiver vendo uma camiseta da última coleção, é possível recomendar outras camisetas da mesma coleção.

(ii) Predição

Outra forma de entrega é predizendo se o usuário gostará ou não do item. Existem algoritmos para cálculo de predição em função do perfil do usuário.

Normalmente, esses algoritmos geram um valor que representa a opinião ou avaliação que o usuário ativo daria em relação a um item do catálogo.

A predição pode ser usada para indicar as chances que um usuário específico tem de gostar ou não de um item com base no que o sistema já sabe sobre ele e sobre o item.

(iii) Avaliações

Uma terceira forma de entrega é exibir as avaliações (*ratings*) e/ou comentários que outros usuários forneceram para um determinado item. Isso contribui para que o usuário que estiver avaliando o item tenha mais informações e conheça-o melhor.

É também um modo sutil de influenciar o usuário que não conhece o produto e que deseja saber mais. A figura abaixo é um exemplo de como as avaliações podem ser apresentadas:



Figura 22 - Exemplo de avaliação de um item.

Fonte: sítio <http://www.submarino.com.br>, acessado em junho de 2009.

C) Operação

Em termos de operação, os sistemas de recomendação podem ser classificados quanto à arquitetura, à localização e ao modo como trabalham.

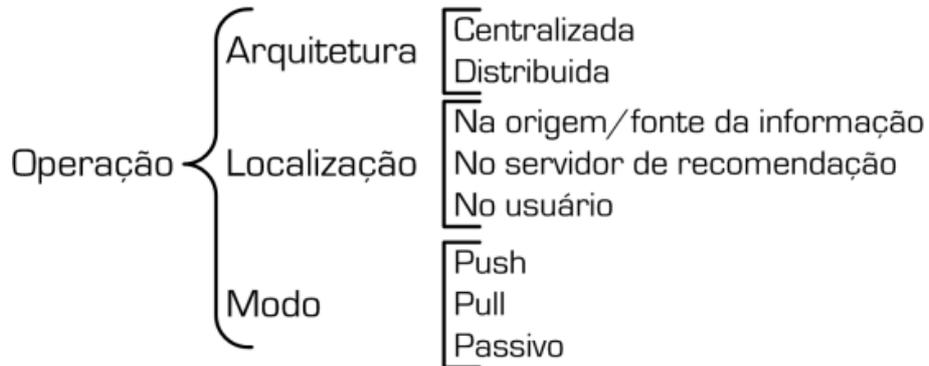


Figura 23 – Classificação da operação do sistema.

Fonte: elaborado pelo autor a partir de Manouselis et al. (2007)

i) Classificação quanto à arquitetura

a) Centralizada

Em um sistema de recomendação de arquitetura **centralizada**, o processo é realizado exclusivamente em um só local.

b) Distribuída

Em um processo de arquitetura **distribuída** as recomendações são geradas em diversos servidores, como observados em aplicações ponto a ponto (*peer to peer* em inglês).

ii) Classificação quanto à localização

Quanto à localização, o sistema pode gerar as recomendações em três locais: na origem da informação; no servidor de recomendações; ou ainda na máquina do usuário.

a) Na origem

Quando a recomendação normalmente é gerada na **origem** em sítios de comércio eletrônico, como forma de oferecer um serviço de valor agregado a seus usuários.

Um exemplo desse tipo de localização é utilizado pela Amazon (<http://www.amazon.com>), que gera suas recomendações a partir de seu catálogo de produtos em seus servidores.

b) No servidor de recomendação

A utilização de um **servidor de recomendação** ocorre por empresas que não dispõem das competências para, ou tenham optado por terceirizar essa função.

A Loomia (<http://www.loomia.com>) é uma empresa que oferece esse tipo de serviço.

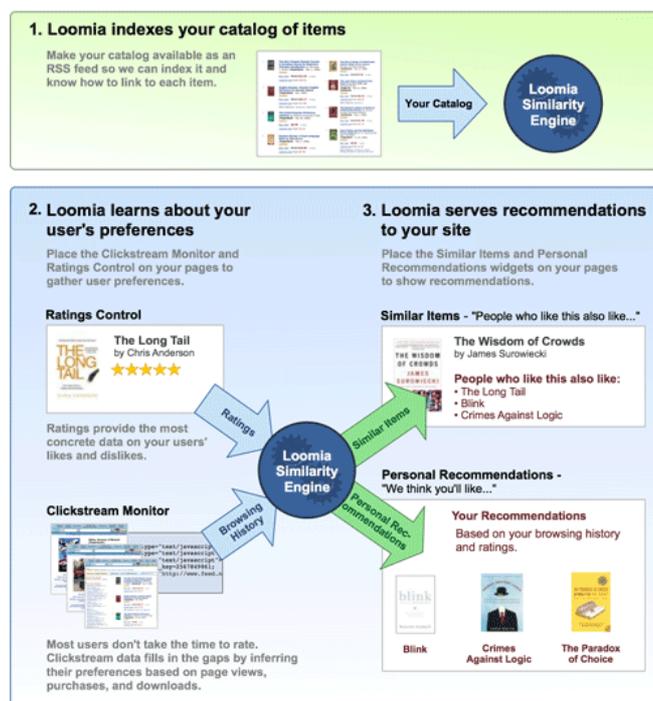


Figura 24 - Diagrama de como o Loomia funciona (original em inglês).

Fonte: sítio <http://www.marketingrev.com/2007/07/>, acessado em junho de 2009.

c) No ambiente do usuário

Por último, pode-se observar a geração de recomendações em aplicativos de correio eletrônico que usam técnicas desse tipo para filtrar mensagens indesejadas (SPAM). Esses casos são exemplo de recomendações geradas no ambiente do **usuário**.

A figura 25 é um exemplo de aplicação de regras para a filtragem de SPAM. Nela observa-se que em função dos filtros que a mensagem correspondeu, são somados pontos à mensagem. Uma vez que esses pontos somem mais de 5.0 (vide o exemplo abaixo), a mensagem é considerada SPAM.

Content analysis details: (5.1 points, 5.0 required)		
pts	rule name	description
1.6	RCVD_IN_SBL	RBL: Received via a relay in Spamhaus SBL [187.16.216.57 listed in zen.spamhaus.org]
3.2	FROM_LOCAL_NOVOWEL	From: localpart has series of non-vowel letters
-2.6	BAYES_00	BODY: Bayesian spam probability is 0 to 1% [score: 0.0000]
0.0	HTML_MESSAGE	BODY: HTML included in message
1.5	MIME_HTML_ONLY	BODY: Message only has text/html MIME parts
1.5	URIBL_SBL	Contains an URL listed in the SBL blocklist [URIs: expansao.com]

Figura 25 - Exemplo de filtro de e-mail baseado em regras.
Fonte: elaborado pelo autor.

iii) Classificação quanto ao modo de ativação

Em relação à ativação do sistema de recomendação, identificam-se os modos ativo e passivo, sendo que o modo ativo pode ser dividido em *push* ou *pull*, o que depende de quem inicia o processo.

- a) No modo **push**, as recomendações são entregues aos usuários mesmo que eles não as tenham solicitado, cabendo a ele usá-las ou não.
- b) Já no modelo **pull**, o usuário deve solicitar formalmente ao sistema que este lhe envie as recomendações existentes para determinado item.

- c) Finalmente, no modo **passivo**, as recomendações fazem parte da filosofia de operação do sistema, e são entregues conforme a especificidade de cada situação.

A seguir serão comentados alguns exemplos de sistemas de recomendação

2.3 Exemplos de aplicações de sistemas de recomendação

Nesta seção, serão exemplificadas algumas das diversas aplicações para os sistemas de recomendação. Ao final deste trabalho, há uma lista completa dos sítios visitados que inclui os endereços para os exemplos aqui citados.

2.3.1 O laboratório GroupLens

O GroupLens é o laboratório de pesquisa do Departamento de Ciência da Computação e Engenharia da universidade de Minnesota nos Estados Unidos. É considerado como um centro de pesquisa na área dos sistemas de recomendação.

O **projeto MovieLens** (disponível em <http://movielens.umn.edu>) é um sistema voltado à recomendação de filmes. Ele foi proposto por Riedl et al. (1999) para sugerir filmes que as pessoas possam gostar em função do que outras pessoas, com gostos similares, tenham visto.

2.3.2 Aplicações de comércio eletrônico

Das aplicações de sistemas de recomendação na área de comércio eletrônico, a **Amazon** constitui-se em um dos casos de sucesso mais notórios. Além da Amazon, há também o sítio **StyleFeeder** que oferece produtos que possam interessar ao usuário tomando por base outros produtos que ele tenha adquirido, criando assim um “estilo”.

2.3.3 Aplicações voltadas ao entretenimento

Na área do entretenimento se encontram diversos exemplos de utilização de sistemas de recomendação. De forma geral, o objetivo tradicional tende a ser a indicação de música, filmes, e fotografias. As seguintes aplicações podem ser consideradas como exemplos: **Collarity; Amie Street; Genius; iLike; Last.fm; Netflix; Pandora; RateYourMusic; Slacker; e Photoree.**

2.3.4 Alguns produtos e serviços de recomendação

Finalmente, cabe também citar produtos e serviços orientados à geração de recomendações para auxiliar os consumidores e aumentar o faturamento das lojas. Como exemplos pode-se citar: **Baynote; ChoiceStream; Criteo; inSuggest; Loomia; Minekey; MyStrands; Jumper 2.0; Peerius; prudsys RE; StumbleUpon; e richrelevance.**

2.4 Outros aspectos associados aos sistemas de recomendação

Existem ainda outros aspectos que devem ser levados em consideração no desenho deste tipo de sistemas.

2.4.1 Volatilidade das informações

A questão da volatilidade de informações pode ser exemplificada quando se fala em moda. Para gerar recomendações satisfatórias, com um grau de relevância satisfatório, faz-se necessário dispor de um volume mínimo de dados. Isso evidencia que as recomendações são geradas com “olhos para o passado”. No entanto, quando a moda muda, as informações do passado perdem parte de seu valor.

Com isso, é preciso criar algum mecanismo de compensação, ou de consolidação para que as informações mais recentes tenham um valor maior do que as informações

antigas. Não se esquecendo de que tal mecanismo deverá considerar o mercado, a massa de dados, e a volatilidade com a qual as informações são alteradas.

2.4.2 Mudanças nas preferências dos usuários

Outro problema dos sistemas de recomendação é considerar um evento no histórico do usuário como sendo algo que ele faça, ou aprecie, com frequência. Por exemplo, ao entrar em uma loja virtual para comprar um presente para o filho de um amigo, é possível que a loja passe a lhe recomendar brinquedos e roupas infantis, ainda que você não tenha filhos.

2.4.3 Recomendações imprevisíveis

É bastante difícil também para que os algoritmos de recomendação existentes consigam compreender o comportamento humano, e reconhecer que alguém pode ser ao mesmo tempo fã de dois artigos completamente diferentes como, por exemplo, livros de ficção científica e livros de auto-ajuda.

Mas isso é inerente do ser humano. Não há um padrão. Somos imprevisíveis, e é fundamental levar isso em consideração.

2.4.4 Peças que ainda não foram vendidas não podem ser comparadas

Usando a filtragem colaborativa, é preciso que ao menos uma pessoa tenha comprado ou avaliado um item, para que ele possa ser considerado nas recomendações futuras. Assim, se uma peça não foi vendida ela não tem histórico de venda, e conseqüentemente não poderá ser comparada a nenhuma outra.

Felizmente, para esses casos há soluções de contorno, como sugerir os novos itens indicando que se trata de novas adições ao acervo, e desta forma surpreendendo os usuários, ou então usar a figura de um curador para criar conjuntos de itens que misturem novos e antigos para promover a venda.

2.5 O que foi visto neste capítulo

Esse capítulo iniciou por apresentar os sistemas de recomendação, sua origem a partir da recomendação boca-a-boca, o surgimento do sistema Tapestry e a diferença entre a recuperação e a filtragem de informações. Isso feito, o capítulo segue com a descrição detalhada da taxonomia adotada, para finalizar com outros aspectos secundários sobre sistemas de recomendação.

3. METODOLOGIA DE PESQUISA

A descrição do sistema de recomendação para redes de varejo de moda proposto neste trabalho é o objeto a ser tratado neste capítulo. Inicialmente será apresentado o procedimento metodológico que foi seguido, esclarecendo o seu processo construtivo. Em seguida, os dados de entrada e saída, assim como os algoritmos utilizados, são apresentados e comentados.

3.1 Procedimentos metodológicos e caracterização da pesquisa

Entende-se que o trabalho científico em questão trata de esclarecer o problema apresentado na seção 1.1, aprimorando idéias já discutidas, e obtendo novas intuições capazes de contribuir para a resolução do problema. Assim, classifica-se o trabalho em pesquisa exploratória conforme Gil (2002).

Os procedimentos metodológicos adotados neste trabalho iniciaram-se pela revisão bibliográfica e contextualização do tema proposto, com vistas a definir o processo construtivo que conduzirá ao alcance dos objetivos estabelecidos.

É possível afirmar que o delineamento metodológico da pesquisa é função do método de abordagem que foi utilizado, e que este, por sua vez, relaciona-se com o tipo de raciocínio utilizado⁶ (SANTAELLA, 2001)

Neste capítulo é estabelecido o procedimento metodológico sugerido para o desenvolvimento de sistemas de recomendação para redes de varejo de moda. Em seguida, será apresentada a construção do protótipo usado filtragem colaborativa e informações contextuais. No processo de construção teórica, usou-se de forma predominante o método indutivo para consolidar a abordagem adotada pelo autor.

⁶ Os tipos de raciocínio, segundo Santaella, são classificados em: indutivo, dedutivo, hipotético-dedutivo e dialético.

No próximo capítulo, que trata da construção do protótipo do sistema de recomendação, para validar a proposta em questão, usou-se uma abordagem dedutiva, uma vez que o principal objetivo da construção é a validação da proposta por meio de resultados mensuráveis.

3.2 Delineando o problema

Atualmente pode-se afirmar o processo de venda executado pelos consumidores nas lojas segue um roteiro relativamente simples. Ele começa pela visita do cliente à loja, que depois seleciona as peças que pretende comprar. Algumas vezes ele experimenta a peça, outras não. Finalmente, ele se dirige ao caixa para realizar o pagamento.

Neste ponto, o caixa da loja registra a operação que poderá ou não estar associada a um dado de identificação do cliente (como o telefone de contato ou o CPF). Os dados dessa transação são então enviados para o escritório central da rede, o que pode ser feito *on-line*, ou ao final do dia dependendo dos recursos de sistema envolvidos.

Uma vez que as transações sejam recebidas pelo escritório central, o sistema de retaguarda⁷ efetua a consolidação dos dados atualizando os estoques e os indicadores de desempenho⁸ das lojas e da rede.

Neste cenário o escritório central da rede dispõe das transações de todos os seus clientes, e é capaz de identificar o que eles compraram, quando e onde. Assim sendo, é perfeitamente factível a criação de relatórios para apoiar estratégias de *marketing* recomendando as peças que os consumidores tenham mais chance de comprar em função de suas compras no passado.

⁷ Ou *BackOffice* como costumam ser chamados tais sistemas em inglês.

⁸ Também conhecidos como *Key Performance Indicators* (KPI) em inglês.

3.3 Etapas da pesquisa

Por meio da adaptação da metodologia proposta por Oliveira (2008), e atualizando-a com a taxonomia proposta por Manouselis e Costopoulou (2007), definem-se quatro passos da pesquisa que servirão de guia para alcançar o objetivo proposto. A figura 25 ilustra estes passos desde sua concepção até o desenvolvimento.

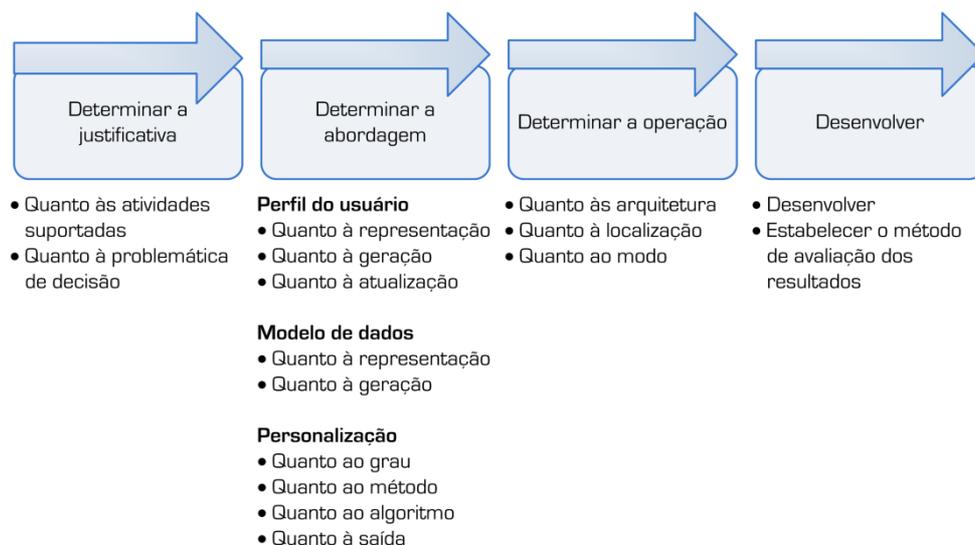


Figura 26 - Passos da pesquisa.
Fonte: elaborado pelo autor.

De acordo com a abordagem metodológica aqui descrita e a taxonomia discutida no item 2.2 (p.12), o primeiro passo consiste em determinar a **justificativa** do sistema; que será determinada tanto em relação às atividades que serão suportadas, quanto à problemática da decisão.

Em seguida, deve ser identificada a **abordagem** a ser dada ao perfil do usuário, às características dos itens e à personalização. O perfil do usuário trata de como os dados do usuário e suas avaliações em relação aos itens serão coletados. As características dos itens correspondem exatamente ao tratamento que será dado ao catálogo de itens disponível para recomendação. E por fim, a personalização compreende o tipo de filtragem a ser empregado, o processo a ser empregado para gerar recomendações, e a forma de apresentar os resultados.

Depois a **operação** do sistema é definida quanto a sua arquitetura, localização e ao modo de operação.

Finalmente é realizado o **desenvolvimento** do sistema e o estabelecimento de uma mecânica para avaliar a qualidade das recomendações geradas.

Nesta seção esse procedimento metodológico será explicado, enquanto que no capítulo 4 (p. 60) serão tratados os aspectos relativos ao desenvolvimento do protótipo.

3.3.1 Justificando o sistema de recomendação para redes de varejo de moda

Em termos das atividades suportadas, o fato de permitir que se encontre itens interessantes será a atividade adotada haja vista que o resultado do sistema proposto tem sua origem na correlação dos produtos da rede com a base de consumidores.

Tal resultado será oferecido por meio da relação ordenada dos produtos com os coeficientes de correlação dos produtos que podem serem comprados pelos clientes, o que leva ao fato de que a ordenação é a problemática da decisão a ser trabalhada.

JUSTIFICANDO O SISTEMA DE RECOMENDAÇÃO	
Atividades suportadas	<i>Encontrar itens interessantes</i>
Problemática da decisão	<i>Ordenação</i>

Tabela 1 – Justificativas do sistema de recomendação.

Fonte: elaborada pelo autor.

3.3.2 Definindo a abordagem do sistema face à taxonomia adotada

Uma vez que a justificativa para o sistema de recomendação proposto foi tratada, deve-se trabalhar de que forma será realizada a abordagem em relação ao perfil do usuário, ao modelo de dados, e à personalização. A taxonomia adotada fornece os meios para a correta identificação e classificação destes elementos.

Perfil do usuário

No caso proposto o perfil do usuário é automaticamente alimentado pelos dados do sistema de gestão de retaguarda usado pela rede de varejo.

A classificação do perfil do usuário relativa à taxonomia adotada (explicada na página 12) pode ser vista na tabela a seguir:

CLASSIFICAÇÃO DO PERFIL DO USUÁRIO		
Representação	Método	<i>Características demográficas</i>
		<i>Modelo baseado em vetores</i>
	Tipo	<i>Mensurável</i>
Geração	Inicial	<i>Manual</i>
	Aprendizagem	<i>Desnecessária</i>
Atualização	Método	<i>Explícito</i>
	Técnica	<i>Manual</i>

Tabela 2 – Classificação do perfil do usuário no protótipo.
Fonte: elaborada pelo autor.

Pode-se dizer que o protótipo adotará as características demográficas (bairro, cidade, estado) e um modelo baseado em vetores (quantidade de itens comprados) para compor o método e que tal representação será do tipo mensurável.

No tocante à geração do perfil do usuário, ela será manual com a entrada dos dados sendo feita nas lojas, no momento da venda. Como o protótipo busca demonstrar a viabilidade na adoção de um sistema de recomendação para a geração de campanhas personalizadas, não será necessário desenvolver a aprendizagem do perfil do usuário.

Finalmente, quanto à atualização dos dados do usuário e de suas avaliações em relação aos itens, será usado o método explícito com técnica manual. Isso ocorre pelo fato do protótipo trabalhar com dados gerados no sistema de retaguarda da rede de varejo.

A adoção do método explícito é factível uma vez que a correlação a ser efetuada será realizada entre os itens e não entre os consumidores. Assim, como a tendência dos itens serem vendidos com mais frequência é maior do que a de um consumidor repetir a compra, a opção é pela filtragem colaborativa.

Características dos itens

Os itens que serão tratados pelo sistema são as peças vendidas nas lojas. Cada registro de peça fica armazenado no sistema de retaguarda.

CLASSIFICAÇÃO DAS CARACTERÍSTICAS DOS ITENS		
Representação	Método	<i>Índice/lista</i>
	Tipo	<i>Mensurável</i>
Geração		<i>Manual</i>

Tabela 3 – Classificação das características dos itens no protótipo.
Fonte: elaborada pelo autor.

Para os itens que compõe o acervo do protótipo, o **método** de representação adotado é o de lista, uma vez que ele representa com facilidade a quantidade de itens que um consumidor adquiriu. A correlação a ser realizada (entre quantidades de peças vendidas), implica na adoção de uma representação do **tipo** mensurável.

No tocante à **geração** do acervo de itens, ela também será manual sendo informada no sistema de retaguarda, na medida em que novas peças são cadastradas ou atualizadas. Essa atividade é função da rede de varejo.

Geralmente as novas peças são cadastradas no lançamento de coleções, o que faz com que a geração ocorra em datas pré-determinadas na maior parte das vezes. Em outras indústrias, os itens do catálogo podem ser cadastrados com uma frequência bastante maior.

Entrega

Quanto à entrega, o protótipo está classificado de acordo com a tabela a seguir:

CLASSIFICAÇÃO EM RELAÇÃO À ENTREGA		
Grau		<i>Persistente</i>
Método		<i>Filtragem colaborativa</i>
Algoritmo	Tipo	<i>Baseado em memória</i>
	Técnica	<i>Item-a-item</i>
Saída		<i>Sugestão</i>

Tabela 4 – Classificação das características do protótipo em relação à entrega.
Fonte: elaborada pelo autor.

Como os dados de venda que são coletados e posteriormente usados na geração de recomendações nunca são descartados, o **grau** é persistente.

Em relação ao **método**, a opção pela filtragem colaborativa decorre dos fatores listados adiante e que foram esclarecidos na página 34.

- A recuperação pura das peças não implica em personalização, desviando assim do objetivo do protótipo.
- A seleção manual das peças não atenderia ao objetivo do protótipo, uma vez que não está prevista a intervenção de um curador indicando quais deveriam ser recomendadas.
- A utilização pura e simples dos dados demográficos dos consumidores não é suficiente, já que o protótipo tem por fim recomendar as peças que os consumidores tenham mais chance de comprar em função de suas compras no passado.

- A filtragem baseada em conteúdo não é uma opção viável já que o sistema não dispõe de informações sobre os gostos dos consumidores para poder cruzá-los com os atributos das peças. Além disso, dada a arquitetura do protótipo, onde a entrada de dados é totalmente realizada em outro sistema (de retaguarda), seria necessário permitir ao consumidor atualizar o seu gosto de alguma forma, o que implicaria em um custo operacional mais elevado.
- Uma vez que se dispõe dos dados históricos de compras das peças pelos consumidores, a filtragem baseada em utilidade também não se faz necessária.

Logo, a opção pela filtragem colaborativa surge como a mais viável dada a possibilidade de cruzar as informações históricas de compra dos consumidores e determinar os itens que eles têm mais chance de adquirir.

No tocante ao **algoritmo**, o protótipo é baseado em memória uma vez que considera todos os dados históricos dos clientes e das peças para a geração das correlações item-a-item (que define a técnica), que produzirão as recomendações.

Por fim, é necessário estabelecer de que maneira as recomendações serão repassadas ao usuário do sistema, ou seja, como deve ser feita a **saída**. Como já citado, o objetivo do protótipo é “recomendar as peças que os consumidores tenham mais chance de comprar em função de suas compras no passado”, logo é possível estabelecer as seguintes saídas:

- Consultar uma peça;
- Consultar um consumidor; e
- Relatório de recomendações.

A descrição detalhada para as saídas acima será vista mais adiante no capítulo que trata do protótipo, mas pode-se adiantar que o modelo adotado foi o de sugestões.

3.3.3 Detalhando a operação da solução proposta

Por fim, a operação da solução proposta no trabalho segue o modelo proposta na taxonomia e que será descrito nesta seção.

CLASSIFICAÇÃO EM RELAÇÃO À OPERAÇÃO	
Arquitetura	<i>Centralizada</i>
Localização	<i>No sistema de recomendação</i>
Modo	<i>Pull</i>

Tabela 5 – Classificação das características do protótipo em relação à operação.

Fonte: elaborada pelo autor.

As recomendações serão geradas exclusivamente em um só local, o que leva à adoção de uma **arquitetura** centralizada em vez de distribuída. Até o fato de tratar-se de um protótipo com uma quantidade controlada de registros permite a adoção desse tipo de solução.

As características do resultado desejado levam a trabalhar a **localização** da operação no sistema de recomendação em si. Não seria viável, dado o cenário explicado na descrição da taxonomia (p. 43), efetuar as correlações nos pontos de venda, nem tampouco no ambiente dos usuários (clientes).

Por fim, as recomendações seguem o **modo pull**, no qual elas devem ser explicitamente solicitadas pelos operadores do sistema. E é de fato o que ocorre quando o operador solicita a consulta a uma peça, a um consumidor ou ao relatório de recomendações.

Os dois outros modos (*push* e passivo) não se aplicam ao cenário descrito dadas as suas características inerentes.

3.4 O que foi visto neste capítulo

Este capítulo apresentou a condução da metodologia que foi usada na construção do protótipo, explicando que o autor iniciou a jornada com um raciocínio indutivo e indicando que o trabalho seguirá para o dedutivo. A taxonomia que havia sido apresentada no segundo capítulo foi aplicada ao cenário em questão, e as opções feitas pelo autor foram justificadas uma a uma. O próximo capítulo tratará da constituição do protótipo, segundo o que acaba de ser descrito.

4. DESENVOLVIMENTO DO PROTÓTIPO

Este capítulo traz a descrição do protótipo construído para processar os dados relativos às compras realizadas e a geração das correlações a fim de obter as recomendações.

A base de dados empregada é formada por registros de venda de uma rede de vestuário no período de agosto a dezembro de 2008. Tal base contém mais de 146 mil registros de vendas.

Como se trata de dados comerciais sensíveis, o nome da rede não será mencionado neste trabalho.

4.1 Visão geral do mercado de varejo de moda

Uma das funções do protótipo é demonstrar a possibilidade de gerar um **mapa das correlações** entre as peças tomando por base o histórico real das vendas de cada uma delas. Em função dessas correlações, o gestor da rede de varejo poderá então realizar promoções personalizadas para seus clientes.

É importante estabelecer alguns conceitos adotados para a construção do protótipo e que são usados relativos ao varejo de moda.

Nesse mercado as empresas são conhecidas como redes, onde uma rede pode, ou não, possuir várias lojas. Cada loja tem seus vendedores e esses vendedores atendem aos clientes. A figura a seguir demonstra a relação entre esses conceitos:

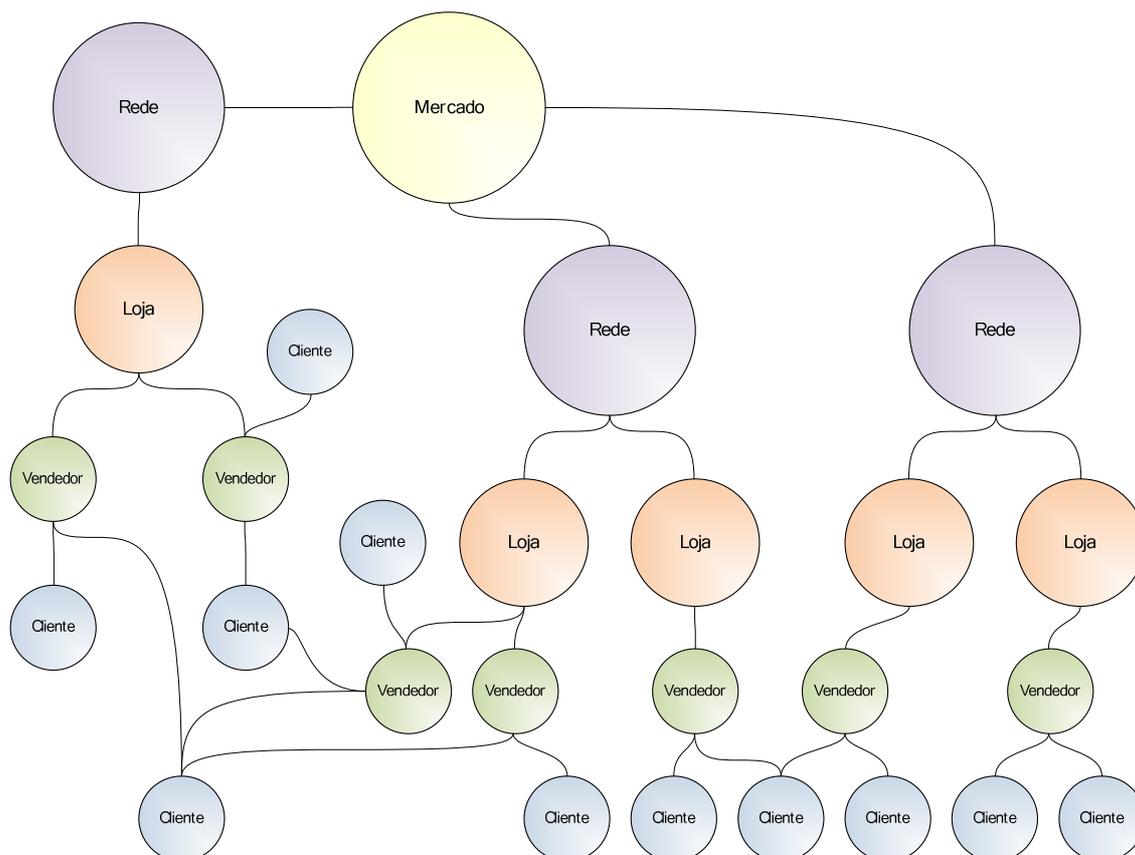


Figura 27 – Relação entre conceitos das redes de varejo de moda.
 Fonte: elaborado pelo autor.

4.2 As funções identificadas no protótipo

Conforme citado anteriormente, uma das funções do protótipo é demonstrar a vantagem em correlacionar as peças efetivamente vendidas em uma rede de lojas, a fim de que a rede possa criar campanhas promocionais personalizadas.

Dessa forma a rede de lojas passa a oferecer descontos com maior chance de ter seus produtos vendidos. O raciocínio é simples: em vez de conceder desconto em toda uma linha de produtos, os descontos são concedidos por produto e somente para os clientes com maior chance de comprá-los.

Na construção do protótipo foram identificadas cinco funções principais que representam o macro-processo do início ao fim: aquisição; entrada; processo; saída; e interpretação.

A figura a seguir ilustra o macro-processo que constitui o protótipo:

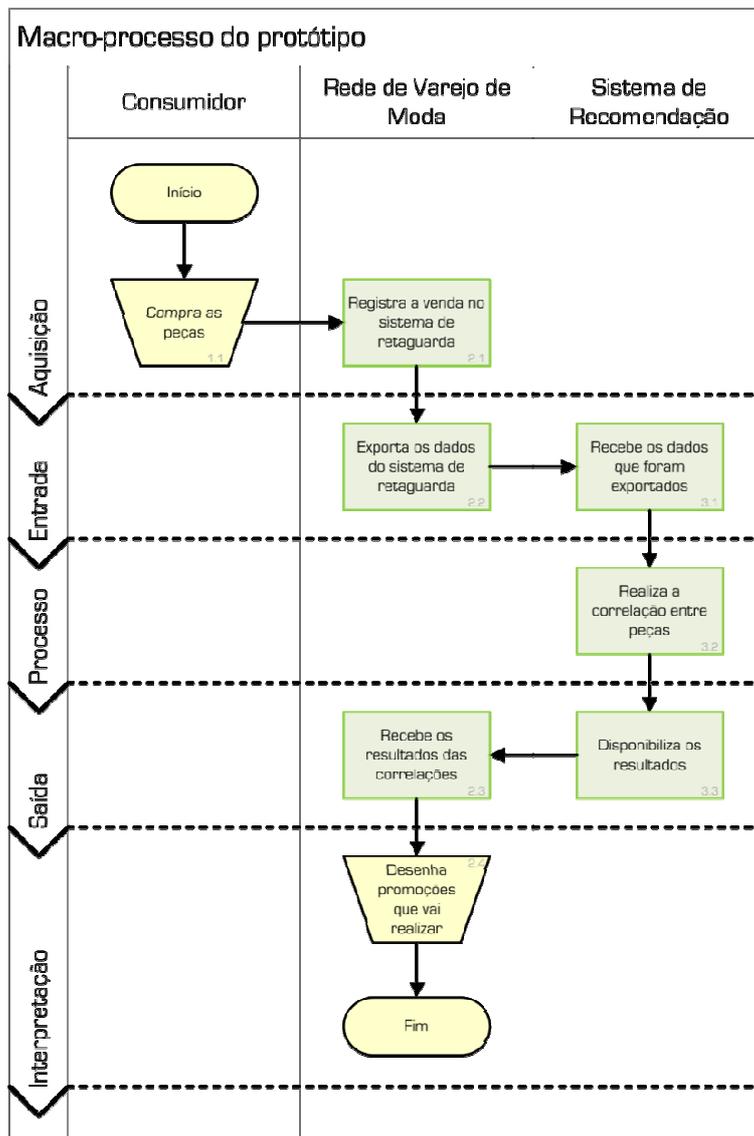


Figura 28 - Diagrama funcional do macro-processo do protótipo.
Fonte: elaborado pelo autor.

4.2.1 Compra as peças

Nesta atividade o cliente vai até uma loja e realiza a compra das peças que deseja. No cenário do protótipo esse é um processo manual uma vez que se trata de um cenário real e não de uma compra em uma loja virtual.

4.2.2 Registra a venda no sistema de retaguarda

Neste momento é realizado o registro da venda no sistema de retaguarda pelo caixa ou vendedor que atendeu ao cliente. A partir deste momento, os dados da compra feita pelo cliente passam a estar disponíveis na base de dados da rede.

4.2.3 Exporta os dados do sistema de retaguarda

De acordo com a periodicidade acordada com a rede, os dados são exportados da base de dados do sistema de retaguarda para arquivos XML. Estes arquivos serão enviados ao sistema de recomendação para serem processados.

O sistema de retaguarda analisado gera seis arquivos no formato texto, a saber:

- Lojas da rede – contém todas as lojas da rede de varejo;
- Vendedores – representa todos os vendedores (ativos e inativos) das lojas;
- Peças – compreende as peças que compõem a grade de produtos da rede;
- Clientes – encerra todos os clientes da rede;
- Notas fiscais – contém as notas fiscais emitidas pelo sistema. Cabe lembrar que notas de devolução costumam estar presentes neste arquivo; e
- Peças vendidas – inclui as peças efetivamente vendidas ou devolvidas associadas às notas fiscais.

4.2.4 Recebe os dados que foram exportados

O sistema de recomendação recebe então os arquivos de dados que foram exportados na atividade anterior.

Esses arquivos são processados usando regras de validação para garantir que os dados que serão considerados nas correlações são válidos.

4.2.5 Realiza a correlação entre peças

Uma vez que os arquivos tenham sido importados para a base de dados relacional do sistema de recomendação, faz-se necessário realizar as correlações entre as peças a fim de determinar as vizinhanças.

Isso pode ser feito tomando-se por base diversos fatores, como por exemplo: a quantidade absoluta de peças vendidas, a quantidade de atendimentos realizados para cada peça, ou o valor médio por peça. Neste protótipo, foi usada a quantidade absoluta de peças vendidas, para estabelecer essas correlações.

A tabela abaixo exemplifica as 5 peças mais vendidas na base de dados de exemplo. Nela se observa que a camiseta de alça foi vendida 1.433 vezes, no entanto, como alguns consumidores comprem mais de uma unidade do mesmo artigo, é importante considerar também a quantidade vendida para efeitos de cálculo de similaridade.

Peças	Operações de venda	Quantidades (em unidades)
Camiseta de alça	1.433	2.468
Camiseta básica	1.147	1.670
Manga comprida	1.058	1.803
T-shirt gola v	580	627
Blusa decotada	551	603

Tabela 6 – Quantidade de peças compradas pelos consumidores.

Os cálculos serão demonstrados para o par de artigos formado pela venda de camisetas básica e as de manga comprida. Estas peças foram selecionadas pela representatividade que possuem no espaço amostral. No entanto, vale ressaltar, que os cálculos foram realizados para os dez pares de peças presentes na tabela acima.

A figura 29 mostra as quantidades compradas dessas peças no plano cartesiano. Acredita-se que a visualização espacial de como é a distribuição quantitativa das peças auxilie na compreensão do método explicado.

Na figura é possível ver que Tatiana comprou duas camisetas básicas e três camisas de manga comprida. Assim como Claudia, Elisa, Julia e Leila que comprara uma unidade de cada uma das duas peças em análise.

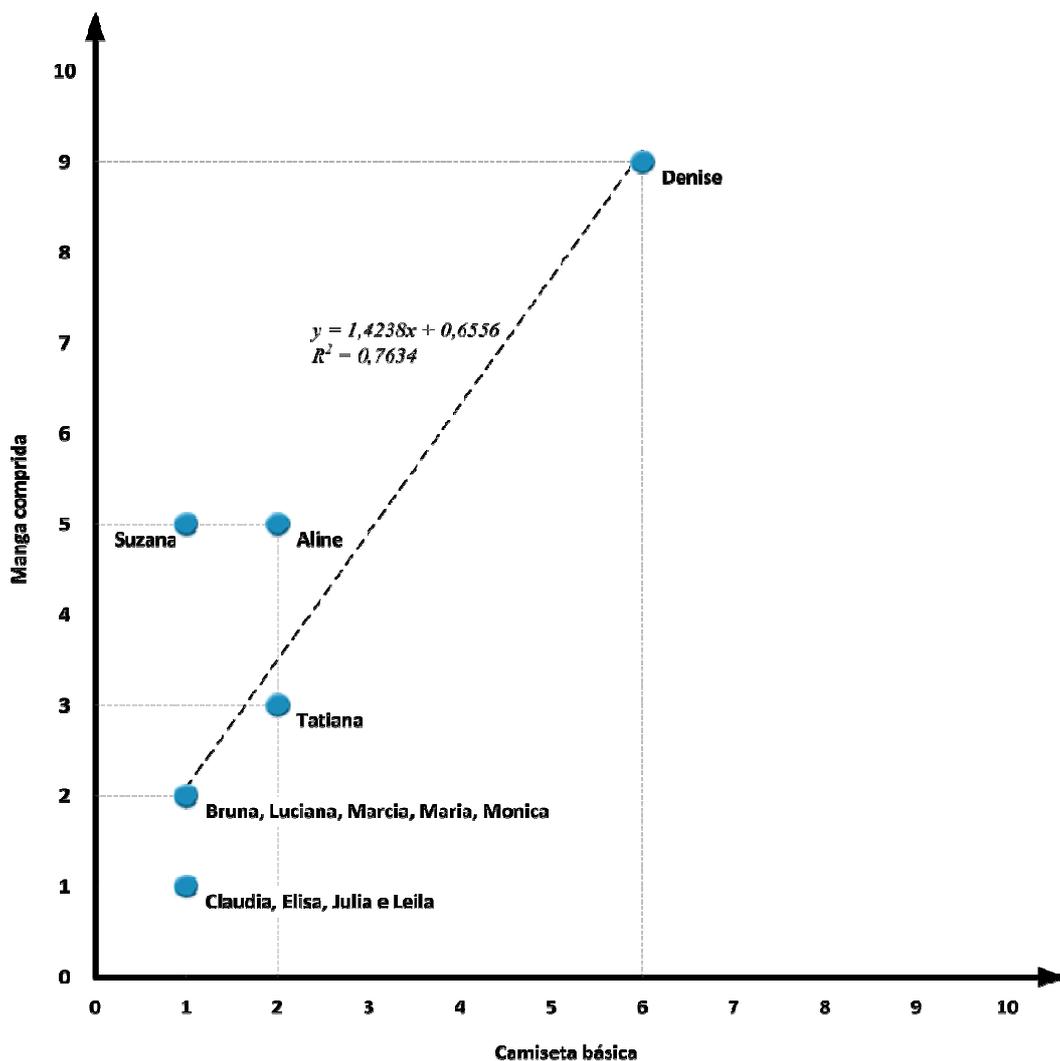


Figura 29 - Quantidades de peças de camisetas básicas e de manga comprida.
Fonte: elaborado pelo autor.

O coeficiente de Pearson é uma medida de correlação entre duas variáveis. No caso deste trabalho, tomou-se como valor de base para o cálculo da correlação a

quantidade de artigos vendidos. Isso foi feito uma vez que, no caso do varejo de moda *off-line*, os consumidores não fornecem uma nota após comprar determinado artigo indicando o quanto gostaram, ou não, do artigo em questão.

Poder-se-ia ter tomado outro tipo de coeficiente de similaridade como o índice de Jaccard ou a similaridade pelo cosseno, no entanto, o autor optou pelo coeficiente de Pearson, uma vez que ele é o mais utilizado na bibliografia tomada por referência.

Calcula-se o coeficiente para cada par de artigos e o resultado é um número no intervalo de [-1;+1], onde -1 indica ausência de correlação e +1 indica uma forte correlação entre os produtos. Para calcular o coeficiente de Pearson a fórmula utilizada foi:

$$r_{a,b} = \frac{\sum_{i=1}^n (q_{a,i} \times q_{b,i}) - \left(\frac{\sum q_a \times \sum q_b}{n} \right)}{\sqrt{\left[\sum_{i=1}^n (q_{a,i}^2) - \frac{(\sum q_a)^2}{n} \right] \times \left[\sum_{i=1}^n (q_{b,i}^2) - \frac{(\sum q_b)^2}{n} \right]}}$$

Equação 1 – O cálculo do Coeficiente de Pearson.

Onde $r_{a,b}$ é a correlação entre o artigo a e o artigo b , $q_{a,i}$ é a quantidade do artigo a comprada pelo cliente i , e n representa a quantidade de clientes que compraram tanto o artigo a quanto o artigo b .

O primeiro passo do cálculo consiste em identificar na amostra de dados, quais consumidores compraram tanto um artigo quanto o outro.

Em seguida, identifica-se a quantidade que cada consumidor comprou dos artigos em questão, preenchendo as colunas (A) e (B) da tabela adiante.

Além disso, também foram calculados os quadrados das quantidades compradas, e o produto, conforme mostram as três últimas colunas da tabela.

Consumidores	Quantidades		(A) ²	(B) ²	(A) x (B)
	Camiseta básica (A)	Manga Comprida (B)			
Aline	2	5	4	25	10
Bruna	1	2	1	4	2
Claudia	1	2	1	4	2
Denise	6	9	36	81	54
Elisa	1	1	1	1	1
Julia	1	1	1	1	1
Leila	1	1	1	1	1
Luciana	1	2	1	4	2
Marcia	1	2	1	4	2
Maria	1	2	1	4	2
Monica	1	2	1	4	2
Suzana	1	5	1	25	5
Tatiana	2	3	4	9	6
Totais	20	37	54	167	90

Tabela 7 – Valores utilizados no cálculo do coeficiente de Pearson para dois artigos selecionados.
Fonte: elaborado pelo autor.

Aplicando-se então a fórmula, temos:

$$r_{a,b} = \frac{90 - \left(\frac{20 \times 37}{13}\right)}{\sqrt{\left(54 - \frac{20^2}{13}\right) \times \left(167 - \frac{37^2}{13}\right)}} = \frac{90 - \frac{740}{13}}{\sqrt{23,23 \times 61,69}} = \frac{90 - 56,92}{\sqrt{1433,16}} = \frac{33,08}{37,86} = 0,873731$$

Equação 2 – Aplicação da fórmula do Coeficiente de Pearson.

O que indica que entre a camiseta básica e a camisa de manga comprida existe uma correlação de 0,8737.

Repetindo-se o cálculo para os demais pares de artigos temos os resultados abaixo:

Camiseta básica	-0,074471			
Manga comprida	-0,005093	0,873731		
T-shirt gola v	0,356637	0,582818	0,926095	
Blusa decotada	-0,095791	0,004298	-0,204112	0,184210
	Camiseta de alça	Camiseta básica	Manga comprida	T-shirt gola v

Tabela 8 – Resultados obtidos usando-se o coeficiente de Pearson.

Fonte: elaborado pelo autor.

4.2.6 Disponibiliza os resultados

Como foi visto na taxonomia, em tipos de saída (p.32), o sistema de recomendação pode fornecer o resultado das correlações de diversas formas. Para o protótipo em questão, foram selecionadas duas funções. Estas funções estão descritas nessa seção.

Consultar uma peça

Dada uma peça, essa função permite consultar seus dados. Ela recebe como parâmetro de entrada o código da peça, e devolve como resultado: o nome da peça e uma lista das peças com seus códigos, nomes, quantidades vendidas, quantidades de atendimento, coeficientes de similaridade, e número de clientes em potencial.

Peça consultada		Qtd Vendida	Atendimentos
Camiseta de alça		2.458 UN.	1.433

Peças relacionadas				
Nome	Qtd Vendida	Atendimentos	Coeficientes	Clientes Potenciais
T-shirt gola v	627 UN.	580	0,3566	148
Blusa decotada	603 UN.	551	-0,0957	29
Camiseta básica	1.670 UN.	1.147	-0,0744	23
Manga comprida	1.803 UN.	1.058	-0,0051	4

Imprimir

Veja os dados da peça selecionada.

Figura 30 - Tela de consulta a uma peça.

Fonte: extraído do protótipo construído.

Consultar um cliente

Da mesma forma que é possível consultar as correlações geradas para uma determinada peça, o protótipo também permite que, dado um cliente, se recupere seus principais atributos, as peças que comprou e quais peças que ele não comprou e tem grande possibilidade de compra.

Relatório de recomendações

Por fim, o protótipo também permite extrair um relatório de recomendações que demonstra claramente quais peças devam ser oferecidas a quais clientes. Como a figura abaixo exemplifica:



Cliente	Oferta 1	Oferta 2	Oferta 3	Oferta 4
Aline	T-shirt gola v	Blusa decotada	-	-
Bruna	Blusa decotada	T-shirt gola v	-	-
Claudia	Blusa decotada	T-shirt gola v	-	-
Denise	T-shirt gola v	Blusa decotada	-	-

4 clientes selecionados.

Figura 31 – Tela do relatório de recomendações.
Fonte: extraído do protótipo construído.

4.2.7 Recebe os resultados das correlações

Uma vez que o sistema gerou os dados para as consultas de peças e clientes, a rede de varejo de moda poderá acessá-los usando a interface do protótipo. Os dados são entregues por meio de uma planilha XLS⁹ ou em um relatório impresso. É o usuário do sistema quem determina a forma como deseja receber os dados.

⁹ Uma planilha XLS é um arquivo de planilha gerado usando o aplicativo Microsoft Excel.

4.2.8 Desenha as promoções que vai realizar

Depois de ter recebido os resultados das correlações, a rede de varejo de moda poderá então desenhar promoções inteligentes e personalizadas.

4.3 O protótipo

O protótipo construído segue o macro-processo descrito na figura 28, principiando pelo registro da venda no sistema de retaguarda e indo até a disponibilização dos resultados.

A tabela 9 a seguir descreve o formato do registro de transação que é enviado pela rede de varejo de moda para o sistema de recomendação. Nele constam a data da transação, o CPF e o nome do cliente, o código e o nome do produto, a quantidade vendida e o valor.

Esses registros de transação são validados seguindo regras pré-estabelecidas, como se o CPF válido, se a data está dentro do período desejado, além de outras que podem ser configuradas pelo operador.

REGISTRO DE TRANSAÇÃO

Data	CPF	Cliente	Código	Produto	Qtd	Valor
01/08/08	089.844.147-19	ALICE SILVA	12.02.01	CAMISETA	1	48,00
01/08/08	089.844.147-19	ALICE SILVA	12.04.02	CAMISA SOCIAL	1	123,00
...		

Tabela 9 – Estrutura do registro de transação.
Fonte: elaborado pelo autor.

Uma vez que os registros de transação tenham sido transferidos para a base de dados do sistema de recomendação, as correlações são efetuadas pelo sistema de acordo com o processo descrito anteriormente na seção 4.2.5.

Finalmente, os resultados ficam disponíveis para que o operador do sistema os extraia da forma desejada, conforme os formatos explicados na seção 4.2.6.

4.4 Dos resultados

A equipe responsável pelo planejamento das campanhas e ações de *marketing* da rede de varejo obtém diversas vantagens com a aplicação de um sistema de recomendação conforme descrito neste capítulo.

A possibilidade de consultar as peças com maior índice de correlação em relação a uma segunda peça permite planejar formas inteligentes de redução dos estoques. Isso pode ser feito sem que seja necessário estabelecer descontos baseados em tentativa e erro, ou ainda a peças que, por si, já vendem o suficiente e não seriam capazes de atrair outras peças.

Este capítulo sugere que o mesmo raciocínio seja aplicado aos clientes da rede para analisar quais clientes têm o padrão de consumo parecido, ou ainda que devam ser impactados por ações de *marketing* orientadas a certos itens do estoque.

4.5 O que foi visto neste capítulo

Este capítulo começou por fornecer uma visão geral do mercado de varejo de moda, para depois descrever as funções que foram selecionadas para o protótipo que foi construído. Em seguida o protótipo foi descrito de forma resumida (uma vez que suas funções já haviam sido), e os resultados capazes de serem obtidos foram explicados, permitindo que o leitor visualize outros, além dos já estabelecidos.

5. CONTRIBUIÇÕES E PERSPECTIVAS FUTURAS

5.1 Contribuições

A principal contribuição deste trabalho é comprovar a vantagem em trabalhar com dados empíricos em um ambiente *off-line*. Nos casos baseados na Internet (como por exemplo na Amazon), se trabalha com o interesse dos clientes, mas isso na verdade não quer dizer que quem viu alguma coisa efetivamente irá efetivamente comprá-la.

A adoção de uma taxonomia já estabelecida, além de sua adequação para o caso proposto, demonstra a viabilidade da construção de um sistema orientado a ele. Tal sistema é capaz de permitir que os investimentos em *marketing*, sobretudo em *marketing* de relacionamento, sejam feitos de forma mais parcimoniosa.

5.2 Perspectivas futuras

O protótipo delineado neste trabalho serviu para demonstrar o valor e a viabilidade em se trabalhar com sistemas de recomendação no ambiente *off-line*. No entanto, alguns aspectos não foram cobertos e poderiam ser mais bem explorados no futuro, tais como:

- A questão da personalização pode ser trabalhada mais a fundo, criando novos tipos de saída específicos para os clientes da rede de varejo de moda. Assim, malas diretas, cupons de desconto, ponto em programas de fidelidade, e outros produtos poderiam ter sua origem nas correlações apontadas pelo sistema de recomendação, tornando esses processos automáticos e com um resultado mais próximo das metas estabelecidas. Além disso, novos indicadores de desempenho poderiam ser criados e acompanhados fornecendo ao lojista novas visões sobre o comportamento de venda de sua rede.
- Outros algoritmos de correlação também podem ser avaliados além da correlação de Pearson. Isso permitirá a avaliação de qual o conjunto de

indicadores e fatores melhor refletem o comportamento de venda, sobretudo em uma indústria tão dependente de tendências e das estações.

- O tratamento da venda via comércio eletrônico não deve ser negligenciado, embora o comércio de roupas pela Internet ainda não seja algo comum no país. A adoção de um sistema de recomendação pode auxiliar na adoção desse canal por parte dos consumidores, reduzindo os custos operacionais da rede.
- Finalmente, o uso de terminais de auto-atendimento também pode trazer os clientes às lojas, e aumentar sua participação nos programas de relacionamento. Os sistemas de recomendação serviriam como mais uma funcionalidade que pode ser explorada nesses terminais, fazendo com que a navegação de cada cliente seja uma experiência eminentemente pessoal.

6. REFERÊNCIAS BIBLIOGRÁFICAS

6.1 Bibliográficas e documentais

BELVIN, N.J.; CROFT, W.B., 1992, "Information Filtering and Information Retrieval: two sides of the same coin?", *Communications of the ACM*, v.35, n.12, p. 29.

DIDEROT; D'ALEMBERT, J., *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*, In: <http://encyclopedie.uchicago.edu/>, University of Chicago: ARTFL Encyclopédie Projet, Winter 2008 Edition, Chicago, Illinois, USA.

GIL, A.C., 2002, *Como elaborar projetos de pesquisa*. 4. ed. São Paulo, Atlas.

GOLDBERG, D.; NICHOLS, D.; OKI, B.M.; et al., 1992, "Using collaborative filtering to weave an information Tapestry", *Communications of the ACM*, v.35, n.12, pp. 61-70.

GRUBER, T., 2008, *Encyclopedia of Database Systems*, In: <http://tomgruber.org/writing/ontology-definition-2007.htm>

HAN, E.H.; KARYPIS, G., 2005, "Feature-Based Recommendation System", *Fourteenth International Conference on Information and Knowledge Management - CIKM'05*, Bremen, Germany, Oct 31–Nov 5.

MALONE, Thomas W.; GRANT, Kenneth; TURBAK, Franklin; et al., 1987, "Intelligent Information Sharing Systems", *Communications of the ACM*, v. 30, n. 5, pp. 390-402.

MANOUSELIS, N.; COSTOPOULOU, C., 2007, "Analysis and Classification of Multi-Criteria Recommender Systems" (draft), *World Wide Web: Internet and Web Information Systems*, In: http://e-services.aua.gr/Papers/WWW_SI_NMKK_draft.pdf.

MONTANER, M.; LOPEZ, B.; ROSA, J.L., 2003, "A Taxonomy of Recommender Agents on the Internet", *Artificial Intelligence Review*, v. 19, pp. 285-330.

MONTANER, M., 2003, *Collaborative Recommender Agents Based On Case-Based Reasoning and Trust*, Ph.D. thesis, Universitat de Girona, Girona, Catalonia, Spain.

MOTTA, C.L.R., *Um Ambiente de Recomendação e Filtragem Cooperativas para apoio à Equipes de Trabalho*, Tese, Engenharia de Sistemas e Computação, COPPE, UFRJ, Rio de Janeiro, RJ, Brasil, 1999.

OLIVEIRA, L.G., *Sistema de recomendação de meios de hospedagem baseado em filtragem colaborativa e informações contextuais*, Dissertação, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, UFSC, Florianópolis, SC, Brasil, 2007.

RESNICK, P.; VARIAN, H.R., 1997, "Recommender Systems", *Communications of the ACM*, v.40, n.3, pp. 55-58.

ROY, B., 1996, *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, Norwell, Massachusetts, USA.

SANTAELLA, L., 2001, *Comunicação e Pesquisa: projetos para mestrado e doutorado*, São Paulo, Hacker Editores.

SCHAFER, J.B.; KONSTAN, J.; RIEDL, J., 2001, "E-commerce recommendation applications". *Data Mining and Knowledge Discovery*, v.5, n. 1-2, p. 115-153.

SCHWARTZ, B., *The Paradox of Choice*, In: http://www.ted.com/talks/barry_schwartz_on_the_paradox_of_choice.html.

SHARDANAND, U.; MAES, P., 1995, *Social Information Filtering: Algorithms for Automating Word of Mouth*, In: <http://citeseer.nj.nec.com/shardanand95social.html>, MIT Media-Lab. Cambridge, Massachusetts, USA.

TORRES, R.D., *Combining Collaborative and Content-based Filtering to Recommend Research Papers*, Dissertação, Programa de Pós-Graduação em Computação, UFRGS, Porto Alegre, RS, Brasil, 2004a.

TORRES, R.D., 2004b, "Sistema de Recomendação Híbridos". In: PRATES, R., São Paulo, *Personalização na Internet*, capítulo 7, São Paulo, SP, Novatec.

6.2 Sítios visitados

<http://www.amazon.com>

<http://amiestreet.com>

<http://www.baynote.com>

<http://www.bee.tv>

<http://www.choicestream.com>

<http://www.collarity.com>

<http://www.criteo.com>

<http://www.ilike.com>

<http://www.insuggest.com>

<http://www.jumpernetworks.com/>

<http://www.last.fm>

<http://www.loomia.com>

<http://minekey.com>

<http://www.movielens.umn.edu>

<http://www.netflix.com>

<http://www.pandora.com>

<http://www.peeries.com>

<http://www.photoree.com>

<http://www.pqmedia.com>

<http://www.prudsys.com>

<http://rateyourmusic.com>

<http://www.richrelevance.com>

<http://www.slacker.com>

<http://www.strands.com>

<http://www.stumbleupon.com>

<http://www.stylefeeder.com>