UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE CIÊNCIA DA COMPUTAÇÃO

DANIELA MESQUITA ZAMBERLAN

Recomendação de Mensagens no Fórum do Moodle

Trabalho de Graduação.

Prof. Dr. Leandro Krug Wives
Orientador

Porto Alegre, dezembro de 2011.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. Carlos Alexandre Netto
Vice-Reitor: Prof. Rui Vicente Oppermann
Pró-Reitora de Graduação: Profa. Valquiria Link Bassani
Diretor do Instituto de Informática: Prof. Luís da Cunha Lamb
Coordenador do CIC: Prof. Raul Fernando Weber
Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS **5**
LISTA DE FIGURAS **6**
ABSTRACT **8**
1 INTRODUÇÃO **9**
1.1 Objetivos **10**
1.2 Contribuição **10**
1.3 Estrutura do texto **11**
2 SISTEMAS DE RECOMENDAÇÃO **12**
2.1 Visão Geral **12**
2.2 Técnicas de Recomendação **13**
2.2.1 Filtragem Baseada em Conteúdo 13
2.2.2 Filtragem Colaborativa 14
2.2.3 Filtragem Híbrida 15
2.3 Exemplos de Sistemas de Recomendação **16**
2.3.1 Amazon.com 16
2.3.2 CDNow 17
2.3.3 eBay 17
2.3.4 Google 18
2.4 Sistemas de recomendação em fóruns eletrônicos **18**
2.4.1 Recomendação baseada em regras 18
2.4.2 Recomendação baseada em categorias 19
3 MOODLE **21**
3.1 Arquitetura do Moodle **21**
3.1.1 Visão Geral de uma Instalação do Moodle 22
3.1.2 Moodle como um Sistema Modular 22
3.1.3 O Núcleo do Moodle 22
3.1.4 Organização do Código 23
3.1.5 A Base de Dados do Moodle 23
3.1.6 Criação de um Módulo no Moodle 24
3.1.7 Pacotes de Linguagem 25
4 IMPLEMENTAÇÃO DO FÓRUM COM RECOMENDAÇÃO **27**
4.1 Visão Geral **27**
4.2 Fórum do Moodle **27**
4.3 Ferramenta Proposta **29**
4.3.1 Gerenciamento de Categorias 30
4.3.2 Recomendação de Mensagens 32
4.3.3 Visualização de Relatórios 37
4.3.4 Principais Alterações Efetuadas no Banco de Dados 38

5 TESTES DA FERRAMENTA DE RECOMENDAÇÃO 40
6 CONCLUSÃO..... 51
REFERÊNCIAS 53

LISTA DE ABREVIATURAS E SIGLAS

API	*Application Programming Interface*
AVA	Ambiente Virtual de Aprendizado
BD	Banco de Dados
CD	*Compact Disc*
CMS	*Course Management System*
CSS	*Cascading Style Sheets*
EAD	Educação a Distância
FBC	Filtragem Baseada em Conteúdo
FC	Filtragem Colaborativa
FH	Filtragem Híbrida
FI	Filtragem de Informações
HTML	*HyperText Markup Language*
LMS	*Learning Management System*
MOODLE	*Modular Object-Oriented Dynamic Learning Environment*
RI	Recuperação de Informações
RSLP	Removedor de Sufixos para a Língua Portuguesa
SGBD	Sistema de Gerenciamento de Banco de Dados
SQL	*Structured Query Language*
SR	Sistemas de Recomendação
UFRGS	Universidade Federal do Rio Grande do Sul
URL	*Uniform Resource Locator*
VLE	*Virtual Learning Environment*
XML	*Extensible Markup Language*

LISTA DE FIGURAS

Figura 2.1: Tela com recomendações da Amazon.com 17
Figura 2.2: Perfil de um vendedor no site eBay 18
Figura 4.1: Diagrama do Banco de Dados do Fórum 28
Figura 4.2: Gerenciamento de Categorias 32
Figura 4.3: Fórmula da Similaridade Cosine 35
Figura 4.4: Fórmula da média por operadores “fuzzy” 37
Figura 4.5: Fórmula do cálculo do grau de igualdade entre os pesoas 37
Figura 4.6: Diagrama com as modificações realizadas no Banco de Dados do módulo “newforum” 39
Figura 5.1: Categorias iniciais do Curso de Teste 40
Figura 5.2: Mensagens inseridas no fórum 41
Figura 5.3: Primeira mensagem sendo inserida no fórum 42
Figura 5.4: Primeira recomendação realizada 43
Figura 5.5: Segunda mensagem sendo inserida no fórum 44
Figura 5.6: Segunda recomendação realizada 45
Figura 5.7: Mensagens do fórum após recusa de recomendação 46
Figura 5.8: Categorias não-gerenciadas 47
Figura 5.9: Categorias do curso após os testes 48
Figura 5.10: Relatório de recomendações 49

RESUMO

Em Educação a Distância, ferramentas de grande importância são os Ambientes Virtuais de Aprendizado, sendo que, nesse cenário, o Moodle é uma das ferramentas que mais tem se destacado. Uma das formas mais importantes de interação com o Ambiente Virtual de Aprendizado é o fórum de discussão, onde os alunos podem enviar suas dúvidas para que sejam respondidas por outros alunos ou pelo professor. Conforme o número de usuários e mensagens enviados ao fórum cresce, mais difícil se torna verificar se uma dúvida já foi solucionada. Além disso, o encaminhamento de mensagens semelhantes faz com que os professores, tutores e alunos tenham que responder diversas vezes as mesmas dúvidas. Com o objetivo de minimizar esse problema, neste trabalho é proposta uma ferramenta de recomendação de mensagens para o fórum do Moodle, utilizando técnicas de Sistemas de Recomendação. Tal ferramenta visa recomendar aos usuários do fórum mensagens semelhantes à mensagem que eles estão enviando. Dessa forma, as dúvidas podem ser solucionadas antes mesmo de serem enviadas ao fórum, poupando tempo e trabalho dos usuários para resolver dúvidas que já foram respondidas. Para isso, foi utilizada a técnica conhecida por Filtragem Baseada em Conteúdo, além da medida de similaridade *Cosine* para determinar o quanto as mensagens estão relacionadas. O documento também apresenta diferentes técnicas utilizadas em Sistemas de Recomendação, além de descrever o Moodle e explicar como um novo módulo pode ser criado para tal ambiente. A ferramenta proposta é detalhada e testada com o intuito de validar seu funcionamento.

Palavras-Chave: Sistemas de recomendação; Filtragem Baseada em Conteúdo; Moodle; Educação a Distância.

Recommendation of Messages in the Moodle's Forum

ABSTRACT

In Distance Education, tools of great importance are the Virtual Learning Environments and, in this scenario, Moodle is one tool that has most distinguished itself. One of the most important ways of interacting with the Virtual Learning Environment is the discussion forum, where the students can send their doubts to be answered by other students or by the teacher. As the number of users and messages sent to the forum grow, more difficult it becomes to check whether a doubt is already answered. In addition, the sending of similar messages makes the teachers, tutors and students have to answer the same questions several times. With the purpose of minimizing this problem, this work proposes a tool for recommending messages to the Moodle's forum, using techniques of Recommender Systems. This tool aims to recommend to Moodles's forum users messages similar to the ones they are sending. In this way, the doubts can be resolved before they are sent to the forum, saving time and effort of users to answer questions that are already answered. For this, the technique known as Content-Based Filtering was used, together with the *Cosine* similarity measure, to determine how the messages are related. The document also presents different techniques used in Recommender Systems, as well as a description of Moodle and a explanation of the creation of a new module to this environment. The proposed tool is detailed and tested in order to validate its operation.

Keywords: Recommender Systems, Content-Based Filtering, Moodle, Distance Education.

1 INTRODUÇÃO

Atualmente, a Educação a Distância (EAD) tem se tornado um mecanismo muito importante de aprendizado, seja como ferramenta principal ou como apoio a cursos presenciais. Como ferramenta principal, ela é utilizada principalmente por pessoas que não tem disponibilidade para comparecer em aulas presenciais, seja por questão de tempo ou de deslocamento. Já em cursos presenciais, ela é utilizada como ferramenta de apoio, seja para distribuição de material, realização de tarefas ou esclarecimento de dúvidas. Segundo Moran (2011), a EAD é uma opção cada vez mais importante para aprender ao longo da vida, para a formação continuada, para aceleração profissional ou para conciliar estudo e trabalho. Ela está modificando todas as formas de ensinar e aprender, inclusive as presenciais, que estão cada vez mais adotando metodologias semipresenciais como forma de apoio ao ensino tradicional.

Segundo o Censo da Educação Superior de 2008 do INEP/MEC, houve um aumento no 96,8% no número de matrículas em cursos a distância no período de 2007 a 2008, fazendo com que essa modalidade passe a representar 14,3% do setor no Brasil. Outro dado que merece destaque é o número de concluintes, que nesse mesmo período cresceu 135%. Isso indica que a EAD não é só um modismo passageiro, mas uma forma de atender às reais necessidades do país (HESPANHA, 2009). No ano de 2010, o Ministério da Educação divulgou dados de um estudo que indicam que um em cada cinco novos alunos de graduação no país ingressa em um curso na modalidade EAD (IEA, 2010).

Na verdade, segundo Oliveira et al. (2009), a EAD existe desde o século XIX, através da utilização do correio para transmitir informação, sendo uma alternativa na educação não formal. Mais tarde, ela foi utilizada na educação convencional, através da sua associação com o rádio ou televisão. Com o surgimento da Internet, criaram-se novas oportunidades para o processo de ensino-aprendizagem na educação a distância. Dessa forma, os professores têm a possibilidade de buscar novas formas de ensino e os alunos novas formas de aprender.

O uso da internet foi uma das principais evoluções que ocorreram na Educação a Distância, pois ela é uma importante ferramenta de facilitação e difusão do ensino e da aprendizagem. (GIRAFFA apud OLIVEIRA et al., 2009). Nesse contexto, há os Sistemas de Gerenciamento de Cursos ou Ambientes Virtuais de Aprendizado (AVA), que fornecem várias formas de interação para facilitar o compartilhamento de informações e comunicação entre participantes de um curso. Eles permitem que os tutores distribuam informações para os estudantes, preparem materiais, tarefas e atividades, realizem aulas a distância e permitem aprendizado colaborativo. Alguns exemplos de sistemas comerciais são Blackboard, WebCT e Top-Class, enquanto exemplo de sistemas gratuitos são Moodle, Ilias e Claroline. Hoje em dia, um dos

sistemas mais usados é o Moodle (ROMERO et al., 2008), e por esse motivo será o sistema utilizado neste trabalho (e detalhado mais adiante). Apesar de serem principalmente considerados como ferramentas para a Educação a Distância, também são usados para dar suporte a salas de aula reais (SIMONSON, 2007).

Uma funcionalidade importante nos AVA são os fóruns, que são atividades de discussão e construção de conteúdo. Grande parte das dúvidas dos alunos são resolvidas nos fóruns, onde também as opiniões podem ser melhor elaboradas e podem ser feitas reflexões mais profundas sobre os assuntos. Além disso, eles são mediados por um tutor, e são o principal canal de reflexão coletivo do curso (SCORSOLINI-COMIN et al., 2009).

Com o uso cada vez maior de EAD, e, consequentemente, fóruns de discussão, surge um problema: a quantidade de mensagens ou dúvidas enviadas aos fóruns é muito grande, logo é difícil descobrir se uma dúvida já foi respondida anteriormente. Dessa forma, o aluno acaba enviando a sua dúvida mesmo sem saber se ela já foi respondida, fazendo com que o tutor perca tempo para respondê-la novamente, tempo esse que poderia ser utilizado para responder questões novas. À medida que o número de alunos e mensagens enviadas ao fórum vai crescendo, esse problema se torna maior. Com isso, muitas vezes, o tutor não consegue responder todas as dúvidas e alguns alunos acabam sendo prejudicados.

1.1 Objetivos

O objetivo principal deste trabalho é, considerando fóruns de discussão do AVA Moodle, utilizar técnicas de análise de similaridade entre textos e recomendação para recomendar mensagens que contenham respostas à mensagem que está sendo enviada por um usuário. Dessa forma, espera-se que uma dúvida possa ser resolvida antes de a mensagem ser postada em um fórum, fazendo com que apenas mensagens não relacionadas a mensagens já existentes sejam postadas no fórum, diminuindo a quantidade de mensagens e, consequentemente, o trabalho do tutor, que passa a atender mais alunos de forma qualificada.

Como objetivos específicos têm-se os seguintes:

- Pesquisar sobre sistemas de recomendação, entendendo os tipos de sistemas existentes e suas diferenças, vantagens e desvantagens de sua utilização, etc., com foco em fóruns de discussão;
- Estudar a API do Moodle, para saber de que forma é possível aplicar técnicas de recomendação nesse ambiente;
- Modelar, propor e definir uma arquitetura de fórum de discussão com recomendação;
- Implementar o fórum, utilizando a API do Moodle;
- Testar a implementação com usuários reais, para determinar se ela cumpriu seu objetivo.

1.2 Contribuição

A contribuição principal deste trabalho é proporcionar uma nova forma de utilização de fóruns de discussão no ambiente do Moodle. Através da recomendação de

mensagens no fórum, tutores e alunos se beneficiarão. Os tutores terão menos mensagens para analisar e responder e, além disso, como mensagens parecidas serão filtradas, eles sentirão que seu trabalho está sendo mais útil e produtivo. Por outro lado, os alunos terão o benefício de poder esclarecer suas dúvidas instantaneamente, sem a necessidade de esperar até que o tutor responda ou de procurar a resposta em um fórum com muitas mensagens. Quanto maior o número de mensagens no fórum, mais esses benefícios serão sentidos.

Como não existe uma funcionalidade assim no Moodle, ela pode ser desenvolvida como um módulo para o mesmo, que pode ser facilmente incorporado e usado. Levando em conta os motivos citados anteriormente, acredita-se que o desenvolvimento desse módulo será de grande utilidade. Após a conclusão do trabalho, o módulo pode ser disponibilizado para download no site oficial do Moodle (várias funcionalidades extras são distribuídas assim), para que outros usuários dessa plataforma também possam usufruir desses benefícios.

1.3 Estrutura do texto

No capítulo 2, serão apresentados os Sistemas de Recomendação, com suas divisões de acordo com o método de filtragem utilizado e exemplos de sistemas que utilizam técnicas de recomendação, além de aplicações de Sistemas de Recomendação em fóruns eletrônicos. O capítulo 3 apresentará as características principais do AVA Moodle, bem como a criação de um módulo básico para ele. No capítulo 4, será apresentada e detalhada a ferramenta desenvolvida, que busca minimizar os problemas existentes nos fóruns, apresentados anteriormente. Com o intuito de validar a ferramenta proposta e desenvolvida, no Capítulo 5 serão apresentados testes realizados com a mesma. Por fim, no Capítulo 6, será apresentada a conclusão do trabalho, com sugestões para trabalhos futuros.

2 SISTEMAS DE RECOMENDAÇÃO

Segundo Sarwar et al. (2001), o avanço da tecnologia facilitou muito a publicação e distribuição de informação, por isso a quantidade de informação no mundo está crescendo mais rápido do que a capacidade humana de processá-la. É cada vez mais necessário criar tecnologias que ajudem as pessoas a descobrir, entre toda a informação disponível, aquela que a interessa mais. Isso é chamado de Sobrecarga de Informação. Por isso, são necessárias ferramentas para realizar a filtragem de informação relevante de forma eficaz. Como uma forma de tentar lidar com esse problema, neste capítulo serão apresentados os Sistemas de Recomendação (SR), que, de acordo com Bezerra (2002) filtram informações relevantes para um usuário com base em seu perfil.

2.1 Visão Geral

Na vida cotidiana, em situações onde não se tem experiência sobre as alternativas existentes, as pessoas baseiam-se em recomendações de outras pessoas para tomar decisões. Os Sistemas de Recomendação auxiliam nessa tarefa (RESNICK, VARIAN, 1997).

Segundo Torres (2004), um Sistema de Recomendação é um sistema que recomenda, entre uma enorme quantidade de itens, aqueles que estão de acordo com as preferências de um usuário. Um item é qualquer coisa que o usuário possa ter interesse, como um filme, um livro ou restaurante. O interesse do usuário em um item pode ser representado de duas formas principais: através de notas, que são avaliações numéricas dos itens ou, no caso de itens textuais, através da similaridade entre um item e itens que o usuário costuma consumir (GARIN et al., 2006).

Os desenvolvedores do primeiro SR, o “Tapestry”, utilizavam o termo “Filtragem Colaborativa”. No entanto, o termo “Sistemas de Recomendação” é atualmente preferido por dois motivos: as pessoas que fazem a recomendação podem não colaborar explicitamente com quem a recebe, já que eles podem ser desconhecidos entre si, e, ainda, recomendações podem sugerir itens particularmente interessantes, além de indicar aqueles que devem ser filtrados (RESNICK, VARIAN, 1997).

Sistemas de Recomendação se tornaram uma importante área de pesquisa desde meados dos anos 90, com o surgimento dos primeiros artigos sobre filtragem colaborativa. O interesse nessa área continua alto por causa da abundância de aplicações práticas que ajudam as pessoas a lidar com a sobrecarga de informação, fornecendo recomendações personalizadas, conteúdo e serviços a elas (ADOMAVICIUS, TUZHILIN, 2005).

Os SRs podem ser classificados em 3 tipos principais, dependendo de como as recomendações são feitas: baseado em conteúdo, colaborativo e híbrido (ADOMAVICIUS, TUZHILIN, 2005). Cada um desses tipos será explicado detalhadamente na próxima seção.

2.2 Técnicas de Recomendação

Cada um dos tipos de principais de SR utiliza uma técnica diferente. Tais técnicas são descritas aqui.

2.2.1 Filtragem Baseada em Conteúdo

A Filtragem Baseada em Conteúdo (FBC) tem suas raízes em Recuperação de Informações (RI) e Filtragem de Informações (FI). Por causa dos significantes avanços feitos pela comunidade de RI e FI e por causa da importância de aplicações baseadas em texto, muitos sistemas baseados em conteúdo se focam em recomendar itens com informações textuais, como documentos e Web Sites. A melhoria sobre as abordagens de RI tradicionais veio com o uso de perfis dos usuários, que contêm informação sobre seus gostos, suas preferências e suas necessidades. Essa informação de perfil pode ser adquirida explicitamente, através de questionários, ou implicitamente, aprendida através do comportamento do usuário ao longo do tempo (ADOMAVICIUS, TUZHILIN, 2005).

Técnicas de FBC recomendam itens similares aos que o usuário gostou no passado (seu perfil). Por exemplo, se um usuário leu vários artigos sobre bioinformática, então essas técnicas irão recomendar outros artigos sobre esse assunto ao usuário (ADOMAVICIUS, TUZHILIN, 2005; CAZELLA, 2006), pois eles terão várias palavras em comum. A ideia é que, se um usuário gostou de um item no passado, ele provavelmente gostará de um item com conteúdo similar no futuro (TORRES, 2004).

Em geral, em se tratando de documentos ou textos, um perfil de usuário é composto de palavras-chave e pesos associados. Esses pesos indicam a força daquela palavra no processo de recomendação. O perfil do usuário é comparado com um grupo de documentos e os mais similares são recomendados. Existem muitas maneiras de calcular essa similaridade (TORRES, 2004) e algumas delas serão explicadas no Capítulo 4.

Para traçar esse perfil, é necessário ter um *feedback* das atividades do usuário. Nesse ponto, muitos Sistemas de Recomendação são intrusivos, já que necessitam de *feedback* explícito do usuário. Isso significa que, para iniciar as recomendações, muitos sistemas precisam de opiniões sobre itens consumidos anteriormente e, muitas vezes, várias opiniões são necessárias. Outros sistemas utilizam *feedback* implícito para saber as opiniões do usuário (ADOMAVICIUS, TUZHILIN, 2005).

Um cenário onde seria viável obter *feedback* implícito seria em um site de comércio eletrônico. Nesse caso, cada vez que um usuário comprasse um item, ele seria armazenado de forma positiva no seu perfil. Para *feedback* explícito, poder-se-ia pensar em um site de recomendação de filmes. Como o site não vende os filmes, ele necessita de uma avaliação do usuário com relação a filmes que ele já viu (BEZERRA, 2002).

As principais vantagens da Filtragem Baseada em Conteúdo, segundo Torres (2004) e Mooney (2010), são:

- Não necessita de outros usuários para recomendar um item, já que as recomendações são feitas se o perfil do usuário e o item têm palavras em comum;
- Possibilidade de recomendar itens para usuários com gostos únicos;
- Possibilidade de recomendar itens novos ou não-populares;
- Podem fornecer explicação sobre as recomendações, listando as características do conteúdo que o fizeram ser recomendado.

Por outro lado, existem algumas desvantagens nessa abordagem. As principais, segundo Torres (2004) e Adomavicius e Tuzhilin (2005), são:

- Só pode analisar itens textuais, não sendo possível analisar o conteúdo de itens como filmes, músicas e imagens com sucesso;
- A maioria das técnicas de análise textual leva em conta apenas as palavras, não considerando o estilo de escrita do autor, a qualidade do conteúdo e experiência do autor no assunto;
- Cria uma superespecialização de conteúdo, pois analisa o conteúdo dos itens e recomenda itens com conteúdo similar, sem considerar outros assuntos. Por exemplo, se um texto contém a palavra "carro" e outro contém a palavra "automóvel", eles não serão considerados similares. Além disso, uma pessoa sem experiência em cozinha grega nunca receberá uma recomendação do melhor restaurante grego na cidade;
- O usuário deve analisar um número suficiente de itens para que o sistema possa entender seus gostos e recomendar itens relevantes.

2.2.2 Filtragem Colaborativa

A Filtragem Colaborativa (FC) é bem diferente da Baseada em Conteúdo. Neste modelo, os usuários podem avaliar itens e registrar esses dados, sendo que essas avaliações serão usadas para a posterior geração de recomendações. Para recomendar itens a um determinado usuário, o sistema identificará na base de dados usuários que realizaram avaliações nos mesmos itens que o usuário em questão (KROHN, 2009). Dessa forma, não se calcula a similaridade entre os itens, mas a similaridade entre os usuários (BALABANOVIC, SHOHAM, 1997). A ideia por trás disso é que, se os usuários tiveram gostos parecidos no passado, eles possivelmente concordarão de novo no futuro (RESNICK et al., 1994). Assim, como em FBC, as opiniões desses usuários podem ser obtidas explicitamente ou através de medidas implícitas (SARWAR et al., 2001).

As vantagens da Filtragem Colaborativa, segundo Torres (2004), são:

- Como as recomendações são baseadas em avaliações feitas por humanos, é possível recomendar qualquer tipo de conteúdo, como filmes, piadas, músicas;
- Podem melhorar o processo de filtragem, analisando coisas que vão além da análise textual. Isso também se deve ao fato de a técnica usar seres humanos no processo de avaliação de itens, já que, para eles, é fácil analisar se um texto está bem escrito, o que é uma tarefa bem difícil para computadores;

- Podem recomendar itens inesperados para os usuários, mas que são boas recomendações. Isso se deve ao fato que a similaridade é entre usuários e não entre itens.

As desvantagens da Filtragem Colaborativa, segundo Mooney (2010) e Adomavicius e Tuzhilin (2005), são:

- Uma certa quantidade de usuários no sistema é necessária para encontrar usuários similares;
- Assim como na Filtragem Baseada em Conteúdo, o sistema precisa aprender as preferências de usuários novos a partir de suas opiniões;
- Como se baseiam nas preferências dos usuários para realizar recomendações, os novos itens acabam não sendo recomendados até que uma quantidade substancial de usuários os avaliem;
- Usuários com gostos únicos não receberão boas recomendações, pois não existirão usuários com gostos similares aos seus e a tendência é que itens populares sejam recomendados.

2.2.3 Filtragem Híbrida

As abordagens de Filtragem Baseada em Conteúdo e Filtragem Colaborativa não são mutuamente exclusivas, e existem várias esforços para combiná-las, com a intenção de obter melhores recomendações (HUANG et al., 2002).

A Filtragem Híbrida (FH) agrega as características de ambas abordagens para solucionar as limitações encontradas e atender melhor as necessidades dos usuários. Usando essa combinação, pode-se alcançar os benefícios da Filtragem Baseada em Conteúdo, que inclui a predição para todos os usuários e itens, independente do número de usuários e avaliações dos itens, além de oferecer exatidão nas predições com Filtragem Colaborativa conforme o número de usuários e recomendações cresce (CLAYPOOL et al., 1999).

Para criar um Sistema de Recomendação Híbrido, os perfis de usuários baseados em análise de conteúdo são mantidos e comparados para determinar usuários similares para recomendação colaborativa. Um usuário recebe uma recomendação quando o item tem uma nota alta em relação ao seu perfil ou quando o mesmo recebeu uma nota alta de um usuário com perfil similar ao seu (BALABANOVIC, SHOHAM, 1997).

Segundo Adomaviciues e Tuzhilin (2005), existem diferentes maneiras de combinar métodos de filtragem colaborativa e filtragem baseada em conteúdo:

- Implementar os métodos colaborativos e baseados em conteúdo separadamente e combinar suas predições;
- Incorporar algumas características das abordagens baseadas em conteúdo em uma abordagem colaborativa;
- Incorporar algumas características das abordagens colaborativas em uma abordagem baseada em conteúdo;
- Construir um modelo de unificação que incorpore as características das duas abordagens.

As principais características herdadas pela Filtragem Híbrida de cada uma das abordagens anteriores são (CAZELLA, REATEGUI apud LOPES, 2007, p.26):

- Descoberta de novos relacionamentos entre usuários;
- Recomendação de itens diretamente relacionados ao histórico;
- Bons resultados para usuários incomuns;
- Precisão independente do número de usuários.

As duas primeiras características são herdadas da Filtragem Colaborativa, pois ela trabalha com a idéia de "perfis similares". Já as duas últimas são herdadas da Filtragem Baseada em Conteúdo, que não leva em consideração qualquer relacionamento entre perfis de usuários e não recomenda itens não relacionados ao perfil do usuário, ou seja, não realiza recomendações "surpresa".

2.3 Exemplos de Sistemas de Recomendação

Existem vários exemplos de Sistemas de Recomendação, a maioria em forma de sites na Internet. Alguns deles serão detalhados a seguir.

2.3.1 Amazon.com

As principais funcionalidades de Sistemas de Recomendação existentes na seção de livros da Amazon.com (www.amazon.com), segundo Schafer et al. (1999), são:

- *Customers who Bougth*: é uma funcionalidade encontrada na página de informações de cada livro de seu catálogo, que apresenta duas listas distintas. A primeira recomenda livros frequentemente comprados por usuários que compraram o livro selecionado. A segunda recomenda autores cujos livros são frequentemente comprados por usuários que compraram obras do autor do livro selecionado. Na figura 2.1, tem-se a tela da Amazon.com com as recomendações criadas pelo sistema;
- *Eyes*: essa funcionalidade permite que os clientes sejam notificados via email sobre novos itens adicionados ao catálogo da Amazon.com. As informações são requisitadas com base em critérios como autor, título, assunto, ISBN ou data de publicação. Também é possível usar critérios baseados em expressões booleanas (AND/OR) para as notificações;
- *Amazon.com Delivers*: é uma variação da funcionalidade Eyes. Os clientes escolhem entre uma lista específica de categorias/gêneros e, periodicamente, recebem anúncios via email, com as últimas recomendações nas categorias selecionadas;
- *Book Matcher*: permite que os clientes deem *feedback* sobre livros que leram, em uma escala de 5 pontos, onde 1 indica que que o usuário não gostou e 5 indica que gostou muito. Assim, o sistema coleta essas avaliações e as usa no futuro para gerar recomendações para os usuários, de acordo com suas preferências;
- *Customer Comments*: permite que os clientes recebam recomendações textuais baseadas nas opiniões de outros usuários. É uma lista de avaliações de 1-5

estrelas e comentários escritos por usuários que leram o livro e enviaram sua avaliação.

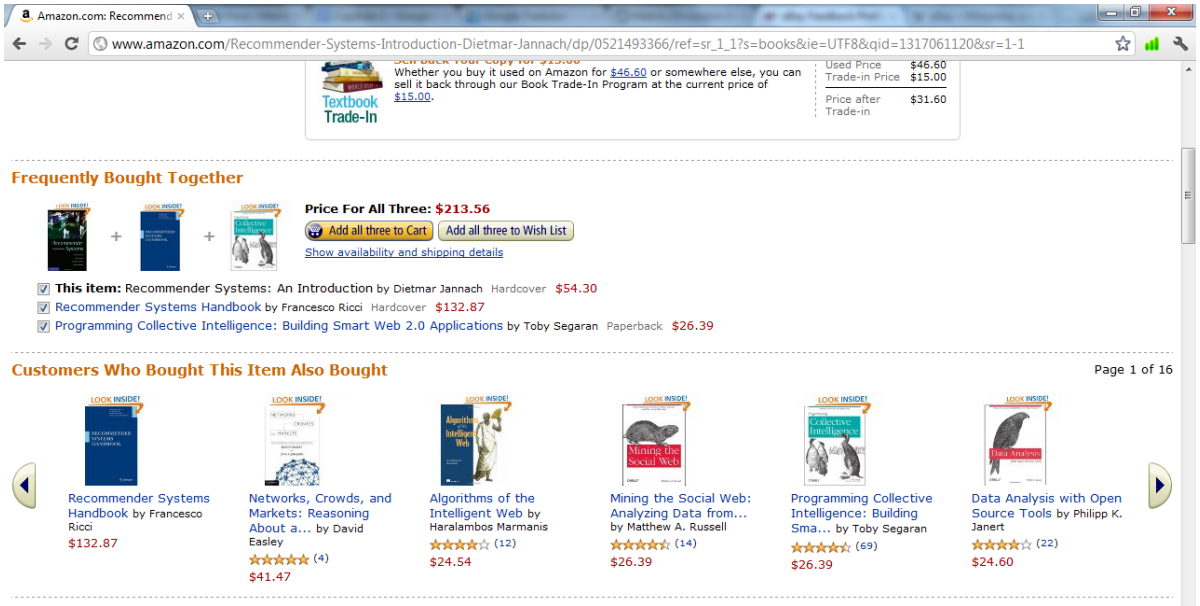

Figura 2.1: Tela com recomendações da Amazon.com

2.3.2 CDNow

A CDNow vendia CDs e era um dos sites mais conhecidos nos anos 90, sendo posteriormente comprada pela Amazon.com (WIKIPEDIA, 2011-a). Suas funcionalidades de Sistemas de Recomendação incluiam (SCHAFER et al., 1999):

- *Album Advisor*: o usuário escolhia um álbum ou até três artistas e o sistema recomendava 10 álbuns relacionados ao álbum ou artista(s) em questão;
- *My CDNOW*: permitia que os clientes criassem sua própria loja musical online, baseado em álbuns e artistas de seu interesse, indicando que álbuns ele tem e que artistas são seus favoritos. Compras do CDNOW eram automaticamente inseridas na lista do usuário, onde ele podia dizer se tinha gostado ou não do álbum comprado. Quando o usuário requisitava recomendações, o sistema recomendava 6 álbuns que o cliente poderia gostar baseado nos que ele já tinha e o usuário poderia dar feedback sobre essas recomendações. Os álbuns recomendados mudavam baseados nesses feedbacks.

2.3.3 eBay

O site de comércio eletrônico eBay (www.ebay.com) tem uma funcionalidade de Sistemas de Recomendação principal. Essa funcionalidade é o Perfil de Feedbacks. Ele permite que compradores e vendedores contribuam com comentários (feedbacks) sobre usuários com os quais fizeram negócios. Esses feedbacks incluem um grau de satisfação (satisfeito/neutro/insatisfeito), além de um comentário específico sobre o usuário. Isso é usado como um sistema de recomendação para compradores, que podem ver o perfil dos vendedores (SCHAFER et al., 1999). Na figura 2.2 é possível ver um exemplo de um desses perfis de vendedores.

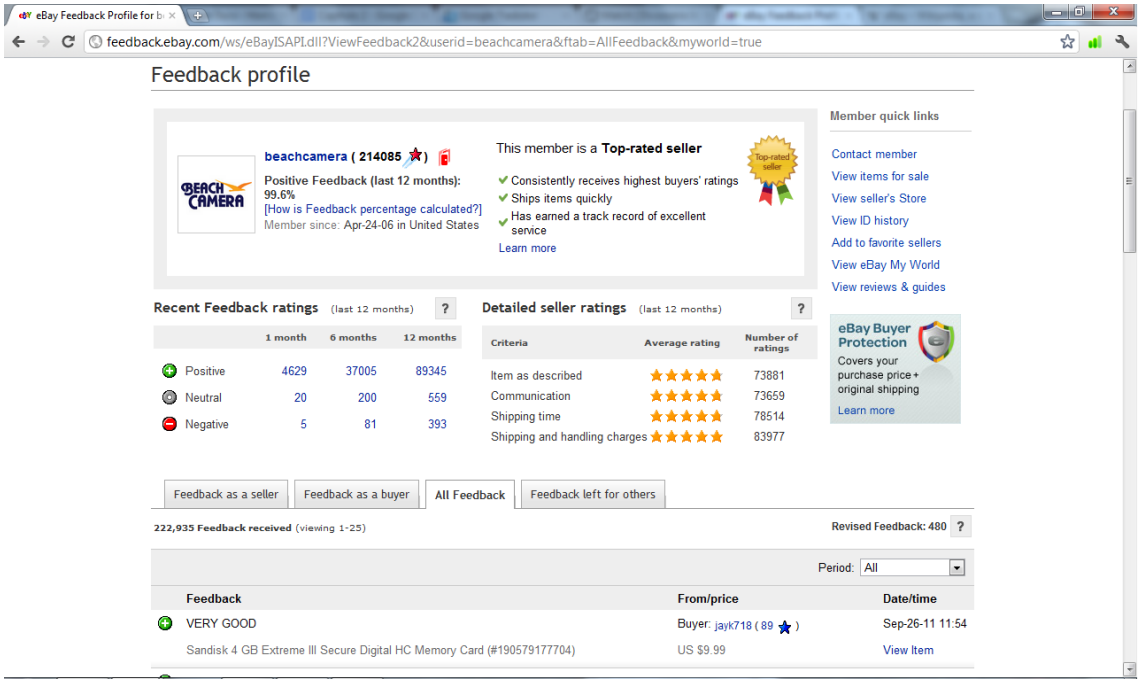

Figura 2.2: Perfil de um vendedor no site eBay

2.3.4 Google

O foco do Google (www.google.com) é em recomendações personalizadas. As principais funcionalidades são (MACMANUS, 2009):

- Os resultados de pesquisa são customizados "quando possível", levando em conta a localização do usuário e/ou suas atividades de busca recentes;
- Quando o usuário estiver conectado à sua conta Google, ele "pode ver resultados ainda mais relevantes e úteis, baseado no seu histórico na Web";
- O algoritmo de busca do Google, o PageRank, é basicamente dependente de recomendações sociais, ou seja, quem referencia a página;
- Ele também realiza recomendação de itens com a sua funcionalidade "Você quis dizer".

2.4 Sistemas de recomendação em fóruns eletrônicos

Dois trabalhos sobre a aplicação de Sistemas de Recomendação em Fóruns Eletrônicos, relacionados com o presente trabalho, serão detalhados a seguir.

2.4.1 Recomendação baseada em regras

O trabalho de Abel et al. (2008) afirma que, em um fórum de discussões com uma estrutura em árvore, os usuários podem navegar facilmente para tópicos nos quais têm interesse. No entanto, isso se torna difícil quando o tópico não está classificado corretamente ou o usuário não tem certeza de como um tópico de seu interesse é classificado na estrutura do fórum. Uma forma de lidar com esse problema é mostrar uma lista de todos os tópicos com modificações recentes. Porém, conforme os fóruns se tornam populares e vários assuntos são discutidos neles, essa lista contém um grande

percentual de tópicos que não interessam ao usuário. Sistemas de Recomendação com Filtragem Colaborativa podem ser usados para lidar com esse problema.

Para tanto, os autores utilizaram um sistema de recomendação denominado Comtella-D, onde um tópico é iniciado quando um estudante posta um link para um artigo relacionado ao fórum, juntamente com uma descrição dele. Dessa forma, se inicia a discussão. O Comtella-D permite que os usuários avaliem os tópicos, adicionando ou removendo "energia" deles (ABEL et al., 2008).

Como com o passar do tempo, muitos tópicos são criados e os mais populares recebem várias respostas. Então ter um overview das discussões é uma tarefa que consome muito tempo, e um usuário que não dedique tempo suficiente a acompanhar o fórum pode perder tópicos de seu interesse. O trabalho explica a criação de um framework de personalização baseado em regras que é usado como Sistema de Recomendação Colaborativa para ajudar a resolver esse problema (ABEL et al., 2008).

Tal framework utiliza dois tipos de dados de entrada: recomendação baseada em *feedback* explícito, através das avaliações baseadas em "energia" ou recomendações baseadas em feedback implícito, obtidas através dos tópicos que o usuário respondeu. Com os testes realizados, os autores observaram que o *feedback* explícito tem um desempenho pior do que o feedback implícito e deve ser usado apenas em casos onde não houver nenhum feedback implícito disponível (ABEL et al., 2008).

Comparado a esse trabalho, o foco do presente trabalho é um pouco diferente. A idéia dos autores é recomendar tópicos do fórum que o usuário pode se interessar, através da observação das ações do usuário no fórum. Apesar de também levar em conta as mensagens que o usuário envia ao fórum, isso não gera recomendações específicas para cada mensagem. As recomendações são genéricas e ocorrem baseadas em toda a interação do usuário com o fórum. Já no presente trabalho, a interação passada do usuário não é levada em conta para recomendar mensagens. Apenas o conteúdo de cada mensagem sendo enviada é analisado e, a partir dele, são geradas as recomendações.

2.4.2 Recomendação baseada em categorias

No trabalho de Fischer (2005), foi proposto um Sistema de Recomendação baseado em Conteúdo para o fórum do sistema Virtuale, o ambiente EAD do Centro Universitário Feevale. Quando o usuário interage com o fórum, suas mensagens são capturadas enquanto o sistema busca mensagens de assuntos semelhantes a serem recomendadas.

O sistema é dividido em quatro módulos principais, que são: Módulo de Definição de Categorias, Módulo de Indexação e Categorização, Módulo de Manutenção do Perfil do Usuário e Módulo de Recomendação. De forma resumida, o funcionamento é o seguinte (FISCHER, 2005):

1. O professor define assuntos (categorias) que serão utilizadas para classificar as mensagens enviadas ao fórum;
2. Quando o usuário envia uma mensagem ao fórum, ela é categorizada e indexada nos assuntos definidos previamente e não é enviada ao fórum ainda;
3. As categorias das mensagens enviadas pelo usuário são gravadas no seu Perfil do Usuário;

4. Mensagens do fórum pertencentes às mesmas categorias contidas no Perfil do Usuário são recomendadas a ele;
5. Caso a mensagem enviada não possa ser classificada em nenhuma categoria existente ou se ainda não houver mensagens dessa categoria no fórum, a mensagem será inserida no fórum para que outros usuários possam lê-la e respondê-la.

Após as recomendações, o usuário deve verificar se elas foram úteis, ou seja, se responderam sua dúvida. Se pelo menos uma das recomendações for útil, elas não são mais mostradas para o usuário, pois a dúvida é considerada respondida e a mensagem não é inserida no fórum. Se nenhuma das recomendações for útil, a dúvida será postada no fórum para que os outros usuários possam respondê-la (FISCHER, 2005).

Analisando esse trabalho e o de Abel et al. (2008), percebe-se que esse é o que mais assemelha-se ao presente trabalho. No entanto, eles têm algumas diferenças principais.

No presente trabalho, as mensagens também são classificadas por categorias. Da mesma forma, as categorias também são definidas pelo professor, mas, antes de enviar a mensagem, o usuário diz em quais categorias a sua mensagem se aplica. Como será detalhado futuramente neste mesmo trabalho, as recomendações são feitas com base nas categorias, mas o conteúdo das mensagens é analisado, e serão recomendadas apenas mensagens similares à que está sendo enviada, desde que sejam da mesma categoria. Isso não ocorre no trabalho de Fischer (2005). Se uma mensagem for da mesma categoria, ela sempre será recomendada, mesmo que não seja tão parecida com a mensagem em questão.

Além disso, o trabalho de Fischer (2005) não utiliza o Moodle como base. Atualmente, o Moodle é o AVA mais utilizado, e o desenvolvimento de módulos desse tipo para ele é uma contribuição relevante.

Outra diferença consiste no fato de que não será criado nem mantido um Perfil de Usuário, pois o Moodle normalmente é usado como apoio para disciplinas que o aluno está cursando. Acredita-se que o usuário pode ter dúvidas únicas sobre assuntos variados ao longo do curso, logo essas dúvidas não necessariamente descrevem o seu perfil e seus interesses.

3 MOODLE

Conforme escrito anteriormente, o Moodle é um Ambiente Virtual de Aprendizagem (AVA) ou *Virtual Learning Environment* (VLE), também conhecido por Sistema de Gerenciamento de Cursos (*Course Management System* ou CMS) ou Sistema de Gerenciamento de Aprendizagem (*Learning Management System* ou LMS). Ele se tornou muito popular entre educadores de todo o mundo como uma ferramenta para criar Web sites dinâmicos para seus estudantes. Para rodar, ele deve ser instalado em algum servidor web (MOODLE, 2011-k).

O Moodle é uma aplicação Web *Open Source*, escrita em PHP, e a companhia Moodle Pty Ltd, do fundador Martin Dougiamas, é quem gerencia o projeto (MOODLE, 2011-g).

O seu foco é fornecer aos educadores as melhores ferramentas para gerenciar e promover aprendizagem, mas existem várias formas de usar o Moodle (MOODLE, 2011-k):

- Ele tem funcionalidades que permitem que seja usado em ambientes muito grandes, com centenas ou milhares de estudantes, mas também pode ser usado em sistemas menores, como de escolas primárias;
- Muitas instituições usam o Moodle como sua plataforma para conduzir cursos totalmente on-line, enquanto algumas o usam apenas como suporte para cursos presenciais;
- Muitos usuários gostam de usar os módulos de atividades (como fóruns, base de dados e wikis) para criar comunidades de aprendizagem colaborativas, enquanto outros preferem usar o Moodle como uma forma de disponibilizar conteúdo aos estudantes e avaliar o conhecimento através de tarefas e questionários.

Existe um site de demonstração do Moodle, disponível em http://demo.moodle.net/, que é um site completo com alguns cursos de exemplo que podem ser testados por professores, alunos ou mesmo administradores. Todo o site é restaurado para um estado limpo a cada hora, então não é necessário se preocupar caso alguma coisa errada seja feita (MOODLE, 2011-k).

3.1 Arquitetura do Moodle

Nesta seção descreve-se a arquitetura do Moodle. Tal descrição é importante para entender como o módulo proposto e apresentado em maiores detalhes no Capítulo 4 integra-se ao Moodle.

3.1.1 Visão Geral de uma Instalação do Moodle

Uma instalação do Moodle inclui o código PHP executando em um servidor Web; uma base de dados de MySQL, PostgreSQL, Microsoft SQL Server ou Oracle; e um repositório de arquivos para arquivos carregados ou gerados (pasta moodledata). Essas três partes podem estar em um único servidor ou separadas para balanceamento de carga (MOODLE, 2011-g).

O Moodle foi desenvolvido para ser simples de instalar em qualquer servidor que atenda esses requisitos. Ele é instalado automaticamente assim que o código é copiado para o servidor e uma base de dados vazia criada. Além disso, também é possível fazer atualização automática de uma versão para outra (MOODLE, 2011-g).

3.1.2 Moodle como um Sistema Modular

Como muitos sistemas *Open Source* de sucesso, o Moodle é estruturado como um sistema central, rodeado por inúmeros *plug-ins* para prover funcionalidades específicas. Os *plug-ins* são de tipos específicos. Isto é, um *plug-in* de autenticação e um módulo de atividades irão se comunicar com o núcleo do Moodle usando diferentes APIs, de acordo com o tipo de funcionalidade que fornecem. Funcionalidades comuns a todos os *plug-ins* (instalação, atualização, permissões, etc) são tratadas de forma consistente em todos os tipos de *plug-ins* (MOODLE, 2011-g).

A distribuição padrão do Moodle inclui o núcleo do Moodle e alguns *plug-ins* de cada tipo, então uma nova instalação pode ser imediatamente usada para iniciar o ensino e aprendizagem. Depois da instalação, um site do Moodle pode ser adaptado para propósito específico, mudando sua configuração *default* e instalando ou removendo *plug-ins* (MOODLE, 2011-g).

Fisicamente, um *plug-in* do Moodle é apenas uma pasta com scripts PHP (e CSS, Javascript, etc., se necessário). O núcleo do Moodle se comunica com o *plug-in* buscando pontos de entrada particulares, muitas vezes definidos no arquivo lib.php do *plug-in* (MOODLE, 2011-g). Esse arquivo lib.php possui as funções que auxiliarão para o correto funcionamento dos *plug-ins*.

3.1.3 O Núcleo do Moodle

O núcleo do Moodle fornece toda a infraestrutura necessária para criar um Ambiente de Aprendizagem Virtual. Ele implementa as principais funcionalidades que todos os *plug-ins* terão de usar. Isso inclui (MOODLE, 2011-g):

- Cursos e Atividades;
- Usuários;
- Grupos;
- Matrículas e Controle de Acesso;
- Navegação e Configuração;
- Biblioteca de Forms;
- Biblioteca de Javascript;
- Instalação e Atualização;

- Logs e Estatísticas.

Para desenvolver funcionalidades customizadas, é sempre preferível escrever um *plug-in* ao invés de modificar o código do núcleo. A maioria das estruturas de *plug-in* do Moodle permitem criar base de dados, controlar o acesso, definir arquivos de linguagem e futuras atualizações de forma bem simples. Além disso, o resultado final pode ser distribuído através de um único arquivo zip. Outra coisa importante é investigar as APIs já existentes no Moodle. Dessa forma, o código escrito será mais confiável, pois provavelmente já foi bastante testado (MOODLE, 2011-b).

3.1.4 Organização do Código

O Moodle segue principalmente uma abordagem baseada no padrão de projeto Transaction Script. A idéia desse padrão é organizar a lógica de negócios por procedimentos, onde cada procedimento (que é considerado uma transação) lida com uma única requisição da camada de apresentação, ou seja, cada procedimento será responsável por validar dados, efetuar cálculos, persistir dados na base de dados e o que mais for necessário (FOWLER et al., 2002). Assim, supondo que se esteja buscando por um fórum. A URL será ../mod/forum/view.php?id=1234 e mod/forum/view.php é o script PHP que gera essa página. Segundo Fowler et al. (2002), esse padrão deve ser usado para aplicações simples, com pouca parcela de lógica, pois conforme a aplicação se tornar mais complexa, mais difícil será manter essa estrutura. Por essa razão, algumas pessoas podem dizer que o padrão Transaction Script não é apropriado para uma aplicação tão complexa quanto o Moodle. No entanto, o Moodle é muito mais um agregado de diferentes *plug-ins* do que uma aplicação única e complexa (MOODLE, 2011-g).

Por trás da abordagem básica de Transaction Script, muitas das funcionalidades principais do Moodle foram refatoradas em bibliotecas (principalmente na pasta lib, que contém as bibliotecas de propósito geral, ou seja, que serão usadas em todo o sistema). O projeto do Moodle iniciou antes de o PHP suportar orientação a objetos, então apenas as partes mais recentes do código usam isso (MOODLE, 2011-g).

Existem duas camadas usadas para separar a apresentação da lógica de negócios. A camada mais externa é o *theme*, que controla os aspectos mais visuais da interface do Moodle. Depois existem classes de renderização que geram HTML baseados nos dados fornecidos pelos scripts. No entanto, nem o PHP nem a arquitetura do Moodle forçam uma separação clara da camada de apresentação. Desenvolvedores desleixados podem fazer uma bagunça, e isso realmente ocorreu no passado. O código na distribuição padrão do Moodle está gradualmente sendo colocado em ordem (MOODLE, 2011-g).

3.1.5 A Base de Dados do Moodle

A base de dados do Moodle tem muitas tabelas (mais de 250), porque ela é um agregado das tabelas do núcleo do sistema com as tabelas de cada *plug-in*. Felizmente, essa grande estrutura é compreensível, porque as tabelas de um *plug-in* em particular normalmente se relacionam apenas entre si e com poucas tabelas do núcleo (MOODLE, 2011-g).

Existem sete tabelas chamadas *forum_algumacoisa*. Se o programador estiver interessado no fórum, deve entendê-las, caso contrário não. E isso vale para os outros *plug-ins* do sistema. Sem levar em consideração as tabelas relacionadas a *plug-ins*, existem aproximadamente 50 tabelas. Esse número ainda é razoavelmente grande, mas

essas tabelas também podem ser divididas em grupos que podem ser entendidos em conjunto ou ignorados (MOODLE, 2010).

A estrutura da base de dados é definida em arquivos install.xml dentro da pasta db de cada plugin. Por exemplo, mod/forum/db/install.xml contém a definição da base de dados do fórum. lib/db/install.xml define a estrutura das tabelas do núcleo do sistema (MOODLE, 2011-g).

Os principais grupos de tabelas do núcleo do sistema são (MOODLE, 2010):

- Configuração;
- Usuários e Perfis de Usuários;
- Papéis e Capacidades (definem o que os usuários podem ou não fazer no sistema);
- Cursos e sua organização em categorias;
- Atividades e sua organização nos cursos;
- Grupos;
- Logs;
- Blocos;
- Eventos;
- Backup;
- Estatísticas;
- Módulos de atividades.

Como essa base de dados pode ser MySQL, PostgreSQL, Microsoft SQL Server ou Oracle, alguma pessoa que não conhecesse a API do Moodle poderia dizer que isso é um problema, já que a sintaxe varia um pouco entre cada Sistema de Gerenciamento de Banco de Dados (SGBD). Para evitar esse problema, existe uma classe chamada *moodle_database*, que pode ser usada através da "importação" do objeto global $DB. Essa classe possui funções para manipulação de dados que devem ser usadas para garantir abstração e portabilidade para os diferentes SGBDs. Maiores informações sobre essas funções e alguns exemplos de uso podem ser encontrados em (MOODLE, 2011-c). A criação da base de dados através dos arquivos install.xml também auxiliam nessa questão de portabilidade.

3.1.6 Criação de um Módulo no Moodle

A principal forma de contribuir e adicionar novas funcionalidades ao Moodle é através de *plug-ins*. Existem vários tipos de *plug-ins* que podem ser criados, e será explicada a criação de um módulo de atividades, pois esse é o tipo de *plug-in* que será utilizado para a implementação do fórum com recomendação, que será explicado no Capítulo 4. Outros exemplos desse tipo de módulo são questionários, tarefas, chats, etc. (MOODLE, 2011-a).

Para iniciar, devemos fazer download do pacote *newmodule*, disponível em (MOODLE, 2011-j). Esse pacote já tem os arquivos indispensáveis para a criação de um novo módulo, nos locais corretos. A explicação mais detalhada de cada arquivo pode ser

encontrada em (MOODLE, 2011-h). Após o download do pacote, seguir os seguintes passos (MOODLE, 2011-i):

- Escolher um nome para o módulo, que não pode conter letras maiúsculas e números;
- Usar um editor de texto para realizar uma substituição de todas as ocorrências de "newmodule" em cada arquivo com o nome escolhido para o módulo;
- Mudar o nome da pasta para o nome do módulo;
- Mudar o nome do arquivo nomedomodulo/lang/en/newmodule.php para "nomedomodulo.php";
- Colocar a pasta criada dentro de moodle/mod;
- Acessar a página de "Notificações" do Moodle (em Moodle -> Administração do Site) para iniciar a instalação do módulo;
- Seguir as instruções da tela.

No final da execução desses passos, tem-se um novo módulo criado e disponível para uso. Mas esse módulo é apenas um esqueleto inicial, ele deve ser modificado para atender as necessidades do desenvolvedor. A primeira modificação importante a ser feita é no arquivo install.xml, conforme explicado na seção 3.1.5. Um detalhe importante sobre a base de dados do módulo é que ela deve conter pelo menos uma tabela com o mesmo nome do módulo. Essa tabela deve ter no mínimo três colunas: id, course e name (MOODLE, 2011-f).

Uma forma talvez um pouco mais simples de criar um módulo é usar como base um módulo que tenha funcionalidade similar à desejada. Dessa forma, o código do módulo deveria ser pego, seu nome substituído pelo novo nome e, então, começar-se-ia a realização das modificações necessárias. Essa foi a forma escolhida para o desenvolvimento do fórum com recomendação neste trabalho. Foi usado como base o fórum original e então foram feitas as modificações necessárias, conforme será explicado no capítulo 4.

A documentação do Moodle para desenvolvedores é bem completa e, além disso, a comunidade colabora bastante nos fórum disponíveis no próprio site do Moodle (MOODLE, 2011-d). Além disso, existe o curso "*Introduction to Moodle Programming*" (MOODLE, 2011-e), disponível no próprio Moodle e acessível a qualquer pessoa que realize o cadastro. Nesse curso, aprende-se como desenvolver um bloco para o Moodle, que é um tipo de *plug-in* que normalmente fica localizado nas laterais dos cursos e alguns *plug-ins* desse tipo são: calendário, pesquisar nos fóruns, descrição do curso, usuários online, etc.

3.1.7 Pacotes de Linguagem

O portal de linguagens do Moodle (http://lang.moodle.org/) permite que tradutores contribuam com pacotes de linguagens e enviem traduções usando uma ferramenta especial do Moodle chamada *AMOS*. Para isso, é necessário criar uma conta no portal, informando o nome completo para receber créditos pela tradução, ler as informações para iniciantes e acessar a ferramenta *AMOS* através do link no bloco de navegação. Se houver alguma dúvida, é possível ainda participar do curso "Translating Moodle course", disponível no próprio Moodle (MOODLE, 2011-l).

A distribuição padrão do Moodle vem apenas com o pacote de linguagem em Inglês. Todas as outras linguagens devem ser adicionadas à instalação. As strings em Inglês para os componentes do núcleo do Moodle podem ser encontradas na pasta *lang/en* da instalação do Moodle. Todos os outros componentes definem suas string na pasta *lang/en* dentro da sua pasta do plugin. As traduções para todos os componentes, inclusive do núcleo, ficam localizadas na pasta moodledata/lang/xxx (MOODLE, 2011-l).

Todos os arquivos de linguagem possuem a extensão .php. Esses arquivos contêm frases curtas, normalmente chamadas de "strings". Essas strings podem conter *placeholders*, que são espaços reservados para posterior substituição de variáveis. Esse *placeholder* é substituído com um valor (passado através do código) quando a string é mostrada. As strings são guardadas em um *array* PHP chamado $string. As chaves desse *array* são chamadas de *identificadores* ou *nomes*. Um exemplo para uma tradução em português seria: $string['addnewcourse'] = 'Adicionar um novo curso'; (MOODLE, 2011-l).

Esses pacotes de linguagens servem para facilitar o desenvolvimento e tradução do Moodle. Por isso, nunca se deve usar mensagens embutidas no código, ou seja, nunca usar "echo 'Fórum com Recomendação'", pois dificultaria a tradução, já que as mensagens deveriam ser procuradas no próprio código e mudadas. Além disso, não seria possível ter duas linguagens "ao mesmo tempo" no Moodle, para que o usuário pudesse escolher a sua linguagem preferida, funcionalidade existente graças a esses pacotes de linguagem. Dessa forma, deve-se sempre utilizar as funções *get_string* ou *print_string*, que esperam como parâmetro o nome da chave no *array* $string e o nome do módulo. Por exemplo, para mostrar a string "Tópicos" do módulo "Fórum", usar-se-ia print_string('discussions', 'forum') (MOODLE, 2011-h). Se uma string não existir em uma linguagem específica, o Moodle tenta achar a string na linguagem pai, quando definida. Se a linguagem pai não tiver sido definida ou a string não existir nessa linguagem, será usada a versão padrão em Inglês (MOODLE, 2011-l).

4 IMPLEMENTAÇÃO DO FÓRUM COM RECOMENDAÇÃO

Neste capítulo será explicada a ferramenta implementada para realizar recomendações no fórum do Moodle com o objetivo de diminuir os problemas relatados no Capítulo 1.

4.1 Visão Geral

Os três componentes principais do sistema são: os usuários, o fórum e a ferramenta de recomendação.

Quando o usuário envia uma mensagem para o fórum, essa mensagem não é automaticamente inserida nele, ela é capturada e analisada. Se houverem mensagens semelhantes no fórum, o usuário receberá essas recomendações. Dessa forma, o usuário pensa que está interagindo diretamente com o fórum, mas a sua interação é com a ferramenta desenvolvida e não com o fórum. Quem interage com o fórum é a ferramenta.

4.2 Fórum do Moodle

O Fórum do Moodle foi o escolhido para esse trabalho, pois o Moodle é uma ferramenta muito usada na UFRGS, tanto como ferramenta de apoio aos cursos presenciais como ferramenta para EAD.

Cada curso no Moodle (que normalmente equivale a uma disciplina de algum curso) tem o seu próprio fórum. Além disso, o Moodle permite que sejam adicionados vários fóruns a cada curso, mas isso não é uma prática comum.

Nesse fórum, cada usuário pode iniciar uma nova discussão (ou tópico), visualizar os tópicos existentes e responder esses tópicos. É possível ainda realizar buscar nas mensagens existentes no fórum. Isso pode ser útil quando houver muitas mensagens, pois fica difícil analisar todas as mensagens. Além disso, é possível assinar um fórum, ou seja, o usuário é notificado via email de cada nova mensagem enviada ao fórum. A forma de assinatura pode ser configurada: o usuário pode decidir assinar o fórum, o tutor pode decidir que o usuário é assinante ou ainda impedir assinaturas. O tutor pode também monitorar a leitura do fórum, como forma de avaliar se os estudantes estão realmente utilizando-o.

Como citado no capítulo 3, o Fórum é um tipo de módulo de atividades no Moodle. A estrutura do seu banco de dados pode ser visualizada na Figura 4.1.

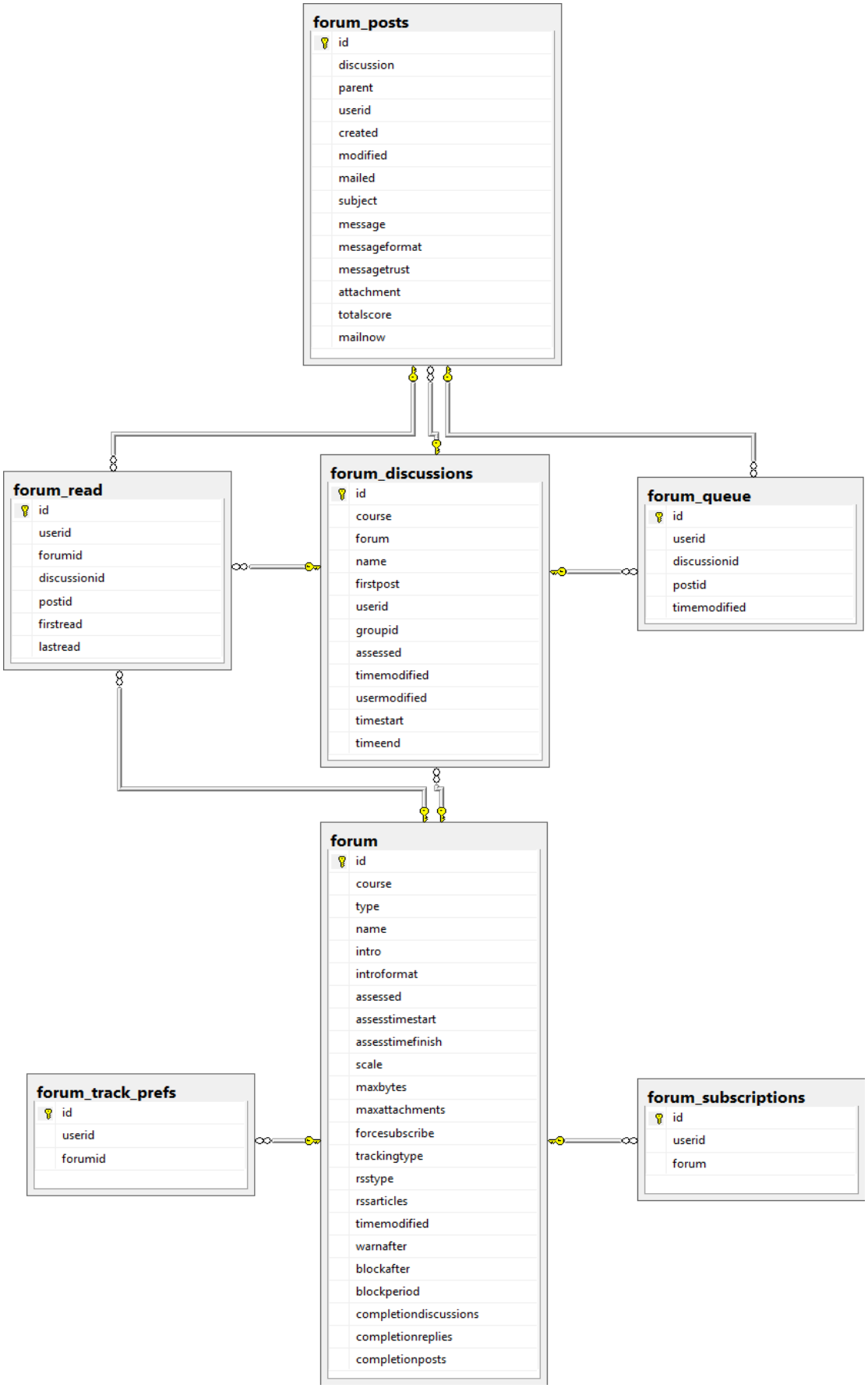

Figura 4.1: Diagrama do Banco de Dados do Fórum

As principais tabelas desse módulo são: forum, forum_discussions e forum_posts. Conforme explicado no Capítulo 3, cada módulo deve ter pelo menos uma tabela com o seu próprio nome. Por isso, tem-se a tabela forum, que contém as principais informações e configurações de cada Fórum. A tabela forum_discussions contém as principais informações de cada tópico iniciado, como título, curso, usuário que o postou, data de modificação, etc. E a tabela forum_posts contém as mensagens propriamente ditas, com informações sobre mensagem pai (quando a mensagem for uma resposta a um tópico), a própria mensagem, o usuário que a postou, a data de criação e modificação, assunto da mensagem, etc.

Além dos relacionamentos apresentados na Figura 4.1, algumas tabelas também se relacionam com as tabelas "course" e "user", através dos campos "userid" e "course". Como citado no Capítulo 3, se o foco fosse o Fórum, seria preciso se preocupar apenas com as tabelas mostradas na Figura 4.1 e essas outras duas tabelas. É isso que torna o Moodle e sua estrutura fáceis de entender.

4.3 Ferramenta Proposta

Para realizar recomendação de mensagens no fórum do Moodle, existiam três opções principais:

- Modificar o código original do Fórum, para interceptar as mensagens e realizar as recomendações;
- Monitorar o Banco de Dados (BD) do Moodle, caso não fosse possível realizar a implementação usando a sua API;
- Criar um novo fórum na forma de um plug-in, parecido com o original, mas com as funcionalidades de recomendação.

Inicialmente, a primeira opção parecia boa, pois o fórum já estava implementado e funcionando e o foco poderia ficar na ferramenta de recomendação. Porém, após analisar melhor a situação, foi percebido que ela seria inviável, pois se a intenção fosse utilizar o Fórum com Recomendação em um ambiente real, como, por exemplo, em um curso no Moodle da UFRGS, todos os outros cursos do sistema seriam obrigados a utilizar esse fórum também. Com a segunda opção, o BD seria monitorado e, quando o usuário mandasse uma mensagem, seriam buscadas recomendações para ela. Caso houvessem recomendações a serem feitas, elas seriam inseridas no BD também, como uma mensagem para o usuário. Essa abordagem certamente seria a mais fácil de ser desenvolvida e poderia ser implementada no próprio BD, através de *stored procedures*. No entanto, a interação seria prejudicada, pois não seria instantânea e, mesmo que as recomendações fossem úteis, as mensagens já teriam sido enviadas ao fórum. Por isso, foi escolhida a terceira opção.

Para criação desse *plug-in*, que é do tipo módulo de atividades, foi utilizado como base o código do módulo fórum. Dessa forma, seria possível manter o foco maior na ferramenta de recomendação, que é a idéia principal do trabalho. Além das funcionalidades do fórum, foram adicionadas funcionalidades para Gerenciamento de Categorias, Recomendação de Mensagens e Visualização de Relatórios, que serão explicadas a seguir. O módulo desenvolvido chama-se *newforum* e é um fórum com recomendação.

Além desse módulo de atividades, foi criado um bloco para realizar pesquisas no fórum com recomendação. Esse bloco é semelhante ao "Pesquisar nos Fóruns" citado anteriormente e foi baseado nele. A única alteração realizada foi uma mudança na URL chamada quando uma busca é realizada: no bloco original, a URL era /mod/forum/search.php, e, no novo bloco, é /mod/newforum/search.php. Essas páginas chamadas são as páginas de busca do módulo de atividades "forum" ou "newforum". Dessa forma, tanto buscas no próprio fórum quanto no bloco são redirecionadas para o mesmo local, sendo realizadas de forma idêntica (evitando assim duplicação de código).

O módulo foi desenvolvido utilizando a versão 2.1 do Moodle, e ele funciona apenas nessa versão e posteriores.

4.3.1 Gerenciamento de Categorias

Para facilitar na busca de mensagens a serem recomendadas, a funcionalidade de envio de mensagens ao fórum foi um pouco modificada. Além dos campos normais existentes na tela de envio de mensagens (assunto, mensagem, tipo de subscrição, anexos, enviar email em seguida), foram adicionados dois campos extra para seleção de categorias para as mensagens. Dessa forma, quando o usuário vai postar sua mensagem no fórum, ele pode escolher até duas categorias nas quais sua mensagem se encaixa. Uma delas é obrigatória e a outra é opcional. Existe uma categoria denominada "Outra". Dessa forma, quando o usuário acreditar que sua mensagem não se enquadra em nenhuma das categorias existentes, ele pode digitar o nome da categoria desejada. Através dessa seleção de categorias, é possível realizar uma filtragem inicial de mensagens a serem recomendadas, diminuindo assim o número de mensagens a serem analisadas e aumentando a eficiência da ferramenta. Isso é feito porque assume-se que uma mensagem da categoria "futebol" não estará relacionada com uma mensagem da categoria "natação", por exemplo.

O Gerenciamento das Categorias deve ser feito pelo tutor do curso no momento da criação do fórum, para que as categorias estejam disponíveis quando o usuário for enviar sua mensagem. As categorias estão relacionadas com os cursos, então, se houver mais de um fórum no mesmo curso, não será necessário cadastrar novamente as categorias. Essa escolha foi feita pois, mesmo que exista mais de um fórum no curso, o curso é o mesmo, então as categorias nas quais ele pode se dividir serão as mesmas.

No momento da criação das categorias, é possível escolher a "categoria-pai" (se for uma subcategoria), o nome da categoria e uma descrição para a mesma. As categorias cadastradas podem ser visualizadas na tela de Gerenciamento de Categorias, de forma hierárquica, mostrando os relacionamentos entre as categorias, como pode ser visto na Figura 4.2. Após a inserção, é possível editar ou remover categorias. Na edição, é possível modificar os campos inseridos anteriormente. Já na exclusão, seleciona-se uma categoria para excluir e é realizada uma verificação. Se essa categoria não possuir nenhuma categoria subordinada ou nenhum tópico existente no fórum for dessa categoria, é solicitada uma confirmação e a categoria é excluída. Caso a categoria possua categorias subordinadas, pode-se escolher entre excluir também as categorias subordinadas ou escolher uma nova "categoria-pai" para elas. E, caso existam tópicos com essa categoria, deve-se escolher uma nova categoria para eles, pois os tópicos não devem ser excluídos e não podem ficar sem categoria.

Além disso, existe a possibilidade de exportar e importar as categorias. Isso pode ser útil para que os professores compartilhem suas categorias de curso, sem a necessidade

de cadastrar novamente coisas que já foram cadastradas e também para uso próprio do professor, no mesmo curso, mas em semestres ou anos diferentes. Quando a opção de exportar categorias é selecionada, as categorias são gravadas em um arquivo com a extensão .xml. Esse formato foi escolhido pois ele é um formato para a criação de documentos com dados organizados de forma hierárquica (WIKIPEDIA, 2011-b), o que é o caso das categorias. Quando a opção de importar categorias é escolhida, é necessário fazer upload de um arquivo .xml, cuja estrutura é validade através de um XML Schema (arquivo com a extensão .xsd). Se o arquivo carregado for válido, as categorias são inseridas no curso.

As categorias selecionadas pelos alunos para as mensagens como "Outra", são inseridas no sistema como "Categorias Não-Gerenciadas", já que não foram inseridas pelo professor e elas não aparecem na lista de categorias a serem escolhidas para as mensagens. Na página de Gerenciamento de Categorias, o professor pode adicionar ou remover essas categorias do sistema. Se ele adicionar, a categoria simplesmente passa de "Não-Gerenciada" para "Gerenciada". Se ele desejar excluir, é necessário selecionar uma nova categoria para as discussões dessa categoria, já que toda mensagem necessita de uma categoria. A inserção de uma categoria não-gerenciada pode ser útil quando o professor considerar que é uma categoria realmente relevante para o curso. Já a exclusão pode acontecer quando, por exemplo, duas categorias forem muito semelhantes. Um aluno pode inserir uma categoria chamada "Notícia" e outro uma categoria chamada "Notícias", que na verdade se referem a mesma coisa mas escrita de formas diferentes. Como o aluno não sabe as categorias que os outros alunos inseriram, esse erro pode acabar acontecendo. Dessa forma, o professor pode unificar as duas categorias em uma só. Outro cenário para a exclusão de categorias não-gerenciadas pode ser quando o professor considerar que uma categoria inserida pelo aluno já existe nas categorias gerenciadas. Na Figura 4.2 podemos ver como as categorias não-gerenciadas são mostradas ao tutor do curso.

Figura 4.2: Gerenciamento de Categorias

4.3.2 Recomendação de Mensagens

Essa funcionalidade é sem dúvida a mais importante da ferramenta. Para efetuar as recomendações, foi usada a abordagem de Filtragem Baseada em Conteúdo, pois ela é mais adequada para uso em fóruns eletrônicos, devido às seguintes características:

- É mais adequada para itens textuais, o que é o caso das mensagens do fórum;
- Não necessita de outros usuários para realizar recomendações, já que as recomendações são feitas com base no conteúdo dos itens e não nas avaliações dos usuários, o que é importante nesse caso, pois pode haver poucos usuários no fórum;
- Pode recomendar itens novos ou não populares, ou seja, pode recomendar um tópico que nunca foi visto ou que acabou de ser inserido, se ele for semelhante ao que o usuário está inserindo, já que a recomendação não depende da opinião de outros usuários.

Quando um usuário envia uma mensagem ao Fórum, as mensagens da mesma categoria são analisadas e suas similaridades são calculadas, conforme será explicado na Seção 4.3.2.1. Após esse cálculo, mensagens cuja similaridade ultrapassar um limiar pré-estabelecido (esse valor é um número entre 0 e 1, que é definido pelo administrador do Moodle no momento em que o módulo é adicionado e cujo valor padrão é 0.5) são recomendadas ao usuário. Essas mensagens são ordenadas em ordem decrescente de

similaridade, para induzir o usuário a olhar primeiro a mensagem mais similar. Essas recomendações são gravadas no BD, para posterior geração de relatórios para análise.

Após receber as recomendações, é tarefa do usuário verificar se elas foram úteis e responderam a sua dúvida. Se alguma delas for útil e solucionar sua dúvida, então a mensagem não é inserida no fórum e a recomendação é considerada aceita. Uma coisa importante a ser observada é que as recomendações são agrupadas, isto é, mensagens que foram recomendadas juntas fazem parte do mesmo grupo. Dessa forma, se alguma das mensagens foi útil e o usuário aceitou a recomendação, então o grupo é considerado aceito. Essa decisão se deve ao fato de que, por exemplo, se a primeira mensagem já for útil, o usuário provavelmente não irá ler as outras e aceitará a recomendação. Se cada mensagem fosse considerada de forma individual, as outras mensagens recomendadas seriam consideradas recusadas, mas na verdade elas nem foram lidas e poderiam ser úteis também, caso a primeira mensagem não existisse. Por outro lado, se o usuário ler todas as mensagens e nenhuma for útil, ele irá recusar a recomendação e o grupo será considerado recusado. Nesse caso, a mensagem é inserida no fórum para que os outros usuários possam lê-la e respondê-la.

Quando uma recomendação é recusada, a nova mensagem recém inserida no fórum é associada com as recomendações já existentes no BD. Dessa forma, o tutor do curso pode analisar as mensagens recusadas para verificar se elas realmente não eram úteis ou se o aluno recusou as recomendações de forma errada. Uma melhoria que poderia ser feita seria comparar as mensagens enviadas com mensagens inseridas após recomendações recusadas. Dessa forma, se uma mensagem fosse muito similar a uma que foi inserida após a recusa de uma recomendação, então provavelmente as recomendações efetuadas para aquela mensagem também não seriam úteis para a nova mensagem.

4.3.2.1 Similaridade entre Documentos

Para calcular a similaridade entre os documentos, os seguintes passos foram seguidos:

1. Definição da Coleção de Textos: consiste em definir a coleção que será utilizada no processo (RIZZI et al., 2000). No presente trabalho, essa coleção é composta das mensagens presentes no fórum;
2. Identificação de Termos: o objetivo é identificar as palavras dos documentos, excluindo do texto símbolos, figuras, caracteres especiais, etc., com o intuito de facilitar a análise do conteúdo (RIZZI et al., 2000). Neste trabalho, como se está lidando com informações textuais, foram removidos os símbolos de pontuação e também os acentos das palavras. Isso foi feito pois considerou-se que algumas pessoas acabam não usando a acentuação corretamente, então isso poderia atrapalhar o processo. Por exemplo, se alguma pessoa escrever “fórum” e outra escrever “forum”, as duas estão se referindo à mesma coisa. Além disso, todo o texto é convertido para letras minúsculas, para garantir que as comparações efetuadas vão retornar resultados realmente corretos, independente de terem sido escritas com letras maiúsculas ou minúsculas.
3. Remoção de Stopwords: consiste em ignorar as palavras que costumam aparecer em todos os documentos. Essas palavras são denominadas stopwords, uma expressão que pode ser traduzida como “palavras-negativas”, “palavras-ferramenta” ou “palavras-vazias”. Essas palavras são preposições, artigos,

conjunções e demais palavras utilizadas para auxiliar na construção sintática das orações (WIVES, 2001). Neste trabalho, foi utilizada uma lista de stopwords para português, disponível em (SNOWBALL, 2011).

4. Seleção de Termos ou Características: busca identificar, através de métodos específicos (como Frequência Absoluta, Frequência Relativa, etc) os termos que mais caracterizam e representam o conteúdo do texto em questão (RIZZI et al., 2000). Esses métodos serão detalhados em 4.3.2.1.1;
5. Método de Similaridade: aplicação de algum método para o cálculo da similaridade entre os documentos (como Cosine ou medidas Fuzzy). Isso será detalhado em 4.3.2.1.2.

Outro passo que poderia ser aplicado seria o *stemming*. *Stemming* é o processo de reduzir as variações de uma palavra a uma representação comum, o radical (ORENGO; HUYCK; 2001). A ideia por trás disso é que variações de palavras com o mesmo radical são similares e devem ser tratadas como uma única palavra (HUANG, 2008). Por exemplo, as palavras "recomendação", "recomendações" "recomendar" e "recomendou" seriam reduzidas para "recomend" e seriam consideradas uma única palavra, ou seja, se uma pessoa estivesse buscando por "recomendação", textos contendo as outras palavras também seriam retornados, pois também seriam interessantes para a pessoa.

A maioria dos experimentos utilizando *stemming* foram realizados em textos da língua inglesa, e idiomas com mais flexões gramaticais (como o português) podem obter benefícios maiores desse processo. Para isso, um dos algoritmos desenvolvido foi o "Removedor de Sufixos para a Língua Portuguesa" (RSLP), que é composto de oito passos, cada um com um conjunto de regras. Esse algoritmo apresentou um resultado melhor do que a versão do algoritmo de Porter (um dos mais usados e conhecidos) para o português (ORENGO; HUYCK; 2001). Essa técnica não foi utilizada neste trabalho, pois poderia acabar sendo muito custosa e diminuir o desempenho da ferramenta. Além disso, como as mensagens normalmente não serão muito grandes, o benefício pode acabar não sendo notado.

4.3.2.1.1 Medidas de Frequência

Grande parte das formas encontradas na literatura para se calcular o peso ou a importância de uma palavra em um documento é baseada na frequência das palavras no documento. A seguir são apresentadas as principais formas de se calcular a importância de uma palavra em um documento, segundo estudo de Wives (2004).

1. Frequência Absoluta de Termos (term frequency ou absolute frequency): indica o número de ocorrências de uma palavra em um documento, ou seja, a quantidade de vezes que ela aparece no documento. Em cada documento, a palavra pode possuir uma frequência diferente, dependendo do seu número de ocorrências.
2. Frequência Relativa de Termos (relative frequency): é muito utilizada e consiste da frequência absoluta normalizada pelo tamanho do documento (número de palavras contidas nele). Essa normalização é feita dividindo-se o número de ocorrências da palavra pelo total de palavras presentes no documento. O resultado é um número entre 0 e 1, que indica a importância relativa da palavra no documento. Apesar de muito utilizada, essa medida não leva em consideração

os demais documentos da coleção. Com isso, palavras que apareçam em muitos documentos têm a mesma importância de palavras que apareçam em poucos documentos (e são mais discriminantes).

3. Frequência Inversa de Documentos (inverse document frequency): leva em conta tanto a frequência absoluta quanto a frequência relativa (número de documentos em que a palavra aparece). Dessa forma, palavras que aparecem em muitos documentos são penalizadas. Apresenta duas variações: log N / n ou log ((N - n) / n), onde N é o número total de documentos e n é o número de documentos em que a palavra aparece. O valor resultante encontra-se sempre no intervalo [0, 1]. No primeiro caso, quanto maior o número de documentos em que a palavra aparece, mais próximo de zero fica o valor resultante. Já no segundo caso, quanto maior o número de documentos em que a palavra aparece, mais próximo de um fica esse valor. A escolha de qual variação da fórmula usar fica a cargo do usuário.

Na ferramenta proposta, os documentos são as mensagens enviadas ao fórum. Existem muitas maneiras de representá-las e a escolhida foi como um "saco de palavras" (expressão "bag of words"), onde se assume que as palavras aparecem independentemente e a ordem não é relevante. O "peso" de cada palavra no "saco" é a frequência com que esse termo aparece no texto e, nesse caso, foi escolhida a frequência relativa para a representação desse "peso" (HUANG, 2008).

4.3.2.1.2 Medidas de Similaridade

Algoritmos que calculam o valor de similaridade entre documentos usam métricas que revelam o quanto um documento é similar a outro. Existem diversos algoritmos e métricas utilizados para diversos fins. Um algoritmo desses pode, por exemplo, ser utilizado na grade de programação digital da televisão, para recomendar ao usuário programas similares ao seu gosto (MAIA, 2008).

Neste trabalho, serão apresentados dois algoritmos para calcular a similaridade: o Cosine e o baseado em Medidas Fuzzy. Ambos algoritmos foram implementados e testados na ferramenta. Nos testes realizados, o Cosine normalmente apresentou uma similaridade maior, mas o Fuzzy tende a ser mais preciso. Optou-se por usar a medida Cosine, pois como não é feita uma análise mais detalhada do texto, foi considerado melhor utilizar essa medida, para que mais mensagens fossem recomendadas e o usuário tivesse mais chance de encontrar sua dúvida respondida.

A similaridade Cosine é uma das medidas de similaridade mais populares (HUANG, 2008). Essa função calcula a similaridade entre objetos como sendo a distância vetorial euclidiana entre os dois objetos, utilizando valores contidos no intervalo [0, 1]. Dessa forma, a similaridade entre os objetos é dada pela fórmula (WIVES, 2004):

$$similaridade(Q,D) = \frac{\sum_{k=1}^{n} w_{qk} \times w_{qd}}{\sqrt{\sum_{k=1}^{n} \left(w_{qk}\right)^2 \times \sum_{k=1}^{n} \left(w_{qd}\right)^2}}$$

Figura 4.3: Fórmula da Similaridade Cosine

Nessa fórmula, Q é o vetor de termos do documento X, D é o vetor de termos do documento Y, w_{qk} são os pesos dos termos do documento X e w_{dk} são os pesos dos termos do documento Y. (WIVES, 2004).

Uma propriedade interessante da similaridade Cosine é a sua independência de tamanho. Se combinarmos duas cópias idênticas do documento d para gerar um novo documento *d'*, a similaridade entre os documentos *d* e *d'* será 1, o que significa que esses documentos são considerados idênticos. Além disso, se considerarmos outro documento *f*, a similaridade entre *d* e *f* e entre *d'* e *f* será a mesma (HUANG, 2008).

Com relação às medidas Fuzzy (difusas), O termo *fuzzy* foi introduzido por volta dos anos sessenta em um estudo realizado por *Latfi Zadeh*. Podemos dizer que a lógica *fuzzy* está para o raciocínio aproximado assim como a lógica tradicional está para o raciocínio preciso (OLIVEIRA apud WIVES, 1999, p. 27).

A função *fuzzy* mais simples que pode ser utilizada é a de **inclusão simples**, que avalia a presença de palavras nos dois documentos comparados. Se o termo aparece nos dois documentos, é somado o valor 1 ao contador, caso contrário, o valor 0 é somado. O valor de similaridade é um valor *fuzzy* entre 0 e 1, calculado através da média, ou seja, o valor total do contador de termos comuns dividido pelo número total de termos nos dois documentos (sem contar repetições) (WIVES, 1999).

Esse valor *fuzzy* representa o grau de igualdade entre os elementos. O problema dessa função é que ela só dá importância para o fato de uma palavra aparecer em ambos os documentos. O fato de uma palavra ser mais importante em um ou outro documento não é considerado (WIVES, 1999).

Esse problema pode ser resolvido através de outra função, que realiza a média por operadores *fuzzy*, semelhante à Inclusão Simples, mas utilizando pesos para os termos. Dessa forma, o fato de termos aparecerem com importâncias diferentes é considerado (OLIVEIRA apud WIVES, 1999, p. 28).

Neste caso, o peso dos termos pode ser a Frequência Relativa ou um valor de discriminação. A similaridade é calculada pela média entre os pesos médios dos termos comuns. Dessa forma, quando o termo aparece nos dois documentos, a média dos seus pesos é somada e não mais o valor 1. Ao final, a média é calculada sobre o total de termos nos dois documentos (WIVES, 1999).

Baseando-se na idéia anterior, foi criada a seguinte função para cálculo da similaridade entre documentos:

$$gs(X,Y) = \frac{\sum_{h=1}^{k} gi_h(a,b)}{n}$$

onde:

gs é o grau de similaridade entre os documentos ***X*** e ***Y***;

gi é o grau de igualdade entre os pesos do termo ***h*** (peso ***a*** no documento ***X*** e peso ***b*** no documento ***Y***);

h é um índice para os termos comuns aos dois documentos;

k é o número total de termos comuns aos dois documentos;

n é o número total de termos nos dois documentos (sem contagem repetida).

Figura 4.4: Fórmula da média por operadores "fuzzy" (WIVES, 1999)

Basicamente, essa função utiliza um contador que vai acumulando pontos toda vez que um termo aparece em ambos os documentos. O valor utilizado para atualizar esse contador é dado por outra fórmula, que identifica o grau de igualdade entre termos comuns. Essa outra fórmula é necessária para considerar os graus de importância dos termos nos documentos. Essa função, que pode ser vista na figura 4.4, leva em conta a média da frequência relativa dos termos em ambos os documentos (WIVES, 1999).

$$gi(a,b) = \frac{1}{2}\left[(a \to b) \wedge (b \to a) + (\bar{a} \to \bar{b}) \wedge (\bar{b} \to \bar{a})\right]$$

$onde,$

$\bar{x} = 1 - x$

$a \to b = \max\{c \in [0,1] \mid atc \leq b\}, t = produto$

$\wedge = \min$

Figura 4.5: Fórmula do cálculo do grau de igualdade entre os pesoas (WIVES, 1999)

Assim, se um termo aparece em ambos os documentos com pesos muito diferentes, o grau de igualdade torna-se baixo. Por outro lado, se os valores dos pesos são próximos, o grau de igualdade torna-se alto. Caso um termo apareça em apenas um dos documentos, ele contribui com um grau zero e a similaridade diminui. Assim, obtém-se um valor mais realista para ser considerado na similaridade final (WIVES, 1999).

4.3.3 Visualização de Relatórios

Conforme citado anteriormente, as recomendações efetuadas são inseridas no BD. Como a ideia é que o módulo desenvolvido seja disponibilizado para *download* no próprio site do Moodle e assim seja usado por pessoas das mais variadas áreas, optou-se por uma maneira simples de mostrar as recomendações.

Para uma pessoa da área da Computação, seria relativamente simples realizar consultas no BD do Moodle para saber dados sobre as recomendações. No entanto, para

alguém da área da Saúde essa tarefa seria praticamente impossível. Por isso, o administrador do fórum tem a possibilidade de visualizar relatórios sobre as recomendações. Isso inclui dados como: total de mensagens recomendadas, total de grupos recomendados, total de grupos aceitos e recusados, médias de similaridades aceitas e recusadas e percentual de grupos aceitos e recusados.

Ainda assim, pessoas com conhecimento maior de BD podem analisar os dados gravados para obter maiores detalhes. Através dos dados inseridos no BD, é possível visualizar por exemplo quais mensagens foram recomendadas em conjunto e ver as relações entre mensagens recomendadas e recomendações recusadas.

4.3.4 Principais Alterações Efetuadas no Banco de Dados

Para suportar as mudanças efetuadas durante o desenvolvimento do módulo, algumas mudanças no Banco de Dados do módulo também foram necessárias. A tabela de tópicos (newforum_discussions) teve dois campos inseridos: “category” e “category2”, para guardar as categorias das mensagens. Foram adicionadas ainda tabelas para guardar as categorias em si e para guardar informações sobre as recomendações efetuadas.

Na figura 4.6 são mostradas as tabelas que foram adicionadas ou alteradas no modelo original, além da tabela newforum que é a tabela central do plugin. Como pode ser visto, foram adicionados os campos “category” e “category2” na tabela de discussões, para guardar as categorias das discussões do fórum. Foram inseridas ainda tabelas para gerenciar as categorias do sistema (newforum_categories) e para guardar os dados relacionados às recomendações efetuadas pela ferramenta (newforum_recommendations).

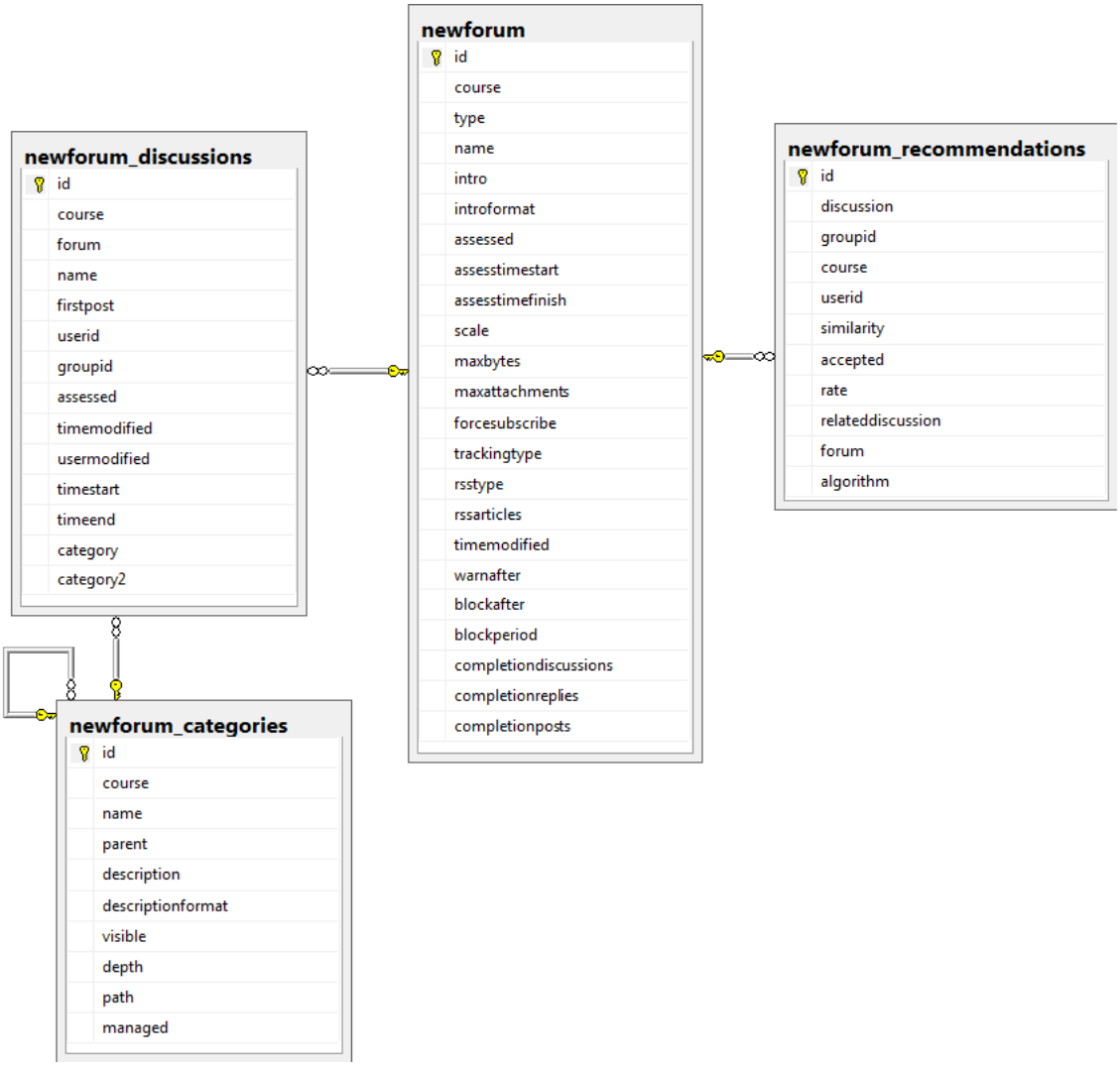

Figura 4.6: Diagrama com as modificações realizadas no Banco de Dados do módulo “newforum”

5 TESTES DA FERRAMENTA DE RECOMENDAÇÃO

Com o intuito de validar a ferramenta proposta, foram realizados testes com a mesma. Para isso, foi criado um curso de teste no Moodle e o módulo implementado foi adicionado a ele. Após a inclusão e configuração do módulo, foram criadas algumas categorias para que os testes pudessem ser iniciados. Na Figura 5.1 pode-se ver as categorias inicialmente criadas:

Categorias Gerenciadas

Categorias	Tópicos	Editar
Hardware	0	
Memórias	0	
Placas de Vídeo	0	
Placas Mãe	0	
Processadores	0	
Sistemas Operacionais	0	
Mac OS	0	
Microsoft	0	
Open Source	0	
Software	0	
Desenvolvimento	0	
Segurança	0	

Acrescentar uma nova categoria

Figura 5.1: Categorias iniciais do Curso de Teste

Após a criação das categorias acima descritas, mensagens que se encaixavam nelas foram extraídas de outros fóruns da internet e manualmente inseridas no curso de Teste, como tópicos do fórum. Ao total, foram inseridos 18 tópicos, a fim de simular 18 perguntas, e alguns deles foram respondidos. Na Figura 5.2 pode ser vista a tela principal do fórum, com algumas das mensagens inseridas.

Este é um fórum para realização de testes na ferramenta desenvolvida como Trabalho de Conclusão de Curso, que é um fórum com recomendação de mensagens.

Acrescentar um novo tópico de discussão

Tópico	Autor	Categoria(s)	Comentários	Última mensagem
Drivers Sony VAIO	Admin User	Drivers	0	Admin User sex, 4 nov 2011, 21:19
Backup Windows Server 2008 em HD USB	Admin User	Microsoft, Windows Server	1	Admin User sex, 4 nov 2011, 21:17
Google e Facebook pedem para instalar Setup_xxxx.exe	Admin User	Segurança	3	Admin User qui, 3 nov 2011, 18:57
Atualizando Leopard para Snow Leopard	Admin User	Mac OS	2	Admin User qui, 3 nov 2011, 18:53
Incompatibilidade entre marcas de memórias DDR3	Admin User	Hardware, Memórias	0	Admin User qui, 3 nov 2011, 18:39
Dúvida sobre limpeza do processador	Admin User	Processadores	1	Admin User qui, 3 nov 2011, 18:37
Processador esquentando	Admin User	Processadores	1	Admin User qui, 3 nov 2011, 18:36
GeForce 8400GS : Compatibilidade	Admin User	Hardware, Placas de Vídeo	1	Admin User qui, 3 nov 2011, 18:28
Problema instalando o Microsoft Visual C++ 2008 Express Edition	Admin User	Software, C++	2	Admin User qui, 3 nov 2011, 18:26
Coloração de Grafos em C	Admin User	Desenvolvimento, C	0	Admin User qui, 3 nov 2011, 18:13
Instalar JavaDoc no Eclipse	Admin User	Java, Eclipse	1	Admin User qui, 3 nov 2011, 18:06
Problema para instalar o Windows 8	Admin User	Microsoft	1	Admin User seg, 31 out 2011, 20:29
Acesse seu Ubuntu do iPad, iPhone e Android	Admin User	Open Source	0	Admin User seg, 31 out 2011, 20:22
Consultas Case Insensitive e Accent Insensitive	Admin User	Desenvolvimento, Banco De Dados	0	Admin User

Figura 5.2: Mensagens inseridas no fórum

Em seguida, foi inserida uma mensagem, conforme apresentado na Figura 5.3. Tal mensagem simula o encaminhamento de uma nova pergunta no fórum.

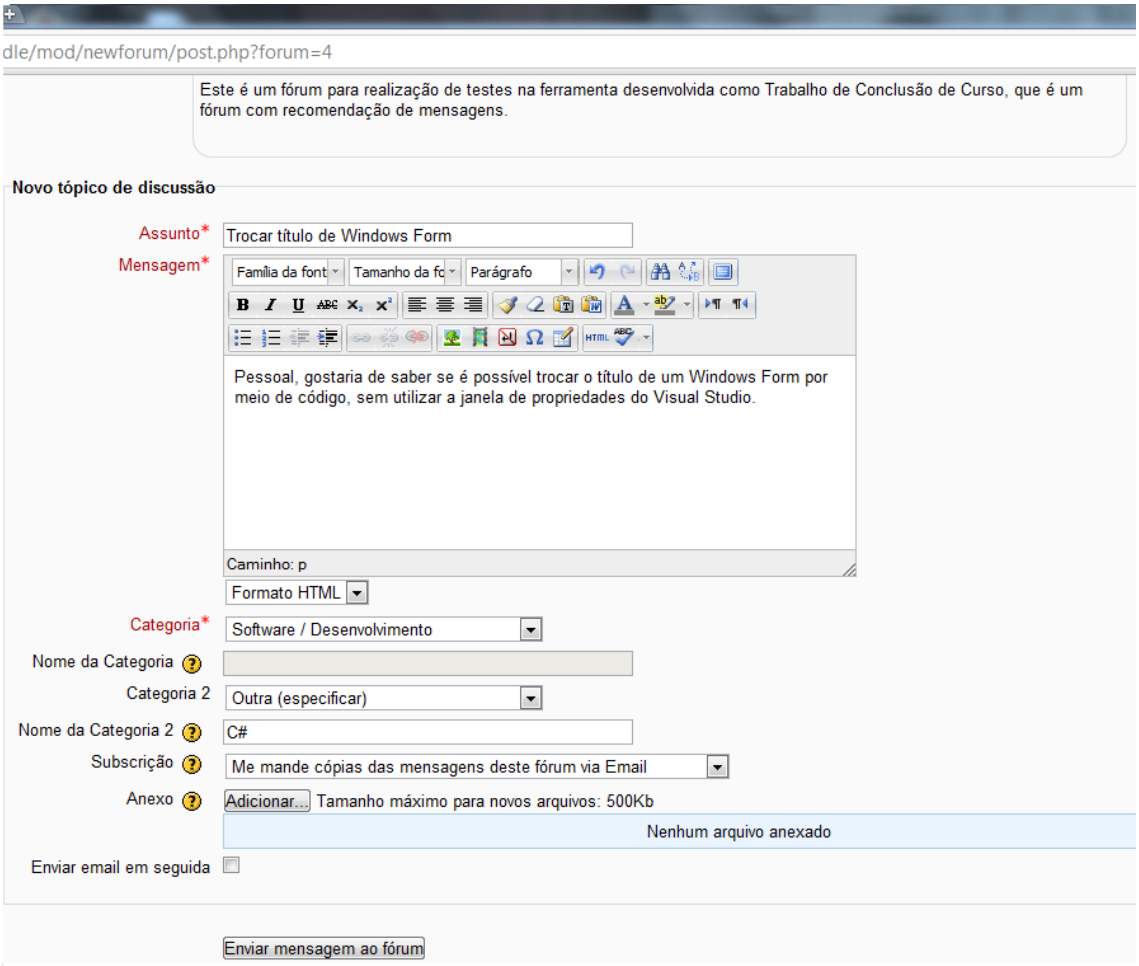

Figura 5.3: Primeira mensagem sendo inserida no fórum

Após o envio da mensagem, a ferramenta buscou mensagens pertencentes à categoria “Software / Desenvolvimento”. Conforme explicado no Capítulo 4, se o usuário selecionar a categoria “Outra”, ele deve inserir o nome da categoria desejada. Nesse caso, o usuário desejava que a categoria de sua mensagem fosse “C#” e essa categoria não existia ainda, por isso foi selecionada a opção “Outra”. Se algum outro usuário enviou uma pergunta ao fórum com essa mesma categoria, então mensagens dela também serão buscadas. Caso contrário, apenas as mensagens da primeira categoria serão comparadas com a mensagem enviada. Após o cálculo da similaridade entre as mensagens, foi realizada uma recomendação. Várias mensagens poderiam ser recomendadas, mas, considerando as mensagens existentes no fórum, apenas uma poderia ser útil para o usuário, que então visualizou a recomendação, como pode ser visto na Figura 5.4.

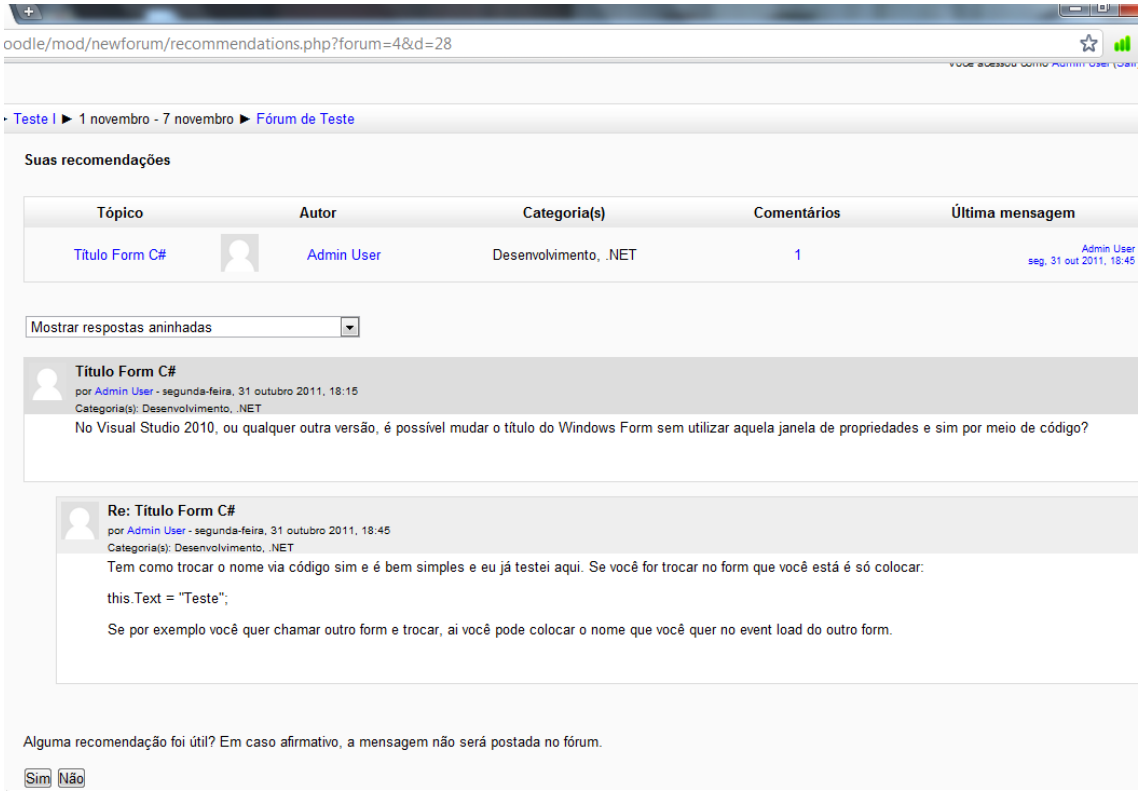

Figura 5.4: Primeira recomendação realizada

Após visualizar a mensagem, o usuário pôde perceber que ela era muito parecida com a sua e resolvia a sua dúvida. Dessa forma, ele selecionou a opção “Sim”, dizendo que a recomendação foi útil. Conforme explicado no Capítulo 4, quando uma recomendação é aceita, a ferramenta assume que a dúvida foi respondida, e a mensagem em questão não é inserida no fórum. Além disso, a recomendação é marcada como aceita no BD, para ser visualizada futuramente no relatório.

Após esse, foi realizado mais um teste, com a mensagem que pode ser visualizada na Figura 5.5.

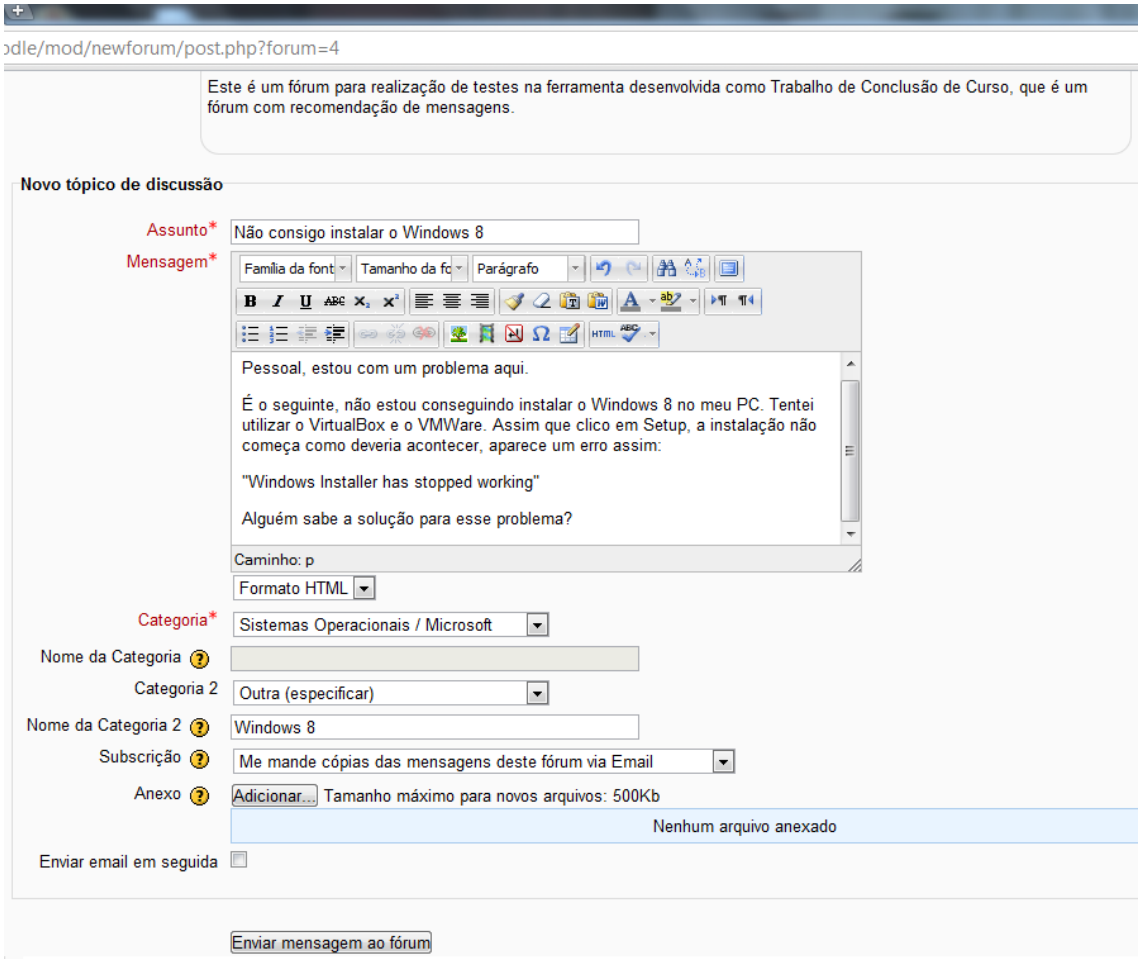

Figura 5.5: Segunda mensagem sendo inserida no fórum

Nesse caso, as mensagens das categorias Sistemas Operacionais / Microsoft e Windows 8 foram analisadas e isso gerou uma recomendação. O usuário então visualizou a recomendação, como pode ser visto na Figura 5.6.

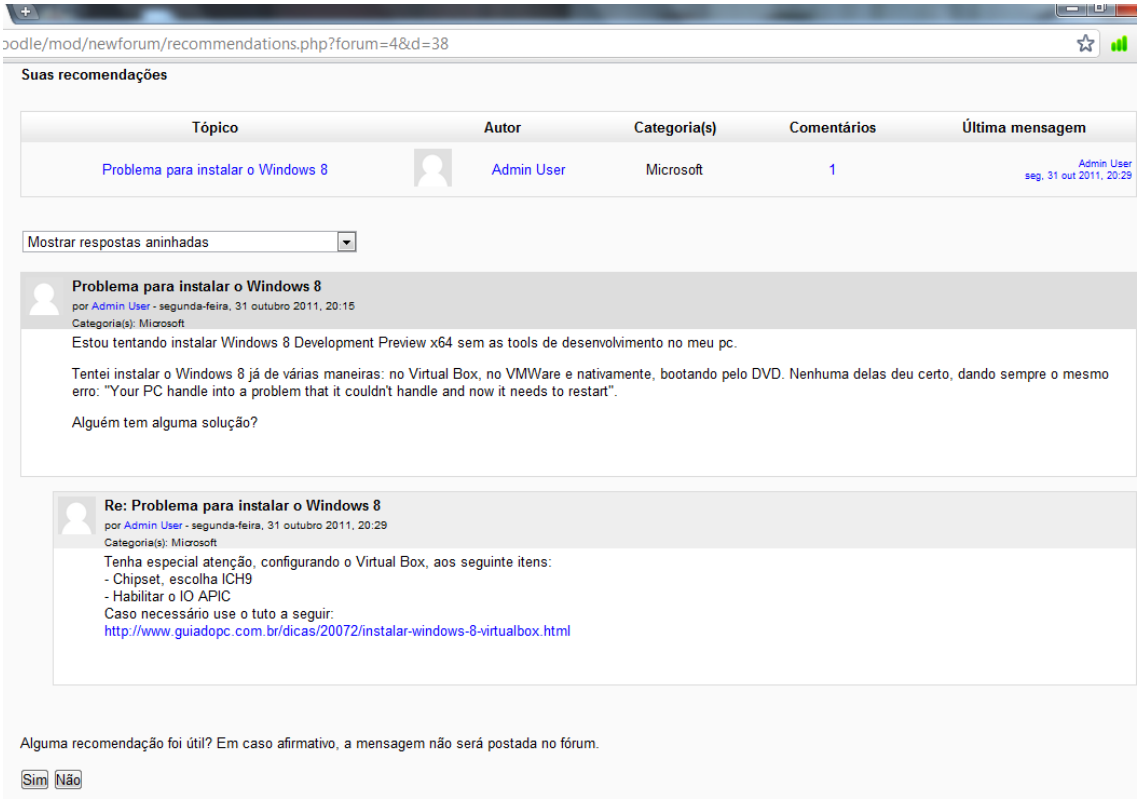

Figura 5.6: Segunda recomendação realizada

Após visualizar a mensagem recomendada, o usuário percebeu que, apesar de ambas falarem sobre um erro durante a instalação do Windows 8, o erro era diferente. Dessa forma, a opção escolhida foi “Não”, pois a solução para os erros seria diferente e a dúvida sendo postada não estava respondida ainda. Como explicado no Capítulo 4, se uma recomendação é recusada, a mensagem é então inserida no fórum para que outros usuários possam visualizá-la e respondê-la. Isso pode ser visto na Figura 5.7. Nessa figura, a mensagem enviada é a primeira da lista, pois as mensagens são ordenadas de forma que as mais recentes apareçam primeiro. Pode-se perceber também que a mensagem enviada anteriormente, cuja recomendação foi aceita, não aparece na lista.

Este é um fórum para realização de testes na ferramenta desenvolvida como Trabalho de Conclusão de Curso, que é um fórum com recomendação de mensagens.

Acrescentar um novo tópico de discussão

Tópico	Autor	Categoria(s)	Comentários	Última mensagem
Não consigo instalar o Windows 8	Admin User	Microsoft, Windows 8	0	Admin User seg, 7 nov 2011, 13:34
Drivers Sony VAIO	Admin User	Drivers	0	Admin User sex, 4 nov 2011, 21:19
Backup Windows Server 2008 em HD USB	Admin User	Microsoft, Windows Server	1	Admin User sex, 4 nov 2011, 21:17
Google e Facebook pedem para instalar Setup_xxxx.exe	Admin User	Segurança	3	Admin User qui, 3 nov 2011, 18:57
Atualizando Leopard para Snow Leopard	Admin User	Mac OS	2	Admin User qui, 3 nov 2011, 18:53
Incompatibilidade entre marcas de memórias DDR3	Admin User	Hardware, Memórias	0	Admin User qui, 3 nov 2011, 18:39
Dúvida sobre limpeza do processador	Admin User	Processadores	1	Admin User qui, 3 nov 2011, 18:37
Processador esquentando	Admin User	Processadores	1	Admin User qui, 3 nov 2011, 18:36
GeForce 8400GS : Compatibilidade	Admin User	Hardware, Placas de Vídeo	1	Admin User qui, 3 nov 2011, 18:28
Problema instalando o Microsoft Visual C++ 2008 Express Edition	Admin User	Software, C++	2	Admin User qui, 3 nov 2011, 18:26
Coloração de Grafos em C	Admin User	Desenvolvimento, C	0	Admin User qui, 3 nov 2011, 18:13
Instalar JavaDoc no Eclipse	Admin User	Java, Eclipse	1	Admin User qui, 3 nov 2011, 18:06

Figura 5.7: Mensagens do fórum após recusa de recomendação

Durante esses testes, algumas categorias não-gerenciadas foram inseridas, ou seja, foi selecionada a opção “Outra” e foi digitado o nome da categoria. Elas não aparecem para os outros usuários escolherem quando vão enviar mensagens ao fórum, mas são inseridas como categorias não-gerenciadas para que o professor possa visualizá-las e decidir o que fazer com as mesmas. Essas categorias podem ser vistas na Figura 5.8.

Categorias Não-Gerenciadas

Nome da Categoria	Tópicos	Adicionar/Remover	Categoria-pai
.NET	1	☐	Nenhuma
VB.NET	1	☐	Nenhuma
Banco De Dados	2	☐	Nenhuma
Windows 8	1	☐	Nenhuma
Java	1	☐	Nenhuma
Eclipse	1	☐	Nenhuma
C	1	☐	Nenhuma
C++	1	☐	Nenhuma
Windows Server	1	☐	Nenhuma
Drivers	1	☐	Nenhuma

Adicionar categorias selecionadas | Remover categorias selecionadas

Figura 5.8: Categorias não-gerenciadas

Nesse caso, foi decidido unificar as categorias ".NET" e "VB.NET" em uma única categoria (".NET"), filha de "Software / Desenvolvimento". Para isso, a categoria ".NET" foi selecionada, a categoria "Software / Desenvolvimento" foi escolhida como "Categoria-pai" e foi escolhida a opção "Adicionar categorias selecionadas". Após isso, a categoria "VB.NET" foi selecionada, tendo como "Categoria-pai" a categoria recém adicionada "Software / Desenvolvimento / .NET" e foi escolhida a opção "Remover categorias selecionadas". Dessa forma, os tópicos da categoria "VB.NET" passaram a ser da categoria ".NET". As categorias "Banco de Dados", "Windows 8", "Java", "C", "C++" e "Windows Server" também foram adicionadas, pois considerou-se que elas poderiam ser relevantes para o curso. As categorias "Eclipse" e "Drivers" continuaram como não-gerenciadas, pois se acreditou que não existiriam outras dúvidas nessas categorias. Na Figura 5.9, pode-se ver as categorias do curso após essas alterações.

Categorias Gerenciadas

Categorias	Tópicos	Editar
Hardware	2	
Memórias	1	
Placas de Vídeo	1	
Placas Mãe	0	
Processadores	2	
Sistemas Operacionais	0	
Mac OS	1	
Microsoft	3	
Windows 8	1	
Windows Server	1	
Open Source	1	
Software	2	
Desenvolvimento	6	
.NET	2	
Banco De Dados	2	
C	1	
C++	1	
Java	1	
Segurança	1	

Figura 5.9: Categorias do curso após os testes

Considerando os testes aqui apresentados, foi gerado o seguinte relatório de recomendações.

Figura 5.10: Relatório de recomendações

Além desse relatório, se o professor tiver conhecimento em Banco de Dados, poderá visualizar as recomendações recusadas, para ter certeza que elas realmente não solucionavam as dúvidas enviadas. Isso pode ser feito acessando-se o BD do Moodle, através de uma consulta SQL na tabela *newforum_recommendations*. No caso de recomendações recusadas, é possível saber qual foi a mensagem recomendada e qual a mensagem inserida no fórum após essa recomendação.

Conforme citado no Capítulo 4, é possível exportar e importar as categorias, para que elas possam ser compartilhadas entre diferentes cursos, sem a necessidade de adicionar manualmente todas as categorias em cada um deles. O formato utilizado para isso é o XML e um exemplo desse arquivo pode ser visualizado abaixo.

```
<?xml version="1.0" encoding="UTF-8"?>
<categories>
 <category>
  <name>Hardware</name>
  <description></description>
  <childs>
   <category>
    <name>Memórias</name>
    <description>Descrição da categoria memórias</description>
    <childs/>
   </category>
   <category>
    <name>Placas de Vídeo</name>
    <description></description>
    <childs/>
   </category>
   <category>
    <name>Placas Mãe</name>
    <description></description>
    <childs/>
   </category>
   <category>
    <name>Processadores</name>
    <description></description>
    <childs/>
   </category>
```

```
  </childs>
 </category>
</categories>
```

Conhecendo a estrutura do arquivo XML, é possível também modificá-lo, se for considerado conveniente, para não precisar excluir ou inserir categorias através da interface da ferramenta. No exemplo acima, supõe-se que a categoria “Placas Mãe” não é considerada relevante, então ela pode ser excluída diretamente no XML, ao invés de ser inserida e depois excluída. Da mesma forma, se desejar-se adicionar uma categoria, é possível adicioná-la no XML antes de importá-lo.

Além dos testes aqui apresentados, outros testes serão realizados na disciplina de “Tópicos Especiais em Computação IX: Engenharia e Desenvolvimento de Aplicações Web”, ministrada pelo professor Leandro Krug Wives. O fórum com recomendação será utilizado para discussões sobre os Seminários de Tecnologias de Desenvolvimento, que acontecerão no período de 11/11/2011 a 02/12/2011. Por questões de tempo, esses testes não serão apresentados aqui.

Além disso, durante a fase de desenvolvimento da ferramenta, foram realizados testes com notícias existentes em sites. Para isso, eram buscadas notícias iguais de lugares diferentes, pois elas eram similares, mas não exatamente iguais. Dessa forma, foi possível verificar se as recomendações estavam sendo realizadas corretamente. Após a conclusão do desenvolvimento, foi realizado também um teste de desempenho, utilizando também notícias de sites. Nesse caso, foram buscadas notícias grandes e inseridas sempre na mesma categoria, para que fossem sempre analisadas. Foi inserido um total de 32 mensagens e verificou-se que o tempo de processamento para analisar as mensagens e realizar as recomendações continuava praticamente o mesmo.

6 CONCLUSÃO

A Educação a Distância vem sendo cada dia mais utilizada, tanto para atividades totalmente a distância quanto como apoio para cursos presenciais. Nesse cenário, uma ferramenta muito importante é o Moodle. Para esclarecer dúvidas, normalmente é utilizado o fórum. Dessa forma, os alunos enviam suas dúvidas para que o professor, ou mesmo os colegas, respondam. No entanto, devido à grande quantidade de mensagens, isso pode acabar não acontecendo. Também devido ao número de mensagens, o usuário pode acabar enviando novamente dúvidas que já foram respondidas, porque não as encontra no fórum.

Como forma de tentar diminuir esses problemas e ajudar tanto professores quanto alunos, foi proposta uma ferramenta de Recomendação de Conteúdo no Fórum do Moodle, que utiliza a técnica Filtragem Baseada em Conteúdo, adotada em Sistemas de Recomendação, além dos métodos de Frequência Relativa de Termos e Similaridade Cosine, utilizados para determinar a similaridade entre documentos.

A principal contribuição deste trabalho para a área acadêmica é criar uma nova forma de interação com o fórum do Moodle, que visa melhorar a experiência dos usuários do mesmo. O desenvolvimento da ferramenta demonstrou que ela pode facilitar muito a interação com os fóruns, principalmente à medida que o número de usuários e mensagens no fórum aumenta.

Inicialmente, pensou-se que interagir com o Moodle e usar sua API seria uma tarefa complexa, mas observou-se que não é tão complicado assim. A documentação é bastante completa e, conforme mostrado no Capítulo 3, para desenvolver um módulo para o mesmo basta seguir alguns passos básicos. Após a criação do módulo, sua instalação também é bastante simples. Uma vantagem dessa abordagem de módulos é que o núcleo do Moodle não é modificado, então os seus administradores não precisam ficar receosos ao instalar novas funcionalidades. Essa também é uma contribuição importante, pois se comprovou que é possível desenvolver um módulo para o Moodle e, assim, várias outras funcionalidades podem ser desenvolvidas para ele.

Nos testes realizados, a ferramenta se comportou de forma satisfatória, recomendando sempre mensagens similares e que, na maioria das vezes, solucionavam as questões.

Como forma de melhorar a ferramenta, poderiam ser testadas outras funções de similaridade para definir qual delas apresenta melhor resultado. Neste trabalho foram testadas as medidas Cosine e Fuzzy, mas existem várias outras medidas que poderiam ser consideradas. Além disso, como as recomendações realizadas são armazenadas no Banco de Dados, elas poderiam ser analisadas ao realizar as recomendações, para alcançar um nível de recomendações ainda melhor. Poderia também ser implementada

uma funcionalidade de sugestão de categorias, para que, assim que o usuário terminasse de escrever sua mensagem, ela fosse analisada e fosse sugerida uma categoria para a mesma. As mensagens poderiam ser classificadas através de TAGS também, ao invés de categorias, como uma maneira mais informal de classificação. Isso poderia ser implementado como outro módulo e, então, seus resultados poderiam ser comparados.

Uma opção interessante seria armazenar também as mensagens cuja recomendação foi aceita, relacionando-as com a recomendação realizada. Isso poderia ser feito através da criação de uma nova mensagem no fórum, com a mensagem recomendada, tendo como resposta as mensagens que receberam e aceitaram essa recomendação. Outra maneira de fazer isso poderia ser através de uma nova funcionalidade no fórum, para que as recomendações aceitas não se misturassem com as mensagens enviadas pelos usuários. Dessa forma, as recomendações não seriam mais agrupadas e o usuário aceitaria ou recusaria cada recomendação individualmente. No caso de uma mensagem receber várias recomendações e a primeira delas já ser útil, as outras não seriam consideradas aceitas ou recusadas, haveria uma nova classificação que seria "não visualizada".

Além das mensagens enviadas ao fórum, outras formas de interação com o Moodle poderiam ser consideradas para a realização das recomendações no fórum. As mensagens enviadas no chat, mensagens enviadas pelo professor aos alunos e até materiais disponibilizados pelo professor poderiam ser fontes de recomendação. Assim a chance de uma dúvida ser respondida antes de ser postada no fórum é ainda maior. Se os históricos de versões passadas do mesmo curso forem importados para os novos cursos, as recomendações poderão ser ainda melhores, pois elas dependem muito dos dados já existentes nos fóruns.

Através do processamento de linguagem natural seria possível obter recomendações melhores, que levariam em consideração também o contexto das palavras e frases utilizadas, ao invés de apenas suas similaridades baseadas nas palavras do texto isoladamente. Como isso é bastante custoso computacionalmente, deveriam ser consideradas também questões de tempo de resposta da ferramenta para saber se valeria a pena utilizar esse tipo de processamento.

Outra idéia que surgiu durante o desenvolvimento do trabalho, e que poderá ser utilizada para novos trabalhos, é desenvolver a ferramenta como um WebService. Esse WebService poderia receber uma mensagem e as credenciais de acesso ao banco como parâmetro e retornar os identificadores das mensagens a serem recomendadas. A estrutura básica das tabelas do banco de dados deveria ser previamente definida, para que o WebService pudesse executar sua tarefa. Isso poderia ser útil para criar um fórum mais genérico, que pudesse ser utilizado em qualquer ambiente sem muita dificuldade.

REFERÊNCIAS

ABEL, F. et al. A Rule-Based Recommender System for Online Discussion Forums. **Adaptative Hypermedia and Adaptative Web-Based Systems**. Heidelberg: Springer Berlin, 2008. p. 12-21. Disponível em: < http://www.springerlink.com/content/2863118582q74534/>. Acesso em 26 set. 2011.

ADOMAVICIUS, G.; TUZHILIN, A. Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. **IEEE Transactions on Knowledge and Data Engineering**, v. 17, n. 6, p. 734-749, jun. 2005. Disponível em: < http://www.computer.org/portal/web/csdl/doi/10.1109/TKDE.2005.99>. Acesso em 23 set. 2011.

BALABANOVIC, M.; SHOHAM, Y. Fab: Content-Based, Collaborative Recommendation. **Communications of the ACM**, Nova Iorque, v. 40, n. 3, p. 66-72, mar. 1997. Disponível em: <http://dl.acm.org/citation.cfm?id=245124>. Acesso em 23 set. 2011.

BEZERRA, B. L. D. **Estudo de algoritmos de filtragem de informação baseados em conteúdo**. 2002. 45f. Trabalho de graduação em Inteligência Artificial – Centro de Informática, Universidade Federal de Pernambuco, Recife. Disponível em: <http://www.cin.ufpe.br/~tg/2001-2/bldb.pdf>. Acesso em 23 set. 2011.

CAZELLA, S. C. **Aplicando a relevância da opinião de usuários em sistemas de recomendação para pesquisadores.** 2006. 180f. Tese (Doutorado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre. Disponível em: < http://hdl.handle.net/10183/8424>. Acesso em 13 dez. 2011.

CLAYPOOL, M. et al. Combining Content-Based and Collaborative Filters in an Online Newspaper. In: ACM/SIGIR WORKSHOP ON RECOMMENDER SYSTEMS, 1999, Berkeley. **Proceedings...** Nova Iorque: ACM, 1999. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.45.5230>. Acesso em 26 out. 2011.

FISCHER, C. T. **Recomendação de Conteúdo em Fóruns Eletrônicos**. 2005. 82f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Instituto de Ciências Exatas e Tecnológicas, Centro Universitário Feevale, Novo Hamburgo.

FOWLER, M. et al. **Patterns of enterprise application architecture**. Boston: Addison-Wesley, 2008.

GARIN, R. S. et al. O uso de técnicas de recomendação em um sistema para apoio à aprendizagem colaborativa. **Revista brasileira de informática na educação**, v. 14, n.3,

p. 49-59. Disponível em: <http://hdl.handle.net/10183/23202>. Acesso em 13 dez. 2011.

HESPANHA, D. Novos dados da EAD no Brasil. **Blog "Educação & Tecnologia"**, 2009. Disponível em: <http://ticsnaeducacao.wordpress.com/2009/12/23/novos-dados-do-ead-no-brasil/>. Acesso em 26 out. 2011.

HUANG, A. Similarity Measures for Text Document Clustering. In: NEW ZEALAND COMPUTER SCIENCE RESEARCH STUDENT CONFERENCE, 2008, Christchurch. **Proceedings...** New Zealand: University of Canterbury, 2008. p. 49-56. Disponível em: <http://nzcsrsc08.canterbury.ac.nz/site/proceedings/Individual_Papers/pg049_Similarity_Measures_for_Text_Document_Clustering.pdf>. Acesso em 28 set. 2011.

HUANG, Z. et al. A Graph-Based Recommender System for Digital Library. In: 2nd ACM/IEEE-CS JOINT CONFERENCE ON DIGITAL LIBRARIES, 2002, Portland. **Proceedings...** Nova Iorque: ACM, 2002. Disponível em: <http://dl.acm.org/citation.cfm?id=544231>. Acesso em 26 out. 2010.

IEA. Dados comprovam forte ritmo de expansão da EAD. **Site do Instituto de Estudos Avançados**, 2010. Disponível em: <http://www.iea.org.br/noticias/dados-comprovam-forte-ritmo-de-expansao-da-ead>. Acesso em 27 set. 2011.

KROHN, F. **Framework para Recomendação de Novos Relacionamentos em uma Rede Social a partir do Uso de Técnicas de Folksonomia**. 2009. 92f. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Unidade Acadêmica de Graduação, Universidade do Vale do Rio dos Sinos, São Leopoldo.

LOPES, G. R. **Sistemas de Recomendação para Bibliotecas Digitais sob a Perspectiva da Web Semântica**. 2007. 69f. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre. Disponível em: < http://hdl.handle.net/10183/10747>. Acesso em 26 set. 2011.

MACMANUS, R. A Guide to Recommender Systems. **Site Read Write Web**, 2009. Disponível em: <http://www.readwriteweb.com/archives/recommender_systems.php>. Acesso em 26 set. 2011.

MAIA, L. C. G. **Uso de Sintagmas Nominais na Classificação Automática de Documentos Eletrônicos**. 2008. 158f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, UFMG, Belo Horizonte. Disponível em: <http://www.luizmaia.com.br/ogma/TeseFinal.pdf>. Acesso em 27 set. 2011.

MOODLE. 2010. **Database schema introduction**. Disponível em: <http://docs.moodle.org/dev/Database_schema_introduction>. Acesso em 07 out. 2011.

MOODLE. 2011. **Developer documentation**. Disponível em: <http://docs.moodle.org/dev/Developer_documentation>. Acesso em 07 out. 2011.

MOODLE. 2011. **Development hints and tips**. Disponível em: <http://docs.moodle.org/dev/Development_hints_and_tips>. Acesso em 13 out. 2011.

MOODLE. 2011. **DML functions**. Disponível em: <http://docs.moodle.org/dev/DML_functions>. Acesso em 13 out. 2011.

MOODLE. 2011. **Forums**. Disponível em: <http://moodle.org/forums/>. Acesso em 07 out. 2011.

MOODLE. 2011. **Introduction to Moodle Programming**. Disponível em: <http://dev.moodle.org/login/index.php>. Acesso em 07 out. 2011.

MOODLE. 2011. **Modules**. Disponível em: <http://docs.moodle.org/dev/Modules>. Acesso em 07 out. 2011.

MOODLE. 2011. **Moodle architecture**. Disponível em: <http://docs.moodle.org/dev/Moodle_architecture>. Acesso em 06 out. 2011.

MOODLE. 2011. **NEWMODULE Documentation**. Disponível em: <http://docs.moodle.org/dev/NEWMODULE_Documentation>. Acesso em 07 out. 2011.

MOODLE. 2011. **NEWMODULE Tutorial**. Disponível em: <http://docs.moodle.org/dev/NEWMODULE_Tutorial>. Acesso em 07 out. 2011.

MOODLE. 2011. **NEWMODULE.zip**. Disponível em: <http://download.moodle.org/plugins/mod/NEWMODULE.zip>. Acesso em 07 out. 2011.

MOODLE. 2011. **O que é o Moodle?** Disponível em: <http://moodle.org/about/>. Acesso em 06 out. 2011.

MOODLE. 2011. **Translation**. Disponível em: <http://docs.moodle.org/20/en/Translation>. Acesso em 13 out. 2011.

MOONEY, R. J. Collaborative Filtering & Content-Based Recommending. 2010. Disponível em: <http://www.cs.utexas.edu/~mooney/ir-course/slides/Recommenders.ppt>. Acesso em: 23 set. 2011.

MORAN, J. M. A educação a distância como opção estratégica. **Site Educação Humanista Inovadora**, 2011. Disponível em: <http://www.eca.usp.br/prof/moran/estrategica.html>. Acesso em 27 set. 2011.

OLIVEIRA, A. de M.; CARNEIRO, M. L. F.; MUNHOZ, A. M. Análise do Ambiente Virtual do Moodle como Tecnologia de Apoio aos estudantes de Biblioteconomia. In: ENCONTRO NACIONAL DE ESTUDANTES DE BIBLIOTECONOMIA, DOCUMENTAÇÃO, CIÊNCIA E GESTÃO DA INFORMAÇÃO, 2009, Rio de Janeiro, **Anais eletrônicos...**, Rio de Janeiro: Universidade Estadual do Rio de Janeiro, 2009. Disponível em <http://www.bibliotecadigital.ufrgs.br/da.php?nrb=000730984&loc=2011&l=63799a04083366e7>. Acesso em 21 set. 2011.

ORENGO, V. M.; HUYCK, C. A Stemming Algorithm for the Portuguese Language. In: EIGHTH INTERNATIONAL SYMPOSIUM ON STRING PROCESSING AND INFORMATION RETRIEVAL, 2001, Laguna de San Rafael. **Proceedings...** Los Alamitos: IEEE Computer Society, 2001. p. 186-193. Disponível em: < http://www.inf.ufes.br/~claudine/courses/ct08/transparencias/a_stemming_algorithm_for_the_portuguese_language.pdf>. Acesso em 17 out. 2011.

RESNICK, P. et al. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In: ACM CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK, 1994, Chapel Hill. **Proceedings...**, Nova Iorque: ACM, 1994. Disponível em: <http://dl.acm.org/citation.cfm?id=192905>. Acesso em: 26 out. 2011.

RESNICK, P.; VARIAN, H. R. Recommender Systems. **Communications of the ACM**, Nova Iorque, v. 40, n. 3, p. 56-58, mar. 1997. Disponível em: < http://dl.acm.org/citation.cfm?id=245121>. Acesso em 23 set. 2011.

RIZZI, C. et al. Fazendo uso de Categorização de Textos em Atividades Empresariais. In: INTERNATIONAL SYMPOSIUM ON KNOWLEDGE MANAGEMENT/DOCUMENT MANAGEMENT (ISKM/DM 2000), III, 2000, Curitiba. **Proceedings...** Curitiba: PUC-PR, 2000. p. 125-147. Disponível em: < http://www.leandro.wives.nom.br/pt-br/publicacoes/iskmdm2000-2.pdf>. Acesso em 28 set. 2011.

ROMERO, C.; VENTURA, S.; GARCÍA, E. Data mining in course management systems: Moodle case study and tutorial. **Computers & Education**, v. 51, i. 1, p. 368-384, ago. 2008. Disponível em: <http://150.214.190.154/docencia/doctoM6/Romero-Ventura-Garcia-CE.pdf>. Acesso em 21 set. 2011.

SARWAR, B. et al. Item-Based Collaborative Filtering Recommendation Algorithms. In: 10th INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, 10., 2001, Hong Kong. **Proceedings...**, Nova Iorque: ACM, 2001. Disponível em: < http://dl.acm.org/citation.cfm?id=372071>. Acesso em 22 set. 2011.

SCHAEFER, J. B.; KONSTAN J.; RIEDL, J. Recommender Systems in E-Commerce. In: 1st ACM CONFERENCE ON ELECTRONIC COMMERCE, 1999, Denver. **Proceedings...**, Denver: ACM, 1999. Disponível em: < http://dl.acm.org/citation.cfm?id=337035 >. Acesso em 26 set. 2011.

SCORSOLINI-COMIN, F. et al. A interação em um ambiente virtual de aprendizagem por meio do fórum de discussão: uma leitura bakhtiniana. **Revista PerCursos**, Florianópolis, v. 10, n. 1, p. 62-75, jan. 2009/ jun. 2009. Disponível em: < http://www.periodicos.udesc.br/index.php/percursos/article/viewFile/1582/1456>. Acesso em 21 set. 2011.

SIMONSON, M. Course Management Systems. **The Quarterly Review of Distance Education**, North Miami Beach, v. 8, n. 1, p. vii-ix, 2007. Disponível em: <http://web2integration.pbworks.com/f/COURSE+MANAGEMENT+SYSTEMS.pdf>. Acesso em 21 set. 2011.

SNOWBALL. 2011. **Lista de stop-words para Português**. Disponível em: <http://snowball.tartarus.org/algorithms/portuguese/stop.txt>. Acesso em 11 out. 2011.

TORRES, R. D. **Combining Collaborative and Content-based Filtering to Recommend Research Papers**. 2004. 71f. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre. Disponível em: < http://www.lume.ufrgs.br/handle/10183/5887>. Acesso em 23 set. 2011.

WIKIPEDIA. 2011. **CDNow**. Disponível em: <http://en.wikipedia.org/wiki/CDNow>. Acesso em 26 set. 2011.

WIKIPEDIA. 2011. **XML**. Disponível em: <http://pt.wikipedia.org/wiki/XML>. Acesso em 10 out. 2011.

WIVES, L. K. **Um estudo sobre Agrupamento de Documentos Textuais em Processamento de Informações não Estruturadas Usando Técnicas de "Clustering".** 1999. 84f. Dissertação (Mestrado em Ciência da Computação) –

Instituto de Informática, UFRGS, Porto Alegre. Disponível em: < http://www.leandro.wives.nom.br/pt-br/publicacoes/dissertacao.pdf>. Acesso em 28 set. 2011.

WIVES, L. K. **Utilizando *conceitos* como descritores de textos para o processo de identificação de conglomerados (*clustering*) de documentos.** 2004. 136f. Tese (Doutorado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre. Disponível em: <http://www.leandro.wives.nom.br/pt-br/publicacoes/Tese.pdf>. Acesso em 28 set. 2011.