

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**Modelo I2P: Recomendação de Recursos baseando-se em
Preferências, Interesses e Popularidade**

Reginaldo Aparecido Gotardo

**São Carlos
Agosto/2008**

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

G683mi

Gotardo, Reginaldo Aparecido.

Modelo I2P : recomendação de recursos baseando-se em preferências, interesses e popularidade / Reginaldo Aparecido Gotardo. -- São Carlos : UFSCar, 2008.
95 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2008.

1. Sistemas de recomendação. 2. Privacidade e personalização. 3. Sistemas de ensino. I. Título.

CDD: 003.7 (20^a)

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

“Modelo I2P: Recomendação de Recursos baseando-se em Preferências, Interesses e Popularidade”

REGINALDO APARECIDO GOTARDO

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação

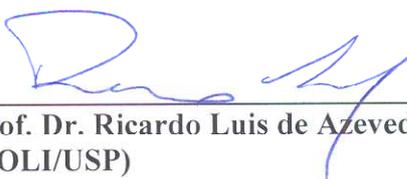
Membros da Banca:



Prof. Dr. Sérgio Donizetti Zorzo
(Orientador - DC/UFSCar)



Prof. Dr. César Augusto Camillo Teixeira
(DC/UFSCar)



Prof. Dr. Ricardo Luis de Azevedo da Rocha
(POLI/USP)

São Carlos
Agosto/2008

Epígrafe

*A tarefa não é contemplar o que ninguém ainda contemplou, mas meditar como
ninguém ainda meditou sobre o que todo mundo tem diante dos olhos.
(Arthur Schopenhauer)*

Agradecimentos

*“Todos juntos somos fortes
Somos flecha e somos arco
Todos nós no mesmo barco
Não há nada pra temer
- ao meu lado há um amigo
Que é preciso proteger
Todos juntos somos fortes
Não há nada pra temer”
(Chico Buarque)*

À Deus pela criação de tudo o que há de mais maravilhoso em nossas vidas e por permitir que eu pudesse concluir mais este trabalho.

Aos amigos e colegas do mestrado, em especial aqueles que me acompanharam na jornada deste curso: Paulinho, Ricardinho e Bruno Kimura. Que nossa troca de experiências tenha trazido tantos benefícios a eles quanto trouxe a mim.

Aos Professores do programa de mestrado e todas as pessoas que atuam no departamento de computação, desde a limpeza até a chefia, todos temos um importante papel a cumprir para que tudo fique em ordem.

Ao meu orientador, professor Dr. Sérgio Zorzo, pela paciência, pela dedicação, pela confiança, pela ajuda nas horas difíceis e pelas risadas nas horas fáceis. Não foi por acaso meu orientador e meu pai terem o mesmo primeiro nome!

Ao meu co-orientador, Dr. Cesar Teixeira, pela ajuda, pelas dicas, pelas correções!

Às instituições CAPES e FAPESP pelo apoio financeiro ao projeto Tidia-Ae o qual este trabalho esteve inserido desde o início.

Aos meus pais e minha família por compreenderem minhas ausências, principalmente nos momentos de conclusão de trabalhos, como este.

À minha noiva e amiga, Érica, pelo auxílio, pelo carinho, pelo amor dedicado.

Resumo

*“É no problema da educação que assenta o grande segredo do aperfeiçoamento da humanidade.”
(Immanuel Kant)*

O desenvolvimento de tecnologias que auxiliem no processo de ensino-aprendizagem é assunto discutido em várias áreas do conhecimento. A grande difusão de Sistemas Educacionais baseados nas tecnologias existentes na Web (também chamados de Sistemas Educacionais baseados na Web – Web-based Educational Systems WbE-S) demonstra a popularização da educação a distância e das ferramentas de suporte a esta. O projeto Tidia-Ae, financiado pela FAPESP visa, sobretudo, o desenvolvimento de um WbE-S que possa explorar os conceitos da internet de alta velocidade. Os WbE-S, comumente, não possuem um tratamento personalizado das ações dos usuários no sistema. Assim, a oferta de recursos de personalização de sistemas visa melhorias no processo de ensino-aprendizagem através do tratamento das necessidades reais e pessoais de cada aluno. A recomendação de conteúdo é uma das possíveis técnicas para oferta de personalização. Trata-se de uma forma não intrusiva de auxiliar o processo de escolha dos usuários num sistema com grande conjunto de informações. Esta técnica foi amplamente explorada e, junto com o projeto Tidia-Ae, serviu como base para a criação do modelo I2P.

Este trabalho define e propõe o modelo I2P baseado em métricas de Interesses, Preferências e Popularidade obtidas no relacionamento entre os usuários e os recursos do sistema. Estas métricas fornecem o embasamento para oferta de recursos adequados às necessidades dos usuários num WbE-S. O cálculo para oferta de recomendação é realizado com a técnica de Filtragem Colaborativa e, assim, a personalização é oferecida de forma colaborativa, considerando o aprendizado em grupo.

Palavras-Chave: Sistemas de Recomendação, Personalização, Sistemas Educacionais baseados na Web

Abstract

"It is in the education problem that the great secret of the humanity improvement is based."

(Immanuel Kant)

The development of technologies that assist in the teach-learning process is an argued subject in some areas of knowledge. The great diffusion of Web-based Educational Systems (WbE-S) has been shown the popularization of distance learning and its support tools. The Tidia-Ae project, support by FAPESP, aim at the development of a WbE-S that can use the concept about high velocity internet. But, the WbE Systems don't have a personal treatment of user's necessities. So, the offers of personalization resources for systems aim at improving the teach-learning process using the treatment of real necessities of each user. The content recommendation, more specifically a recommendation system, is one of several techniques for that and it is a non-intrusive meaning of help user's in a system with a lot of information. This technique was used in Tidia-Ae environment to development of this thesis.

This thesis presents the I2P model based on metrics of Interests, Preferences and Popularity which are acquired by the measuring of the relationship of users and system resources. These metrics provide a form to calculate the recommendation offers of resources. The calculation is done using Collaborative Filtering technique and the personalization is offered in collaborative form, considering the group learning.

Keywords: *Recommendation Systems, Personalization, Web-based Education Systems*

Sumário

1. Introdução	1
1.1 A influência das Tecnologias de Informação no contexto educacional	2
1.2 Objetivos	3
1.3 Motivação	4
1.4 Organização do Trabalho	5
2. Personalização na <i>Web</i>	7
2.1 A Personalização na <i>Web</i>	8
2.2 Hipermídia Adaptativa.....	10
2.3 Sistemas de Recomendação	12
2.4 Sistemas Tutores Inteligentes.....	14
2.5 Outras Classificações para Personalização na <i>Web</i>	16
2.6 Considerações	17
3. Sistemas de Recomendação	18
3.1 Personalização na <i>Web</i> através de Sistemas de Recomendação.....	19
3.2 Estratégias de Recomendação	20
3.2.1 Listas de Recomendação.....	21
3.2.2 Avaliações de usuários	22
3.2.3 Associação por conteúdo	23
3.2.4 Recomendações Pessoais.....	25
3.2.5 Recomendações Off-line: E-mails	26

3.3	Entrada de Dados, Apresentação e Grau de Personalização.....	27
3.4	Métodos de Recomendação - Filtragem de Informações	29
3.4.1	<i>Filtragem Baseada em Conteúdo (FBC)</i>	30
3.4.2	<i>Filtragem Colaborativa (FC)</i>	30
3.5	Descoberta de Conhecimento	31
3.5.1	<i>Uso de Regras de Associação para Filtragem Baseada em Conteúdo</i>	31
3.5.2	<i>A vizinhança próxima e a Filtragem Colaborativa</i>	33
3.6	Considerações	36
4.	E-Learning e EAD	37
4.1	Projeto Instrucional e Medidas de Aprendizagem.....	38
4.2	A Ensino a Distância (EAD) e o Ensino Eletrônico (<i>e-Learning</i>)	39
4.3	O Ambiente de Ensino à Distância Tidia-Sakai	41
4.3	Considerações	43
5.	O Modelo I2P para Recomendação de Conteúdo.....	44
5.1	Modelagem de Usuários e Sistemas de Recomendação	45
5.2.	O Modelo I2P	48
5.3	Preferências dos Usuários	52
5.4	Interesses dos Usuários	52
5.5	Popularidade dos Recursos	55
5.6	Considerações do Capítulo.....	57
6.	Avaliação.....	59
6.1	Comparação com outras abordagens para Personalização na Educação	60
6.2	Estudo de Caso 1 – Erro Médio na Predição de Interesses	61
6.2	Estudo de Caso 2 – Comparação na Predição de Preferências Tradicional e usando valores Implícitos	63
6.3	Estudo de Caso 3 – Formação de Grupos ou Comunidades de Interesse	66
6.3	Estudo de Caso 4 – Estudo da Implementação num Ambiente WbE-S real....	70
6.3.1	<i>Desenvolvimento de Ferramentas para o Tidia-Sakai</i>	72
6.5	Considerações	75
7.	Conclusões	76
7.1	Discussão sobre o Modelo e Resultados.....	77

7.2 Trabalhos futuros	78
Bibliografia.....	79
Apêndice A – Sobre Aprendizagem: Teorias Educacionais	85
A.1 A Educação como Necessidade Social	85
A.2 Teorias Educacionais	87
A.2.1 <i>Inatismo</i>	88
A.2.2 <i>Comportamentalismo</i>	89
A.2.3 <i>Construtivismo</i>	91
A.2.4 <i>Sócio-Construtivismo</i>	92
A.2.5 <i>Comparação entre as Teorias Educacionais</i>	94

Lista de Figuras

Figura 2.1 – Divisão dos Sistemas computacionais.	10
Figura 2.2 – Classificação das Tecnologias adaptativas segundo Brusilovsky.	12
Figura 2.3 – Recomendação em <i>websites</i> de <i>e-commerce</i> : a) classificação geral de determinado produto; b) opinião emitida por usuário ou cliente do site	14
Figura 2.4 – Arquitetura Básica de um STI	16
Figura 3.1 Taxonomia de Classificação	20
Figura 3.2 Exemplo de Listas de Recomendação	21
Figura 3.3 Exemplo de Avaliação dos Usuários	23
Figura 3.4 Exemplo de Avaliação dos Usuários	24
Figura 3.5 Exemplo de Avaliação dos Usuários	24
Figura 3.6 Exemplo de recomendações baseada no perfil do usuário	25
Figura 3.7 O usuário pode acompanhar um item específico	26
Figura 3.7 Exemplo de Avaliação dos Usuários	27
Figura 3.8 Regras de Associação resultantes dos Dados da tabela 3.2	32
Figura 4.1 Tela Inicial do Tidia-Sakai	42
Figura 4.2 Ferramenta DropBox	43
Figura 5.1 Arquitetura I2P	49
Figura 5.2 Formação de Vizinhança e cálculo de Predição	51
Figure 5.3 Possíveis Classificações fornecidas pelo usuário	52
Figura 5.4 Cálculo de Interesse.....	55
Figura 5.5 Recomendação baseada na Popularidade dos Recursos.....	57

Figura 6.1 Comparação de acertos entre a predição ponderada com valores implícitos e a predição tradicional.	65
Figura 6.2 Comparação entre valores altos e baixos para medidas de preferência..	65
Figura 6.3 Arquitetura com a formação de Grupos	66
Figura 6.4. a) Resultados do algoritmo de agrupamento k-means aplicado aos dados de Interesse b) Distribuição dos grupos de acordo com o algoritmo aplicado. ..	68
Figura 6.5) Resultados do algoritmo de agrupamento k-means aplicado aos dados de Interesses e Preferências b) Distribuição dos grupos de acordo com o algoritmo aplicado	69
Figura 6.6 DER do Modelo I2P.....	70
Figura 6.7 Tecnologias Utilizadas pelo Sakai (BASMAN, 2007b).	73
Figura 6.8 Recomendação baseada na Popularidade dos Recursos (BASMAN, 2007b)	74
Figura A.1 Pirâmide das Necessidades segundo Maslow (PISANDELLI, 2003).....	85
Figura A.2 Relacionamento entre as necessidades (PISANDELLI, 2003)	87
Figura A.3 Relacionamento entre o comportamentalismo e o construtivismo (KAMII e DEVRIES, 1980)	95

Lista de Tabelas

Tabela 3.1: Matriz de avaliações dos usuários para itens de consumo.	22
Tabela 3.2: Dados para descoberta de regras de associação	32
Tabela 3.3: Avaliação Usuários X Itens.....	35
Tabela 3.4: Avaliações em comum dos usuários U1 e U2	35
Tabela 3.5: Coeficiente de Person dos usuários em relação ao usuário U2	36
Tabela 5.1 Matriz de Interesses/Preferências - usuários X recursos	50
Tabela 6.1 Matriz de Interesses para o Estudo de Caso apresentado	61
Tabela 6.2 Correlação dos usuários com o usuário-alvo 1	62
Tabela 6.3 Matriz de Preferências para o Caso de Estudo 2	63
Table 6.4 Cálculo da Matriz de Correlação de Interesses	67
Tabela 6.5 Matriz de Correlação de Interesses e Preferências Ponderados por um Tutor – IP(UxU)	69
Tabela A.1: Teorias Educacionais - Baseado em (LEIDNER e JARVENPAA, 1995) apud (MARTINS, 2002).....	94

Siglas e Abreviaturas

STI – Sistemas Tutores Inteligentes

SHA – Sistemas Hipermídia Adaptativa

WbE – Web based Education

WbE-S – Web based Educational Systems

CAI – Computer Aided Instruction

ICAI – Intelligent Computer Aided Instruction

FC – Filtragem Colaborativa

FBC – Filtragem Baseada em Conteúdo

KDD – Knowledge Discovery in Database

CSCW – Computer Supported Cooperative Work

CSCL – Computer Supported Cooperative Learning

LMS – Learning Management Systems

MFU – Most Frequently Used

MRU – Most Recently Used

TAT – Total Access Time

Introdução

*“Conhecimento é poder”
(Sir Francis Bacon)*

A geração, organização, descoberta e utilização do conhecimento pode ser tão valiosa quanto o próprio capital. O tratamento das informações em ambientes comerciais pode resultar em ganhos de competitividade, desempenho e financeiros. O tratamento em ambientes educacionais visa, sobretudo, a melhoria no processo de ensino-aprendizagem. Este capítulo apresenta a descrição do trabalho desenvolvido, abordando, na seção 1.1, a relevância da Tecnologia da Informação para a Educação e o tratamento dos dados pessoais coletados durante o processo de ensino-aprendizagem; na seção 1.2 são apresentados os objetivos deste trabalho; na seção 1.3 são discutidos os fatores que motivam este trabalho e a seção 1.4 apresenta como está organizado o restante do trabalho.

1.1 A influência das Tecnologias de Informação no contexto educacional

Com o advento da globalização e a consecutiva “explosão” da internet o mundo tem passado por profundas mudanças nas áreas econômicas, sociais, políticas e científicas. A invenção de novas tecnologias participa ativamente na promoção dessas mudanças. Dentre as tecnologias criadas, algumas afetam profundamente a educação, como, por exemplo, as redes de comunicação (correios, telefonia, internet) que alteraram a percepção do processo de ensino-aprendizagem individual ou condicionado para um processo coletivo e dinâmico. Redes de relacionamentos, grupos de discussão, comunidades na internet, alteram o modelo regional de socialização e trazem à tona o modelo globalizado. Aprende-se italiano e, rapidamente, têm-se diversos amigos e contatos na Itália sem nunca tê-los visto. Estas mudanças podem afetar não só as relações sociais, mas também relações econômicas, onde a informação detém papel de destaque. (TOFFLER, 2001)

A sociedade de consumo (BAUDRILLARD, 1985) e a produção em série modificaram o que antes era uma relação mais individualizada. As pessoas, muitas vezes, são tratadas como iguais e “categorizados” ou estereotipados, de acordo com regras como classe econômica ou nacionalidade. Antigamente, o artesanato, por exemplo, permitia a venda de itens de consumo exclusivos e o contato produtor-consumidor era mais direto, havendo maior satisfação das necessidades pessoais. Com o avanço das redes de comunicação, a proliferação das redes sociais e as mudanças nas regras de consumo, este tratamento tende a mudar.

Notam-se grandes investimento em sistema de CRM (Customer Relationship Management), em DataWarehouse e em Data Mining que, juntos, visam permitir às empresas mudarem o foco de seus negócios para os clientes. O CRM, também denominado Marketing de Relacionamento, é uma nova ferramenta para negócios usada na expansão e manutenção da carteira de clientes de uma empresa. O Princípio de Pareto¹ (REH, 2002) justifica a importância de se conquistar a fidelidade dos clientes de uma empresa. A personalização busca esta conquista tornando o relacionamento da empresa com o cliente muito mais fácil, mais rápido e

¹ Este princípio afirma que apenas 20% dos clientes geram 80% da receita de uma empresa, ilustrando a importância da conquista da fidelidade dos clientes.

valorizado. O cliente percebe que a empresa foi criada para atendê-lo. Esta abordagem oferece aos clientes serviços agregados aos produtos vendidos pela empresa, na medida que visa atender necessidades especiais do mesmo e isto torna a personalização na Web, principalmente em Comércio Eletrônico, um fator de competitividade.

Além do uso comercial da Personalização, seu uso específico na educação é tema de estudos para sistemas de Ensino a Distância (EAD) e Ensino Eletrônico (*e-learning*). O EAD e o *e-learning* geram condições de acesso ao ensino às pessoas de forma complementar aos meios tradicionais. O advento das redes de computadores e da internet, principalmente, trouxe uma nova forma de comunicação para o ensino não presencial.

Para que seja possível complementar o modelo de ensino presencial ou substituí-lo, oferecendo cursos em que professor e aluno estejam distantes e também cursos auxiliados pela tecnologia, são necessários recursos computacionais que auxiliem na interpretação das dificuldades individuais e coletivas e promovam a melhor estratégia de ensino-aprendizagem. Estes recursos devem não apenas promover o melhor aproveitamento do conteúdo, mas auxiliar na aprendizagem colaborativa, que ganha cada vez mais enfoque na área educacional e empresarial.

Neste contexto aborda-se as principais áreas deste trabalho como o desenvolvimento de Sistemas Educacionais baseados na Web e os Sistemas de Recomendação.

1.2 Objetivos

Neste trabalho são analisados e apresentados vários aspectos relacionados à personalização na *Web*, a importância da personalização na Educação e formas para recomendação de informações que auxiliam os usuários a encontrar informações relevantes. Dentro deste âmbito, apresenta-se um método para tratamento da personalização em Sistemas Educacionais Web.

Este trabalho define e propõe um Método para Recomendação de conteúdo aos usuários de Sistemas Educacionais baseados na Web direcionado a auxiliar o processo de ensino-aprendizagem através de métricas bem definidas sobre o uso dos recursos do sistema.

Este Método pode ser utilizado em diversos Modelos Computacionais e por esta razão foi denominado Modelo I2P: Recomendação de Recursos baseando-se em Preferências, Interesses e Popularidade.

Através de medidas implícitas e explícitas o modelo apresentado visa representar o comportamento do usuário e a análise deste comportamento resulta em recomendações de informações aos usuários. O comportamento do usuário é medido considerando sua interação com os recursos do sistema e esta interação é descrita em métricas bem definidas, apresentadas ao longo deste trabalho.

1.3 Motivação

A motivação para este trabalho reside na sua característica multidisciplinar, abrangendo estudos na área de Ciência da Computação, Informática na Educação e Psicopedagogia, e na possibilidade de contribuição nas áreas científica e social, com os resultados a serem obtidos. Buscando melhora no tratamento das informações pessoais, aliada ao ganho de aprendizagem promovido pelas técnicas de recomendação de informações e conteúdos.

A proposta inicial deste trabalho surgiu no contexto do Projeto Tidia-Ae (05/60653-1). Trata-se de um projeto para o aprendizado eletrônico, financiado pela FAPESP, que visa o desenvolvimento de tecnologias para exploração deste novo paradigma de aprendizagem. Neste interim, o Modelo I2P surgiu inicialmente como proposta de estudo de caso e viabilidade de inserção da personalização de conteúdo num ambiente educacional web.

Ocorreram algumas mudanças ao longo da fase II do projeto Tidia-Ae como a utilização e parceria do Projeto Sakai² e a migração das ferramentas do projeto Tidia-Ae para o Sakai. O Projeto Sakai (*Sakai Project*), que tem entre os seus parceiros tecnológicos a Sun e a Oracle, desenvolveu a plataforma Sakai com o objetivo de criar um sistema de gestão de cursos capaz de concorrer com as ferramentas comerciais existentes. A plataforma Sakai inclui a maioria das funcionalidades existentes num sistema de gestão de aprendizagem como a gestão e distribuição de conteúdos, fóruns de discussão, salas de conversação, upload de trabalhos, e exercícios online. Como plataforma de aprendizagem colaborativa, inclui, também, correio eletrônico, sistema de wikis e RSS.

² Disponível em www.sakaiproject.org

O Modelo I2P foi então desenvolvido considerando-se as principais características do Projeto e ambiente de ensino Tidia-Sakai. O trabalho foi realizado em parceria com o Laboratório de Inovação em Computação e Engenharia da Universidade Federal de São Carlos (Lince³ – UFSCar).

Foram realizados estudos de caso sobre a aplicação do Modelo e um estudo de caso específico que mostra como seria a implementação do método no ambiente Tidia-Sakai.

A construção de um perfil que é representado pelo modelo de usuário buscará compreender as necessidades e preferências dos alunos em relação a utilização da ferramenta educacional. A construção do modelo é motivada por sua característica inerente de ser adaptável a diversas situações e ferramentas.

1.4 Organização do Trabalho

Este trabalho está dividido ao longo de 7 capítulos, descritos da seguinte forma:

- No Capítulo 2, Personalização na Web, são apresentadas as características de sistemas personalizáveis, além de tecnologias para tal fim;
- No Capítulo 3, Sistemas de Recomendação, é descrita uma das tecnologias para personalização de grande abrangência e foco de estudo neste trabalho;
- No Capítulo 4, E-learning e EAD, são apresentados o relacionamento do EAD com a Web e a necessidade de personalização;
- No Capítulo 5, O Modelo I2P para Recomendação de Conteúdo, está dedicado a apresentação do modelo criado;
- No Capítulo 6, Avaliação, tem-se a descrição dos estudos de caso realizados;
- Por fim, no Capítulo 7, tem-se a conclusão do trabalho.

³ <http://lince.dc.ufscar.br>

Além disto, há ainda o apêndice A com maiores informações sobre os temas educacionais envolvidos no trabalho.

Personalização na Web

*"Nossa personalidade é uma invenção dos outros."
(Marcel Proust)*

As possibilidades geradas com tecnologias que fornecem a aproximação dos recursos oferecidos aos interesses dos usuários são inúmeras. É possível melhorar a exploração econômica disto, visando por exemplo aumentar as vendas num site de comércio eletrônico. Também possibilita a distinção de características de aprendizagem entre grupos de alunos, dentre outros. As seções a seguir apresentam como a personalização é explorada e como é classificada. A seção 2.1 apresenta a personalização na Web; a seção 2.2 a Hipermedia Adaptativa. A seção 2.3 apresenta os Sistemas de Recomendação. A seção 2.4 apresenta os Sistemas Tutores Inteligentes. A seção 2.5 apresenta outras classificações para sistemas de recomendação. A seção 2.6 apresenta as considerações deste capítulo.

2.1 A Personalização na *Web*

A personalização de sistemas tem sido amplamente utilizada em diversas aplicações como em sistemas de informação, *help desks*, sistemas de recuperação de informação e sistemas educativos (DE-BRA *et al.*, 2000).

De maneira geral, a personalização de sistemas ocorre em duas etapas⁴:

1. A interação do usuário no sistema é monitorada. As informações coletadas são utilizadas para manutenção de modelos de usuários.
2. A apresentação dos conteúdos do sistema ao usuário pode ser modificada, sugerindo para o usuário qual o próximo caminho ou quais opções de caminhos tem o usuário.

O conteúdo em sistemas personalizados é direcionado aos interesses do usuário e/ou aos interesses do domínio em questão (pode ser uma empresa, com interesses comerciais, ou uma instituição de ensino com interesses pedagógicos)⁵. Por isso, a chance do usuário conseguir encontrar as informações mais relevantes é maior. O sistema personalizado também pode inferir relacionamentos entre produtos (ou assuntos) e sugerir novas abordagens ao usuário

Existem diversas abordagens dadas à personalização que, de forma positiva, contribuem para as pesquisas nesta área de serviços e de ambientes *Web*. Por exemplo, um atendente seja pessoal ou virtual pode conhecer o usuário pelo seu nome e, de acordo com seus padrões de consumo, pode recomendar produtos e serviços em algumas situações corriqueira, como ir ao supermercado. Situações como esta, comuns no passado, que foram descaracterizadas pelo processo de “massificação⁶”, estão sendo novamente retomadas e tratadas com o surgimento de novas tecnologias e áreas de estudo como a Hipermídia Adaptativa e a Interface Humano Computador (IHC).

Os sistemas *Web* tradicionais que apresentam conteúdos e ligações estáticas oferecem o mesmo conteúdo e conjunto de nós para todos os usuários, sem considerar características individuais (inerentes a cada usuário). Caminhos de

⁴ Alguns autores como Kobsa preferem a divisão em três etapas, separando a segunda etapa em: modificação da apresentação de *links* e modificação na apresentação de conteúdos (KOBASA, KOENEMANN e POHL, 2001)

⁵ Neste caso o usuário também pode ser chamado de cliente ou aluno.

⁶ Processo de “massificação” é considerado como a mudança acontecida na sociedade onde a produção em série é fator predominante, gerando produtos iguais para pessoas diferentes e enquadrando os consumidores em categorias criadas pelos “produtores”.

navegação (*hiperlinks*) que não foram previamente criados (pelo desenvolvedor do *website*), também não podem ser utilizados e podem comprometer a entendimento do conteúdo pelos usuários. Esta situação problema reafirma a questão da massificação, a não observação de preferências individuais.

Representar a informação relacionando-a ao contexto de navegação, às preferências do usuário ou a adequações necessárias do próprio site (como por exemplo, o uso em filtragem colaborativa) apresentam ganhos significativos tanto aos usuários, pois agilizam o processo de navegação, quanto aos sites, que tornam a comercialização de produtos e serviços mais rápida e viável. Auxiliam, também, nos serviços de ensino, quando o sistema tratar-se de um Sistema Tutor Inteligente (BRUSILOVSKY, 2001) (KOBASA, 1994).

A personalização de sistemas possui diversas outras nomenclaturas que alteram algumas especificidades a respeito das técnicas e resultados. Entretanto, os objetivos restringem-se a oferecer informações para as pessoas que realmente necessitam delas (BELVIN e CROFT, 1992). Algumas das denominações são Sistemas Hiperídia Adaptativa (SHA) (BRUSILOVSKY, 2001), Sistemas de Recomendação (RESNICK e VARIAN, 1997), ou Sistemas Tutores Inteligentes. Estas são denominações de sistemas adaptativos, onde há algum mecanismo transparente ao usuário que realiza a adaptação de conteúdo, de *links* ou sugestões. Existem também os sistemas adaptáveis, onde o poder de personalização é controlado pelo próprio usuário. Sistemas adaptáveis e adaptativos são tipos de sistemas interativos, ou seja, sistemas que lidam com as respostas dos usuários. Os sistemas hiperídia tradicionais, apesar de permitirem a navegação não linear, são sistemas não interativos. Estas definições podem ser combinadas e estão representadas na figura 2.1.

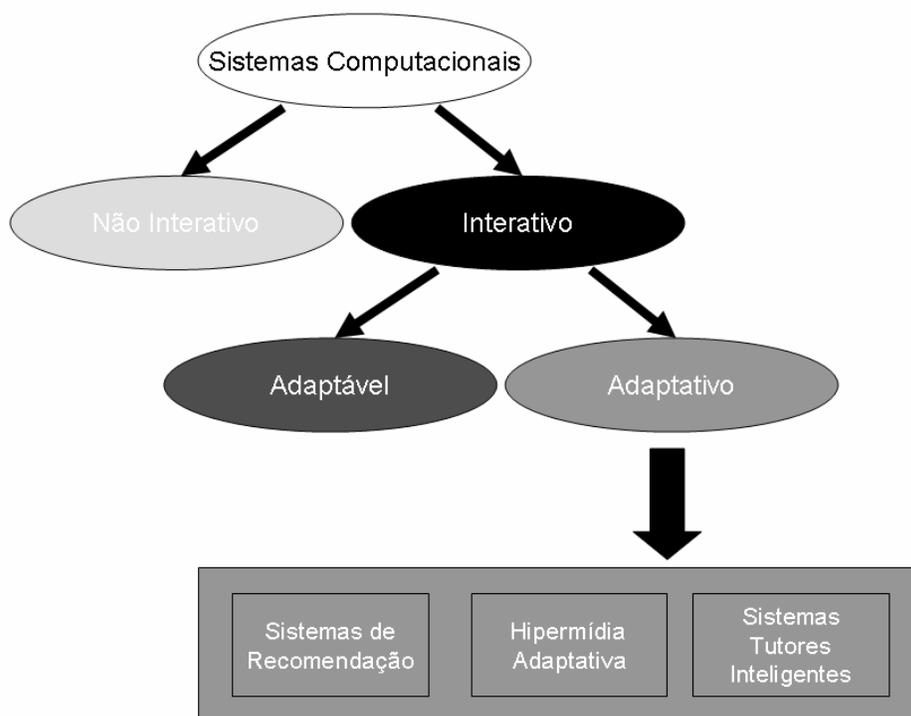


Figura 2.1 – Divisão dos Sistemas computacionais.⁷

Uma denominação mais abrangente é proposta por Kobsa (1994) a chamada Customização. Trata-se da observância do modelo de usuário e também do contexto onde está inserida e informação (KOBASA, 1994).

Além das técnicas citadas, pode-se classificar a personalização quanto ao local onde ocorre a personalização, quanto ao usuário estar *on-line* ou *off-line* (MURUGESAN e RAMANATHAN, 2001), ou quanto ao domínio onde é aplicada (SAE-TANG e ESICHAIKUL, 2001).

2.2 Hipermedia Adaptativa

A área de Hipermedia Adaptativa relaciona-se ao estudo de sistemas hipermedia onde o conteúdo possa alterar-se dinamicamente, de acordo com alguma função pré-estabelecida. Na concepção de sistemas hipermedia, apesar das ligações

⁷ Adaptada de (ARAGÃO, 2004)

(os chamados *links*) entre conteúdos (também chamados de nós) e da possibilidade do usuário percorrer diversos caminhos, não há alteração destas ligações ou da forma como são apresentadas ao usuário.

Existem duas linhas principais de pesquisa para SHA: a navegação adaptativa e a apresentação adaptativa. A navegação adaptativa visa auxiliar o usuário a encontrar seu caminho “ideal” num sistema hipermídia, modificando a maneira como os “*links*” estão organizados, considerando-se características particulares deste usuário. Já apresentação adaptativa é utilizada para modificar a apresentação do conteúdo de uma página acessada por um determinado usuário, considerando-se, novamente, algumas características particulares deste usuário. A figura 2.2 ilustra esta divisão e suas subdivisões, proposta por Brusilovsky. (BRUSILOVSKY, 2001)

A navegação adaptativa é subdividida de acordo com a maneira de utilização dos *links*. Essa subdivisão dá-se em: Orientação Direta, Ordenação Adaptativa de *links*, Ocultação Adaptativa de *links*, Anotação Adaptativa de *links*, Geração Adaptativa de *links* e Adaptação de Mapa. Na orientação direta o objetivo é destacar visualmente o melhor nó ou apresentar um *link* dinâmico adicional que conduzirá o usuário ao “melhor nó”⁸. A ordenação classifica os conteúdos e *links* de uma página e quanto mais próximo do melhor conteúdo está o *link*, mais relevante este *link* será. A ocultação consiste em restringir o espaço de navegação do usuário, retirando *links* para conteúdos menos relevantes, podendo ocultar, desabilitar ou remover tais *links*. A anotação adicional aos *links* mais relevantes insere algum tipo de informação (seja de forma textual ou com indicações visuais) que relacionam os próximos nós (menos relevantes) aos atuais. Por fim, a adaptação dos mapas compreende as possibilidades de apresentação dos mapas hipermídia (percursos possíveis no sistema web) aos usuários. Menus de classificação de informações são exemplos de mapas.

A apresentação adaptativa considera várias modalidades de conteúdo hipermídia, porém o conteúdo mais expressivo e estudado é o textual. Observa-se pela figura 2.2 que a adaptação textual divide-se na Adaptação de Linguagem Natural e na Adaptação de fragmentos de texto. Nestes é exigido o apoio de outra

⁸ O “melhor nó” é aquele que, dada uma função classificadora de conteúdos de acordo com o perfil do usuário, resulta no melhor (ou maior) valor de classificação.

área de pesquisa em computação: Processamento de Linguagem Natural (NLP – do inglês *Natural Language Processing*).

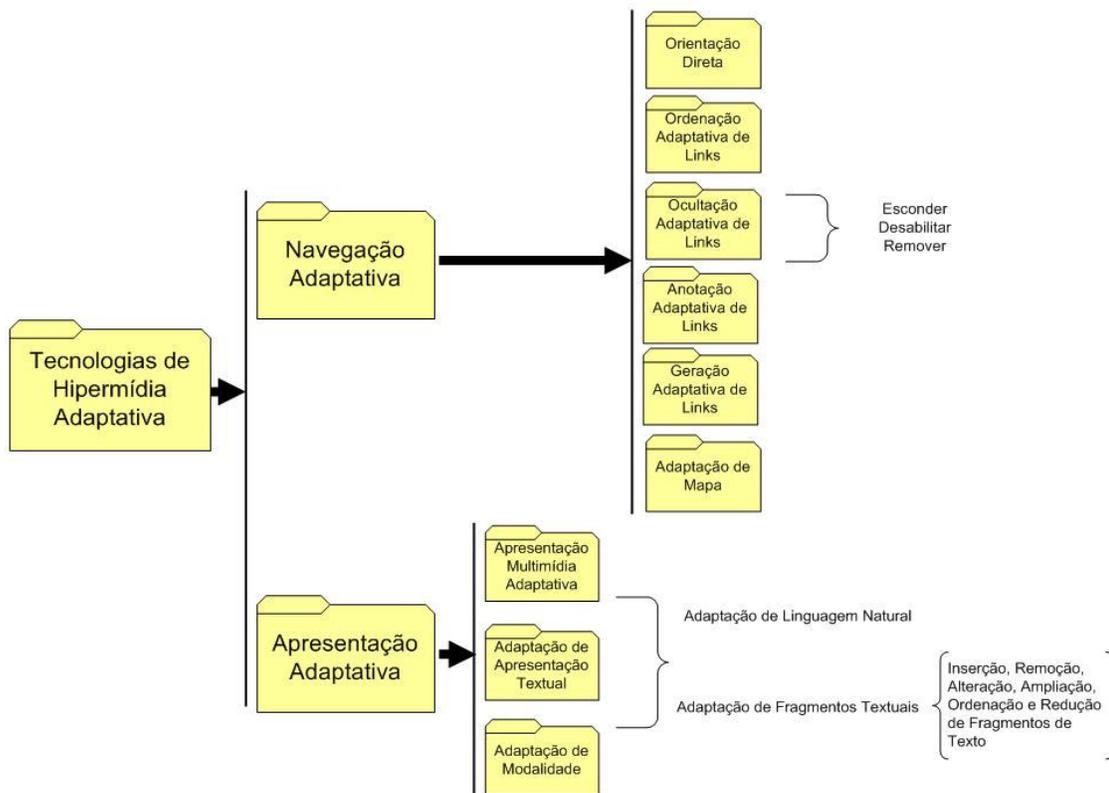


Figura 2.2 – Classificação das Tecnologias adaptativas segundo Brusilovsky.⁹

Nota-se, pelo exposto, que o grande foco da hipermedia adaptativa é a forma de guiar o usuário através do hiperespaço, considerando suas preferências ou características pessoais.

2.3 Sistemas de Recomendação

As recomendações recebidas por outras pessoas são formas de abreviação de escolhas. Uma carta de recomendação é uma forma de apresentar um candidato a algum cargo. Para minimizar as dúvidas e necessidades que temos frente à escolha entre alternativas, geralmente confiamos nas recomendações que são passadas por outras pessoas, as quais podem chegar de forma direta (*word of mouth*) (SHARDANAND e MAES, 1995), cartas de recomendação, opiniões de

⁹ Figura baseada em (BRUSILOVSKY, 2001), com tradução independente.

revisores de filmes e livros, impressos de jornais, entre outros. Os sistemas de recomendação auxiliam no aumento da capacidade e eficácia deste processo de indicação já bastante conhecida na relação social entre seres humanos (RESNICK e VARIAN, 1997).

Os sistemas de recomendação fornecem um tipo personalização mais indireta, pois alteram o ambiente onde o usuário está inserido, mas não o obriga a seguir um único caminho, sugerindo alterações que, não necessariamente, serão aceitas pelo usuário. A utilização de sistemas de recomendação requer estudos em diversas áreas como a Interface Humano-Computador (IHC), devido à composição da interface distinta diferente a cada usuário, e como a Inteligência Artificial – que auxilia na descoberta de informações no sistema.

O primeiro sistema de recomendação foi denominado Tapestry e trouxe com sua criação a expressão “filtragem colaborativa”, designando uma aplicação específica dos sistemas de recomendação. Neste tipo de sistema há o auxílio humano explícito no fornecimento de informações, gerando, mesmo indiretamente, a criação de grupos de interesses (REATEGUI, CAZELLA e OSÓRIO, 2006). Diversos *websites* comerciais utilizam tal técnica com muito sucesso. Este é o caso do site brasileiro Submarino. Na figura 2.3, pode-se observar a classificação de um determinado item baseado nas votações dos usuários que o adquiriram.

O princípio do algoritmo da filtragem colaborativa considera que o usuário ativo possui maior probabilidade de se interessar por itens que usuários semelhantes preferem ou preferiram. Para isto, calcula-se um grau de similaridade entre o usuário ativo (alvo) e os outros usuários. Os itens com maior grau de similaridade são recomendados ao usuário alvo. (GOLDBERG *et al.*, 1992) (RESNICK e VARIAN, 1997)

O foco dos sistemas de recomendação são os *websites* de *e-commerce* (comércio eletrônico). A justificativa do emprego de tais sistemas é o aumento da lucratividade, como já dito, através da melhor adequação de consumo dos clientes. Esta adequação é realizada ofertando-se ao cliente os produtos mais recomendados ao seu perfil.



Figura 2.3 – Recomendação em *websites* de *e-commerce*: a) classificação geral de determinado produto; b) opinião emitida por usuário ou cliente do site.¹⁰

As técnicas de personalização utilizadas são as mais diversas. Algumas trazem ao usuário “ofertas casadas”, “itens preferenciais”, “os mais vendidos” em determinadas categorias, os que possuem melhores classificações (como visto na figura 2.3), entre outros.

O capítulo 4 tratará, exclusivamente, dos sistemas de recomendação, técnicas empregadas, formas de classificação dos mesmos, entre outras informações importantes.

2.4 Sistemas Tutores Inteligentes

¹⁰ Neste exemplo, o site avaliado é o Submarino, um site brasileiro de *e-commerce*. Seu endereço eletrônico é <http://www.submarino.com.br>

Os Sistemas Tutores Inteligentes são programas de computador que utilizam técnicas de inteligência artificial para auxílio à aprendizagem. Trata-se de uma derivação dos programas de Instrução Assistida por Computador (CAI – *Computer Aided Instruction*). Os programas CAI foram a primeira modalidade de software educacional e utilizavam como teoria de aprendizagem o comportamentalismo. Não havia, no entanto, mecanismo inteligente e adaptável nestes sistemas. A partir dos programas de Instrução Assistida por Computador Inteligente (ICAI – *Intelligent Computer Aided Instruction*) foi iniciada a inclusão de mecanismos inteligentes para melhoria no processo de aprendizagem.

Uma diferença determinante entre os STI's e os ICAI's é a existência de uma base de conhecimento e não apenas uma base de dados convencional. Dentro desta perspectiva, o processo de aprendizagem pode ser caracterizado como o mapeamento do conhecimento a ser ensinado – modelado segundo as informações do professor/tutor - para a estrutura de conhecimento do estudante – o modelo do estudante.

Um STI é, mais precisamente, um sistema computacional onde o estudante é tutorado num dado domínio (por exemplo, na disciplina de computação ou matemática). Há a modelagem do estudante, do seu conhecimento e entendimento sobre determinado tópico. À medida que o estudante interage com o sistema, realizando tarefas, seu modelo de entendimento é modificado. O STI compara o entendimento novo com o que seria ideal (relacionando ao modelo do especialista ou tutor) e auxilia o estudante na compreensão para alcançar a aprendizagem ideal, constante no modelo do especialista/tutor. A forma como o STI auxilia o estudante é descrita no modelo pedagógico. Assim, a arquitetura de um STI é definida na figura 2.4. (GIRAFFA, 1999; JAQUES, 1999; SELF, 1999; ROSATELLI e TEDESCO, 2003)

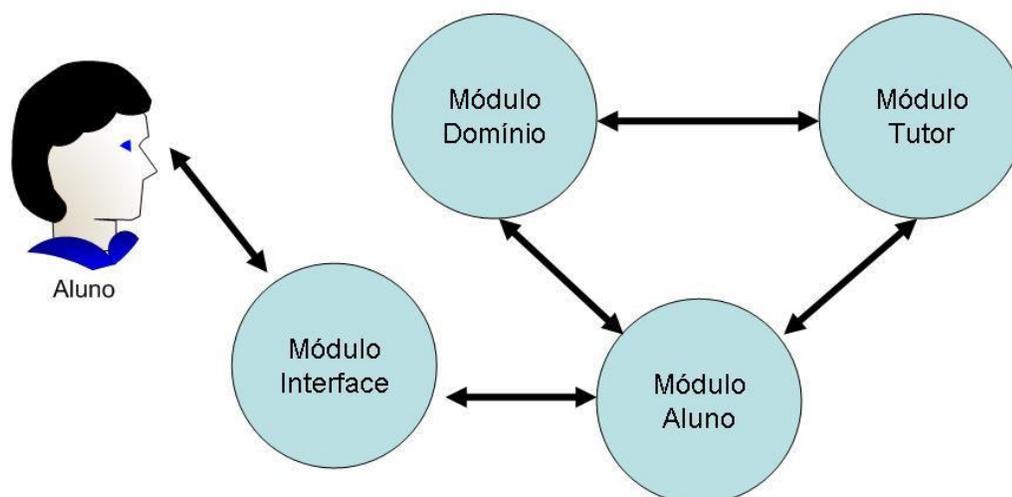


Figura 2.4 – Arquitetura Básica de um STI

2.5 Outras Classificações para Personalização na Web

Uma forma de caracterizar as técnicas de personalização é com relação ao local onde ocorre o processamento dos dados. Esta classificação é dividida em Lado do Cliente (*client-side*) e Lado do Servidor (*server-side*). No lado do cliente são dadas ferramentas ao usuário para que este possa adaptar a interface do sistema. Neste caso, exige-se um usuário mais experiente. Além disto, em aplicações *Web* pode haver sobrecarga na rede, pois é necessário fazer o *download* do sistema. Já no lado do servidor, o processamento para personalização na apresentação do conteúdo é transparente ao usuário. (MURUGESAN e RAMANATHAN, 2001)

A classificação proposta por (MURUGESAN e RAMANATHAN, 2001) assemelha-se à vista anteriormente, adaptada de (ARAGÃO, 2004), pois nos sistemas adaptáveis a personalização é feita pelo próprio usuário (lado cliente), enquanto que em sistemas adaptativos a personalização é automática (ou no lado servidor).

Também é possível separar as técnicas em *online* e *offline*. Através da divisão das técnicas de personalização em “profunda” e “superficial” os autores caracterizam técnicas mais simples (pois trabalham com padrões estáticos e pré-definidos) e técnicas mais complexas (que lidam com fatores dinâmicos). Nas técnicas superficiais são exemplos os e-mails com propaganda ou informações, ou

mesmo notícias enviadas aos usuários após a análise de sua navegação. Já nas técnicas profundas, são realizadas análises em tempo de navegação do usuário e a adaptação é dinâmica, requerendo maior “esforço” do sistema (MURUGESAN e RAMANATHAN, 2001).

Existe ainda outra classificação, considerando o escopo da divisão das tecnologias: direcionadas ao cliente – maior relevância para o perfil e padrões de navegação do usuário – e direcionadas ao negócio – maior relevância ao conteúdo, visando modelar e implementar os processos de negócio específicos do domínio. Esta classificação mais voltada ao comércio eletrônico (*e-commerce*) (SAE-TANG e ESICHAIKUL, 2001). A partir desta abordagem deve-se estabelecer a melhor utilização das tecnologias disponíveis. Ao tratar-se de elementos *Web* que podem ser personalizados têm-se a seguinte divisão (GODERIS *et al.*, 2001):

- Ligações ou *Links*: os destinos dos *links* são calculados de acordo com as preferências do usuário.
- Estrutura: Pode-se modificar a aparência do *website* como cores, posições de menus, disposição de informações, etc.
- Conteúdo: as informações disponíveis são alteradas de acordo com o usuário. Por exemplo, os níveis de acesso a informações, ou o próprio conteúdo do *website*, utilizando-se Processamento de Linguagem Natural (PLN).
- Comportamento: os *links* são combinados com outras funções, aparentando a modificação da navegação para o usuário.

2.6 Considerações

Das técnicas apresentadas neste capítulo os Sistemas de Recomendação, que serão descritos no capítulo seguinte, possuem larga aplicação comercial, mas não muitas aplicações educacionais. Por isto, este trabalho propõe a utilização de um sistema de recomendação com adaptações feitas para atender necessidades específicas em sistemas educacionais. Tudo isto estará descrito nos capítulos seguintes.

Sistemas de Recomendação

“To recommend thrift to the poor is both grotesque and insulting. It is like advising a man who is starving to eat less.”¹¹
(Oscar Wilde)

Uma das formas de personalização é através de Recomendações ao usuário ou ao grupo de usuários de um sistema. Para que seja possível recomendar-se itens, produtos ou conteúdo a um usuário são necessárias diversas técnicas e estratégias para coleta, representação e apresentação de informações e conteúdos. Além disso, a recomendação está intrinsecamente relacionada com áreas como Inteligência Artificial e Descoberta de Conhecimento em Bases de Dados. Na seção 3.1 apresentam-se questões sobre a personalização através de sistemas de recomendação e uma classificação que será descrita nas seções seguintes; a seção 3.2 apresenta estratégias para a recomendação; a seção 3.3 apresenta questões sobre a entrada de dados, apresentação das recomendações e o grau de personalização de um sistema de recomendação; a seção 3.4 descreve métodos para filtragem de informações; a seção 3.5 associa os métodos a processos de descoberta de conhecimento e apresenta exemplos de aplicação; por fim, na seção 3.6 são apresentadas as considerações finais deste capítulo.

¹¹ “Recomendar economia para o pobre é, ao mesmo tempo, grotesco e insultuoso. É o mesmo que aconselhar um faminto a comer menos”.

3.1 Personalização na *Web* através de Sistemas de Recomendação

Como visto no capítulo 2, figura 2.1, os sistemas de recomendação são uma das formas para personalização na *Web*. O principal objetivo é conduzir o usuário, através de sugestões, conselhos e advertências, a atingir seus objetivos. Isto difere da hipermídia adaptativa, que modifica a estrutura dos “*links*” e, consecutivamente do sistema, para adaptar o ambiente ao usuário. Difere também dos Sistemas Tutores Inteligentes, que possuem abordagem e interesse educacionais, movendo-se pela estrutura de aprendizagem dos usuários e não apenas pelos seus interesses.

Os sistemas de recomendação possuem larga aplicação comercial. Os exemplos utilizados neste capítulo ilustrarão isto. Porém, possuem características que podem ser aproveitadas no processo de ensino-aprendizagem e justificadas por teorias educacionais. Por isto, este capítulo apresentará uma revisão sobre os sistemas de recomendação, formas de classificação e exemplos de uso, e por fim, a forma de utilização destes recursos na proposta deste trabalho.

Através da classificação de sistemas de recomendação representada no trabalho de Venson (VENSON, 2002), que baseia-se no trabalho de Schafer, Konstan e Riedl (SCHAFER, KONSTAN e RIEDL, 2001), serão apresentados os principais componentes observados nos sistemas de recomendação (figura 3.1). Além destes dois trabalhos também foi utilizado para complementar informações na classificação o trabalho de Reategui, Cazella e Osório (REATEGUI, CAZELLA e OSÓRIO, 2006).

De acordo com esta classificação, os sistemas de recomendação ficam divididos em 4 componentes:

1. A entrada de dados do usuário alvo e da comunidade;
2. O método de recomendação (descoberta ou associação de informações) utilizado;
3. A estratégia para apresentação das recomendações;
4. Como será realizada a apresentação de recomendações e em que grau;

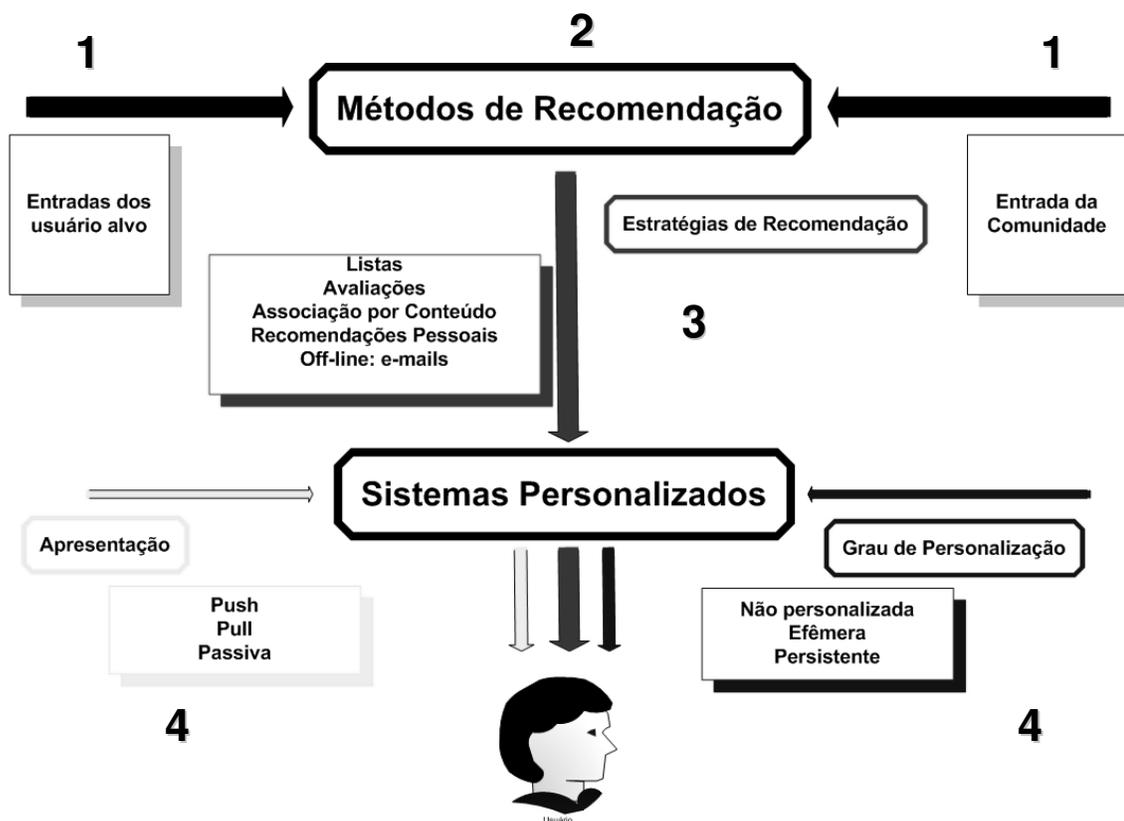


Figura 3.1 Taxonomia de Classificação¹²

Como pode ser observado na figura 3.1, as entradas do usuário alvo e da comunidade são armazenadas e tratadas segundo um método de recomendação. Após isto são definidas as estratégias de apresentação da recomendação. Necessita-se também a definição do grau de personalização adotado pelo sistema e de como o resultado será apresentado ao usuário. Nas seções seguintes serão descritas cada etapa desta taxonomia.

3.2 Estratégias de Recomendação

Tratam-se das formas de apresentação das recomendações aos usuários, estipuladas de acordo com o contexto do sistema.

¹² Baseada em (SCHAFER, KONSTAN e RIEDL, 2001; VENSON, 2002)

3.2.1 Listas de Recomendação

As listas de recomendação são agrupamentos de itens mais populares, não havendo necessidade de análise mais profunda para a criação destas. Esta estratégia consiste em manter listas de itens organizados por tipos de interesses.

Este tipo de abordagem estratégica é de implementação trivial, pois as listas são organizadas e mantidas de acordo com os interesses do domínio em questão. Caso seja um site de vendas de livros (comércio eletrônico) o interesse do domínio é o lucro e, assim, as listas podem ser organizadas de acordo com os itens mais vendidos numa categoria, por exemplo. Se o contexto for educacional, pode-se manter uma lista dos assuntos mais relevantes para o grupo de estudo de uma determinada disciplina. Um problema deste tipo de estratégia é que as listas representam os interesses do domínio e não do usuário.

The screenshot shows the website interface for 'livrariacultura.com.br'. The main content area is titled 'Mais Vendidos' and is filtered for 'Ficção' from '19/02 a 25/02 de 2007'. It lists three books:

Nacionais	
Ficção	Não-Ficção
Administração	Esotéricos/Auto-ajuda
	Informática
	Infanto-Juvenis

Ficção 19/02 a 25/02 de 2007	
1.º	<p>CAÇADOR DE PIPAS, O HOSSEINI, KHALED</p> <p>'O caçador de pipas' conta a história de Amir, um afegão há muito imigrado para os Estados Unidos, que se vê obrigado a acertar as contas com o passado e retorna a seu país de origem. O ponto de partida do livro é a infância do protagonista, quando Cabul ainda não era a capital do país que foi invadido pela União Soviética, dominado pelos talibãs e subjugado pelos Estados Unidos.</p> <p>Preço = R\$ 39,90 Mais Cultura = R\$ 33,92</p> <p>comprar</p> <p>Veja mais detalhes</p>
2.º	<p>CALVIN E HAROLDO - MUNDO E MAGICO, O WATTERSON, BILL</p> <p>Criada em 1985, a tirinha 'Calvin & Haroldo' foi publicada diariamente, durante dez anos, em mais de 2.400 jornais ao redor do mundo. Os álbuns publicados por Bill Watterson, criador da dupla, venderam mais de 30 milhões de cópias. A tirinha conta a história de Calvin, um garoto hiperativo de seis anos cujo maior amigo é o tigre de pelúcia Haroldo - que ganha vida quando não existe nenhum adulto por perto. Ao lado das fantasias e brincadeiras da dupla, surgem questões sobre política, cultura, soc...</p> <p>Preço = R\$ 44,90 Mais Cultura = R\$ 35,92</p> <p>comprar</p> <p>Veja mais detalhes</p>
3.º	<p>VIAGENS NO SCRIPTORIUM AUSTER, PAUL</p> <p>Fechado num pequeno quarto e vigiado o tempo todo por câmeras e microfones, um homem de certa idade busca reconstituir sua memória. Ele não sabe onde</p>

Figura 3.2 Exemplo de Listas de Recomendação

A figura 3.2 apresenta um exemplo de lista de recomendações da Livraria Cultura, que resulta nos livros mais vendidos na categoria ficção no período de 19 a 25 de fevereiro de 2007. Neste tipo de abordagem também estão inclusas a chamada avaliação Top N, onde são demonstrados os mais vendidos ou mais interessantes numa categoria. A métrica, neste caso, é apenas a relação de itens vendidos no total de vendas.

3.2.2. Avaliações de usuários

As avaliações dos usuários consistem na classificação do usuário dada a um determinado item. Num site de *e-commerce*, esta classificação pode ser após a compra de um produto. Num site educacional, pode ser a avaliação de um conteúdo apresentado.

Este tipo de estratégia fornece aos futuros usuários a opinião dos anteriores que tiveram algum tipo de relacionamento com o item em questão. Isto assegura a novos consumidores (ou alunos) a melhoria no processo de escolha, visto que pode ser comparado com escolhas anteriores. A avaliação dada pelos usuários precisa ser verídica para garantir a “honestidade” da recomendação.

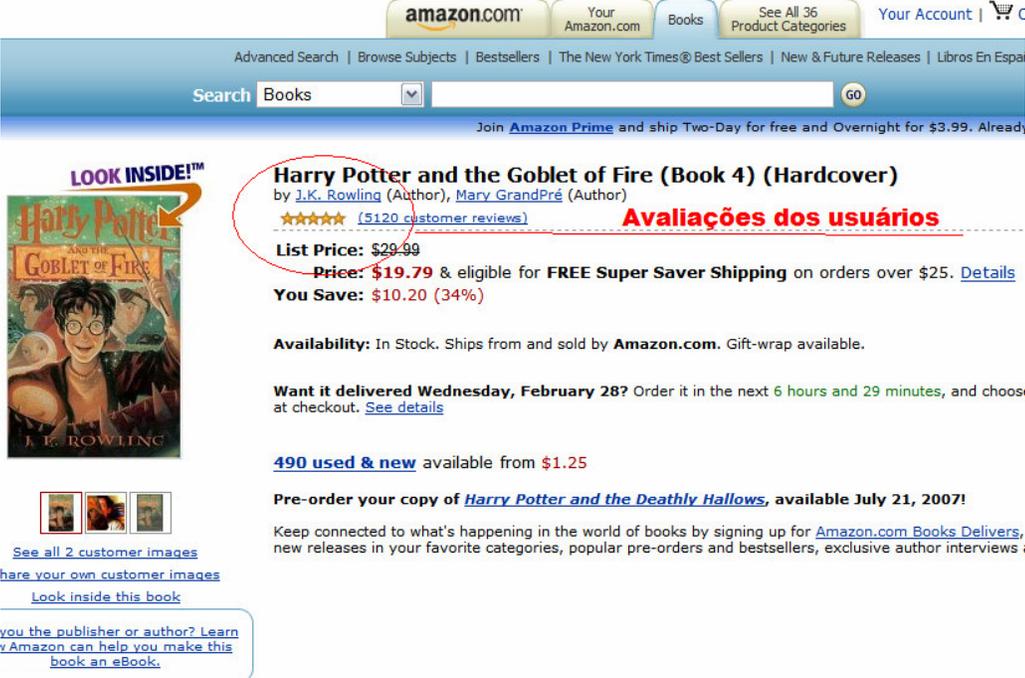
As avaliações dos M usuários para os N itens avaliados ou de consumo em que cada elemento da matriz contém o valor atribuído, denotando a avaliação feita, como por exemplo, a nota concedida, o grau de importância considerado, etc, podem ser armazenadas numa matriz $M \times N$, como a tabela 3.1 a seguir.

Tabela 3.1: Matriz de avaliações dos usuários para itens de consumo.

	Item1	Item2	...	ItemN
Usuário1	1	5	...	3
Usuário 2	Muito Bom	Bom	...	Médio
...
...
UsuárioM-1	Muito Bom	Bom	...	Médio
UsuárioM	Muito Difícil	Difícil	...	Muito Fácil
	Valor Atribuído			

O armazenamento destas informações necessita de mecanismos simples quanto à implementação. Posteriormente, com a aplicação de técnicas para descoberta de conhecimento é possível a inferência para que recomendações possam ser geradas e apresentadas aos usuários.

A figura 3.3 apresenta um exemplo de classificação de acordo com as avaliações dos usuários que compraram o livro Harry Potter no site da Amazon.com.



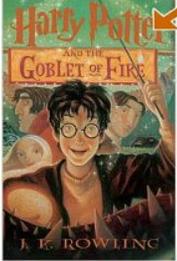
amazon.com Your Amazon.com Books See All 36 Product Categories Your Account | 

Advanced Search | Browse Subjects | Bestsellers | The New York Times® Best Sellers | New & Future Releases | Libros En Español

Search Books

Join [Amazon Prime](#) and ship Two-Day for free and Overnight for \$3.99. Already

LOOK INSIDE!™



Harry Potter and the Goblet of Fire (Book 4) (Hardcover)
by [J.K. Rowling](#) (Author), [Mary GrandPré](#) (Author)

★★★★ (5120 customer reviews) **Avaliações dos usuários**

List Price: ~~\$29.99~~
Price: **\$19.79** & eligible for **FREE Super Saver Shipping** on orders over \$25. [Details](#)
You Save: **\$10.20 (34%)**

Availability: In Stock. Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it delivered Wednesday, February 28? Order it in the next **6 hours and 29 minutes**, and choose at checkout. [See details](#)

490 used & new available from **\$1.25**

Pre-order your copy of *Harry Potter and the Deathly Hallows*, available July 21, 2007!

Keep connected to what's happening in the world of books by signing up for [Amazon.com Books Delivers](#), new releases in your favorite categories, popular pre-orders and bestsellers, exclusive author interviews

[See all 2 customer images](#)
[Share your own customer images](#)
[Look inside this book](#)

[Do you the publisher or author? Learn how Amazon can help you make this book an eBook.](#)

Figura 3.3 Exemplo de Avaliação dos Usuários

3.2.3 Associação por conteúdo

Neste tipo de recomendação considera-se o conteúdo dos itens, como por exemplo, autoria do conteúdo, contexto ou, no caso de um sistema de vendas de livros, quais os itens que possuem semelhança com o comprado, visitado ou apresentado.

Diversas técnicas podem ser usadas para esta recomendação. Podem ser inseridas, manualmente, informações como *tags* que, posteriormente permitirão a identificação da associação. A associação por conteúdo também pode ser realizada manualmente, mas trata-se de um serviço oneroso.

A figura 3.4 apresenta um exemplo deste tipo de estratégia na forma: “quem comprou este item X também comprou o item Y”, utilizada pelo site da editora Campus Elsevier.

The screenshot shows the Elsevier website interface. At the top, there is a navigation bar with a search box labeled 'Busca Avançada', a shopping cart icon showing 'Total: R\$ 0.00', and a 'Ver Sacola' button. Below the navigation bar, the main content area is titled 'Série Concursos' and 'Serie Concursos > Direito'. A sidebar on the left contains various user options like 'Livros Importados Elsevier', 'Seja um Autor', and 'Blogs dos Autores'. The main content features a book cover for 'Direito Tributário Vol. III' by Claudio Borba. To the right of the cover, the book's details are listed: 'Direito Tributário Vol. III Livro Segundo do Código Tributário Nacional: Título I, Legislação Tributária, e Título II, Obrigação Tributária'. Below the book details, there is a price tag of 'R\$ 45,90' and a 'Comprar' button. A 'Compartilhar' button is also visible. At the bottom of the page, there is a 'LEIA TAMBÉM' section with several related book titles like 'Direito Tributário Vol. II', 'Direito Processual Penal - Vol I', etc.

Figura 3.4 Exemplo de Avaliação dos Usuários

Já a figura 3.5 apresenta um exemplo de associação por conteúdo na forma de “Aproveite Também”, realizada pelo site Submarino.

The screenshot shows the Submarino website interface. The main content area features a book cover for 'Cavalo de Tróia 8' by J. J. Benitez. To the right of the cover, the book's details are listed: 'Cavalo de Tróia 8: Jordão J.J. BENITEZ'. Below the book details, there is a price tag of 'R\$ 31,90' and a 'COMPRE' button. A 'Sinopse' section is also visible. At the bottom of the page, there is an 'Aproveite Também' section with several related book titles like 'Cavalo de Tróia 8: Jordão - J.J. BENITEZ + Testamento de São João, O - J.J. BENITEZ'. The sidebar on the left contains various user options like 'Destacados', 'Livros Importados', and 'Literatura Estrangeira'. The top navigation bar includes a search box and a shopping cart icon.

Figura 3.5 Exemplo de Avaliação dos Usuários

3.2.4 Recomendações Pessoais

Através da análise detalhada das ações do usuário pode-se traçar um perfil e representá-lo num modelo de usuário (MU), que servirá para futuras recomendações ao usuário. Esta análise é utilizada amplamente em Sistemas Tutores Inteligentes, visto que o modelo de aluno, um dos principais componentes do sistema, é predominantemente direcionado para a estratégia de ensino-aprendizagem e personalização do conteúdo.

Além da análise de ações comportamentais, o usuário pode informar explicitamente suas preferências ou o sistema pode manter informações sobre as consultas realizadas pelo usuário, caso o sistema disponha de mecanismo de consultas. O usuário também pode marcar um determinado item como interessante e contribuir para a formação de seu perfil¹³, que será guardado no MU.

Na figura 3.6 é apresentado um caso de recomendações pessoais, direcionadas ao cliente de acordo com seu perfil de navegação

Figura 3.6 Exemplo de recomendações baseada no perfil do usuário

Na figura 3.7 o usuário tem a opção de acompanhar um item específico e assim, explicitamente, declara-se interessado pelo *Mp4 Player* e, possivelmente, pela categoria de Eletrônicos, Áudio e Vídeo, a categoria que abrange o produto em questão.

¹³ Neste caso o usuário não refere-se a categorias de itens ou conteúdos, mas a itens específicos sobre os quais detêm algum tipo de interesse.

Possibilidade de "Acompanhar" item específico

mercado Livre associado ao eBay

Vendedor | Seguidos | Nome | Vendedor | Minha Conta | Mapa do Site

Navegar | Vender | Buscar | Meu MercadoLivre | MercadoPago | Comunidade | Conta

O que está buscando? >> Buscar [Como buscar](#)

Mp4 Player 1Gb Mp3 + Vídeo + Grav Voz + Radio FM + Pen Drive

Eletrônicos, Áudio e Vídeo → MP3 e MP4 Players → Multifunção → 1 GB → MP4 Produto: #53596614

Preço Fixo: **R\$ 150.00** unit. [Compare](#)
ou 6 parcelas de R\$ 29.00 (com MercadoPago)
[Ver meios de pagamento com MercadoPago](#)

Vendedor: **BRUNAFURLAN (1528)**
Pontos do vendedor: 1528
100% qualificações positivas (0% negativas)
Membro desde: 20/05/2006 | [Ver reputação](#)
[Ver os produtos deste vendedor](#)

Tipo de produto: Novo
Localização: SÃO PAULO
Finaliza em: 28h 4h (27/03/2007 19:34)
Quant. de Ofertas: 793 [Ver compradores](#) Visitas: 28656

Quantidade: de 206 disponíveis

Sua Oferta: R\$ 150.00 unit.

Programa de Proteção ao Comprador [Ver requisitos](#)

[Cadastre-se é grátis!](#) [Como comprar?](#) [Como vender?](#) [Ver perquisitas ao vendedor](#) [Acompanhar este produto](#) [Enviar este link a um amigo](#)

▶ [Veja os conselhos para uma negociação segura e com êxito. Clique aqui.](#)
▶ O vendedor assume total responsabilidade pela publicação do produto. Você acha que este produto é contrário às políticas do MercadoLivre? Informe [aqui](#).

Figura 3.7 O usuário pode acompanhar um item específico

Neste tipo de estratégia devem ser utilizadas técnicas capazes de relacionar numa base de dados informações do usuário e armazená-las num perfil de usuário definido previamente. São necessários meios para coletar a informação do usuário (como cookies ou clickstream) e o tratamento destas informações para cada usuário específico.

3.2.5 Recomendações *Off-line*: E-mails

Para atrair a atenção dos usuários enquanto não estão no sistema, a personalização vale-se também de técnicas *off-line*, já discutidas no capítulo 2. As recomendações realizadas após a navegação do usuário e saída do sistema, podem fazer uso das estratégias anteriores. Através dos e-mails de ofertas, pode-se recomendar itens ou conteúdos aos usuários, estruturando-os conforme seu perfil de navegação ou de acordo com as estratégias do modelo de negócios do sistema.

Num sistema comercial podem ser oferecidos novos itens para compra. Já num ambiente educacional podem ser sugeridos conteúdos curriculares ou extra-curriculares baseados no perfil do usuário.

A figura 3.7 apresenta um exemplo de envio de ofertas pelo site MercadoLivre, baseadas na navegação do usuário.



Figura 3.7 Exemplo de Avaliação dos Usuários

3.3 Entrada de Dados, Apresentação e Grau de Personalização

Segundo a taxonomia apresentada, os dados de entrada do usuário alvo da comunidade são a base para os sistemas de recomendação. No trabalho de Venson (VENSON, 2002) são apresentados tipos destes dados que podem ser coletados.

As entradas do usuário alvo podem ser dados vindos de sua navegação implícita, de sua opinião e interesses (coleta explícita), através de seu histórico de compras e através de avaliações (que além da opinião do usuário incluem algum tipo de classificação). Além destas informações, a criação de categorias de produtos e itens pode incluir atributos-chave que, associados à coleta explícita ou implícita,

fornecem dados relevantes. Por exemplo, saber que um usuário navega pela categoria de livros num *website* de comércio eletrônico, através da inclusão de informações especiais (atributos-chaves) que identificarão esta preferência de usuário pela sua navegação.

As entradas da comunidade permitem a criação de um ambiente colaborativo. Incluem todos os itens de entrada do usuário alvo, porém aplicado a um grupo mais amplo de pessoas que possam ter interesses em comum. Adiciona-se a este tipo de entrada os comentários textuais, encorajados em alguns sites, para que os clientes dêem explicitamente sua opinião.

A apresentação das recomendações de maneira adequada para os usuários (clientes, por exemplo) é decisão importante nos sistemas computacionais. Por isto, classifica-se a apresentação da seguinte forma:

- Tecnologia Push: recomenda itens sem que o usuário esteja interagindo com o sistema;
- Tecnologia Pull: as recomendações só são apresentadas quando o usuário deseja e solicita explicitamente;
- Apresentação passiva: as recomendações estão inseridas no contexto da aplicação;

Através de dois fatores importantes na recomendação, a precisão desta e sua utilidade, é possível determinar a eficiência do sistema de personalização. Faz-se necessário observar o comprometimento na oferta de recomendações que podem estar “erradas” ou fazerem parte de um processo de inferência não muito preciso. Por isto, os sistemas podem não usar todos os recursos disponíveis na oferta de recomendações, atenuando o problema da imprecisão, mas com a contrapartida de que haverá perdas na oferta de serviços personalizados. Assim, os níveis de personalização para sistemas de recomendação podem ser agrupados da seguinte forma:

- Não-personalizado: permitem a seleção manual através de categorias estabelecidas no projeto do sistema.
- Efêmero: Não usam dados persistentes para as recomendações, apenas dados da sessão de navegação atual do usuário.

- Persistente: oferecem as recomendações mais personalizadas e necessitam de grande quantidade de informações para isto.

3.4 Métodos de Recomendação - Filtragem de Informações

A quantidade de informações disponível na web, atualmente, é imensa. Mesmo restringindo-se a análise a um *website* específico tem-se um volume grandioso de informações. Cada usuário, com características e preferências de consumo distintas optará sempre por uma parte destas informações. Mas, o tempo despendido até encontrar o que realmente se deseja pode ser relevante. A necessidade de tecnologias para filtragem de informação é algo que remonta a década de 1990. Já existia a preocupação com a quantidade de informação que estava sendo gerada pelos sistemas de informação e demasiada carga de informações que chegava até os usuários (LOEB e TERRY, 1992).

A expressão “filtragem de informação” para Belvin e Croft (BELVIN e CROFT, 1992) é o nome dado a uma variedade de processos, visando a entrega de informação para as pessoas corretas, ou seja, aquelas pessoas que realmente necessitam destas informações.

Existe outra expressão, “recuperação de informação”, que, freqüentemente, era confundido com a filtragem de informação. No entanto, apesar das duas expressões descreverem soluções que visam auxiliar na solução de problemas de sobrecarga de informações, a recuperação de informação envolve armazenamento, criação de índices e recuperação de documentos textuais. A partir destas etapas, o usuário poderá consultar explicitamente as informações que deseja, realizando a interação com o sistema.

Já a filtragem de informação é uma abordagem voltada ao perfil de usuário, constituído de preferências de navegação e outras informações diversas, utilizando este perfil para fornecer informações aos usuários. Este fornecimento de informações, muitas vezes, é automático, tornando a interação do usuário com o sistema mais dinâmica (FOLTZ e DUMAIS, 1992). A seguir, são descritas algumas subdivisões da filtragem de informação.

3.4.1 Filtragem Baseada em Conteúdo (FBC)

A filtragem baseada em conteúdo consiste em categorizar as informações e relacioná-las com os interesses dos usuários. Também é chamada de correlação item-a-item e, freqüentemente, encontra itens similares aos de interesse do usuário ou cliente para sugestão.

Este tipo de abordagem procura padrões de compra ou consumo semelhantes para recomendar produtos ou informações aos usuários. Vale-se de mecanismos para Descoberta de Conhecimento em Bases de Dados, como regras de conhecimento e Mineração de Dados. Como exemplo de aplicação, um livro pode ter sido muito vendido com uma camiseta, ou seja, boa parte das pessoas que compraram este livro também levaram a camiseta. Tem-se aqui, uma correlação item-a-item, ou por conteúdo.

Pode ser realizado em conjunto de informações, baseando-se em palavras-chave para relacionar documentos textuais, ou qualquer outro tipo de dado.

3.4.2 Filtragem Colaborativa (FC)

Esta abordagem é baseada na correlação existente entre os itens comprados ou visitados por outros usuários e pelo usuário em questão (alvo). A correlação pode ser explícita, pelo histórico ou freqüência de utilização, ou ainda, pelo emprego de técnicas que descubram tais relações.

A FC reforça o conceito de conhecimento comunitário, mesmo que as implicações de um usuário sobre outro não seja explícita. A idéia é aprender (e prever) os comportamentos de um usuário baseado em usuários com características de comportamento semelhantes.

Segundo Herlocker (HERLOCKER, 2000) a FC pode ser descrita em três fases:

- Calcula-se a métrica de similaridade dos usuários com o usuário alvo;
- Os usuários com maior similaridade são selecionados;
- As características dos usuários selecionados predizem de forma ponderada o comportamento do usuário alvo.

3.5 Descoberta de Conhecimento

A Descoberta de Conhecimento em Bases de Dados (KDD - *Knowledge Discovery in Database*) visa a descoberta de padrões existentes entre dados coletados, representados e relacionados. As informações existentes, coletadas dos usuários em sistemas personalizáveis, precisam ser tratadas de modo a resultar em conhecimento aplicável a personalização. Tanto a filtragem baseada em conteúdo como a filtragem colaborativa necessitam de técnicas de KDD. A Mineração de Dados (*Data Mining*) é uma das etapas do KDD e vale-se de mecanismos de aprendizagem de máquina (MA - *Machine Learning*) para efetivar-se. A KDD possui várias etapas de realização e foge ao escopo deste trabalho lidar com estas etapas e todos os métodos disponíveis.

Esta seção apresentará duas aplicações de descoberta de conhecimento para sistemas de recomendação. Uma através do uso de regras de associação, para descoberta de correlação item-a-item e outra através da proximidade de vizinhança (usando-se correlação de *Person*) para filtragem colaborativa.

3.5.1 Uso de Regras de Associação para Filtragem Baseada em Conteúdo

A aplicação de regras de associação considera entradas numa base de dados estruturada, onde cada tupla é tratada como um conjunto de itens que podem assumir um valor verdadeiro ou falso. A regra de associação consiste no relacionamento $X \Rightarrow Y$ (X e Y são conjuntos de itens com intersecção vazia), associada a um valor de suporte (Fator de Suporte - FSup) e a um valor de confiança (Fator de Confiança - FConf).

O valor do Fator de Suporte representa a razão de representatividade da associação entre X e Y , pois consiste na razão de tuplas que satisfazem X e Y sobre o total de tuplas. Desta forma:

$$\text{FSup} = \frac{|X \Rightarrow Y|}{N}, \text{ onde } N \text{ é o número total de tuplas.}$$

O valor do Fator de Confiança visa garantir que a relação realmente ocorre é a medida da frequência em que a relação acontece sobre o item em questão. Assim:

$$FConf = \frac{|X \Rightarrow Y|}{|X|}$$

A extração de regras significativas ocorre estipulando-se valores para FSup e FConf, encontrando regras na base de dados que satisfaçam estes valores.

A tabela 3.2 demonstra um conjunto de dados que após passar por um processo de descoberta usando regras de associação apresenta os resultados gerados na figura 3.8. A coluna ID representa os identificadores dos usuários e as demais colunas são os itens existentes no sistema. A informação de cada linha é se um determinado usuário comprou ou não determinado item.

Tabela 3.2: Dados para descoberta de regras de associação¹⁴

ID	Item1	Item2	Item3	Item4	Item5	Item6	Item7
1	Não	Sim	Não	Sim	Sim	Não	Não
2	Sim	Não	Sim	Sim	Sim	Não	Não
3	Não	Sim	Não	Sim	Sim	Não	Não
4	Sim	Sim	Não	Sim	Sim	Não	Não
5	Não	Não	Sim	Não	Não	Não	Não
6	Não	Não	Não	Não	Sim	Não	Não
7	Não	Não	Não	Sim	Não	Não	Não
8	Não	Não	Não	Não	Não	Não	Sim
9	Não	Não	Não	Não	Não	Sim	Sim
10	Não	Não	Não	Não	Não	Sim	Não

<p>Conjunto de Itens freqüentes: Item2, Item4. FSup = 0,3</p> <p>Regra: Se (Item2) então (Item4), FConf = 1</p> $FSup = \frac{ 3 }{ 10 } \text{ e } FConf = \frac{ 3 }{ 3 }$
<p>Conjunto de Itens freqüentes: Item4, Item5. FSup = 0,4</p> <p>Regra: Se (Item4) então (Item4), FConf = 0,8</p> <p>Regra: Se (Item5) então (Item4), FConf = 0,8</p> $FSup = \frac{ 4 }{ 10 } \text{ e } FConf = \frac{ 4 }{ 5 }$

Figura 3.8 Regras de Associação resultantes dos Dados da tabela 3.2

Neste exemplo de aplicação, a primeira coluna da tabela representa um identificador para cada transação e as demais indicam se um item foi ou não adquirido na transação. A suposição para os valores de FSup e FConf são,

¹⁴ Extraído de (REATEGUI, CAZELLA e OSÓRIO, 2006)

respectivamente 0,3 e 0,8. As regras descobertas, apresentadas na figura 3.8 servirão para a recomendação de itens de forma conjugada, devido a similaridade de consumo descoberta entre eles. Este exemplo de aplicação foi extraído de (REATEGUI, CAZELLA e OSÓRIO, 2006).

3.5.2 A vizinhança próxima e a Filtragem Colaborativa

De maneira mais formal, a pontuação $p(a,i)$ (ilustrado na equação 1) do item “i” pelo usuário alvo (a), é estimada com base na pontuação $p(u,i)$ fornecida pelos usuários vizinhos (u) ao mesmo item “i”. Os usuários vizinhos (u) são similares ao usuário alvo em questão, ou seja, possuem comportamento similar. Este valor $p(a,i)$ trata-se de uma agregação de pontuações de outros usuários para o mesmo item.

$$p_{a, i} = \text{aggr} \underset{u \in \hat{U}}{p_{u, i}} \quad (1)$$

O símbolo \hat{U} representa o conjunto dos N usuários “u” mais similares ao usuário “a” que classificaram o item “i”.

Este modelo de FC utiliza contribuições dos usuários para a classificação realizada pelo sistema. A entrada do sistema é uma matriz cujas linhas representam os usuários e as colunas são itens. Os dados na matriz são as avaliações (*rating*) que cada usuário forneceu aos itens. Esta técnica também é chamada de técnica dos vizinhos mais próximos ou “*k-nearest-neighbor*” e a definição de similaridade pode utilizar coeficientes, como a correlação de Pearson (SHARDANAND e MAES, 1995; HERLOCKER, 2000; REATEGUI, CAZELLA e OSÓRIO, 2006)

A equação (2) apresenta $W_{a,u}$ como a correlação do usuário alvo (a) com um usuário vizinho (u). Há a necessidade de mais de uma avaliação em comum para que a correlação seja viável, e os resultados variam entre 1 para similaridade total, e -1 para total dissimilaridade. O cálculo da predição pode ser efetuado através da equação (3).

$$W_{a,u} = \frac{\sum_{i=1}^m [(r_{a,i} - \bar{r}_a) * (r_{u,i} - \bar{r}_u)]}{\sqrt{\sum_{i=1}^m (r_{a,i} - \bar{r}_a)^2 * \sum_{i=1}^m (r_{u,i} - \bar{r}_u)^2}} \quad (2)$$

a – usuário alvo

u – usuário “vizinho”

$W_{a,u}$ – correlação do usuário “a” com o usuário “u”

$r_{a,i}$ – avaliação do usuário “a” para o item “i”

–

\bar{r}_a - média de todas as avaliações do usuário “a” comuns com o usuário “u”

$r_{u,i}$ – avaliação do usuário “u” para o item “i”

–

\bar{r}_u - média de todas as avaliações do usuário “u” comuns com o usuário “a”

A predição $P_{a,i}$ da classificação do item “i” para o usuário alvo “a” é a média ponderada das avaliações que os N vizinhos (u) do usuário alvo (a) deram ao item “i”. O valor N pode ser determinado pelo próprio sistema, de acordo com critérios particulares.

Por exemplo, de acordo com a tabela 3.3 o usuário U2 avaliou 10 itens (item1 até item 10), segundo uma escala de classificação de 1, para menor interesse, até 5, para maior interesse. O item7 não foi avaliado pelo usuário U2.

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n [(r_{u,i} - \bar{r}_u) * W_{a,u}]}{\sum_{u=1}^n (W_{a,u})} \quad (3)$$

Tabela 3.3: Avaliação Usuários X Itens

Usuários	Itens									
	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10
U1	5	-	-	4	5	1	-	-	4	3
U2	5	1	4	4	4	1	-	1	5	4
U3	-	4	1	1	1	4	3	4	-	-
U4	5	1	3	3	4	1	4	-	-	-
U5	2	2	1	5	4	1	-	-	-	-
U6	3	5	2				4	4	4	4
U7	2	2	2	3	3	3	-	-	-	-
U8	4	1	4	3	-	-	5	-	4	2

De acordo com a matriz apresentada, o usuário U2 não gostou do item6, avaliando-o com o valor 1. No entanto, este usuário gostou do item1, avaliando com o valor máximo (5). Verifica-se também, que o usuário U2, nestes dois itens, concorda com a avaliação dada pelo usuário U1. Para medir a similaridade entre estes dois usuários utiliza-se o coeficiente de Pearson. Os itens em comum estão destacados na tabela 3.4.

Tabela 3.4: Avaliações em comum dos usuários U1 e U2

Usuários	Itens									
	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10
U1	5	-	-	4	5	1	-	-	4	3
U2	5	1	4	4	4	1	-	1	5	4
U3	-	4	1	1	1	4	3	4	-	-
U4	5	1	3	3	4	1	4	-	-	-
U5	2	2	1	5	4	1	-	-	-	-
U6	3	5	2				4	4	4	4
U7	2	2	2	3	3	3	-	-	-	-
U8	4	1	4	3	-	-	5	-	4	2

As avaliações dos itens em comum entre os usuários ficaram da seguinte maneira:

U1: [5; 4; 5; 1; 4; 3]

U2: [5; 4; 4; 1; 5; 4]

Nota-se a concordância entre estas avaliações. A média das avaliações foi de 3,67 para U1 e 3,83 para U2. Aplicando a equação (2) obtém-se o valor 0,87. Desta forma, U1 e U2 são bastante similares, concordando em suas avaliações.

Aplicando agora o coeficiente de Pearson para o usuário U2 em relação a todos os outros usuários têm-se os resultados na tabela 3.5.

Tabela 3.5: Coefeciente de Person dos usuários em relação ao usuário U2

Usuário	Pearson (U2)
U1	0,87
U3	-1
U4	0,95
U5	0,39
U6	-0,55
U7	-0,11
U8	0,86

De acordo com os resultados da tabela 3.5 observa-se que o usuário U2 é bastante similar a U4, U1 e U8, relativamente similar a U5 e não é similar a U3, U6 e U7 em suas preferências. A partir da descoberta dos usuários com similaridade, parte-se para o próximo passo. A predição para os novos itens deste usuário.

Aplicando-se a equação (3) para prever a nota que o usuário U2 atribuiria ao item item7, levando em consideração todos os itens que os usuários vizinhos, com limiar superior a 0,9 (limite assumido), pontuaram em comum com o usuário alvo, tem-se o resultado de 4,33. Este resultado significa que caso o usuário U2 consumisse o item7, daria como uma nota igual a 4,33 (predição) para tal item, considerando a avaliação dos mais próximos, ou seja, uma boa recomendação.

3.6 Considerações

Este capítulo apresentou alguns conceitos sobre os Sistemas de Recomendação, sua classificação, aplicações e métodos para geração de recomendações baseados no perfil dos usuários do sistema. O próximo capítulo apresenta as abordagens educacionais estudadas.

E-Learning e EAD

"A principal meta da educação é criar homens que sejam capazes de fazer coisas novas, não simplesmente repetir o que outras gerações já fizeram. Homens que sejam criadores, inventores, descobridores. A segunda meta da educação é formar mentes que estejam em condições de criticar, verificar e não aceitar tudo que a elas se propõe."

(Jean Piaget)

O processo de ensino está calcado em séculos de reflexões sobre a arte de educar. Mesmo os profissionais atuantes no ensino que não conhecem as obras de Platão, Aristóteles, Piaget, Vygotsky, estão sob forte influência destes, através da incorporação das idéias destes pensadores na prática pedagógica. Este capítulo visa relacionar a importância do Projeto Instrucional e das formas de medir a aprendizagem, na seção 4.1 e conceituar o Ensino a Distância e o Ensino Eletrônico, na seção 4.2. Para mais informações sobre a importância social que

representa a educação e sobre as teorias educacionais que sustentam a prática veja o Apêndice A.

4.1 Projeto Instrucional e Medidas de Aprendizagem

As teorias educacionais dão suporte às práticas de ensino, e à avaliação destas práticas. No apêndice A, tem-se um detalhamento mais completo sobre as teorias educacionais estudadas. O projeto instrucional define o ciclo de vida de um curso. Em outras palavras, define toda a prática de ensino que será adotada para um determinado curso, seguindo metodologia apropriada que, normalmente, é embasada numa dada teoria educacional.

Um projeto instrucional é descrito em diversas etapas a serem cumpridas e componentes a serem organizados. Os componentes integrantes na organização de um curso são: a informação – que constitui o conteúdo do curso -, os resultados – que são os objetivos esperados -, os métodos instrucionais – o processo pelo qual o curso será aplicado, qual teoria educacional seguirá - e os meios instrucionais – a forma pela qual o curso será realizado, ferramentas de apoio, etc. Assim, o projeto instrucional envolve não apenas o conteúdo programático, mas também as ações pedagógicas que serão adotadas. A união entre a teoria e a prática.

As medidas de aprendizagem são agrupadas em três grandes contextos: as medidas de capacidade (cognitivas), as medidas afetivas (comportamento) e as medidas comportamentais (psicomotoras). Através destas medidas o aprendizado pode ser guiado para que os alunos obtenham desempenhos esperados e construam um modelo de educação apropriado. A cognição compreende o conhecimento, a compreensão e o desenvolvimento de capacidades intelectuais. A afetividade compreende as mudanças de atitudes, de interesses, dos valores individuais. As habilidades psicomotoras tratam das capacidades de percepção física e de resposta a estímulos externos. A psicometria é a área da psicologia responsável por tais medições.

4.2 A Ensino a Distância (EAD) e o Ensino Eletrônico (*e-Learning*)¹⁵

A maneira como se tem aprendido e ensinado foi transformada ao longo dos últimos séculos pelas novas formas de comunicação e pela revolução tecnológica. Dois produtos desta transformação são: o Ensino a Distância e o Ensino Eletrônico.

O ensino a distância, diferentemente do ensino eletrônico, não foi inspirado no surgimento de novas tecnologias, mas sim, pelo enfretamento das dificuldades de locomoção ou disponibilização de tempo das pessoas para a educação. Com o grande desenvolvimento dos serviços postais na Europa, no século XVIII, tem-se a criação da primeira forma de ensino à distância, que se manteve durante o século XIX até meados do século XX. A partir daí, o EAD passou a utilizar-se das novas formas de comunicação de massa, como a televisão e o rádio. Estas formas de comunicação unidirecional dificultavam o papel do professor/tutor, da sua interação com os alunos e da interação entre os alunos. Tal dificuldade foi superada com a chegada das redes telefônicas que passaram a auxiliar o processo de ensino-aprendizagem. Um novo problema foi o alto custo deste tipo de comunicação, devido à baixa popularização. O surgimento das redes de computadores e, posteriormente, da internet¹⁶ trouxe a melhoria na comunicação e a possibilidade de ferramentas para auxiliar no processo de ensino-aprendizagem.

O ensino eletrônico, cujas características fundamentais são a existência de um projeto instrucional bem definido e a dependência de algum meio tecnológico de interação com o usuário, surgiu como ferramenta de auxílio à aprendizagem

¹⁵ Por que “Ensino” e não “Educação”?

Entre educação e aprendizagem há uma conexão conceitual, pois não há educação sem que aconteça a aprendizagem. Desta forma a educação é a construção da aprendizagem. A aprendizagem pode ser fruto de processos externos, do ensino, e também pode acontecer independentemente do ensino, resultando de processos internos. A aprendizagem e o ensino são conceitos com neutralidade moral, pois é possível ensinar e aprender para o bem ou para o mal. Já a educação não possui essa neutralidade. Educar é realizar algo correto e valioso. Assim, a educação é a construção positiva da aprendizagem. A aprendizagem é um processo que ocorre dentro do indivíduo. Mesmo sendo obtida por um processo de ensino bem-sucedido, ocorre internamente ao indivíduo. O ensino aplicado a uma pessoa pode ser ineficaz em outra. O ensino a distância é possível, mas a educação é interna e não à distância. Por isto, a preferência pelo termo “Ensino a Distância” e não “Educação a Distância”. Ao Ensino Eletrônico a justificativa é de que Aprendizagem (ou Educação) Eletrônica tratar-se-ia de processos internos de aprendizagem de mecanismos eletrônicos e não de pessoas, assim, o termo estaria semanticamente incorreto. Alguns autores preferem o termo Ensino Mediado pela Tecnologia (EMT).

¹⁶ Com igual ou maior importância que a internet tem-se a WWW (*World Wide Web*), criação de *Tim Berners-Lee* que popularizou a rede mundial de computadores com a forma de apresentação de conteúdos.

mediada por computador, os sistemas CAI (Instrução Assistida por Computador) tratados na seção 2.2.3. Atualmente, o ensino eletrônico abrange não apenas os STI (Sistemas Tutores Inteligentes, que garantem a adaptação de conteúdos e práticas pedagógicas), mas também toda forma de comunicação que visa a aprendizagem, valendo-se de recursos hipermídia e de interatividade.

O EAD e *e-learning* coexistem e se complementam. Isto é resultado da revolução tecnológica e dos efeitos causados pela globalização. A união destes conceitos promove a criação potencial de ambientes educacionais ricos em informações, divergentes em características pessoais, abrangentes e personalizados. Da fusão entre o EAD e *e-learning* surgem outras classificações, de acordo com a utilização das tecnologias, como por exemplo o EBW (Educação Baseada na Web) e o *Groupware* que pode ser estendido como CSCW (*Computer Supported Cooperative Work* – Trabalho Cooperativo Suportado por Computador) ou CSCL (*Computer Supported Cooperative Learning* – Aprendizagem Colaborativa Suportada por Computador).

A Educação Baseada na Web (WbE do termo inglês *Web-based Education*) popularizou-se no final do ano de 1996. Refere-se à utilização de tecnologias *Web* para auxílio ao processo educativo, ou seja, ao ensino a distância. Estas tecnologias são as ferramentas e os conceitos utilizados na *Web* como a hipermídia, o e-mail, a transmissão de vídeo, os grupos de notícias, os fóruns de discussão e, mais atualmente, o RSS¹⁷ e as comunidades virtuais¹⁸.

O termo *Groupware* refere-se a um processo combinado de ferramentas de software e trabalho em equipe. Estas ferramentas devem permitir a comunicação entre os membros e melhoria na eficácia e eficiência do trabalho. O CSCW é empregado por vezes como sinônimo de *Groupware*, porém trata-se de uma disciplina de pesquisa para estudar técnicas e metodologias de trabalho em grupo e como este trabalho pode ser beneficiado pela tecnologia. Enquanto o CSCW é visto

¹⁷ RSS é um recurso desenvolvido em XML que permite aos responsáveis por sites, canais de notícias e *blogs* divulgarem informações e novidades. Para que isso seja possível, um arquivo com extensão .xml, .rss ou .rdf é enviado com o link e o resumo das informações. Esse arquivo é conhecido como *feed*. A sigla RSS tem mais de um significado: 1) *RDF Site Summary*; 2) *Really Simple Syndication*; 3) *Rich Site Summary*. Isto ocorre, pois é uma tecnologia desenvolvida por mais de um grupo. Para utilização do RSS basta incluir o link do *feed* do site desejado (fornecedor de RSS) num programa leitor de RSS (agregador). Este programa lê o conteúdo dos arquivos recebidos e apresenta-o ao usuário. Uma curiosidade é a tecnologia RSS ter surgido no ano de 1999 como uma criação da Netscape que, pouco tempo depois, abandonou o projeto por não achá-lo viável. Maiores informações em: <http://www.infowester.com/rss.php>

¹⁸ Fonte: <http://www.vdl.ufc.br/catedra/ccwb/intro.htm>

como a pesquisa nesta área, o *Groupware* é tido como o resultado da pesquisa e a sua aplicação. Exemplos de ferramentas que promovem o *Groupware* são: o correio eletrônico, editores colaborativos (texto ou código fonte), videoconferência, entre outros. O CSCL é a utilização do CSCW para a aprendizagem e para o ensino. Assim, o *Groupware* acontece para promover não o trabalho em grupo apenas, seguindo regras de negócios, mas a aprendizagem colaborativa pelos integrantes do grupo.

4.3 O Ambiente de Ensino à Distância Tidia-Sakai

No decorrer deste trabalho, estudou-se a possibilidade de aplicação do método proposto. O ambiente estudado para aplicação foi o Tidia-Sakai. Existem diversos softwares para Educação baseada na Web como como o Moodle¹⁹, o Blackboard²⁰ e o WebCT²¹, mas a escolha do Sakai foi devido ao laboratório de pesquisa Lince²², da UFSCar que atua no projeto de aprendizagem eletrônica do Tidia, o Tidia-Ae, há alguns anos.

Estes softwares são também chamados de Learning Management Systems (LMS). O Sakai é um LMS focado, principalmente, na colaboração online que resultada no aprendizado por intermédio das interações e relações sociais entre os alunos. Os envolvidos são motivados a colaborar entre si para construir competências. Grande parte dos LMS disponíveis foca na realização de tarefas e download de arquivos.

Os recursos disponíveis no Sakai são os mais diversos e parecidos com os recursos dos LMS mais usados, como: Fórum, Chat, Central de Mensagens, Calendário, Apresentações, Repositório de Arquivos, Avisos, Wiki, Dentre outros.

Além disto, existem ferramentas em desenvolvimento como o comunicador instantâneo desenvolvido pelo Lince UFSCar e a ferramenta para gestão de laboratórios remotos.

A figura 4.1 mostra a tela inicial do sistema de Educação baseada na Web Tidia-Sakai. A esquerda tem-se os menus de acesso às ferramentas do usuário, na direita o calendário e ao centro informações sobre o contexto do usuário. Cada

¹⁹ Disponível em www.moodle.org

²⁰ Disponível em www.blackboard.com

²¹ Disponível em www.webct.com

²² Informações disponível em www.lince.dc.ufscar.br

usuário pode participar de contextos ou “workspaces”. Um workspace contém um curso ou conteúdo relacionado a este e ferramentas de acesso.

The screenshot displays the TIDIA Sakai workspace interface. At the top left is the TIDIA logo. The top navigation bar includes 'My Workspace', 'LARC - POLI/USP', and 'Teste'. A 'Logout' link is in the top right. A left sidebar menu lists: Home, Profile, Membership, Schedule, Resources, Announcements, Worksite Setup, Preferences, Account, and Help. The main content area is divided into three sections: 'Message of the Day' (with an 'Options' link and a red arrow pointing to it labeled 'funções permitidas de configuração'), 'Calendar' (showing October 2007 with a table of dates), and 'My Workspace Information' (with an 'Options' link). Below this is a welcome message for 'Luiza González' explaining the purpose of a workspace. At the bottom right is a 'Recent Announcements' section.

Figura 1

Figura 4.1 Tela Inicial do Tidia-Sakai

Outro exemplo de ferramenta (ilustrado na figura 4.2) é o DropBox que é uma área para armazenamento de recursos. É utilizado para troca destes recursos entre os usuários.

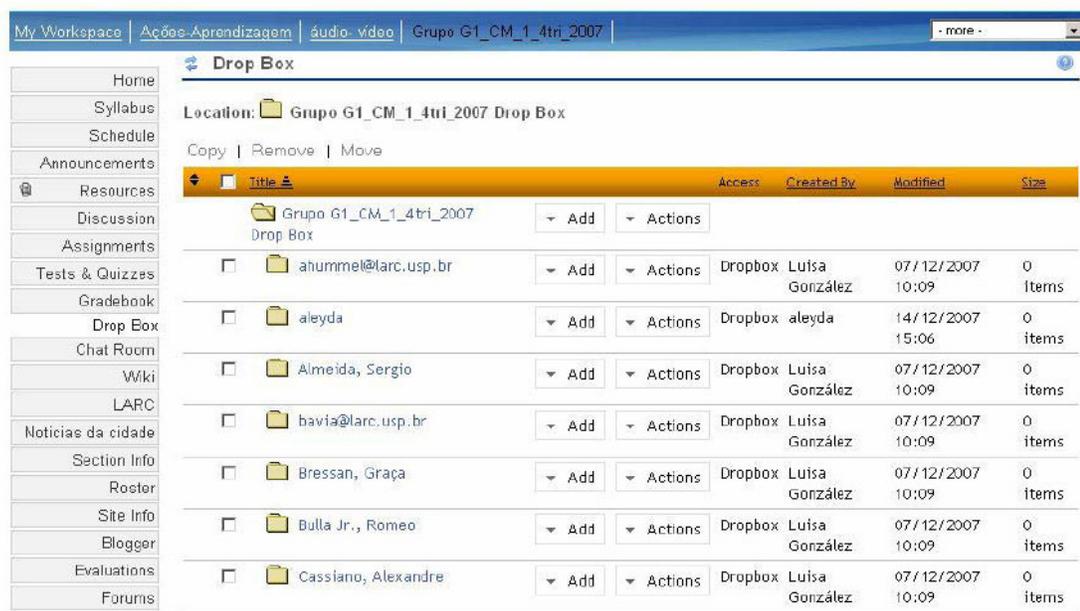


Figura 4.2 Ferramenta DropBox

4.3 Considerações

Este trabalho busca auxiliar o processo de ensino-aprendizagem, trabalhando com preferências individuais e coletivas (através dos sistemas de recomendação), oferecendo também novas possibilidades de interação e avaliação. Interação porque usará a filtragem colaborativa num modelo sócio-interacionista e avaliação, pois os interesses do aluno, através de técnicas de aprendizagem de máquina, serão representados num perfil que, posteriormente, será usado para personalização do sistema.

O Modelo I2P para Recomendação de Conteúdo

*“A força não provém da capacidade física e sim de
uma vontade indomável.”
(Mahatma Gandhi)*

A Educação Baseada na Web (Web-based Education - WbE) é utilizada para educação à distância, fornecendo facilidade para acesso e reposição de conteúdos e melhorando a comunicação entre estudantes e professores. Também contribui, de forma significativa, com o processo de aprendizagem. Nesta seção é descrito o método criado e denominado I2P, utilizado para descrever um modelo de usuário. Este método utiliza informações implícitas e explícitas para a formação do modelo de usuário, coletadas durante a utilização do sistema pelo usuário. Também é utilizado um método linear para recomendação de recursos aos usuários de um WbE-S. Esta seção está organizada como segue: a seção 5.1 trata da modelagem de usuários; a seção 5.2 trata do modelo I2P; nas seções 5.3, 5.4 e 5.5 são apresentados os componentes do modelo I2P, respectivamente as preferências dos usuários, seus interesses e a popularidade dos recursos no sistema; na seção 5.6 tem-se as considerações deste capítulo.

5.1 Modelagem de Usuários e Sistemas de Recomendação

Sistemas Educacionais Baseados na Web (WbE-S), como visto no capítulo anterior, têm ganhado mais importância, principalmente com a grande difusão da internet. No entanto, estes sistemas não fornecem facilidades para personalização de conteúdos. Algumas destas facilidades são usualmente restritas à Sistemas Tutores Inteligentes (Intelligent Tutoring Systems – ITS) e exigem grande dedicação dos tutores para modelagem do conhecimento e, assim, permitir que o sistema adapte-se de acordo com o conhecimento adquirido pelo aluno.

A personalização de sistemas na Web, como visto no capítulo 2, é por isto outra área de relevante estudo neste trabalho. A personalização, aplicada a WbE-S, é tema de vários outros trabalhos e diversas pesquisas e campos como Mineração de Dados, Mineração na Web, Modelagem de Usuário, Hipermídia Adaptativa, Sistemas Tutores Inteligentes e Sistemas de Recomendação (RESNICK e VARIAN, 1997; PALAZZO, 2006; BRUSILOVSKY e MILLÁN, 2007).

A modelagem de usuário, ou do estudante para este trabalho, é uma área de grandes estudos e desafios para WbE-S. Para a modelagem e para a adaptação do perfil do usuário, são necessárias diversas informações sobre o comportamento do mesmo. Estas informações podem ser observadas implicitamente ou coletadas de forma explícita, “perguntando-se” ao usuário.

A construção de sistemas personalizados tem ganhado grande importância social e econômica, além de ser tópico de diversas pesquisas científicas. Para a construção destes sistemas é necessária a criação e manutenção de perfis (ou modelos) de usuários (User Model – UM). Diversas áreas são envolvidas na criação destes perfis, desde Psicologia e Pedagogia até áreas como Inteligência Artificial, Interação Humano-Computador e Recuperação de Informações, que são áreas relacionadas à Computação.

Um modelo de usuário representa explicitamente as propriedades de um determinado usuário e as informações necessárias para que o sistemas possa adaptar-se a estas propriedades. De acordo com Brusilovsky e Milán (BRUSILOVSKY e MILLÁN, 2007), este modelo é a representação de informações que são essenciais para sistemas adaptativos poderem prover a personalização necessária, como por exemplo, comportar-se diferente para diferentes usuários.

Quando o usuário realiza pesquisas, por exemplo, o sistema pode adaptar a prioridade dos itens retornados, considerando a relevância destes no perfil do usuário. O sistema também pode “manipular” a navegação do aluno de acordo com seu estilo cognitivo, visando a melhoria do processo de ensino-aprendizagem.

Temos ainda como exemplo, sites de comércio eletrônico, como a Amazon, que oferecem recomendações de compra personalizadas aos consumidores, e assistentes para auxílio em dúvidas como Assistente do Microsoft Office, base da pesquisa do Projeto Lumiere²³ nos laboratórios da Microsoft.

Nestes exemplos, demonstra-se a necessidade da construção de um Modelo de Usuário efetivo para oferecer personalização adequada. Na educação, um modelo de usuário e sua utilização são descritos no trabalho de Andrade, Giraffa e Vicari (ANDRADE, GIRAFFA e VICARI, 2003), onde é apresentado um modelo que visa descrever os aspectos cognitivos do perfil de estudantes.

As pesquisa sobre modelagem de usuário teve início com os Sistemas Tutores Inteligentes e visava a construção de modelos capazes de identificar aspectos cognitivos e a evolução da aprendizagem dos estudantes. Os modelos iniciais tinham problemas quanto a eficiência computacional, pois as inferências necessitavam de grande esforço para serem calculadas. Isto era causado, basicamente pela grande complexidade do modelo.

A manutenção dos modelos de usuário com precisão era outro problema, pois sempre era necessário um modelo adequado e dinâmico para o usuário. Assim, o modelo precisava, constantemente, ser atualizado.

Os avanços na pesquisa em aprendizado de máquina revelaram novos modelos, mais relacionados aos domínios de aplicação. Os novos modelos ganharam em eficiência e reduziram a complexidade de representação e facilitaram a adaptação constante. (ANDRADE, GIRAFFA e VICARI, 2003).

Como visto no Capítulo 2, a personalização de sistemas possui diversas abordagens que divergem quanto às técnicas usadas e resultados obtidos. Entretanto, os objetivos das diversas abordagens restringem-se a oferecer a informação que as pessoas mais precisam (BELVIN e CROFT, 1992). Basicamente, um sistema de recomendação tenta identificar os itens mais importantes para os usuários e então os recomenda (SHARDANAND e MAES, 1995).

²³ Mais informações em <http://research.microsoft.com/~horvitz/lumiere.htm>

Como visto no capítulo 3, os sistemas de recomendação pode realizar essa tarefa com técnicas distintas: Filtragem Colaborativa, Filtragem Baseada em Conteúdo e uma combinação das duas. A filtragem colaborativa utiliza informações de grupos de usuário para geração de recomendações. Desta maneira, os usuários se auxiliam na construção de manutenção de seus perfis e suas preferências.

Apesar do caráter extremamente comercial dos sistemas de recomendação, diversos trabalhos vêm utilizando-os em educação (GEYER-SCHULZ, HAHSLER e JAHN, 2001). Na área de marketing, estes sistemas identificam produtos e serviços que oferecem maior satisfação e maiores possibilidades de compra durante navegação do cliente (SHARDANAND e MAES, 1995). Já em sistemas educacionais, estes métodos visam identificar padrões de preferência e seqüências de navegação para cada estudante no sistema, como a técnica de raciocínio baseado em casos usada em ambientes de aprendizagem interativos (REATEGUI, CAZELLA e OSÓRIO, 2006).

Para recomendar itens para cada usuário, primeiro o sistema precisa coletar diversas informações deste. Estas informações são representadas num modelo de usuário para tentar identificar o que é relevante sobre as preferências deste usuário e os conteúdos do sistema. Estas informações podem ser coletadas de maneira explícita (usuário podem prove-las deliberadamente) ou implícita (o comportamento do usuário é coletado).

A Filtragem Colaborativa pode ser separada em três passos (após a coleta de informações):

- 1 – O cálculo da semelhança entre os indivíduos (formação da vizinhança);
- 2 – A seleção dos vizinhos mais próximos;
- 3 – O cálculo da predição, ponderando os valores das avaliações dos vizinhos pelo valor de semelhança deste vizinho com o usuário-alvo (calculado no passo 1).

O coeficiente de correlação de Pearson foi utilizado para cálculo da semelhança por dois motivos: 1 – os valores abordados nas medições são contínuos; 2 – é uma medida que tem oferecido bons resultados em Sistemas de Recomendações (HERLOCKER, 2000).

O coeficiente de Pearson, usado no passo 1 da filtragem colaborativa, mede a “força” do relacionamento entre duas variáveis. Esta força é representada por um número que varia entre -1, indicando que há uma correlação inversa, e 1, indicando uma forte correção positiva.

5.2. O Modelo I2P

Para descrever o relacionamento entre o sistema e os usuários num Sistema Educacional Baseado na Web e prover recomendação de conteúdo associado ao perfil do usuário, foi desenvolvido o Modelo I2P – Interesses, Preferências e Popularidade -, apresentado neste trabalho e descrito nesta seção. Neste modelo tem-se:

- Interesses: são valores que descrevem o comportamento dos usuários em relação aos recursos do sistema. Estes valores são coletados implicitamente.
- Preferências: são valores de classificação fornecidos pelos usuários ao utilizarem os recursos do sistema.
- Popularidade: são valores que relacionam os dois valores anteriores e denotam quais os recursos com maior grau de interesse e/ou preferência.

Para classificar o relacionamento entre o usuário e os objetos do sistema, todos os objetos são considerados como recursos. Assim, em WbE-S convencionais, os recursos são agrupados em: objetos de aprendizagem, ferramentas, usuários e grupos (ou comunidades). Um recurso R pode ser um usuário (U), uma comunidade (C), uma ferramenta do sistema (T), ou um objeto de aprendizagem (O).

Como descrito na figura 5.1, o usuário explicitamente classifica os recursos acessados. Além disto, a interação entre o usuário e o recurso é medida implicitamente. Esta medição é realizada considerando todos os acessos aos recursos do sistema.

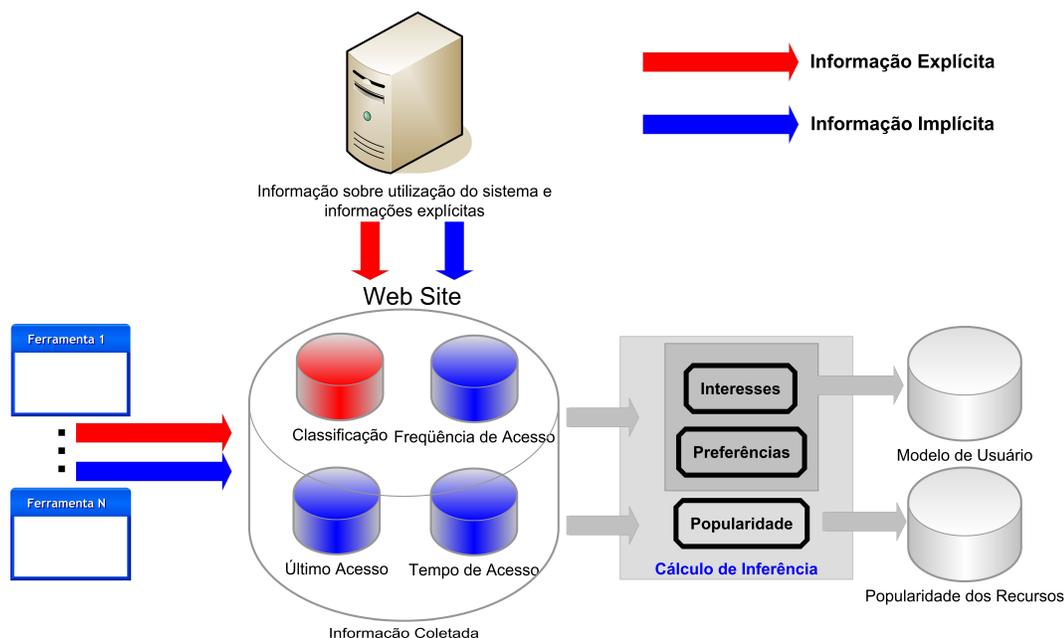


Figura 5.1 Arquitetura I2P

Os valores da medição implícita e os valores de classificação são armazenados numa base de dados que, posteriormente, será usada como entrada para cálculo dos valores de interesse e preferência. Além disto, estes valores serão usados para o cálculo da popularidade dos recursos.

Os valores de interesse e/ou preferência serão usados também para o cálculo da vizinhança do usuário-alvo, visando descobrir quais vizinhos possuem recursos de interesse/preferência em comum.

A entrada do sistema é uma matriz M cujas linhas representam os usuários e as colunas representam os recursos, como apresentado na tabela 5.1. A informação na posição $M_{u,r}$ representa o valor de interesse ou preferência do usuário “ u ” no recurso “ r ”. O cálculo do interesse é descrito na seção 5.4 e a preferência é descrita na seção 5.3.

O princípio do algoritmo da Filtragem Colaborativa considera que o usuário alvo tem maior probabilidade de interessar-se por itens que seus vizinhos mais próximos se interessaram. Logo, calcula-se os valores de similaridade entre o usuário-alvo e demais, usando o coeficiente de Pearson e posteriormente, é formada a vizinha deste usuário alvo (GOLDBERG *et al.*, 1992; RESNICK e VARIAN, 1997).

Tabela 5.1 Matriz de Interesses/Preferências - usuários X recursos

	Recurso 1	Recurso 2	...	Recurso N
Usuário 1				
Usuário 2				
	Interesse Medido ou Valores de Preferência			
Usuário N				

Este método é baseado nas medidas de covariância entre os valores de interesse/preferência para cada usuário. Além disto, a correlação de Pearson atenua as diferenças entre valores extremos de interesses/preferências dos usuários.

A correlação de Pearson está representada na equação 5.1, $W_{a,u}$ como a correlação do usuário-alvo “a” com o vizinho “u”. Há a necessidade de mais que uma avaliação em comum para que o cálculo da correlação seja viável. O valor pode variar entre total dissimilaridade (-1) e total similaridade (1). A equação 5.1 descreve o cálculo da correlação de Pearson considerando como métrico o interesse dos usuários.

$$W_{a,u} = \frac{\sum_{i=1}^m [(i_{a,i} - \bar{i}_a) \times (r_{u,i} - \bar{i}_u)]}{\sqrt{\sum_{i=1}^m (i_{a,i} - \bar{i}_a)^2 \times \sum_{i=1}^m (i_{u,i} - \bar{i}_u)^2}} \quad (5.1)$$

- a – usuário alvo
- u – usuário vizinho
- $W_{a,u}$ – correlação entre usuário “a” e usuário “u”
- $i_{a,r}$ – interesse/classificação do usuário “a” no recurso “r”
- \bar{i}_a – média de todos os interesses em comum entre o usuário “a” e o usuário “u”
- $i_{u,r}$ – interesse do usuário “u” no recurso “r”
- \bar{i}_u – média de todos os interesses em comum entre o usuário “u” e o usuário “a”

A geração de predições pode ser feita avaliando-se e pesando-se todas as avaliações feitas pelos vizinhos, independente do método usado para a geração da vizinha (RESNICK e VARIAN, 1997). O cálculo das predições está na equação 5.2.

$$P_{a,r} = \bar{i}_a + \frac{\sum_{i=1}^n [(i_{u,r} - \bar{i}_r) * w_{a,r}]}{\sum_{i=1}^n (w_{a,u})} \tag{5.2}$$

A predição $P_{a,r}$, novamente considerando-se medidas de interesse, para o recurso "r" pelo usuário-alvo "a" é a média ponderada das avaliações feitas pelos N vizinhos "u" do usuário-alvo "a" ao recurso "r". O valor de N pode ser determinado de acordo com critérios particulares, ou testado por alguma função de otimização.

Estes passos estão descritos na figura 5.2, onde os valores de interesse ou preferências dos usuários são usados para a formação de vizinhanças. Após isto, é separada a quantidade adequada de vizinhos para cálculo da predição. Os recursos com melhores predições são recomendados ao usuário alvo.

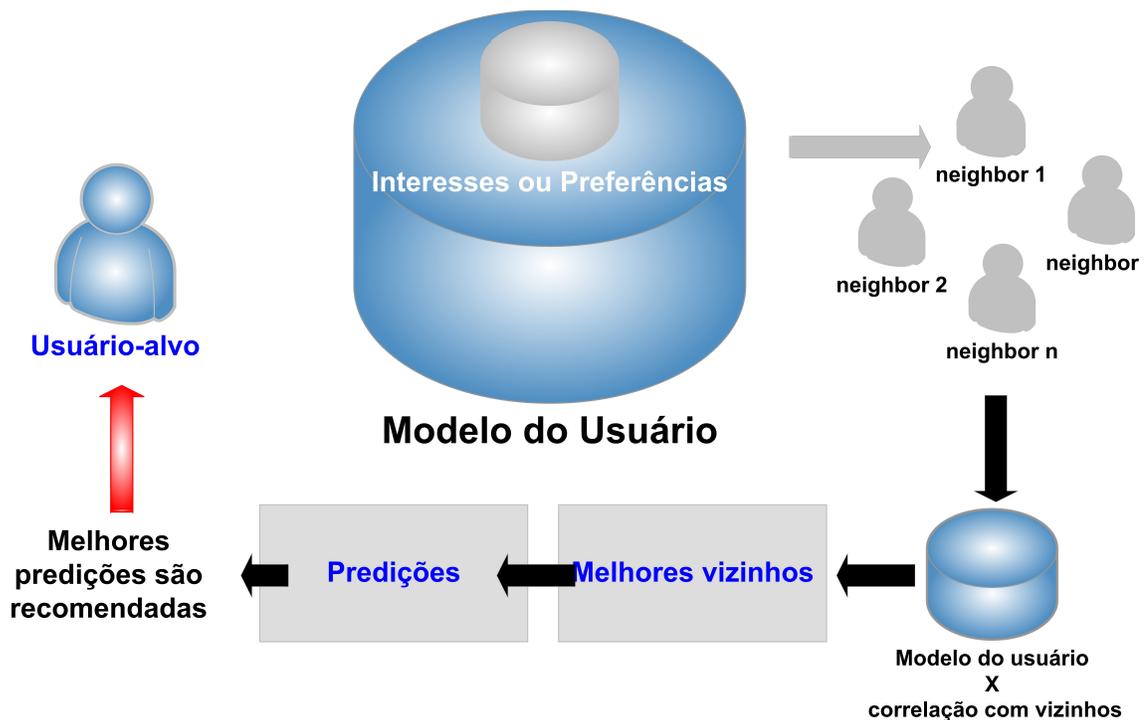


Figura 5.2 Formação de Vizinhança e cálculo de Predição

5.3 Preferências dos Usuários

A Filtragem Colaborativa usa contribuições dos estudantes para classificação no sistema. A entrada do sistema é, novamente, uma matriz de usuários X recursos, como a tabela 5.1. Os valores da matrix representam as classificações dadas pelos usuários nos respectivos recursos. Após isto, o cálculo da vizinhança e a predição são realizados da maneira já apresentada na seção 5.2. Esta técnica de formação de vizinhança também pode ser chamada de “k-nearest-neighbor” (SHARDANAND e MAES, 1995; HERLOCKER, 2000; REATEGUI, CAZELLA e OSÓRIO, 2006).

As preferências medidas são calculadas considerando a classificação que o usuário forneceu explicitamente ao recurso acessado.

Neste trabalho, utilizou-se cinco níveis de preferências, relacionados à aprovação do usuário quanto ao recurso visitado, como demonstrado na figura 5.3. Por exemplo, um usuário que visita uma página web com conteúdo que ele acredita ser interessante, poderia classificá-la com o valor máximo 5. Caso este conteúdo não fosse interessante, ele poderia atribuir o valor mínimo 1.



Figure 5.3 Possíveis Classificações fornecidas pelo usuário

5.4 Interesses dos Usuários

As informações explícitas que os usuários fornecem sobre os recursos possuem um problema: o usuário pode acessar o recurso, mas deixar de classificá-lo. Desta forma, o sistema não teria informações sobre o relacionamento deste usuário com o recurso já acessado.

Para resolver este problema, foi proposta uma abordagem para medição da interação do usuário com os recursos do sistema, além da classificação explícita.

A interação entre o usuário e os recursos é medida usando-se de características de algoritmos bem conhecidos e usados em sistemas operacionais para troca de páginas de memória.

As medidas usadas são três: Mais Frequentemente Acessado (Most Frequently Used – MFU), Mais Recentemente Acessado (Most Recently Used – MRU) e Tempo Total de Acesso (Access Total Time – ATT).

O algoritmo para cálculo do valor de MFU, uma inversão do algoritmo LFU (Menos Frequentemente Acessado - Least Frequently Used) é usado para verificar itens cujo acesso pelo usuário é mais frequente, medindo isto através do número de visitas ao recurso.

$$MFU(u, Ri_u) = \frac{\text{freqüência}(u, Ri_u)}{\max \text{Freqüência}(Ri)} \quad (5.3)$$

Assim, o algoritmo MFU que relaciona o usuário “u” ao recurso “Ri” é calculado pelo valor da freqüência de visitação de “u” ao recurso “Ri”. Dividindo este valor pela freqüência máxima de visitação ao mesmo recurso (este valor é calculado considerando todos os usuários e suas respectivas freqüências de visitação ao mesmo recurso), tem-se um MFU já normalizado, relacionando todos os usuários do sistema. Esta normalização é de grande importância para utilização futura do MFU.

O algoritmo MRU (Most Recently Used), uma inversão do algoritmo LRU (Menos Recentemente Usado - Least Recently Used), é usado para verificar os recursos cujo acesso dos usuários é mais recente, usando para isto o valor de último acesso ao recurso. A medida MRU é obtida pelo valor do último acesso ao recurso “Ri” pelo usuário “u”. Esta medida é normalizada, dividindo-a pelo valor máximo de MRU para o mesmo recurso, como descrito na equação 5.4

$$MRU(u, Ri_u) = \frac{\text{últimoAcesso}(u, Ri_u)}{\max \text{últimoAcesso}(Ri)} \quad (5.4)$$

Além destes, o tempo de acesso, na equação 5.5, é usado como referência do interesse do usuário, pois, por exemplo, múltiplos acessos com pouca duração não necessariamente significam grande interesse.

$$ATT(u, Ri_u) = \frac{ATT(u, Ri_u)}{\max ATT(Ri)} \quad (5.5)$$

Desta forma, o tempo total de acesso num recurso auxilia o cálculo e descoberta do interesse do usuário neste recurso.

Após a medição e cálculo destes três valores, de acordo com o comportamento dos usuários no sistema (navegação do usuário pelo site, já que é um WbE-S), os interesses dos usuários são calculados através da soma ponderada de MRU, MFU e ATT. A ponderação da soma é feita pelo tutor, ou por um especialista do domínio, de modo a otimizar ao máximo a função de interesse do usuário.

$$\begin{aligned} Interesse(u, Ri_u) = & MFU(u, Ri_u) \times \alpha + \\ & MRU(u, Ri_u) \times \beta + ATT(u, Ri_u) \times \gamma \end{aligned} \quad (5.6)$$

Onde

α Peso de MFU

β Peso de MRU

γ Peso de ATT

e

$$\alpha + \beta + \gamma = 10$$

De acordo com a equação 5.6, que descreve a função interesse de um usuário "u" num recurso "Ri" acessado por ele, a soma ponderada dos valores medidos deve variar de 0 a 10. Cabe ressaltar que as medidas já estão normalizadas.

A figura 5.4 ilustra o processo de inserção de pesos por um tutor e a geração dos valores de interesse. Após isto, os valores de interesse são armazenados no modelo de usuário, numa base de dados.

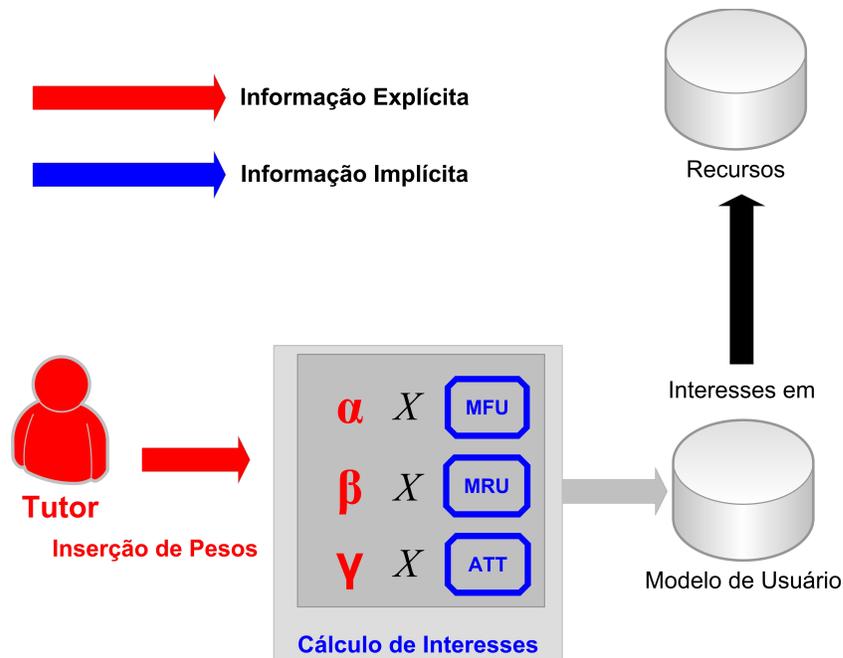


Figura 5.4 Cálculo de Interesse

5.5 Popularidade dos Recursos

A popularidade dos recursos é calculada usando a soma dos valores MFU, MRU e ATT para cada recurso. A equação 5.7 descreve a popularidade de um recurso R com base no seu valor de frequência. Trata-se da soma de cada valor MFU de todos os usuários que acessaram o recurso R.

$$PopFreq(R) = \sum_{u=1}^N MFU(u, R) \quad (5.7)$$

O mesmo é válido para o valor de MRU. Assim, a popularidade relacionada a quão recentes são os acessos num determinado recurso é calculada na equação 5.8.

$$PopÚltimoAcesso(R) = \sum_{u=1}^N MRU(u, R) \quad (5.8)$$

Na equação 5.9 é descrito o cálculo da popularidade baseado no tempo de acesso ao recurso “R”.

$$PopTempoAcesso(R) = \sum_{u=1}^N ATT(u, R) \quad (5.9)$$

Além destes valores, também pode-se utilizar outros dois valores para o cálculo da Popularidade Total de um recurso R. Um deles é a Popularidade da Classificação deste recurso, que mede os valores fornecidos explicitamente pelos usuários, classificando tais recursos e caracterizando os valores de preferência destes usuários sobre os recursos. Isto está descrito na equação 5.10.

$$PopClassif(R) = \sum_{u=1}^N classif(u, R) \quad (5.10)$$

O outro valor é atribuído pelo próprio tutor, ou especialista do domínio. Este valor visa iniciar a popularidade dos recursos para que a recomendação seja realizada no começo da utilização do sistema, sem a necessidade de coleta de informações para isto.

$$Popularidade(R) = \frac{PopFreq(R)}{MaxPopFreq(R)} \times \alpha + \frac{PopTempoAcesso(R)}{MaxPopTempoAcesso(R)} \times \beta + \frac{PopÚltimoAcesso(R)}{MaxPopÚltimoAcesso(R)} \times \gamma + \frac{PopClassif(R)}{MaxPopClassif(R)} \times \varphi + T \quad (5.11)$$

Onde:

- α Peso para Frequencia
- β Peso para Tempo de Acesso
- γ Peso para Último Acesso
- φ Peso para Classificação
- T Valor inserido pelo Tutor

e

$$\alpha + \beta + \gamma + \varphi + T = 10$$

Após o cálculo dos valores de popularidade para um recurso, o valor da popularidade total é calculado. Cada valor anteriormente obtido recebe um peso de acordo com a experiência do tutor. A soma dos pesos precisa ser igual a 10, como descreve a equação 5.11.

Desta maneira, o valor da popularidade é uma soma ponderada entre a Popularidade de Frequência, Popularidade pelo Último Acesso, Popularidade pelo Tempo de Acesso, Popularidade pela Preferência e Popularidade pela Experiência do Tutor.

Após isto, como descreve a figura 5.5, os recursos com os melhores valores de popularidade são recomendados aos usuários.

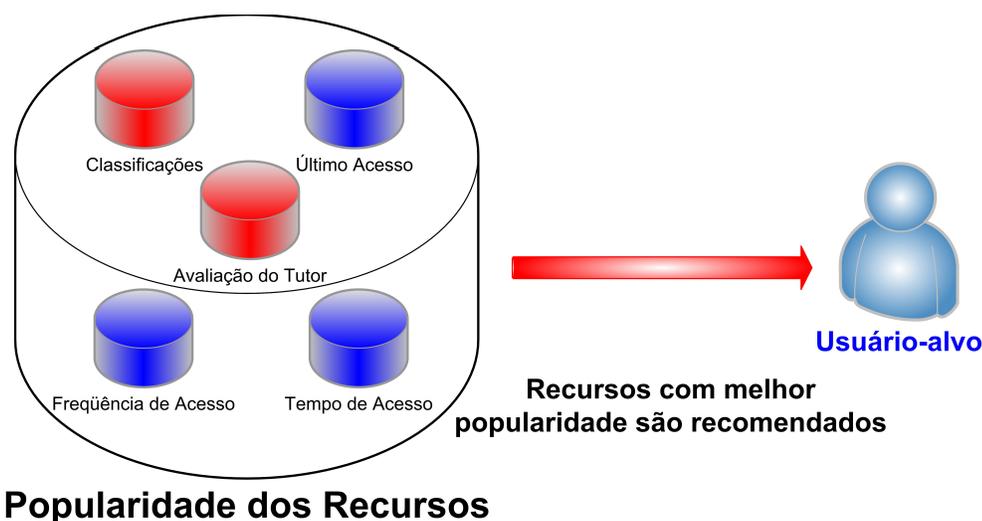


Figura 5.5 Recomendação baseada na Popularidade dos Recursos

5.6 Considerações do Capítulo

Neste capítulo o modelo I2P (Interesses, Preferências e Popularidade) para recomendação de conteúdo em Sistemas Educacionais baseados na Web, através da técnica de Filtragem Colaborativa, foi descrito. A geração dos valores de Interesse e Preferência são obtidos por métricas inseridas neste capítulo (MFU, MRU e ATT) e obtidas de maneira implícita, observando o comportamento do usuário, e também pela classificação explícita deste após fazer uso dos recursos no WbE-S.

Além disto, a união destes dois tipos de informação permite o cálculo de uma terceira métrica, a popularidade dos recursos. Esta métrica visa fornecer suporte na recomendação, principalmente no início da utilização do sistema quando não se dispõe de muitas informações sobre o sistema.

Avaliação

*“A grande finalidade da vida não é conhecimento,
mas ação.”
(Thomas Huxley)*

Este capítulo de avaliação visa descrever as atividades realizadas para validar a proposta apresentada. São descritos 3 estudos de caso realizados que foram apresentados em artigos internacionais. Além disto, apresenta-se os resultados obtidos e a comparação com outros modelos para personalização.

6.1 Comparação com outras abordagens para Personalização na Educação

Como exemplos de ambientes de educação à distância baseada na Web tem-se o Web-CT²⁴, o Sakai²⁵ e Moodle²⁶ que incluem serviços para compartilhamento de conteúdos, comunicação síncrona (chats) ou assíncrona (fóruns), módulos para avaliação com perguntas e respostas, “*white boards*”, repositório de dados, entre outras.

Numa classe de aula virtual os educadores provem recursos como textos, multimídias, simulações e discussões. Faz-se necessário estimular ao máximo os estudantes, pois a abordagem não presencial dificulta a visibilidade do educador sobre a satisfação do aluno nas tarefas e no curso.

Através de medidas de similaridades entre os estudantes e na utilização do sistema pelos mesmos fez-se possível recomendar atividades, tarefas ou conteúdos para estes (o que foi caracterizado como recurso), baseando-se no modelo I2P apresentado. No trabalho de (ZAIANE, 2002) é apresentado um sistema de recomendação usando agentes para aprendizado eletrônico. A principal diferença no modelo I2P é a utilização de métricas implícitas com possibilidade de ponderação pelo educador ou tutor, inserindo sua experiência e opinião no valor de tais métricas.

A diferença entre o modelo apresentado e outras abordagens como Hipermídia Adaptativa ou Sistemas Tutores Inteligentes, descritos no capítulo 2 está na fácil implementação e portabilidade do modelo I2P, pois não necessita de grandes adaptações no ambiente de software, podendo funcionar como um modelo separado. Na hipermídia adaptativa é necessário conhecer o sistema e os recursos e ponderar a navegação do usuário. Além disto, trata-se de um método de personalização intrusiva, que altera a navegação do usuário. Nos sistemas tutores inteligentes é necessário modelar o conhecimento do tutor ou da disciplina e compará-lo com o desenvolvimento do aluno, além de também ser uma técnica intrusiva.

²⁴ Maiores informações em www.webct.com

²⁵ Maiores informações em www.sakaiproject.org

²⁶ Maiores informações em www.moodle.org

A abordagem apresentada, além de não intrusiva, pode ser portada para outros sistemas e estudos de caso, bastando implementar as funções de cálculo para recomendação e geração de recomendação ou apenas utilizar interfaces pré-definidas. A vantagem de uma técnica não intrusiva consiste nas adaptações que esta permite, tornando-a mais flexível.

A seguir, estão descritos alguns estudos de caso utilizando a abordagem apresentada. Em algumas situações a abordagem não foi utilizada em sua totalidade, permitindo assim, novos estudos de caso.

6.2 Estudo de Caso 1 – Erro Médio na Predição de Interesses

Para testar o modelo realizou-se uma validação cruzada para cálculo do erro de predição na Recomendação do Sistema baseando-se em Interesses do Usuário. Foi avaliado um cenário hipotético com 10 estudantes e 8 recursos, usando valores fornecidos por um especialista de domínio. (GOTARDO, TEIXEIRA e ZORZO, 2008b)

Os valores de interação do usuário com os recursos foram medidos segundo as métricas propostas: MFU, MRU e TAT. Os pesos respectivos para as variáveis foram: 1, 3 e 6. Estes valores descrevem um cenário onde a métrica mais importante é o tempo total que um usuário gasta acessando um recurso.

O objetivo da validação cruzada foi testar todas as possíveis predições para um usuário e o erro desta predição. A tabela 6.1 demonstra os valores de Interesse dos usuários já calculados pela ponderação das três variáveis anteriores.

Tabela 6.1 Matriz de Interesses para o Estudo de Caso apresentado

	Recursos							
	1	2	3	4	5	6	7	8
Usuário 1	5,8	6,3	6,5	7,1	10,0	2,8	3,1	9,4
Usuário 2	3,3	4,0	1,8	1,6	5,2	6,3	7,3	1,8
Usuário 3	9,1	9,4	8,5	9,1	8,8	3,2	3,3	10,0
Usuário 4	3,1	2,0	2,9	3,0	6,6	5,3	3,7	3,6
Usuário 5	6,0	6,0	7,7	8,3	10,0	2,5	3,7	8,5
Usuário 6	4,5	2,3	2,6	3,5	7,0	2,7	3,2	3,0
Usuário 7	3,2	4,3	2,6	3,4	6,4	3,8	3,0	3,7
Usuário 8	1,1	1,3	1,3	1,5	5,4	8,8	6,6	0,9
Usuário 9	0,0	2,3	2,9	1,2	5,0	9,3	7,3	1,5
Usuário 10	1,0	0,0	2,3	0,8	4,6	8,3	5,6	1,7

Para este estudo de caso foi calculado o erro médio para todas as possíveis recomendações neste cenário, da seguinte maneira:

- Remove-se um valor de interesse para algum usuário;
- Calcula-se o interesse dos vizinhos deste usuário;
- Escolhe-se os vizinhos mais apropriados;
- Calcula-se a predição de interesse do usuário-alvo, usando-se os valores dos vizinhos, para o interesse removido;
- Calcula-se o erro de predição, considerando o interesse removido.

Após isto, estes passos foram repetidos para todos os usuários e todos os interesses na matriz para cada item. Por fim, calculou-se o erro médio das predições.

Por exemplo, para o usuário 1, removendo-se o interesse no recurso 6 cujo valor é 2,8, o cálculo da vizinhança de acordo com a equação 5.1 seria:

Tabela 6.2 Correlação dos usuários com o usuário-alvo 1

Usuários	Correlação com Usuário 1
Usuário 2	-0.447
Usuário 3	0.745
Usuário 4	0.535
Usuário 5	0.944
Usuário 6	0.481
Usuário 7	0.671
Usuário 8	-0.233
Usuário 9	-0.307
Usuário 10	-0.155

Para este experimento, escolheu-se os melhores vizinhos usando a heurística: “os dois melhores vizinhos serão necessários para predição”. Neste exemplo, os dois melhores vizinhos são os usuários 3 e 5.

O valor de predição estimado foi de 2,656. O valor real era 2,8. Logo o erro para este caso fica em 0,144.

Existem diversas métricas para avaliação da precisão da predição estatística como a Média de Erro Absoluto (Mean Absolute Error - MAE) e a Raiz Quadrada de Erro Absoluto (Root Mean Squared Error - RMSE), mas experiências mostram que estas métricas podem oferecer resultados similares (GOOD *et al.*, 1999). O valor MAE é a medida de desvio do valor de recomendação real para todos

os usuários. Cada par Interesse e Predição $\langle p_i, q_i \rangle$ trata do erro absoluto entre eles. O valor MAE é a soma de todos os erros absolutos para N pares $\langle p_i, q_i \rangle$ divididos pelo valor N. Formalmente:

$$MAE = \frac{\sum_{i=1}^N [p_i - q_i]}{N} \quad (6.1)$$

O resultado obtido no experimento foi um MAE de aproximadamente 1,00. Considerando que a medida de interesse varia entre 0,00 e 10,00 e são usadas medidas implícitas, tem-se um erro médio de aproximadamente 10% na predição.

6.2 Estudo de Caso 2 – Comparação na Predição de Preferências Tradicional e usando valores Implícitos

Neste outro Estudo de Caso, avaliou-se o mesmo cenário, considerando as três variáveis anteriores (MFU, MRU and TAT), e realizou-se uma comparação entre os valores de Predição para Classificação obtidos com o método tradicional de Filtragem Colaborativa e com a utilização de ponderação pelas métricas propostas. (GOTARDO, TEIXEIRA e ZORZO, 2008a)

Foram selecionadas aleatoriamente duas classificações numa matriz

Tabela 6.3 Matriz de Preferências para o Caso de Estudo 2

	Recursos							
	1	2	3	4	5	6	7	8
Usuário 1	5,0	3,0	4,0	2,0	4,0	2,0	2,0	1,0
Usuário 2	2,0	1,0	2,0	1,0	3,0	1,0	5,0	1,0
Usuário 3	5,0	3,0	4,0	2,0	4,0	2,0	2,0	1,0
Usuário 4	3,0	2,0	2,0	1,0	3,0	1,0	3,0	1,0
Usuário 5	5,0	3,0	4,0	2,0	4,0	2,0	2,0	1,0
Usuário 6	3,0	2,0	2,0	1,0	3,0	1,0	3,0	1,0
Usuário 7	3,0	2,0	2,0	1,0	3,0	1,0	3,0	1,0
Usuário 8	2,0	1,0	2,0	1,0	3,0	1,0	5,0	1,0
Usuário 9	2,0	1,0	2,0	1,0	3,0	1,0	5,0	1,0
Usuário 10	2,0	1,0	2,0	1,0	3,0	1,0	5,0	1,0

O teste realizado consistiu em excluir 2 valores de classificações da matriz para um determinado usuário (escolhido aleatoriamente). Após isto, calculou-se a vizinhança para este usuário e a predição de interesse nos recursos removidos. O recurso com melhor predição era recomendado ao usuário. Este passo consiste em efetuar a predição tradicional de Sistemas de Recomendação com Filtragem Colaborativa.

Para comparar, foram realizados outros 3 tipos de testes: a exclusão de dois valores altos para classificação de um determinado usuário em recursos vistos, mas com valores distintos (por exemplo, recursos classificados como 4 e 5); a exclusão de dois valores de baixa classificação (1 e 2 por exemplo) para um determinado usuário; e a exclusão de um valor de alta e outro de baixa classificação (1 e 5 por exemplo).

Após isto, comparou-se o resultado da predição tradicional em Sistemas de Recomendação e da predição usando valores implícitos de ponderação, analisando as divergências.

A figura 6.1 apresenta as divergências encontradas no experimento. Observou-se que a predição sem a utilização de valores implícitos acertou todas as recomendações. Na utilização do Tempo Total de Acesso (TAT) como ponderação, o acerto foi de 89%. Para a ponderação usando a Frequência de Acesso (MFU) os acertos foram de 78% e para a utilização dos valores de acessos mais recentes (MRU) o acerto foi de 44%.

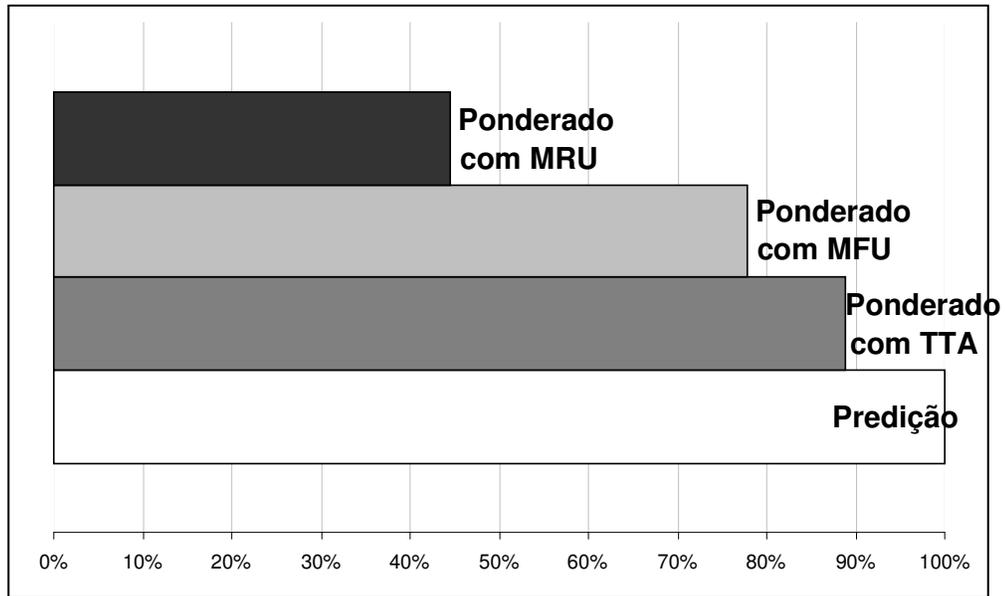


Figura 6.1 Comparação de acertos entre a predição ponderada com valores implícitos e a predição tradicional.

A predição para valores de classificação baixos acertou 56% dos testes, enquanto que a predição para valores altos acertou em 67%. Já a predição com a exclusão de um valor alto e baixo obteve acertos de 89% nos testes.

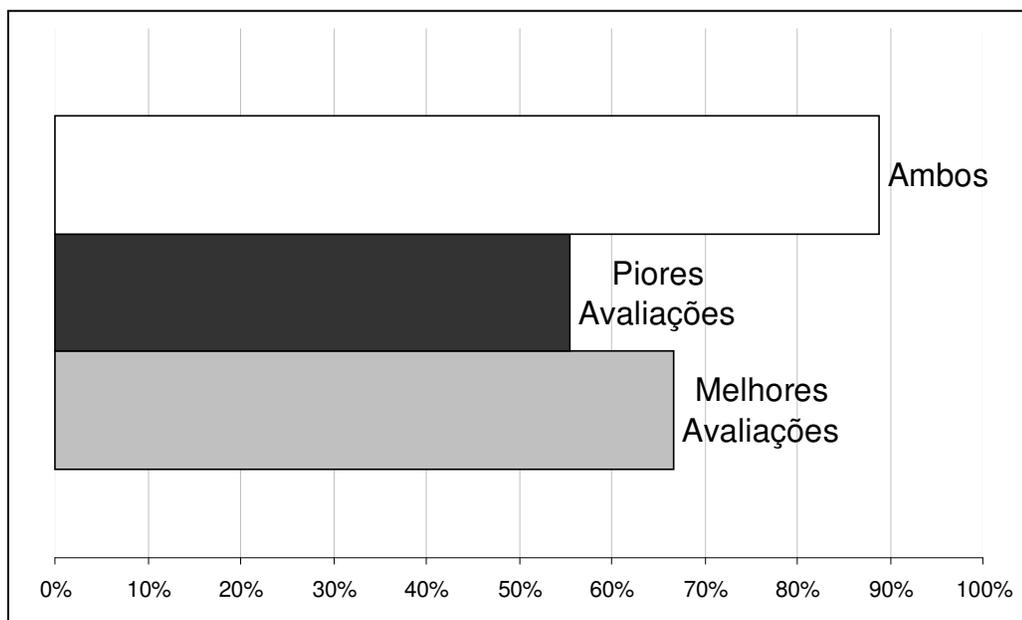


Figura 6.2 Comparação entre valores altos e baixos para medidas de preferência

Estes resultados demonstram que a predição de preferências obtidas por valores explícitos de classificação não refletem exclusivamente a real opinião do usuário. Faz-se necessário a avaliação de métricas implícitas para comparação com o real comportamento do usuário no sistema e não apenas sua classificação.

6.3 Estudo de Caso 3 – Formação de Grupos ou Comunidades de Interesse

Neste estudo de caso realizou-se a aplicação do modelo proposto na formação de comunidades ou grupos de usuários com interesses em comum o que ajuda no compartilhamento de conhecimento entre seus membros. O modelo usa informações implícitas e explícitas coletadas dos usuários. A correlação destas informações e a análise posterior da correlação são usadas para a formação de comunidades com características distintas (GOTARDO, TEIXEIRA e ZORZO, 2008c). No modelo, descrito na figura 6.3 os recursos são acessados e classificados e isto fica registrado na base dados (1), como ocorre no modelo deste trabalho, já descrito no capítulo 5. O cálculo do interesse do usuário é realizado (2) de acordo com a inserção de pesos do tutor (3).

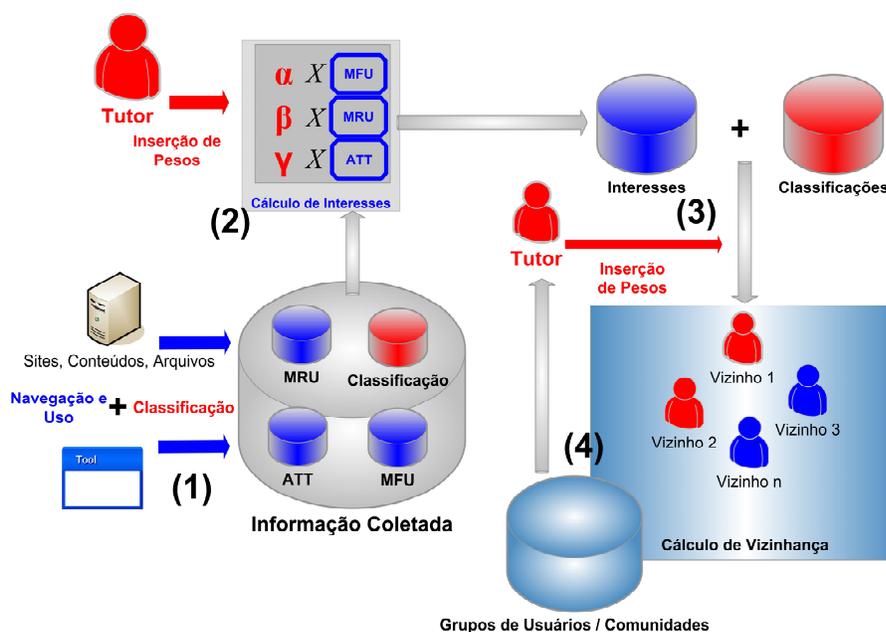


Figura 6.3 Arquitetura com a formação de Grupos

Após isto, o tutor insere novamente os pesos para interesse e classificação que serão usados na criação dos grupos. E então, o cálculo da vizinhança é realizado usando-se o algoritmo k-means (4), implementado pela plataforma WEKA (inserir nota). Para concluir, o tutor pode usar os grupos formados para futuros trabalhos, aplicação de seminários, etc.

O estudo de caso avaliou as mesmas matrizes de dados de entrada 6.1 e 6.3, respectivamente matrizes de interesse e de preferência. Assim, tem-se o mesmo cenário com 10 alunos e 8 recursos, usando os valores providos por um especialista de domínio. Os valores para interesse foram obtidos pela ponderação de MFU, MRU e TAT e os pesos considerados foram, respectivamente, 1, 3 e 6.

A tabela 6.4 mostra a matriz de correlação de interesses dos usuários. O valor de cada conjunto linha e coluna expressa a correlação entre os usuários respectivos da linha e coluna. Quanto mais próximo de 1, maior a correlação entre eles, quanto mais próximo de -1, correlação inversa.

Table 6.4 Cálculo da Matriz de Correlação de Interesses

Usuários	1	2	3	4	5	6	7	8	9	10
1	1,0	0,7	0,8	0,7	0,9	0,7	0,7	0,4	0,5	0,4
2	0,7	1,0	0,5	0,8	0,5	0,8	0,9	0,7	0,7	0,6
3	0,8	0,5	1,0	0,3	0,9	0,4	0,3	-0,1	0,0	-0,2
4	0,7	0,8	0,3	1,0	0,4	1,0	0,9	0,9	0,8	0,9
5	0,9	0,5	0,9	0,4	1,0	0,4	0,3	0,0	0,1	0,0
6	0,7	0,8	0,4	1,0	0,4	1,0	0,9	0,8	0,6	0,8
7	0,7	0,9	0,3	0,9	0,3	0,9	1,0	0,9	0,8	0,8
8	0,4	0,7	-0,1	0,9	0,0	0,8	0,9	1,0	0,9	1,0
9	0,5	0,7	0,0	0,8	0,1	0,6	0,8	0,9	1,0	0,8
10	0,4	0,6	-0,2	0,9	0,0	0,8	0,8	1,0	0,8	1,0

O algoritmo utilizado para a formação dos grupos foi o K-means implementado pela ferramenta WEKA²⁷. K-means é um algoritmo para agrupar n objetos baseando-se em atributos de k partições, onde $k < n$. Assume-se que os atributos dos objetos forma um vetor espacial. O objetivo, então, é tentar minimizar a variância “intra-grupo” ou o erro quadrático.

²⁷ www.cs.waikato.ac.nz/ml/weka

$$V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2 \quad (6.2)$$

Na função 6.2 há “k” clusters S_i , $i = 1, 2, \dots, k$ e μ_i é o centróide ou ponto médio de todos os pontos $x_j \in S_i$

Como apresentado na figura 6.4, foi utilizada a matriz de entrada, com os valores de interesse dos usuários, no algoritmo k-means e o resultado foi a distribuição em grupos da turma de alunos avaliada.



Figura 6.4. a) Resultados do algoritmo de agrupamento k-means aplicado aos dados de Interesse b) Distribuição dos grupos de acordo com o algoritmo aplicado.

Neste caso, foram obtidos 3 grupos distintos, como mostrado na figura 6.4 (b).

Tabela 6.5 Matriz de Correlação de Interesses e Preferências Ponderados por um Tutor – IP(UxU)

Usuários	1	2	3	4	5	6	7	8	9	10
1	1,0	0,2	0,8	0,1	0,6	0,4	0,2	-0,6	-0,6	-0,4
2	0,2	1,0	-0,1	0,1	-0,5	0,0	0,0	0,2	0,4	0,3
3	0,8	-0,1	1,0	-0,1	0,7	0,4	0,1	-0,7	-0,7	-0,6
4	0,1	0,1	-0,1	1,0	0,4	0,7	0,9	0,7	0,4	0,7
5	0,6	-0,5	0,7	0,4	1,0	0,8	0,6	-0,2	-0,6	-0,3
6	0,4	0,0	0,4	0,7	0,8	1,0	0,9	0,2	-0,2	0,0
7	0,2	0,0	0,1	0,9	0,6	0,9	1,0	0,5	0,2	0,4
8	-0,6	0,2	-0,7	0,7	-0,2	0,2	0,5	1,0	0,8	0,8
9	-0,6	0,4	-0,7	0,4	-0,6	-0,2	0,2	0,8	1,0	0,9
10	-0,4	0,3	-0,6	0,7	-0,3	0,0	0,4	0,8	0,9	1,0

No entanto, realizou-se outra análise a partir dos dados de interesse e preferências dos usuários. Tem-se na figura 6.5 uma nova matriz, desta vez reunindo estes dois valores ponderados pelo tutor.

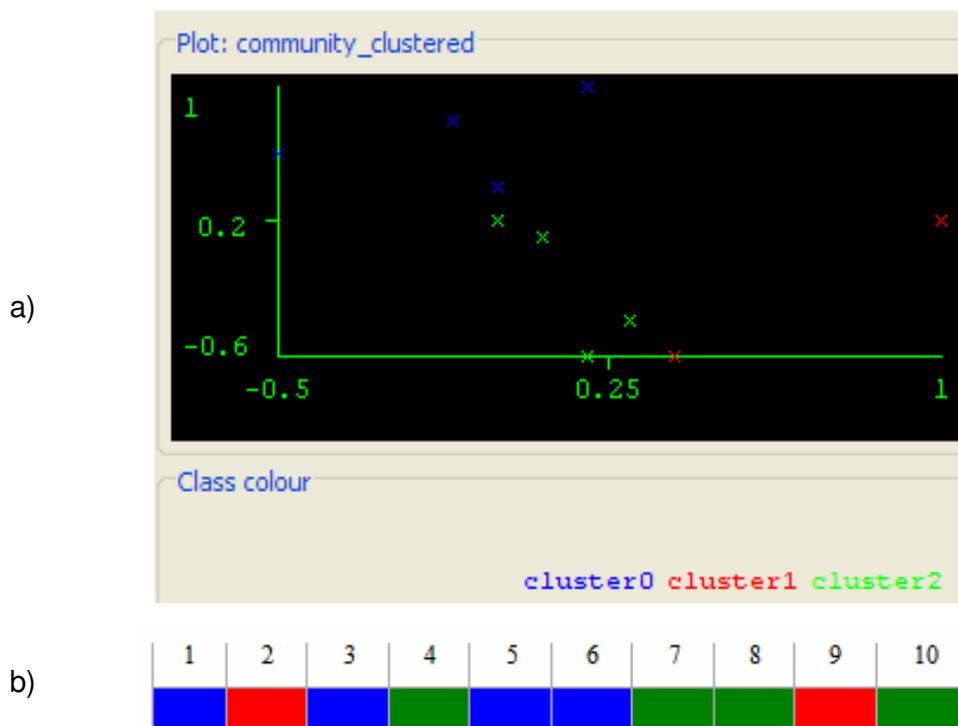


Figura 6.5) Resultados do algoritmo de agrupamento k-means aplicado aos dados de Interesses e Preferências b) Distribuição dos grupos de acordo com o algoritmo aplicado

Os pesos utilizados foram de 6 para preferência e 4 para interesses. O resultado dos novos grupos é apresentado na figura 6.5.

Como é possível observar, novos grupos foram formados, pois os valores de interesse e preferências interferiram-se e causaram uma aproximação mais real do comportamento dos usuários no sistema.

6.3 Estudo de Caso 4 – Estudo da Implementação num Ambiente WbE-S real

Para implementação do Modelo proposto, utilizou-se o Modelo de Dados abaixo como base de dados para armazenar os valores necessários para cálculo e geração de recomendações.

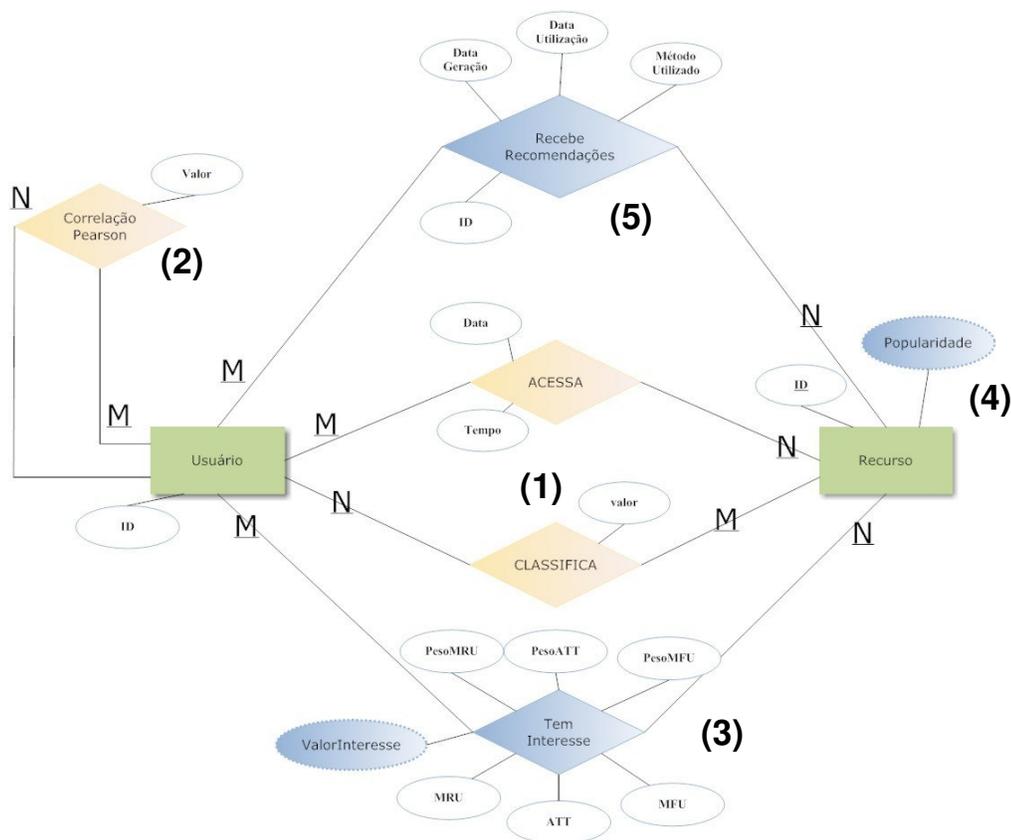


Figura 6.6 DER do Modelo I2P

Pode-se verificar, pelo diagrama ilustrado na figura 6.6 (1) que são necessárias apenas as informações de acesso (implícitas) dos usuários aos

recursos e as informações de classificação (explícitas). Após o cálculo da correlação de Pearson, guarda-se esta informação na própria base de dados (2). Com as informações de acesso e classificação são calculados os valores de interesses e preferências dos usuários (3) e os valores de popularidade dos recursos (4). Além disto, armazena-se na base de dados as informações sobre as recomendações realizadas para fins de análise e feedback (5).

O algoritmo abaixo deve ser implementado para cálculo da correlação de Pearson do usuário alvo com os demais usuários.

Algoritmo para Cálculo de Correlação do usuário alvo e demais usuários

```

Determinar usuário alvo
Recuperar todas as avaliações feitas pelo usuário alvo (usuário alvo,
idNotícia, Valor_Aval)
//Recuperar as avaliações feitas por outros usuários para
//os mesmos recursos do usuário alvo
    /*Teremos neste momento uma matriz i x j, onde i são os usuários
    e j são os recursos avaliados e i,j(1,1) são as avaliações do
    usuário alvo.*/
//Calcular a média de avaliações para cada usuário
somar i(usuário, 1... j) e dividir por j
PARA X = 1 até i
Carrega usuário vizinho com i

    PARA Y = 1, até j
        SOMA1 = SOMA1 + (avaliação(Y) Alvo - Média(Y) Alvo) X
        (avaliação(Y) vizinho i - Média(Y) Vizinho i)

        SOMA2 = SOMA2 + (avaliação(Y) Alvo - Média(Y) Alvo)2
        SOMA3 = SOMA3 + (avaliação(Y) vizinho i - Média(Y) Vizinho i)2
    FIM PARA
    Correlação entre Alvo e Usuário i = SOMA1 / [RAIZ ( SOMA2 X
    SOMA3) ]

    Incrementar i
    Calcular correlação para próximo vizinho i
FIM PARA

```

Abaixo, apresenta-se o algoritmo implementado para cálculo da predição de classificação do usuário e recomendação de conteúdo com base nestas predições.

Algoritmo para Cálculo da Predição de Conteúdo e posterior recomendação.

```
Verificar um número de recursos que o usuário alvo ainda não
classificou.

Calcular a predição para estes recursos.

Escolher usuários com correlação calculada e que classificaram o
recurso não visto pelo alvo.

Verificar o valor da predição e indicar os valores com maiores
predições.
```

Todos os cálculos para implementação já foram descritos nas seções anteriores (3, 4 e 5).

6.3.1 Desenvolvimento de Ferramentas para o Tidia-Sakai.

Para criação de uma ferramenta no ambiente Sakai, pode-se seguir o tutorial disponível em:

<http://bugs.sakaiproject.org/confluence/download/attachments/18439/SakaiPersistenceAndHibernate.ppt?version=1>

O esqueleto de uma aplicação com acesso a persistência pode ser construído com o plugin do Sakai para o Eclipse, escolhendo a criação de uma aplicação CRUD.

Os serviços e componentes do Sakai são um conjunto de implementações de alguma API do Sakai. A persistência não é feita diretamente, pois as ferramentas do Sakai servem apenas como interface gráfica (GUI). Existe um componente próprio para isto. Já os componentes não fazem qualquer tipo de serviço de interface gráfica como HTML, por exemplo, porém podem acessar outros serviços por meio das APIs disponíveis. O Sakai também possibilita que serviços e componentes sejam registrados para que possam fornecer funcionalidades comuns sem conhecer o domínio de seus objetos (BASMAN, 2007a).

Cada componente do Sakai publica os seus serviços usando o arquivo "components.xml". Estes componentes são iniciados quando o Sakai é iniciado e

suas dependências são resolvidas automaticamente e ao mesmo tempo (BASMAN, 2007a).

As APIs do Sakai são definidas como interfaces Java, existindo uma única implementação para cada interface da API no Sakai.

A figura 6.7 ilustra as tecnologias utilizadas pelo Sakai em camadas. As tecnologias envolvidas para o funcionamento do Sakai (versão 2.3) são:

- Java: versão 5.0 ou superior. Java Runtime Environment (JRE) para instalação e uso e Java Development Kit (JDK) para o desenvolvimento de ferramentas Sakai;
- Apache HTTP e/ou Tomcat: servidor Apache HTTP em conjunto com os módulos mod_jk para o Container Tomcat 5.5 ou superior;
- JSF, Velocity e RSF: para o desenvolvimento de GUI; e MySQL ou Oracle: para a persistência de dados.

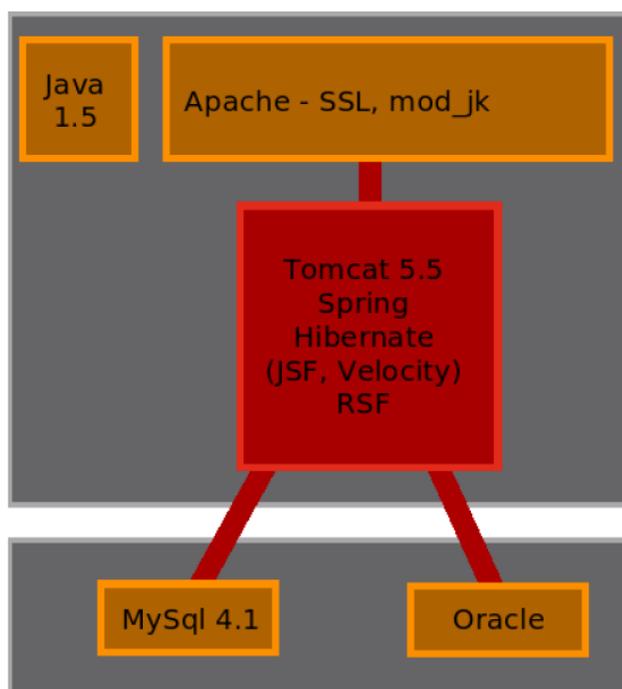


Figura 6.7 Tecnologias Utilizadas pelo Sakai (BASMAN, 2007b).

Com base nas tecnologias envolvidas, a estrutura do Sakai é ilustrada na figura 6.8 seguir.

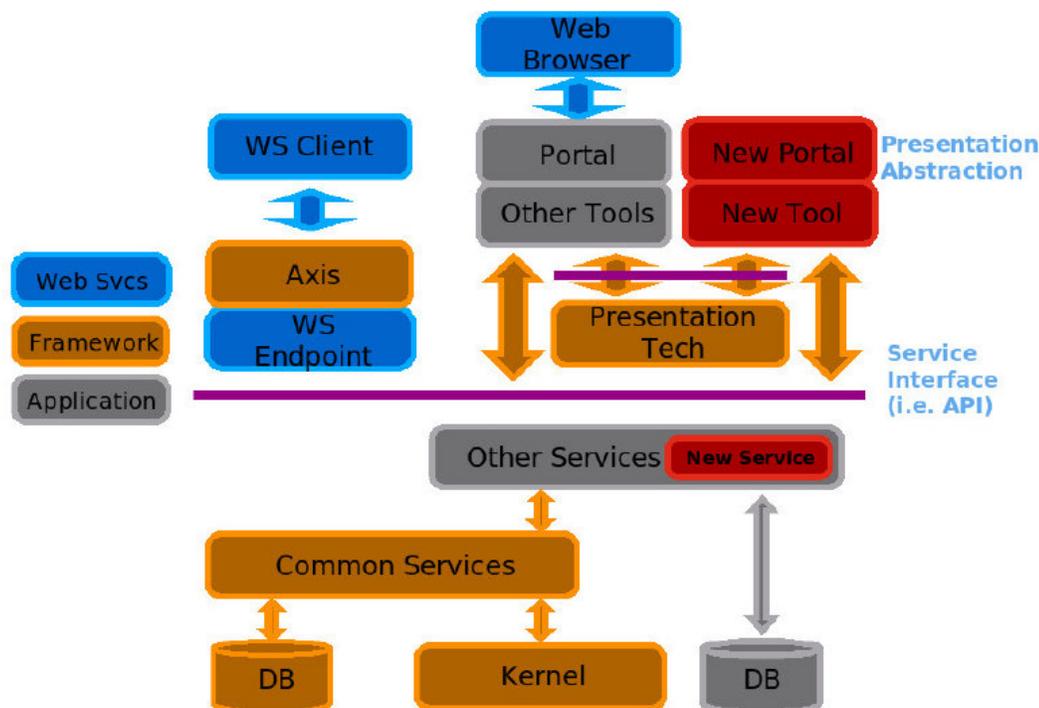


Figura 6.8 Recomendação baseada na Popularidade dos Recursos (BASMAN, 2007b)

O Sakai está estruturado em duas partes principais:

- Apresentação: camada que permite a interação de clientes por meio de tecnologias como, por exemplo, Web Services (Axis) e por meio de browsers via RSF e JSF;
- Serviços: formada por serviços comuns (banco de dados e kernel) além de outros serviços que podem ser adicionados.

Dessa forma, é possível construir ferramentas Sakai com base nas tecnologias e estrutura envolvidas. Uma ferramenta Sakai é algo parecido com um servlet. (BASMAN, 2007a, 2007b)

Existe um único contexto da aplicação Sakai global e compartilhado que possui todos os beans que representam os serviços Sakai. Para o desenvolvimento da interface gráfica da ferramenta Sakai é possível utilizar RSF - *Reasonable Server Faces* -, um framework open-source de programação Web ou mesmo JSF - *Java Server Faces*. O RSF fornece um template em XHTML e um tratamento completo do ciclo de vida. Além disso, o RSF utiliza Ajax e CSS, assim os desenvolvedores não precisam tratar `HttpServletRequest`'s, URLs ou tags HTML. Permite, também,

programação pura baseada no modelo de beans, assim como JSF, além de uma arquitetura altamente modular e plugável.

6.5 Considerações

Foram descritos neste capítulo alguns estudos de caso realizados para a abordagem proposta. Demonstrou-se a utilização de métricas implícitas com resultados interessantes para medição do comportamento dos usuários num sistema educacional web.

Os estudos de caso também mostraram a utilização de métricas explícitas e implícitas em conjunto, apresentando resultados também interessantes.

A metodologia de análise dos resultados foi a validação cruzada. Neste tipo de análise, os dados originais são retiradas em partes e são realizados testes de comparação para verificar com qual grau de consistência o modelo aproxima-se aos valores originais.

Foi apresentado um estudo particular sobre a formação de grupos ou comunidades através da utilização das métricas MFU, MRU, TAT e classificação explícita dos usuários. Neste estudo mostrou-se que a utilização apenas de métricas explícitas diverge do real comportamento do usuário (quando este comportamento é medido por valores implícitos). Os resultados deste estudo foram divulgados num artigo científico no Evento “*Frontiers in Education*”.

O capítulo seguinte apresenta as conclusões obtidas deste trabalho e as apresentações de propostas para trabalhos futuros.

Conclusões

*“Perguntaram a dois pedreiros de cantaria o que estavam fazendo. O primeiro disse: “Estou cortando essa pedra em bloco”. O segundo respondeu: “Faço parte de uma equipe que está construindo uma catedral”.*²⁸

Ao longo deste trabalho diversas áreas de pesquisa foram apresentadas e na interligação entre elas a proposta aqui descrita. A personalização em sistemas educacionais web e a utilização de sistemas de recomendação como base para tal fim foi o tema recorrente deste trabalho. A decisão de utilizar os sistemas de recomendação recaiu no problema de que são vastamente utilizados na área comercial, mas pouco estudados na área educacional. Para suprir a deficiência de usar apenas métricas explícitas, foram propostas métricas implícitas que pudessem refletir o comportamento do usuário no sistema. Além disto, a ponderação destas métricas por um especialista ou tutor, permite adaptar a proposta às situações adversas que costumam ocorrer em classes com aprendizes. A seguir, as conclusões e propostas de trabalhos futuros.

²⁸ Extraído de (BROWN-JR., 2005)

7.1 Discussão sobre o Modelo e Resultados

O Modelo de Recomendação baseado em Interesses, Preferências e Popularidade (I2P) apresenta-se como uma nova abordagem para sistemas de recomendação aplicados em Sistemas Educacionais baseados na Web.

A geração de recomendações assemelha-se a modelos existentes, mas difere quanto a utilização de métricas implícitas, previamente definidas e com a possibilidade da ponderação de tais métricas pelo tutor (ou tutores). Esta característica confere ao modelo adaptação à experiência do tutor e a sua opinião sobre os alunos que acompanha, podendo, assim, utilizar as métricas como melhor entender.

A partir das métricas propostas que são coletadas implicitamente e explicitamente e da ponderação feita pelo tutor, o modelo gera recomendações através da filtragem colaborativa e da correlação de Pearson. Desta maneira, a recomendação baseia-se no comportamento do usuário e dos seus afins no sistema, tratando-se de uma aprendizagem colaborativa, mesmo que isto fique transparente ao usuário.

A utilização das três métricas implícitas apresentadas Mais Frequentemente Utilizado (MRU), Mais Recentemente Utilizado (MFU) e Tempo Total de Acesso (TAT) permite diversas descrições sobre características de comportamento dos usuários. Por exemplo, o acesso mais frequente pode indicar que o usuário tem dificuldade de aprendizagem num determinado conteúdo ou possui muito interesse neste conteúdo, cabendo a inferência ao tutor. Além disto, já que a apresentação do conteúdo é de responsabilidade do tutor, ele pode inserir pesos nas métricas, adaptando o sistema ao comportamento que achar conveniente.

Os casos de estudo apresentados demonstram a importância da utilização de métricas implícitas além de explícitas para medir o comportamento do usuário. No entanto, a quantidade de informações pode ser vasta e este modelo restringiu-se a classificação de apenas três que pudessem ser obtidas facilmente num sistema educacional baseado na web.

Considera-se, ainda, a aplicação de tal abordagem em outros sistemas, como a TV Digital, não restringindo-a ao domínio educacional.

Como subproduto deste trabalho, apresentou-se um estudo de caso onde as informações dos usuários foram usadas para a criação de grupos com características semelhantes (medidas implícita e explicitamente).

7.2 Trabalhos futuros

O trabalho apresentado possui características multidisciplinares e as sugestões de trabalhos futuros podem compreender as diversas áreas envolvidas como a Educação à Distância, Sistemas Educacionais Baseados na Web, Sistemas de Recomendação e Personalização.

Como trabalhos futuros podem ser abordados diversos aspectos tais como:

- A definição e utilização de novas métricas para avaliação implícita do comportamento do usuário no sistema;
- A aplicação da mesma abordagem em diferentes domínios;
- A utilização de novas medidas de similaridade entre usuários e a comparação com a correlação de Pearson para domínios diversos e métricas implícitas ou explícitas;

O estudo e comparação da efetividade de utilização dos sistemas de recomendação, hipermídia adaptativa e sistemas tutores inteligentes, também compreende interesse trabalho de continuidade.

Bibliografia

ABREU, A. **Inatismo**. 2006. Filosofia. Disponível em: <<http://albertoabreu.wordpress.com/2006/07/18/inatismo/>>. Acesso em: 10/01/2007.

ANDRADE, A. F. D.; GIRAFFA, L. M. M. e VICARI, R. M. Uma Aplicação da Teoria Sociointeracionista de Vygotsky para construção de um Modelo de Aluno. *In*: XIV Simpósio Brasileiro de Informática na Educação – SBIE. Universidade Federal do Rio de Janeiro - UFRJ: 2003.

ARAGÃO, A. L. D. **Utilização de Aprendizado de Máquina para adaptação de estruturas em hipermídia adaptativa**. (Mestrado). Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2004. p.

B. F. Skinner - O cientista do comportamento e do aprendizado. *In*: **Revista Nova Escola** 1^a. 2004.

BASMAN, A. **Recap of Sakai Services**. 2007a. Sakai NWU Workshop. Disponível em: <<http://bugs.sakaiproject.org/confluence/download/attachments/32712/Recap+of+Sakai+Services.ppt>>. Acesso em: 10/02/2008.

_____. **A Recap of Sakai Services.** 2007b. Disponível em: <<http://bugs.sakaiproject.org/confluence/download/attachments/32712/Recap+of+Sakai+Services.ppt>>. Acesso em: 10/02/2008.

BAUDRILLARD, J. **A sombra das Maiorias Silenciosas - O fim do social e o surgimento das massas.** Brasiliense. 1985.

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

BROWN-JR., H. J. **De Pai para Filho.** 1. Ediouro. 2005. 152 p.

BRUSILOVSKY, P. Adaptive Hypermedia. **User Modeling and User Adapted Interaction**, v.11, p.87-110. 2001.

BRUSILOVSKY, P. e MILLÁN, E. User Models for Adaptive Hypermedia and Adaptive Educational Systems. *In: The Adaptive Web: Methods and Strategies.* Berlim, Heidelberg, New York: Springer-Verlag, v 4321, 2007

COSTA, R. M. E. M. D.; SANTOS, N. e ROCHA, A. R. C. D. Diretrizes Pedagógicas para Modelagem de Usuário em Sistemas Tutoriais Inteligentes. **TISE - Taller Internacional de Software Educativo.** 1997.

DE-BRA, P. *et al.* Making General-Purpose Adaptive Hypermedia Work. *In: WebNet.* 2000. Pág. 117-123.

FOLTZ, P. W. e DUMAIS, S. T. Personalized information delivery: an analysis of information filtering methods. **Communications of ACM**, v.35, n.12, p.51-60. 1992.

GEYER-SCHULZ, A.; HAHLER, M. e JAHN, M. Educational and Scientific Recommender Systems: Designing the Information Channels of the Virtual University. **International Journal of Engineering Education**, v.17, p.153-163. 2001.

GIRAFFA, L. M. M. **Uma arquitetura de tutor utilizando estados mentais.** (Doutorado). Instituto de Informática, Universidade Federal do Rio Grande do Sul - UFRGS, Porto Alegre, 1999. 177 p.

GODERIS, S. *et al.* Combining meta-level and logic-based constructs in web personalization. *In: Proceedings of the 6 th International Computer Science.* 2001. Pág. 57-64.

GOLDBERG, D. *et al.* Using collaborative filtering to weave an information Tapestry. **Communications of the ACM**, December, p.61-70. 1992.

GOOD, N. *et al.* Combining Collaborative Filtering with Personal Agents for Better Recommendations. *In: Conference of the American Association of Artificial Intelligence.* 1999.

GOTARDO, R. A.; TEIXEIRA, C. A. C. e ZORZO, S. D. An Approach to Predict User's Interests in Web-based Educational Systems using a Collaborative Filtering Weighted Method. *In: The 2008 International Symposium on Scientific and Engineering Computing (SEC-08).*São Paulo, Brasil: 2008a.

_____. An Approach to Recommender System in Web-based Educational Systems using Usage Mining to Predict User s Interests. *In: The 15th International Conference on Systems, Signals and Image Processing, IWSSIP.*Bratislava - Slovak Republic: 2008b.

_____. Forming Communities in Web-based Educational Systems through Users' Preferences and Interest Measuring. *In: (To be Published in) The 38th Annual Frontiers in Education (FIE) Conference.*Saratoga Springs - New York: 2008c.

HERLOCKER, J. L. **Understanding and improving automated collaborative filtering systems.** 2000. 144 p.

JAQUES, P. A. **Agentes de Software na Monitoração da Colaboração em Ambiente Telemáticos de Ensino.** (Mestrado). Departamento de Informática,

Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS, Porto Alegre, 1999. 65 p.

Jean Piaget - O biólogo que pôs o aprendizado no microscópio. *In: Revista Nova Escola* 1^a. 2004.

KAMII, C. e DEVRIES, R. **Jogos em Grupo na Educação Infantil - Implicações da Teoria de Piaget**. São Paulo: Editora Grano EPC. 1980.

KOBSA, A. User Modeling and User-Adapted Interaction. **CHI'94 Tutorial Notes**. 1994.

KOBSA, A.; KOENEMANN, J. e POHL, W. **Personalised hypermedia presentation techniques for improving online customer relationships**. United Kingdom: Cambridge University Press. vol.16:2. 2001. 111-155 p. (The Knowledge Engineering Review).

LEIDNER, D. E. e JARVENPAA, S. L. The use of information technology to enhance management school education: a theoretical view. **Society for Information Management and The Management Information Systems Research Center**, v.19, n.3, p.265-291. 1995.

LEV VYGOTSKY - O teórico do ensino como processo social. *In: Revista Nova Escola* 1^a. 2004.

LOEB, S. e TERRY, D. Information filtering. **Communications of ACM**, v.35, n.12, p.26-28. 1992.

MARTINS, J. G. **Uma arquitetura baseada em agentes inteligentes para ambientes computacionais voltados à educação à distância**. (Mestrado). Pós-Graduação em Engenharia de Produção - Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2002. 199 p.

MURUGESAN, S. e RAMANATHAN, A. Web Personalization - An Overview. *In: Active media technology: 6th International computer science conference. Proceedings* Berlim: 2001. Pág. 65-76.

PALAZZO, L. A. M. Hipermissão Adaptativa. *In: Ambientes Adaptativos: trilhando novos caminhos para a hipermissão*. Rio de Janeiro: Ciência Moderna, v, 2006, p.57-98. (Coleção Hipermissando)

PISANDELLI, G. M. V. L. A Teoria de *Maslow* e sua relação com a educação de adultos. *Psicopedagogia OnLine. Educação e Saúde Mental* 2003.

REATEGUI, E. B.; CAZELLA, S. C. e OSÓRIO, F. S. Personalização de Páginas Web através dos Sistemas de Recomendação. *In: Tópico em Sistemas Interativos e Colaborativos*. São Carlos, v, 2006

REH, J. **Pareto's Principle: The 80-20 Rule**. 2002. Disponível em: <<http://management.about.com/cs/generalmanagement/a/Pareto081202.htm>>. Acesso em: 10/09/2007.

RESNICK, P. e VARIAN, H. R. Recommender Systems. *Communications of the ACM*, v.40, March, p.55-58. 1997.

ROSATELLI, M. C. e TEDESCO, P. A. Diagnosticando o Usuário para Criação de Sistemas Personalizáveis. *In: Anais do XXIII Congresso da SBC - III Jornada de MCIA* Porto Alegre: Sociedade Brasileira de Computação, v vol. VIII, 2003, p.153-201

SAE-TANG, S. e ESICHAIKUL, V. Web Personalization techniques for e-commerce. *In: Active media technology: 6th International computer science conference. Proceedings* Berlim: 2001. Pág. 36-44.

SCHAFER, J. B.; KONSTAN, J. A. e RIEDL, J. E-Commerce Recommendation Applications. *Data Min. Knowl. Discov.*, v.5, n.1-2, 1384-5810, p.115-153. 2001.

SELF, J. The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. **International Journal of Artificial Intelligence in Education**, v.10, p.350-364. 1999.

SHARDANAND, U. e MAES, P. Social Information Filtering: Algorithms for Automating "Word of Mouth". *In*: CHI95 - Conference on Human Factors in Computing Systems. 1995. Pág. 210-217.

TOFFLER, A. **A Terceira Onda**. Record. 2001.

URBAN, P. **Platão e o Conhecimento Inato**. 2006. Disponível em: <<http://www.casadobruzo.com.br/ilustres/platao2.htm>>. Acesso em: 10/01/2007.

VENSON, E. **Um Modelo de Sistema de Recomendação Baseado em Filtragem Colaborativa e Correlação de Itens para Personalização no Comércio Eletrônico** (Mestrado). Sistemas de Computação, Universidade Federal de Santa Catarina - UFSC, Florianópolis, 2002. 132 p.

VYGOTSKY, L. S. **A Formação Social da Mente**. 1^a. São Paulo: Editora Martins Fontes. 1984.

ZAIANE, O. R. Building a recommender agent for e-learning systems. *In*: International Conference on Computers in Education. 2002.

Apêndice A – Sobre Aprendizagem: Teorias Educacionais

A.1 A Educação como Necessidade Social

A educação tem papel importante na denominada “Hierarquia das Necessidades” de Maslow. Nessa teoria é possível o estabelecimento de uma relação entre o processo de aprendizagem e a motivação do indivíduo para, através do aprendizado contínuo, promover seu desenvolvimento. Maslow apresentou uma teoria que organiza as motivações humanas em níveis de necessidades, popularmente representada pela teoria das necessidades de Maslow.

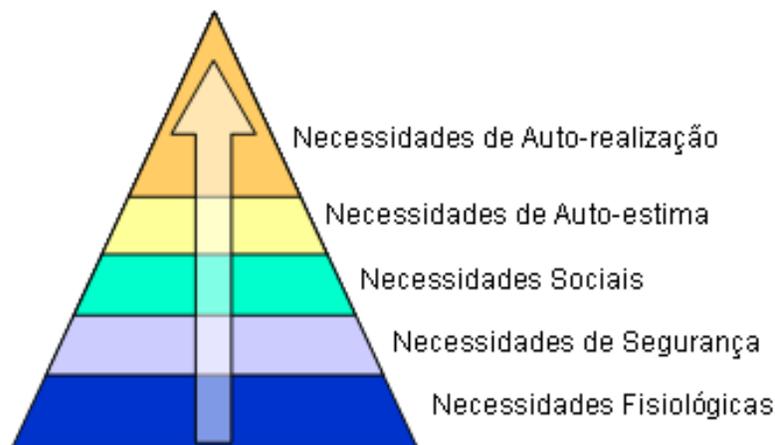


Figura A.1 Pirâmide das Necessidades segundo Maslow (PISANDELLI, 2003)

As necessidades vão desde as mais primitivas (fisiológicas) até as mais refinadas (auto-realização), cada uma delas com significados próprios. A figura A.1 representa esta pirâmide.

As necessidades fisiológicas, o primeiro nível das necessidades humanas, são: a alimentação, o abrigo, o repouso, o sexo, entre outras. São

necessidades relacionadas à sobrevivência e preservação da espécie, que nascem com os indivíduos.

As necessidades de segurança, o segundo nível, são relacionadas à segurança, a busca de proteção contra a ameaça ou privação, contra o perigo. Estas necessidades são apresentadas pelo indivíduo quando as necessidades fisiológicas estão satisfeitas (ou relativamente satisfeitas).

O próximo degrau da pirâmide são as necessidades sociais. Quando as necessidades anteriores estão relativamente satisfeitas, surge a busca pela associação, pela participação e aceitação pela sociedade – ou parte dela – a busca por amizade, afeto, amor. Neste degrau da teoria de Maslow reside a importância da abordagem deste assunto neste capítulo: letramento²⁹. No modelo de sociedade a satisfação desta necessidade é imprescindível. Cada vez mais nota-se a busca por aprimoramento pessoal. Um exemplo é o grande aumento na oferta de cursos a distância devido à procura crescente.

Para Vygotsky (1984) a alfabetização é importante a partir do momento que é possível realizar-se mais do que apenas juntar letras. A aprendizagem (coletiva) é a força motriz do progresso social, pois faz aflorar formas mais sofisticadas de comportamento humano, como raciocínio abstrato, memória ativa e a resolução de problemas (VYGOTSKY, 1984). Neste contexto verifica-se a importância da educação na sociedade e, consecutivamente, a importância do Aprendizado Eletrônico (*e-Learning*) e da Educação à Distância (EAD). É uma atividade de formação social do indivíduo.

A necessidade de auto-estima está relacionada com a frustração das necessidades anteriores. Os sentimentos como incompetência e falta de adaptação social são causados pela insuficiência das necessidades sociais. Para que isto seja evitado, o indivíduo precisa sentir-se respeitado, prestigiado, aprovado pelo mundo, assim, demonstrando um estreito relacionamento com a camada anterior da Pirâmide. (PISANDELLI, 2003)

Por fim, as necessidades de auto-realização permitem o auto-desenvolvimento. O indivíduo consegue enxergar-se como único, mesmo fazendo parte de um contexto social, mas com grande sentimento de unicidade.

²⁹ Segundo o Dicionário *Houaiss*, letramento é a representação da linguagem falada por meio de sinais; escrita; processo de alfabetização (pedagogia); pedagogia; conjunto de práticas que denotam a capacidade de uso de diferentes tipos de material escrito.

Estas necessidades estão intrinsecamente relacionadas, não dependendo da total satisfação de uma para que outra possa surgir. Como mostrado na figura A.2, as necessidades coexistem e dependem da satisfação adequada de um nível para que outro possa ser atendido. A necessidade motivadora é aquela predominante no momento e, assim que atendida, passar-se-á para a próxima na pirâmide de Maslow.

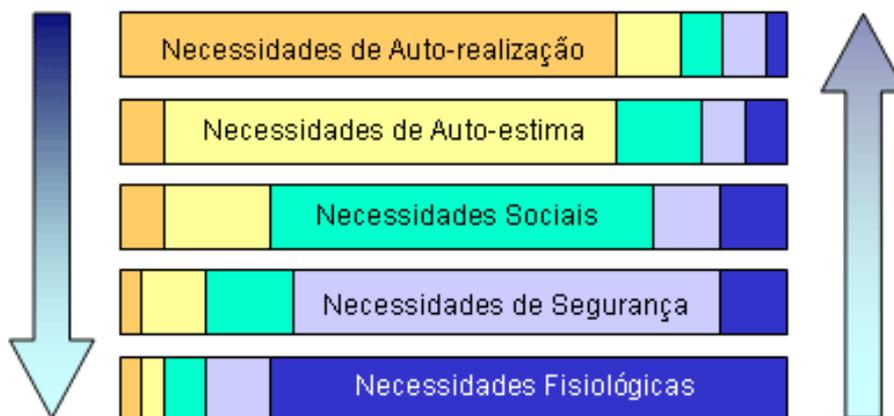


Figura A.2 Relacionamento entre as necessidades (PISANDELLI, 2003)

A importância deste trabalho reside nas necessidades sociais (contribuir com o processo de educação) e nas necessidades de segurança (garantir a inviolabilidade da privacidade dos usuários (alunos) no sistema).

A.2 Teorias Educacionais

O processo de ensino-aprendizagem justificado por diversas teorias educacionais, ao longo da história, tiveram grande e distinta importância. Atualmente, é comum a união entre algumas destas teorias, pois são complementares. O ensino é definido pelo seu momento sócio-histórico e, assim, pode ser explicado por teorias diferentes em momentos diferentes.

Enquanto as teorias educacionais visam explicar os mecanismos envolvidos na aprendizagem o projeto instrucional visa à adequação do que se quer ensinar as diferentes modalidades de ensino e diferentes aprendizes. Dentre

diversas teorias educacionais existentes as mais utilizadas são o inatismo, o comportamentalismo, o construtivismo e o sócio-construtivismo.

A.2.1 Inatismo

“Segundo Platão, conhecer é recordar verdades que já existem em nós - teoria que pode ser atestada sempre que nos deixamos guiar pela voz do inconsciente.”³⁰

O Inatismo se fundamenta numa concepção racionalista e idealista do ser humano. Na concepção racionalista, o único meio para se alcançar o conhecimento é através da razão, que é inata e imutável a todos os seres humanos. Na concepção idealista o mundo das idéias e o real são confundidos. Esta teoria enfatiza fatores hereditários e de maturação, pois entende que o ser humano possui potencialidades, dons e aptidões ao nascer que serão desenvolvidos ao longo de seu amadurecimento biológico. O homem nasce pronto e nada após seu nascimento poderá alterar sua personalidade. Nem mesmo o contato social. Ao aprender o ser humano apenas aprimora aquilo que já sabia (já era inato). Ao aprender, descobre-se o que já era sabido. (MARTINS, 2002; ABREU, 2006)

O maior representante Inatista foi Platão³¹. Segundo ele o aprendizado é um processo natural de descobertas que entrevem as idéias e a racionalidade (latentes) guardadas em nosso mundo interior. No seu livro a República, Platão fundamenta a reminiscência através da alegoria de *Er*, um pastor que, após ser morto em batalha, tem seu corpo encontrado intacto e volta à vida narrando os fatos que viu no além, contando sobre os juízes que separavam as almas boas das ruins, dando-lhes suas sentenças.

³⁰ Fonte: (URBAN, 2006)

³¹ Platão nasceu em 428-7 a.C., na cidade-estado de Atenas. Foi o mais importante discípulo de Sócrates e, aos seus 40 anos de idade, fundou sua academia, local onde passou o resto de sua vida. A academia de Platão estenderia seu funcionamento durante 900 anos, a mais longa existência de instituição educacional até hoje registrada. Aristóteles foi um de seus célebres discípulos. Em sua longa jornada Platão conheceu e foi influenciado por figuras importantes como Euclides de Megara, Heráclito, Parmênides, Teodoro de Cirene e Pitágoras, além de Sócrates, obviamente. Sócrates o “despertou” para o costume de refletir sobre o homem e seus problemas éticos e também para a arte de descobrir a verdade através de perguntas.

Esta metáfora de Platão, que até hoje influencia várias correntes religiosas, visa explicar como o conhecimento pode preexistir nas pessoas. (URBAN, 2006)

O papel da educação é a facilitação promovida pelo professor para que esta essência inata manifeste-se e, assim, o espírito criativo dos alunos possa surgir. A concepção inatista propõe práticas pedagógicas pouco desafiadoras, pois subestima a capacidade intelectual dos seres humanos. Estando, assim, livre da responsabilidade de contribuição e intervenção no processo de desenvolvimento do aluno.

Para Aristóteles, discípulo de Platão, nada é original, tudo é uma composição do que já foi visto, ouvido ou sentido. Toda idéia é uma reinvenção de idéias já pensadas antes, mudando-se apenas o ponto de vista.

A.2.2 Comportamentalismo

“A educação é o estabelecimento de comportamentos que serão vantajosos para o indivíduo e para outros em algum tempo futuro”³²

O comportamentalismo considera o homem como um ser passivo e os estímulos do ambiente externo é que o governam. O comportamento é o que pode ser observado quando se submete o indivíduo a algum estímulo. Para Skinner³³ é possível condicionar o comportamento dos animais, dentre eles o ser humano.

³² Esta citação foi retirada de (LEV VYGOTSKY - O teórico do ensino como processo social, 2004).

³³ O psicólogo *Burrhus Frederic Skinner* é o mais representativo teórico do comportamentalismo (o que chamou de *behaviorismo*), sendo muito influenciado pelo empirismo, atividade onde o conhecimento provém da experiência e que possui representantes como *Sir Francis Bacon, Tomás Hobes, John Locke, Auguste Comte*, entre outros. Nascido no estado norte-americano da Pensilvânia, em 1904, na cidade de *Susquehanna*. Foi criado com disciplina severa, mas foi um rebelde, cujos interesses foram, na adolescência, a poesia e a filosofia. Formou-se em língua inglesa na Universidade de Nova York e, posteriormente, redirecionou sua carreira para psicologia. Em Harvard, onde cursou psicologia, teve seus primeiros contatos com o *behaviorismo*. Foi o responsável pelas chamadas “caixas de skinner”, ambientes fechados onde conduzia seus experimentos de condicionamento com animais (ratos e pombos). Há um boato de que *Skinner* tenha criado um berço climatizado para sua filha e conduzido experiências, semelhantes às de seus animais de laboratório, com ela. Em 1948 tornou-se professor em Harvard e ocupou o cargo até o final de sua vida. Faleceu em 1990, militante sempre a favor do *behaviorismo*. (B. F. Skinner - O cientista do comportamento e do aprendizado, 2004)

Através de técnicas de condicionamento (estímulo – resposta – reforço) buscava prever e controlar o comportamento de animais. O estímulo de uma determinada situação resulta num comportamento, e isto é passível de manipulação (MARTINS, 2002).

Os comportamentalistas acreditam que o único meio pelo qual o conhecimento humano pode ser obtido é através da observação³⁴. O ser imaturo aprende observando aquele que sabe. Assim, o conhecimento só é possível através do que observável e não pela especulação. O conhecimento é um repertório de comportamentos que são manifestados por estímulos distintos e da probabilidade do comportamento especializado. Neste processo de ensino, o reforço tem papel de extrema importância. O professor é o principal responsável por planejar as chamadas “contingências de reforço”, atuando como agente central no processo de ensino-aprendizagem (COSTA, SANTOS e ROCHA, 1997).

Na teoria de Skinner, os principais aspectos são: a aprendizagem pelo ensino programado, estímulos positivos, reforço, observação do aprendizado pelo comportamento apresentado e organização do conteúdo em níveis de dificuldade. O indivíduo não é capaz de intervir no meio e transformá-lo, nem transformar-se. É passivo, condicionado ao sucesso ou fracasso de acordo com seu treinamento.

A extensão do comportamentalismo que é também uma ponte com o construtivismo (que será visto a seguir) é chamada de neo-comportamentalismo. A aprendizagem incorpora teorias de processamentos mentais internos. Há uma visão mais cognitiva dos processos mentais. O processo de aprendizagem apresenta, além da mudança comportamental, a permanência desta mudança. Os estímulos do ambiente são considerados fatores de entrada para a aprendizagem, enquanto as modificações de comportamento (o desempenho) são os fatores de saída.

O maior representante desta concepção é o psicólogo americano Robert Gagné. Para Gagné a aprendizagem envolve quatro elementos: o aprendiz, a situação, o comportamento explícito do aprendiz e a mudança interna. As habilidades intelectuais dos indivíduos são compostas de “habilidades menores” e de forma hierárquica, que são iniciadas com conexões de estímulo-resposta, tornando-se regras e, posteriormente, a solução de problemas.

³⁴ Aristóteles já afirmava que o que está na inteligência precisa passar pelos sentidos “aprioristicamente”. (MARTINS, 2002)

“Quando houver domínio sobre a ciência do comportamento, ela será a única alternativa para a sociedade planejada”³²

A.2.3 Construtivismo

“O conhecimento não pode ser uma cópia, visto que é sempre uma relação entre objeto e sujeito”³²

As interações com o ambiente externo são responsáveis pela construção do conhecimento no Construtivismo. O aluno deixa de ser sujeito passivo e torna-se ativo na aprendizagem, experimentando, pesquisando, duvidando e desenvolvendo o raciocínio. O professor ou tutor é agora o estimulador de experiências e cria as estratégias para a busca de respostas.

Os grandes teóricos do construtivismo são *Jean Piaget*³⁵ (biólogo) e *Jerome Bruner* (psicólogo). Suas formações falam pela abordagem e pela importância de cada um no Construtivismo.

Para Piaget, a inteligência é desenvolvida a partir da maturação do indivíduo ao ambiente. A maturação é formada pela adaptação e pela organização. Na adaptação o indivíduo adquire novos conhecimentos, assimilando-os, e a estrutura do conhecimento é modificada para acomodar o que foi aprendido. A organização das estruturas cognitivas de forma coerente realiza a estruturação da informação, gerando novos elementos internos da inteligência. Assim, a aprendizagem é constituída por ações que um indivíduo realiza sobre os objetos, sobre o meio ambiente. Pela dúvida e interação. E a maturação é formada por sucessos e fracassos, ambos importantes para que se alcance um aprendizado eficaz. O erro não é um tropeço, mas um propulsor na busca da conclusão correta.

Já Bruner contempla a aprendizagem “por descoberta”, explorando alternativas e o chamado currículo espiral. Ao explorar alternativas faz-se necessário que o ambiente (ou o conteúdo) de ensino possa proporcionar novas formas de

³⁵ O suíço, *Jean Piaget* nascido em *Neuchâtel* em 1896 publicou, logo aos 10 anos, seu primeiro artigo científico, sobre um pardal albino. Seu interesse precoce também foi desperto para áreas como filosofia, religião e ciência. Formou-se em biologia na universidade de *Neuchâtel*. *Piaget* publicou mais de 50 livros. Até o final de sua vida recebeu diversos títulos honorários de universidades americanas e européias. Morreu em 1980, em Genebra, Suíça. (Jean Piaget - O biólogo que pôs o aprendizado no microscópio, 2004)

representações, para que o aluno possa descobrir relações e inferir princípios. O currículo em espiral permite ao aluno aprender o mesmo assunto em níveis de dificuldade (profundidade de abordagem) diferente. O desenvolvimento intelectual também depende da maturação na representação e integração de informações. A teoria de Bruner acrescenta ao construtivismo a aprendizagem em espiral e a exploração de alternativas.

A concepção construtivista não releva os aspectos histórico-culturais na construção do conhecimento pelo indivíduo. A ênfase desta concepção está na relação do aluno (aprendiz) com o mundo que o cerca e resolução de problemas deste mundo. A escola deve estimular as dúvidas e não a repetição (como no comportamentalismo). (MARTINS, 2002; Jean Piaget - O biólogo que pôs o aprendizado no microscópio, 2004)

“Se o indivíduo é passivo intelectualmente, não conseguirá ser livre moralmente”³²

A.2.4 Sócio-Construtivismo

“O saber que não vem da experiência não é realmente saber”³²

Também chamado por autores de “Colaborativismo” ou “Interacionismo”, o sócio-construtivismo consiste na premissa básica do aprendizado através da interação entre pessoas. Desta forma, a “colaboração social” busca o encorajamento dentre os estudantes na exploração de suas idéias e a defesa das mesmas. O processo de ensino-aprendizagem torna-se, assim, um conjunto de ações de interação, onde a troca de informações é o maior foco. Há a concepção da construção gradativa em um ambiente histórico e social. O desenvolvimento cognitivo e cultural aparece no nível social e, posteriormente, no nível individual. O principal representante desta teoria é o psicólogo *Lev Vygotsky*³⁶. Sua oposição

³⁶ *Lev Semenovitch Vygotsky*, nascido em 1896, numa pequena cidade perto de *Minsk* - a capital da *Bielo-Rússia* - chamada *Orsha*. *Vygotsky* teve uma sólida formação, pois seus pais eram de uma família judaica culta e com boas condições econômicas. Aos 18 anos, matriculou-se no curso de medicina em Moscou, mas acabou cursando a faculdade de direito. Após formar-se, voltou a *Bielo-Rússia*, em 1917, lecionou literatura, estética e história da arte, e fundou um laboratório de psicologia. Nessa área, rapidamente ganhou destaque, graças ao seu pensamento inovador e sua intensa

teórica ao biólogo *Jean Piaget* surge na ênfase do social e dos processos interpessoais. Segundo *Vygotsky* a ausência do outro torna impossível a construção do homem. Para ele, o processo de aprendizagem humano acontece numa relação dialética entre a sociedade e o sujeito. O homem e o ambiente modificam-se mutuamente. Um conceito chave na teoria é a mediação. Toda interação com o mundo é realizada através de instrumentos técnicos como ferramentas agrícolas e a linguagem. Ambos permitem não apenas a interação do sujeito com o mundo, mas com outros também.

Nesta teoria de aprendizagem torna-se necessária a presença de um mediador, ou de outra forma, de um professor, mas ativo do que prevê *Piaget*. Quando se internaliza um procedimento há a apropriação do mesmo. O aprendizado não fica subordinado às estruturas intelectuais apenas, mas vale-se de outros para que aconteça. As métricas para esta teoria são: o desenvolvimento real e o potencial. A “zona de desenvolvimento proximal”, de *Vygotsky*, trata da distância entre o desenvolvimento real de uma criança e aquilo que tem o potencial de aprender. Este potencial é descrito pela competência que se pode desenvolver com a ajuda de um adulto (ou alguém mais experiente), enquanto que o nível de desenvolvimento real é determinado pela capacidade de solução de problemas individualmente. A interação social é o fator que potencializa a “zona de desenvolvimento proximal” e o desenvolvimento real é tratado como as funções mentais que um indivíduo já tem estabelecido, decorrentes de etapas cumpridas de desenvolvimento. (VYGOTSKY, 1984; COSTA, SANTOS e ROCHA, 1997; MARTINS, 2002; LEV VYGOTSKY - O teórico do ensino como processo social, 2004)

*“O caminho do objeto até a criança e desta até o
objeto passa por outra pessoa”²*

atividade, produzindo mais de 200 trabalhos, dois científicos. Em 1925, publicou *A Psicologia da Arte*, um estudo sobre *Hamlet*, de *William Shakespeare*, cuja origem é sua tese de mestrado. O trabalho de *Vygotsky* foi fortemente influenciado por momentos políticos diferentes em sua vida. Acompanhou de perto os acontecimentos que culminaram na revolução russa de 1917. Após a revolução intensificou seus estudos sobre psicologia. Faleceu em 1934, pelas complicações de tuberculose. Após sua morte, em 1936, teve seu trabalho censurado pela ditadura de *Stalin*, permanecendo assim por 20 anos. (LEV VYGOTSKY - O teórico do ensino como processo social, 2004)

A.2.5 Comparação entre as Teorias Educacionais

O comportamentalismo contrapõe-se ao inatismo, afirmando que o conhecimento só pode ser adquirido através da experiência (um outro nome para condicionamento) e não nasce com o indivíduo. Enquanto isto, para o construtivismo e o sócio-construtivismo o conhecimento é aprendido, de forma interna e dependente de algum tipo de interação.

Nota-se que o construtivismo e o sócio-construtivismo são centrados no aluno e o indivíduo é visto como agente ativo no processo de aprendizagem. Ao contrário, no comportamentalismo o indivíduo é visto como ser passivo, sujeito aos acontecimentos do ambiente e influenciado por eles. Esta mudança de abordagem reflete também mudança nos paradigmas educacionais. A tabela A.1 resume as principais diferenças e relações entre as teorias apresentadas.

Tabela A.1: Teorias Educacionais - Baseado em (LEIDNER e JARVENPAA, 1995) apud (MARTINS, 2002)

Teoria	Definição	Objetivo	Premissas	Instrutor / Professor
Inatismo	Não há aprendizado, apenas o desenvolvimento de aptidões no decorrer do amadurecimento.	Facilitar as manifestações do conhecimento que o aluno possui.	O conhecimento já é inerente ao ser humano, que nasce pronto e imutável.	Tem o papel de auxiliar o aluno a manifestar o conhecimento que possui.
Comportamentalismo	O aprendizado é resultado do condicionamento e absorção não crítica do conhecimento.	Memorização do conhecimento. Transferência do professor para o aluno.	O professor detém todo o conhecimento. Os estudantes aprendem estudando de forma intensiva.	Tem o papel de controlar o material e a velocidade de aprendizado, estimulando o estudante.
Construtivismo	O aprendizado é o processo de construção de conhecimento por um indivíduo.	Formar conceitos abstratos para representar a realidade, dar significado a eventos e informações.	Indivíduos aprendem melhor quando descobrem sozinhos e quando controlam a velocidade do aprendizado.	Aprendizado centrado nas atividades dos alunos. Instrutor mais ajuda do que direciona.
Sócio-Construtivismo	Aprendizado emerge através de entendimento partilhado por mais de um	Promover habilidades grupais, comunicação, participação,	Envolvimento e trocas de saberes são fundamentais no aprendizado.	Orientado para a comunicação. Instrutor atua como questionador e

	aluno	capacidade de ouvir. Promover socialização.	Alunos têm algum conhecimento anterior sobre o assunto.	líder da discussão.
--	-------	---	---	---------------------

Segundo Kamii e Devries (1980) o comportamentalismo é visto, muitas vezes, como uma teoria mutuamente exclusiva quando comparado ao construtivismo. Contudo, estas teorias podem ser complementares, sendo a teoria comportamentalista um subconjunto do construtivismo (KAMII e DEVRIES, 1980). A figura A.3 retrata a abordagem dada por Kamii e Devries.

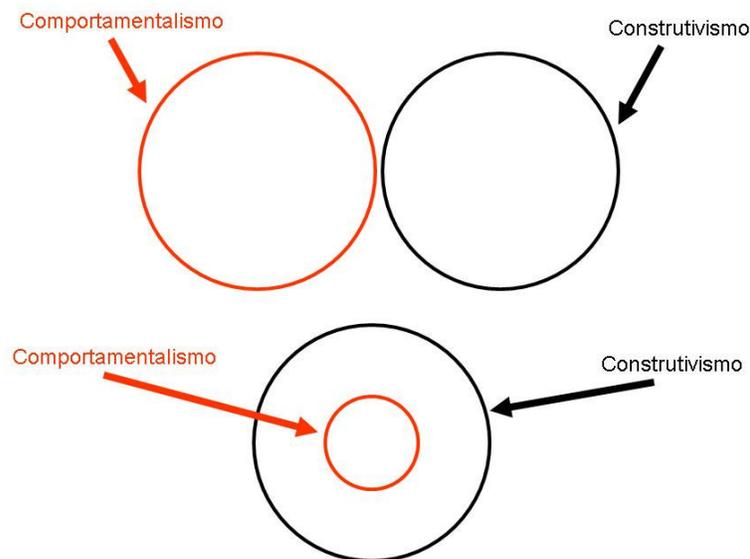


Figura A.3 Relacionamento entre o comportamentalismo e o construtivismo (KAMII e DEVRIES, 1980)

O construtivismo e o sócio-construtivismo possuem bases muito comuns, diferindo, na forma de interação (com o ambiente ou com outros indivíduos) e, assim, também pode ser vistos como complementares.