

# ***SysPaperAdvice: Desenvolvimento de um Protótipo de Sistema de Recomendação de Artigos e Trabalhos Científicos***

**Adriano de Barros<sup>1</sup>, Sidnei Renato Silveira<sup>2</sup>, Solange Pertile<sup>2</sup>**

<sup>1</sup>Curso de Bacharelado em Sistemas de Informação - <sup>2</sup>Departamento de Tecnologia da Informação – Universidade Federal de Santa Maria (UFSM) – Campus Frederico Westphalen – RS - Brasil

adrianob23@hotmail.com, sidneirenato.silveira@gmail.com,  
solangepertile@gmail.com

**Resumo.** *Sistemas de Recomendação são aplicações utilizadas para filtrar informações na Internet, mostrando as principais sugestões encontradas para busca feita pelo usuário que utiliza o sistema. Entretanto, a maioria das pesquisas ainda são realizadas diretamente pelos navegadores ou por meio de ferramentas de busca, fazendo com que o usuário filtre os resultados manualmente, perdendo muito tempo. Nesse contexto, esse trabalho apresenta um protótipo de Sistema de Recomendação de Artigos e Trabalhos Científicos, o SysPaperAdvice, implementado por meio de técnicas de Filtragem Baseada em Conteúdo e de Filtragem Colaborativa, relacionadas com o perfil do usuário cadastrado, para apoiar estudantes, professores e pesquisadores na busca por fontes de informação.*

**Palavras-Chave:** *Sistemas de Recomendação, Sistemas de Informação Inteligentes, Filtragem de Informação.*

**Abstract.** *Systems Recommendation is an application used to filter information on the Internet, showing the main suggestions found for search done by the user who uses the system. However, most searches are still performed directly by browsers or through search tools, causing the user to filter results manually, wasting time. In this context, this work presents a Prototype of a System of Recommendation of Scientific Papers, the SysPaperAdvice is implemented through Content-Based Filtering techniques, Collaborative Filtration, related to the registered user profile, to support students, professors and researchers in the search for information sources.*

**Keywords:** *Recommendation Systems, Intelligent Information Systems, Information Filtering.*

## **1. Introdução**

Sistemas de Recomendação (SRs) são sistemas que produzem recomendações individualizadas como saída, com o objetivo de guiar o usuário de forma personalizada a objetos interessantes e úteis, diante de uma grande variedade de opções [BURKE, 2002citado por JESUS; ESCOBAR,2011]. Nos Sistemas de Recomendação são utilizadas, de forma geral, uma das três técnicas de filtragem de informação citadas a seguir: Filtragem Baseada em Conteúdo (FBC), Filtragem Colaborativa (FC) e

Filtragem Híbrida (FH) [CAZELLA et al, 2010]. No trabalho aqui apresentado utilizamos a filtragem híbrida, combinando técnicas de FBC e FC.

Sistemas de Recomendação (SRs) são utilizados por muitos *sites* e serviços, sendo ferramentas importantes para auxiliar o usuário a encontrar o que é mais relevante diante da imensa quantidade de informação disponível [RODRIGUES JÚNIOR, 2016]. O primeiro Sistema de Recomendação (SR), desenvolvido no Centro de Pesquisa da Xerox em Palo Alto, denominado *Tapestry*, tinha como motivação filtrar a grande quantidade de *e-mails* que começava a incomodar os usuários, devido à expansão do acesso à Internet [RESNICK; VARIAN, 1997] [GOLDBERG et al., 1992].

Atualmente, quando navegamos na Internet, temos nossas informações de pesquisa rastreadas por diversos mecanismos de recomendação. A tarefa de recomendação serve para oferecer produtos dos mais variados tipos. Porém, quando optamos pela alternativa de pesquisar na Internet utilizando buscadores, tais como o *Google*, informações de artigos, livros e trabalhos científicos, acabamos surpreendidos com uma enorme lista de *links* como resultado de nossas buscas.

Para resolver o problema da quantidade de informação disponível e filtrar o conteúdo na Internet, os Sistemas de Recomendação (SRs) são fundamentais, uma vez que permitem aos usuários compartilhar informações sobre materiais consultados, possibilitando também que registrem suas impressões sobre estes materiais [CAZELLA et al., 2010]. Quando existe uma interação com o usuário, um perfil é criado no sistema, formando uma recomendação específica para o próximo acesso ao sistema, contendo uma relação de itens que podem ser músicas, livros, artigos ou outros tipos de produtos de consumo do usuário. O processo de identificação do usuário é opcional em um SR, porém, quando é realizado, esse é o primeiro passo a ser feito, que consiste na coleta dos dados, que são armazenados em uma base, para posteriormente uso na recomendação [BARCELLOS et al., 2007].

Nesse contexto, este artigo apresenta o desenvolvimento de um protótipo de Sistema de Recomendação de Artigos e Trabalhos Científicos, que foi chamado de *SysPaperAdvice*, baseado nos próprios documentos que são armazenados no sistema pelos usuários. A intenção é auxiliar alunos, professores e outros usuários a encontrar o conteúdo com mais facilidade, sem perder muito tempo com pesquisas na *web*, utilizando informações do perfil do usuário e de avaliações realizadas pelos usuários sobre os documentos. O desenvolvimento do protótipo utilizou a filtragem híbrida, combinando técnicas de Filtragem Baseada em Conteúdo e Filtragem Colaborativa.

Para dar conta dessa proposta, o presente artigo está organizado da seguinte forma: a seção 2 apresenta um referencial teórico, detalhando as técnicas escolhidas para elaboração do protótipo do Sistema de Recomendação. Na seção 3 são apresentados os trabalhos relacionados, que compõem o estado da arte. A seção 4 apresenta a modelagem, implementação e validação do protótipo de SR desenvolvido. Finalizando o artigo são apresentadas as considerações finais e as referências empregadas.

## 2. Referencial Teórico

Esta seção apresenta um referencial teórico sobre as áreas envolvidas neste artigo, destacando conceitos de Sistemas de Recomendação (SRs), Perfil de Usuário, Filtragem Baseada em Conteúdo (FBC), Filtragem Colaborativa (FC) e Filtragem Híbrida (FH).

## 2.1 Sistemas de Recomendação

A grande quantidade de conteúdo existente na *web* fez surgir a necessidade da criação de ferramentas de auxílio ao usuário na busca pela informação desejada. Para suprir essa necessidade inicialmente foram utilizados Sistemas de Recuperação de Informação [JESUS; ESCOBAR, 2011]. O problema da técnica de Recuperação da Informação é retornar para o usuário que fez a busca, todo o conteúdo encontrado na *Web*, obrigando que sejam feitos filtros, manualmente, na lista retornada. Para tentar minimizar os problemas de sobrecarga de informação surgiram os Sistemas de Recomendação (SRs) [PRIMO; LOH, 2006 citados por JESUS; ESCOBAR2011].

Para atender à necessidade dos usuários, um SR geralmente possui [BURKE, 2002]:

- Dados prévios ou armazenados: corresponde a toda informação que o sistema armazena para utilizar no processo de recomendação;
- Dados de entrada ou do usuário: tratam-se de informações fornecidas pelo usuário para que o processo de recomendação possa ser iniciado;
- Algoritmo de recomendação: é responsável por combinar dados prévios e dados de entrada para que sejam fornecidas as devidas recomendações.

Um dos grandes desafios dos Sistemas de Recomendação (SRs) é realizar o casamento correto entre os que estão recomendando e aqueles que estão recebendo a recomendação, ou seja, definir e descobrir este relacionamento de interesses [REATEGUI; CAZELLA, 2005 citados por LINHARES, 2012].

## 2.2 Perfil de Usuário

Para fornecer informações personalizadas o SR precisa criar um modelo ou perfil de usuário. O modelo de usuário serve para que o sistema conheça as preferências e restrições de quem se deseja gerar recomendação, para que assim possa apresentar apenas o que for de interesse do usuário [RODRIGUES JÚNIOR, 2016].

As informações contidas no perfil do usuário devem ser relevantes para o SR que se deseja desenvolver. Além disso, a técnica de representação usada na concepção do perfil deve ser adequada para representar as informações relevantes do contexto em que o sistema atua.

Entre as técnicas existentes para criar um perfil de usuário, três delas são citadas por Barcellos et al. [2007], para coletar os dados:

- Explícita: as informações são fornecidas pelo próprio usuário, por meio do preenchimento de um formulário ou inserção de documentos e/ou outros tipos de itens no SR, onde o mesmo informa suas preferências;
- Implícita: as informações são coletadas pelo monitoramento da navegação do usuário com o sistema, utilizando informações tais como as páginas consultadas, histórico de compras e conteúdos selecionados, entre outros;

- Inferência: consiste em descobrir o perfil do usuário com base no comportamento de outros usuários (Grupos), com padrão de comportamento similar ao seu.

Na implementação de um SR, podem ser utilizadas uma ou mais das técnicas apresentadas para criar o perfil do usuário. No sistema desenvolvido foi utilizado a técnica: 1) explícita, contendo um formulário para preenchimento de informações para compor o perfil do usuário e inserção de documento, também utilizando a pontuação feita por outros usuários para os artigos existentes na base do SR.

### 2.3 Tipos de Filtragem de Informações

Para realizar a filtragem das informações que serão apresentadas pelo SR aos usuários, podemos aplicar técnicas baseadas em conteúdo, na similaridade entre os usuários ou combinar estas técnicas, como será visto nas próximas subseções.

#### 2.3.1 Filtragem Baseada em Conteúdo

Nos SRs que utilizam a Filtragem Baseada em Conteúdo (FBC), as sugestões de itens para um usuário são feitas comparando-se o conteúdo de um item com o de outros que o usuário já tenha manifestado interesse anteriormente [NASCIMENTO, 2011]. Um filme pode ter como características, título, diretor, elenco, gênero etc. Já em um texto, as próprias palavras que o compõem são consideradas como suas características. É com base nessas características que os itens podem ser comparados e a semelhança entre eles estabelecida [OLIVEIRA, 2007].

Uma técnica muito utilizada nos sistemas implementados com FBC baseados em textos é a abordagem de indexação de frequência de termos (*term frequency indexing*) [HERLOCKER, 200 citado por CAZELLA et al., 2010]. Neste tipo de indexação, informações dos documentos e necessidades dos usuários são descritas por vetores com uma dimensão para cada palavra que ocorre na base de dados. Cada componente do vetor é a frequência com que a respectiva palavra ocorre em um documento ou na consulta do usuário. Em relação à FBC, temos que observar o pré-processamento de textos, onde ocorrem atividades tais como remoção de *stopwords*, análise léxica e *stemming*, que evitam que palavras que não são relevantes sejam consideradas palavras-chaves ou comparadas às disponibilizadas pelo usuário. Cazella et al. [2010] detalham estas atividades:

- *Stopwords*: Remoção de termos que não são relevantes, como artigos, preposições, conjunções e pronomes, visando eliminar estas palavras dos textos. São eliminados, por exemplo: “mas”, “mesmo”, “qualquer”, “seja”, “ainda”;
- Análise Léxica: Dígitos, sinais de pontuação, acentos e hífens podem ser removidos neste processo, e todas as letras do texto podem ser transformadas para maiúsculas ou minúsculas;
- *Stemming*: É realizado considerando cada palavra isoladamente, tentando reduzir a sua provável raiz. São analisadas características como grau, gênero e número com o objetivo eliminar os sufixos e prefixos das palavras transformando-as em

sua forma primitiva. Por exemplo, informação e informática tem, como raiz, a palavra “informa”.

Dentre as principais vantagens da técnica de FBC, destaca-se o fato de ela não precisar que um item já tenha sido avaliado para gerar a recomendação, além de todos os itens terem a chance de serem recomendados, já que dependem apenas do perfil do usuário [BARBOSA, 2014].

Entre as desvantagens, a FBC deixa o usuário limitado aos itens relacionados ao seu perfil. Este problema é conhecido como superespecialização. Outro problema é que a FBC não funciona bem em domínios que não sejam textuais, como imagens, vídeos e áudios, por ser difícil extrair os atributos relevantes dos mesmos [BARBOSA, 2014].

### 2.3.2 Filtragem Colaborativa

Em SRs, a Filtragem Colaborativa (FC) é usada quando podemos assumir que um usuário se comporta de forma semelhante a outros usuários, ou quando itens possuem a mesma demanda [BARTH, 2010].

O termo Filtragem Colaborativa (FC) foi usado pela primeira vez por David Goldberg em um artigo intitulado *Using Collaborative Filtering to Weave na Information Tapestry*. Goldberg desenvolveu um sistema chamado *Tapestry* que permitia que os usuários fizessem anotações em documentos, marcando-os como interessante ou desinteressante. A partir da análise dessas informações, o sistema conseguia filtrar estes documentos para seus usuários [SEGARAM, 2008 citado por JESUS; ESCOBAR, 2011].

Os primeiros SRs que utilizavam a FC requeriam usuários para especificar o relacionamento de predição entre suas opiniões, ou de modo explícito, indicar os itens de interesse [HERLOCKER et al., 2000 citados por REATEGUI; CAZELLA, 2005]. Posteriormente estes sistemas automatizaram todo o procedimento por meio da coleção das pontuações dos itens pelos usuários. Um usuário de um sistema deve, portanto, pontuar cada item experimentado, indicando o quanto este item casa com a sua necessidade. Estas pontuações são coletadas para grupos de pessoas, permitindo que cada usuário se beneficie das pontuações (experiências) apresentadas por outros usuários. De acordo com Reategui e Cazella [2005] a técnica de FC segue basicamente três passos:

1. Calcular o peso de cada usuário em relação à similaridade ao usuário ativo (métrica de similaridade);
2. Selecionar um subconjunto de usuários com maiores similaridades (vizinhos) para considerar na predição;
3. Normalizar as avaliações e computar as predições ponderando as avaliações dos vizinhos com seus pesos.

No primeiro passo, a definição da similaridade pode ser realizada por meio de diversas técnicas, sendo mais comumente usada a correlação de *Pearson*, cuja fórmula é apresentada na Figura 1.

$$w_{a,u} = \frac{\sum_{i=1}^m (r_{a,i} - \bar{r}_a) * (r_{u,i} - \bar{r}_u)}{\sigma_a * \sigma_u}$$

**Figura 1 – Fórmula da Correlação de Pearson [REATEGUI; CAZELLA 2005]**

De acordo com a fórmula da Figura 1, dado o conjunto de avaliações  $r_{a,i}$  do usuário ativo  $a$ , e o conjunto de avaliações  $r_{u,i}$  do vizinho  $u$ ,  $w_{a,u}$  é o índice de similaridade entre os dois usuários. Devemos notar que é preciso mais de uma avaliação em comum para que o índice seja útil, e os resultados variam entre 1 para similaridade total, e -1 para total dissimilaridade [REATEGUI; CAZELLA, 2005].

No passo três, o cálculo da predição pode ser efetuado por meio da média ponderada, aplicando a equação apresentada na Figura 2.

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_u) * w_{a,u}}{\sum_{u=1}^n w_{a,u}}$$

**Figura 2 –Fórmula do Cálculo de Predição [REATEGUI; CAZELLA 2005]**

Analizando a fórmula da Figura 2, o valor da predição  $p_{a,i}$  do item  $i$  para o usuário ativo  $a$  é a média ponderada das avaliações dadas ao item  $i$  pelos  $n$  vizinhos  $u$  do usuário ativo  $a$ . A quantidade  $n$  de melhores vizinhos – com maiores correlações – é definida pelos responsáveis pela implementação do SR. [REATEGUI; CAZELLA 2005].

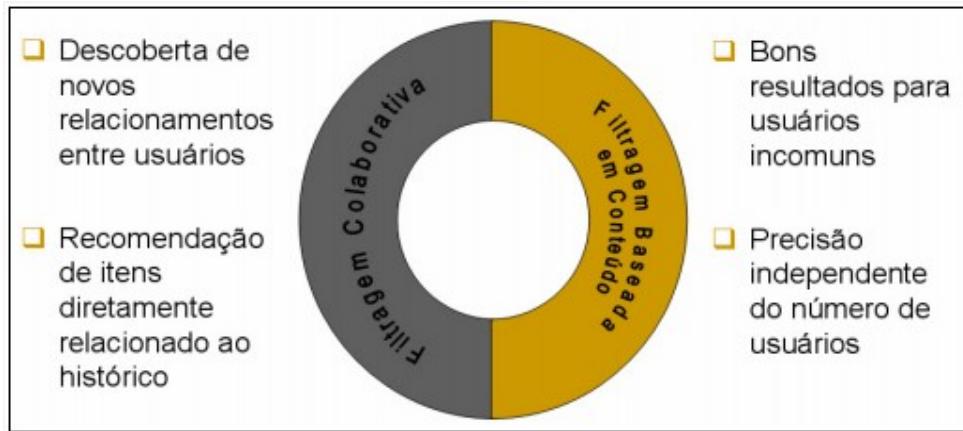
Uma das vantagens da técnica de FC é a de que possui um modelo conceitual de operação de fácil entendimento, possibilitando analisar itens a serem recomendados sem preocupar-se com o conteúdo destes itens, focando nas avaliações dos itens [HERLOCKER et al., 2000]. A FC apresenta outra vantagem, que é a possibilidade de apresentar aos usuários recomendações inesperadas. O usuário poderia receber recomendações de itens que não estavam sendo pesquisados de forma ativa [REATEGUI; CAZELLA, 2005].

Uma das principais desvantagens da Filtragem Colaborativa (FC) é que um usuário recém adicionado ao sistema ou que apresente um número muito pequeno de avaliações, faz com que o sistema não seja capaz de calcular precisamente a sua similaridade com os demais usuários. De maneira similar, um item recém adicionado ou que possua poucas avaliações não será recomendado até que seja avaliado por uma quantidade suficiente de usuários [BARBOSA, 2014].

### 2.3.3 Filtragem Híbrida

A abordagem da Filtragem Híbrida (FH) procura combinar os pontos fortes da Filtragem Colaborativa e da Filtragem Baseada em Conteúdo, visando criar um sistema que possa

melhor atender às necessidades do usuário [HERLOCKER, 2000], unindo as vantagens das duas técnicas e eliminando as fraquezas de cada uma [REATEGUI; CAZELLA, 2005]. A Figura 3 apresenta as vantagens da FH.



**Figura 3 - Filtragem Híbrida [REATEGUI; CAZELLA, 2005]**

Na Filtragem Híbrida (FH), as principais estratégias utilizadas, no que diz respeito à forma como componentes serão combinados para gerar as recomendações segundo Barbosa [2014], são:

- Ponderada: nesta abordagem a FBC e a FC são implementadas separadamente e uma combinação linear é feita com os resultados;
- Mista: nesta abordagem as recomendações geradas pelas duas técnicas são combinadas no processo final de recomendação, de tal forma que as duas recomendações sejam apresentadas ao usuário na mesma lista;
- Combinação sequencial: nesta abordagem a FBC cria os perfis dos usuários e, posteriormente, estes perfis são usados no cálculo da similaridade da FC;
- Comutação: nesta abordagem o sistema utiliza algum critério, como por exemplo a confiança no resultado, para comutar ou chavear entre a FBC e a FC. Pode-se também realizar a comutação de uma técnica nos pontos de desvantagens da outra técnica.

O protótipo de SR aqui apresentado – o *SysPaperAdvice* - foi implementado utilizando a FH com a técnica ponderada, mostrando as recomendações em duas páginas, apresentando ao usuário resultados, independentemente da quantidade de dados inseridos pelo mesmo no sistema.

### 3. Trabalhos Relacionados

Nesta seção são apresentados trabalhos correlacionados ao aqui apresentado, envolvendo a implementação de SRs. Além disso, no final da seção, apresentamos um comparativo entre os trabalhos estudados e o SR implementado.

### 3.1 Sistema SisRecAC

Avila[2006] apresenta o desenvolvimento de um SRdenominado *SiscRecAc*, que permite ao usuário enviar documentos e receber recomendações, podendo avaliar as recomendações recebidas. No sistema acontece a extração das palavras-chave por meio da contagem do número de ocorrências no texto, não considerando as palavras consideradas *stopwords*. A lista das palavras é submetida ao *Google Acadêmico*, que retorna 30 *links* por palavra submetida, recomendando os 5 *links* mais frequentes. Os documentos inseridos no sistema estão no formato *PDF* (*Portable Document Format*) e são convertidos para modo texto pelo *software xpdf* (<http://www.foolabs.com/xpdf/>).

As técnicas de recomendação utilizadas para desenvolver o sistema *SisRecAC* foram a FBC (por meio de palavras-chave no texto) e a FC (por meio da recomendação de documentos para outras pessoas). A Figura 4 apresenta a arquitetura do *SisRecAc*.

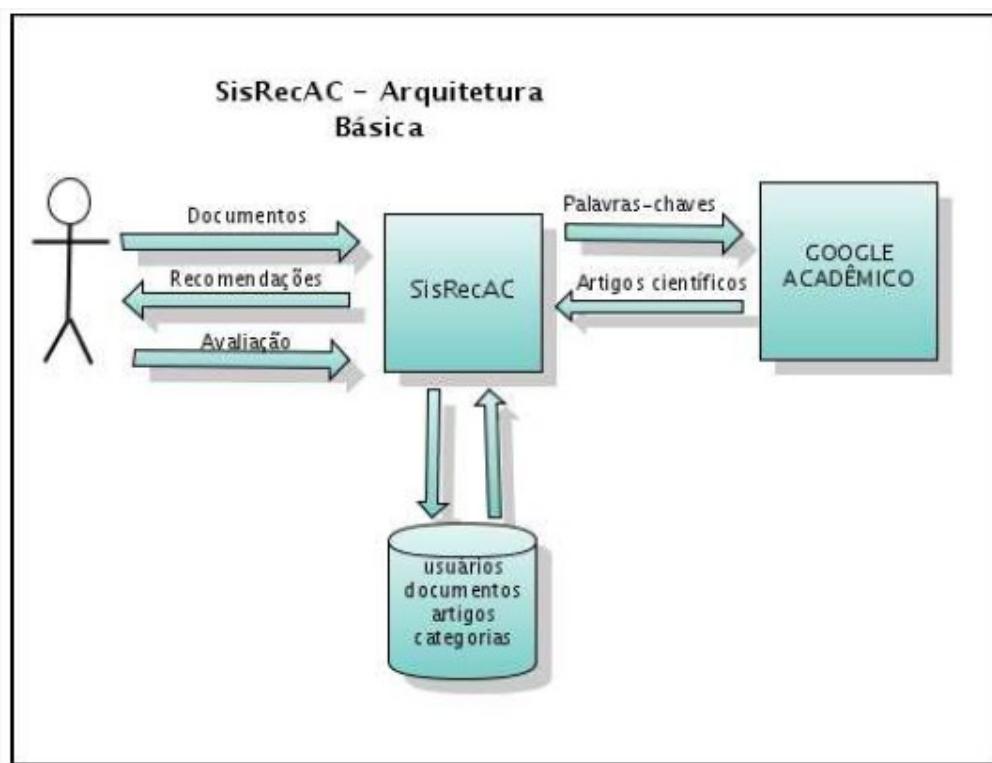


Figura 4 - Arquitetura básica do sistema *SicRecAC*[AVILA, 2006]

No sistema *SisRecAC* deve ser realizado o cadastro inicial do usuário, não sendo permitido fazer consultas sem o cadastro. Para que o usuário do sistema obtenha as recomendações deve ser feito um *upload* de um arquivo que vai servir como base para as mesmas, sendo utilizado também quando são realizadas as avaliações. Conforme o usuário avança com as avaliações das recomendações, o sistema compara o resultado com o arquivo base inserido no sistema pelo usuário. A Figura 5 apresenta a tela de recomendações do SisRecAc.

The screenshot shows the GPSI - SisRecAC interface. On the left, a sidebar has a tree view with categories: "ENVIAR DOCUMENTO", "INTELIGÊNCIA ARTIFICIAL", "JAVA PARA DISPOSITIVOS MÓVEIS" (which is expanded to show "Componentização De Software Em..."), "SISTEMAS DE RECOMENDAÇÃO", and "TESES". Below the sidebar are two buttons: "S A I R" and "DESCRICAO". The main content area has a header "COMPONENTIZAÇÃO DE SOFTWARE EM (105-lmn2.pdf)" with a "BETA" badge. Below the header is the title "Recomendações do SisRecAC". A sub-section titled "Avalie os links abaixo para que possamos aprimorar os algoritmos que geram as recomendações." follows. A link "Um Ambiente de Desenvolvimento de Aplicações Multi-Plataformas e Adaptativas para Dispositivos..." is shown, with a note "... estabelecidas para a criação dos componentes de interface ... engenheiro utiliza a ferramenta, User Interface Generator (UIG ... Figura 2 - Visão geral do processo de ... WV de Carvalho - mcc.ufc.br". Below this is a question "Este documento é relevante no contexto do documento de origem ?" with three radio buttons: "Totalmente relevante", "Parcialmente relevante", and "Irrelevante". A "Avaliar" button is at the bottom of this section.

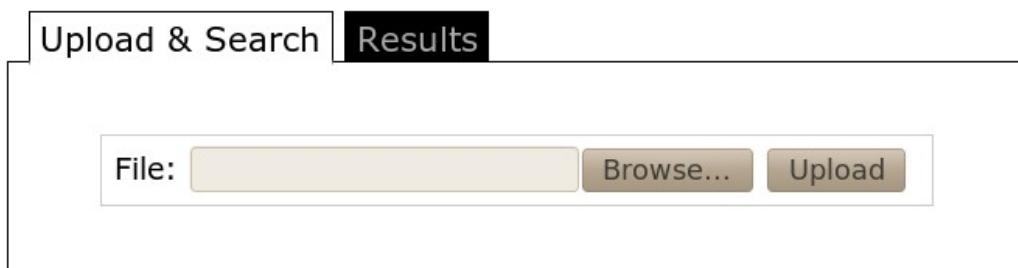
**Figura 5 - Tela de Recomendações do SisRecAC[AVILA, 2006]**

O sistema *SisRecAC* foi desenvolvido utilizando-se a linguagem de programação *PHP 4*. O servidor usado foi *Linux RedHat versão 9* e o banco de dados *PostgreSQL 8.1*.

### 3.2 Um Arcabouço Para Busca e Recomendação de Artigos Científicos

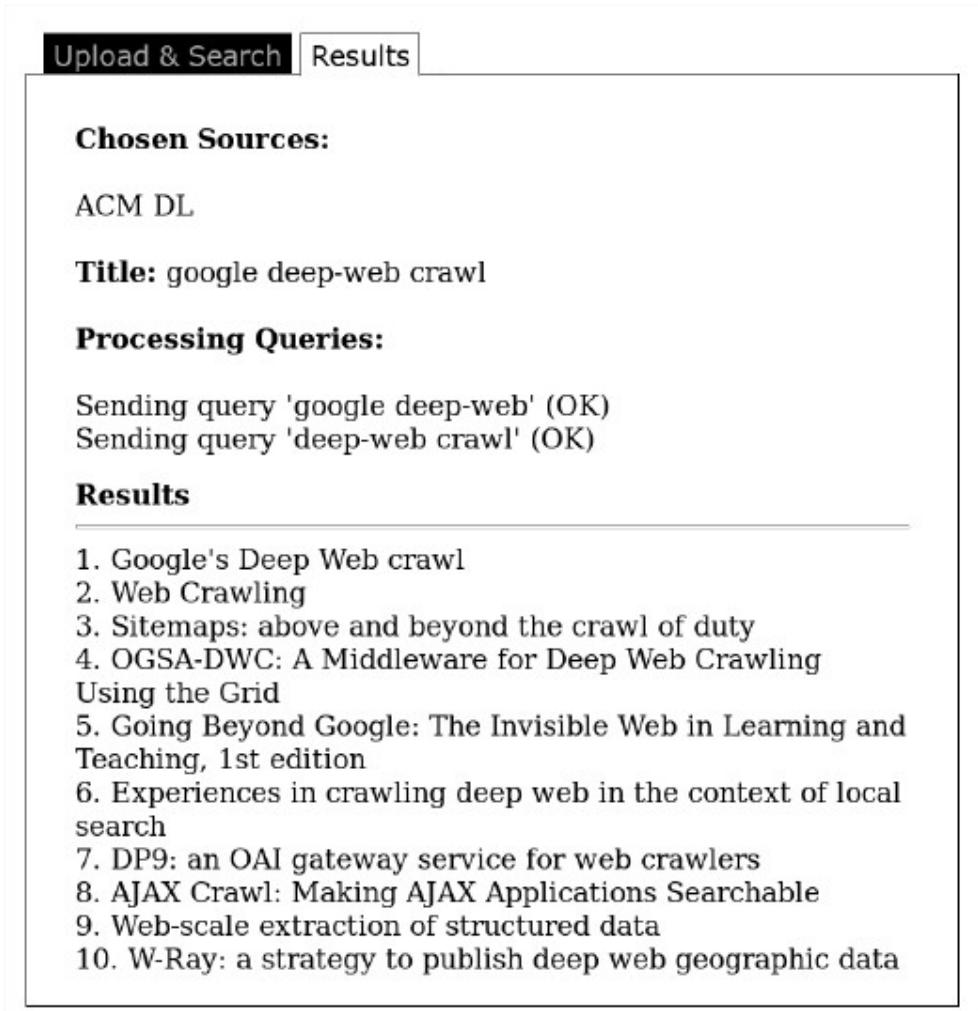
O trabalho apresentado por Nascimento (2011), mostra o desenvolvimento de um SR, onde é proposto um arcabouço para busca e recomendação de artigos científicos, utilizando a Filtragem Baseada em Conteúdo (FBC). O Arcabouço realiza buscas em formulários de consultas em bases de duas bibliotecas digitais *ACM, IEEE Xplore* (<http://ieeexplore.ieee.org>) e *ScienceDirect (SD)*. As consultas são realizadas no ambiente *webde* cada uma das bibliotecas digitais.

O arcabouço gera suas recomendações a partir de informações contidas em um único documento *PDF*, fornecido como entrada pelo usuário. Para realizar a consulta utilizando as palavras-chave, o documento é convertido para os formatos *HTML(HyperText Markup Language)* e *XML (eXtensible Markup Language)*, utilizando a ferramenta *pdf2html* (<http://pdf2html.sourceforge.net/>). A Figura 6 mostra a interface inicial do sistema desenvolvido em uma plataforma *web* e executado por meio de um *browser*.



**Figura 6 – Interface de entrada do protótipo do Arcabouço [NASCIMENTO, 2011]**

Depois da inserção do documento o sistema permite que o usuário escolha a biblioteca desejada, entre *ACM*, *IEEE* ou *ScienceDirect*, para efetuar a consulta. Após a escolha o usuário deve clicar no botão *Search* eo sistema mostra as recomendações. Na Figura 7 podemos verificar a lista de recomendações onde são apresentados os 10 itens.



The screenshot shows a user interface for a prototype system. At the top, there are two buttons: 'Upload & Search' (highlighted in black) and 'Results'. Below these, the 'Results' section is displayed. It contains the following text:

**Chosen Sources:**  
ACM DL

**Title:** google deep-web crawl

**Processing Queries:**  
Sending query 'google deep-web' (OK)  
Sending query 'deep-web crawl' (OK)

**Results**

---

1. Google's Deep Web crawl
2. Web Crawling
3. Sitemaps: above and beyond the crawl of duty
4. OGSA-DWC: A Middleware for Deep Web Crawling Using the Grid
5. Going Beyond Google: The Invisible Web in Learning and Teaching, 1st edition
6. Experiences in crawling deep web in the context of local search
7. DP9: an OAI gateway service for web crawlers
8. AJAX Crawl: Making AJAX Applications Searchable
9. Web-scale extraction of structured data
10. W-Ray: a strategy to publish deep web geographic data

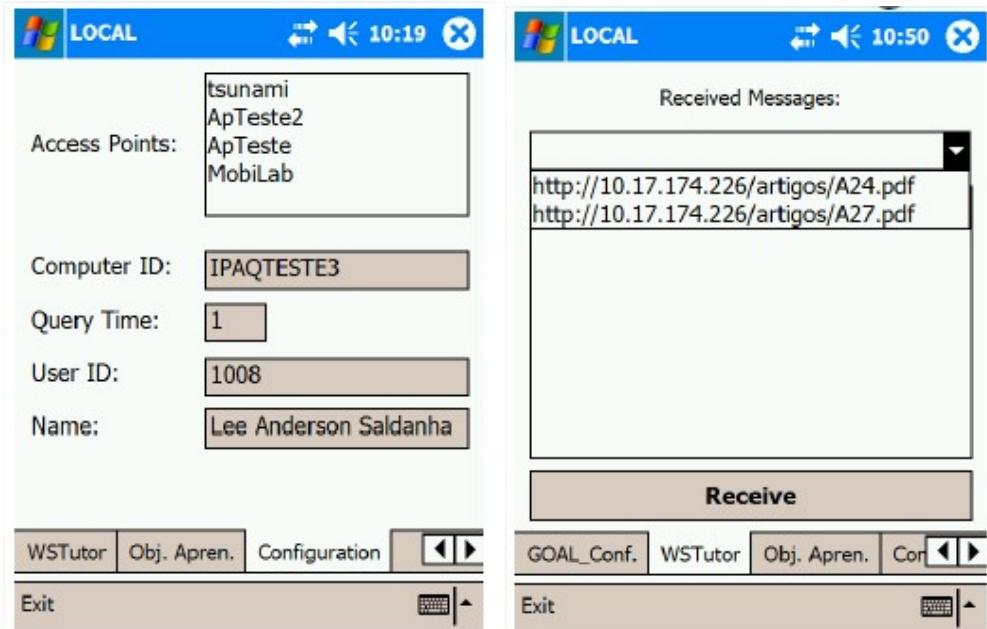
**Figura 7 – Lista de Recomendações [NASCIMENTO, 2011]**

### 3.3MobilRec

No trabalho apresentado por Cazella [2010] temos o desenvolvimento de um SR baseado na Filtragem Baseada em Conteúdo (FBC) para ambientes móveis, denominado *MobilRec*. Cazella [2010] utilizou o modelo Local (*Location and Context Aware Learning*) proposto por Barbosa et al. [2008], que é composto por um Sistema de Perfil de Usuário, Sistema de Localização, Assistente Pessoal, Repositório de Objetos de Aprendizagem, Sistema de Envio de Mensagens Contextuais e Tutor. Alguns agentes de software foram utilizados para o correto funcionando do sistema *MobilRec*. O Agente

*Extrator* foi empregado na extração de conteúdo e na classificação de todos os artigos com extensão PDF, DOC e RTF (*Rich Text Format*) e todos os termos extraídos são inseridos em uma base de dados de termos.

O funcionamento do sistema *MobilRec* é baseado nos seguintes passos: 1) um artigo científico é publicado no repositório de documentos; 2) o agente extrator identifica o novo arquivo e realiza a extração do conteúdo atualizando a base de termos e documentos, e solicita ao modelo LOCAL para que o mesmo insira este artigo em sua base de dados; 3) o perfil do usuário é construído por meio da pontuação atribuída ao artigo; 4) O método *TF-IDF* [Garcia, 2007 citado por Cazella, 2010] é aplicado para a identificação dos artigos de maior interesse para os usuários; 5) o sistema de mensagens do LOCAL envia a recomendação para o usuário, que se conecta e se identifica ao sistema usando um dispositivo móvel (*palm HP iPAQ*). A Figura 8 mostra a interface para que o usuário receba as recomendações.



**Figura 8 – Interfaces do *MobilRec* integrado com o Local [CAZELLA, 2010]**

Para iniciar a experimentação, 27 artigos foram inseridos no repositório do protótipo. Após foram realizadas as tarefas de pré-processamento e cálculo de similaridade entre os artigos, para possibilitar a recomendação por meio da FBC. Uma vez que se obteve a avaliação destes alunos para os artigos selecionados, foi possível iniciar o processo de recomendação por análise de conteúdo.

Para validar a implementação foram realizados dois experimentos. No primeiro experimento (avaliação de artigos pré-selecionados), cada estudante recebeu um *e-mail* no qual foi informado o título de 5 artigos para serem avaliados. Estes artigos foram selecionados de forma aleatória. Os estudantes foram cadastrados no protótipo e receberam suas credenciais (usuário e senha) para efetuarem o *login* na interface e iniciarem as avaliações. As avaliações foram feitas por meio de uma escala *likert* de 5

pontos (1-Péssimo e 5 – Excelente). Todas as avaliações serviram como base para o segundo experimento, que envolveu a geração das recomendações, sendo que os artigos recomendados também foram avaliados com a escala *likert*.

### 3.4 Estudo Comparativo

No Quadro 1 apresentamos um estudo comparativo dos trabalhos apresentados nesta seção e o protótipo de SR desenvolvido (*SysPaperAdvice*), com o objetivo de mostrar as principais características de cada trabalho.

**Quadro 1: Estudo Comparativo**

	<i>SisRecAc</i>	Arcabouço	<i>MobilRec</i>	Protótipo Desenvolvido ( <i>SysPaperAdvice</i> )
Técnicas de Filtragem	FC e FBC	FBC	FBC	FC e FBC
Tecnologias Empregadas	Linguagem de programação PHP e SGBD <i>PostgreSQL</i>	Plataforma <i>Web e bibliotecas digitais ACM, IEEEExplore, ScienceDirect</i>	Location and <i>Context Aware Learning</i>	Linguagem de Programação PHP e SGBD <i>MySQL</i>
Métrica de Similaridade para a FC	Não foi empregada	Não foi empregada	Não foi empregada	Correlação de <i>Pearson</i>
Cálculo da Predição	Não foi empregado	Não foi empregado	Não foi empregada	Média Ponderada
Tipos de Documentos	PDF	PDF	PDF, DOC e RTF	PDF
Tipo de Coleta de Dados	Explícita	Explícita	Explícita	Explícita

Observando o descrito no Quadro 1, constata-se que o protótipo *SysPaperAdvice* desenvolvido contém semelhanças aos demais projetos, todavia o presente trabalho utiliza as FC e FBC simultaneamente. Apesar do *SisRecAC* também empregar a FBC e a FC, cabe destacar que a FC desenvolvida é mais simples do que a implementada neste protótipo já que, no *SisRecAC* a FC ocorre quando um usuário decide recomendar e compartilhar os documentos com outras pessoas, sem aplicar nenhum tipo de cálculo matemático. O *SysPaperAdvice* também conta com a coleta de dados diferenciada pois utiliza as formas Explícita, aplicada no momento do cadastro do usuário no sistema, através de um formulário para inserção das áreas de interesse e também utiliza da coleta de dados para controlar as avaliações realizadas por outros usuários.

#### **4. Solução Implementada**

Neste trabalho foi implementado um protótipo de Sistema de Recomendação de artigos e trabalhos científicos, o *SysPaperAdvice*, utilizando dois tipos de filtragem, a FBC e a FC, visando aumentar a qualidade das recomendações e auxiliar o usuário a encontrar os artigos e trabalhos mais adequados.

As principais atividades desenvolvidas para dar conta deste trabalho foram:

1. Estudo de conceitos, técnicas e ferramentas para o desenvolvimento de Sistemas de Recomendação (SR), visando compor o referencial teórico e o estado da arte;
2. Estudo de técnicas para modelar o perfil do usuário em Sistemas de Recomendação;
3. Estudo de técnicas para realizar a Filtragem Baseada em Conteúdo (FBC) e a Filtragem Colaborativa (FC) em Sistemas de Recomendação;
4. Levantamento dos principais requisitos do (SR) proposto;
5. Modelagem do Sistema de Recomendação;
6. Implementação de um protótipo do SR, utilizando a linguagem de programação PHP eo Sistema Gerenciador de Banco de Dados *MySQL*;
7. Testes e validação do protótipo de Sistema de Recomendação implementado.

Nas próximas seções são apresentadas a modelagem, implementação e validação do *SysPaperAdvice*.

##### **4.1 Levantamento de Requisitos**

Com base no estudo realizado sobre SR e na definição das técnicas de filtragem que seriam aplicadas no desenvolvimento do *SysPaperAdvice*, foram definidos os seguintes requisitos:

- Os usuários poderão realizar o *upload* de documentos (artigos e trabalhos científicos no SR), para que seja possível criar um repositório de documentos que serão avaliados e recomendados. Tendo-se em vista a questão do espaço

para armazenamento, optamos por armazenar a localização do documento na web, por meio de sua URL (*Uniform Resource Locator*);

- Os usuários, ao efetuarem seu cadastro no SR, definirão uma lista de palavras-chave dos assuntos de seu interesse. Essa lista de palavras será utilizada para realizar a FBC, o usuário pode realizar a inserção de até quatro palavras chaves. Para realizar a filtragem foi utilizado o Índice de *Jaccard*, método estatístico que determina a similaridade entre dois itens a partir da intersecção de seus conjuntos [SANTOS, 2014];
- Os usuários receberão uma lista de artigos e trabalhos científicos para avaliar, dando notas de 1 a 5. Com base nas avaliações será aplicada a correlação de *Pearson* (para definir a vizinhança) e a predição dos artigos a serem recomendados será calculada por meio da média ponderada. Estes cálculos serão utilizados para realizar a FC;
- Os usuários poderão visualizar as recomendações por meio das opções disponíveis no menu, sendo uma opção para cada uma das filtragens FBC e da FC (técnica mista da FH, destacada na seção 2 deste artigo).

## 4.2 Modelagem

Nesta seção apresentamos os diagramas de Entidade Relacionamento (ER) e o diagrama de Caso de Uso que foram construídos com base na proposta do *SysPaperAdvice*. Na Figura 9 apresentamos o modelo (ER) que retrata as tabelas que fazem parte do banco de dados do SR.

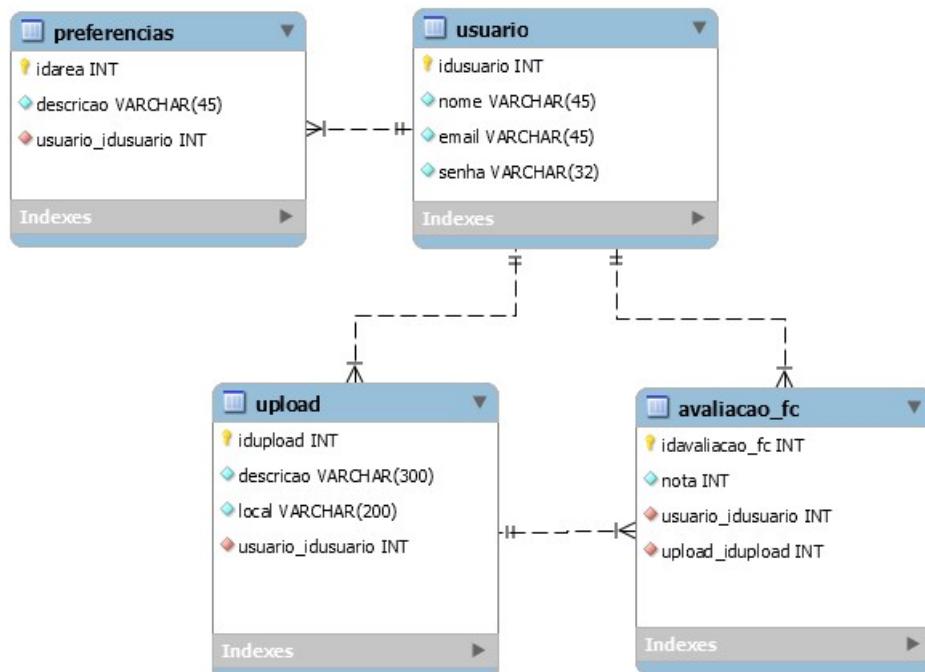
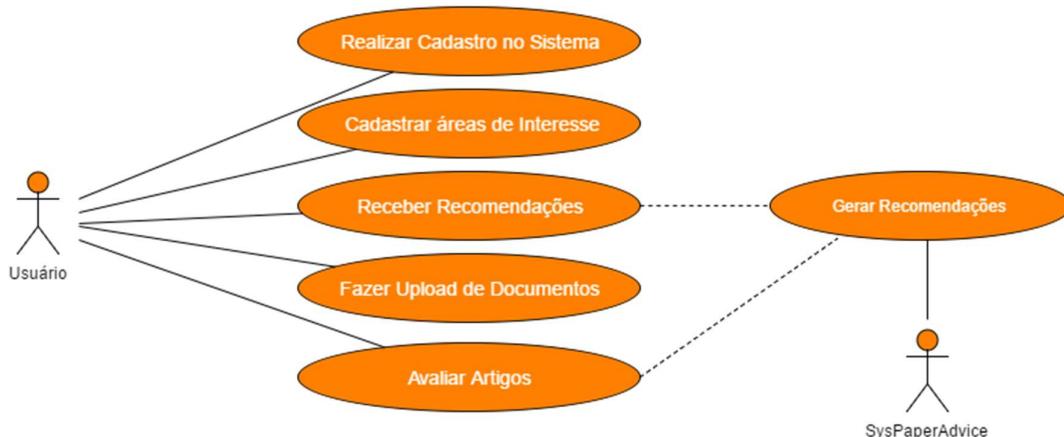


Figura 9 – Modelo Entidade Relacionamento (E-R). Fonte: Dos autores

Analisando o modelo ER da Figura 9, temos as seguintes tabelas no banco de dados do SR:

- *usuario*: tabela onde são armazenados todos os dados de usuários cadastrados no Sistema de Recomendação;
- *preferencias*: tabela onde são armazenadas as áreas de interesse do usuário cadastrado. Estas áreas de interesse são utilizadas na FBC;
- *upload*: tabela onde são armazenados os dados do *upload* de arquivo, contendo o local(URL)em que foi armazenado o arquivo;
- *avaliacao\_fc*: tabela onde são armazenadas as avaliações dos usuários para os documentos existentes no sistema, para realizar a FC;

Na Figura 10 temos o diagrama de Casos de Uso, que mostra as ações que o usuário do sistema pode executar.



**Figura 10 – Diagrama de Casos de Uso. Fonte: Dos autores**

Analisando os Casos de Uso da Figura 10, temos as seguintes funcionalidades no SR proposto:

- Realizar Cadastro no Sistema: o usuário deve realizar o cadastro no sistema informando todos os dados solicitados no formulário, assim o usuário cadastrado pode utilizar os recursos do sistema;
- Cadastrar Áreas de Interesse: o usuário já cadastrado no sistema pode cadastrar áreas de interesse, que serão utilizadas na FBC; caso seja efetuado este cadastro o sistema já direcionará o usuário para a página de recomendação;
- Fazer *Upload* de Documentos: o usuário cadastrado no sistema pode cadastrar documentos, que serão recomendados a outros usuários e estes documentos serão disponibilizados para que os usuários façam a avaliação dos mesmos, que serão utilizadas na FC;

- Avaliar Artigos: o usuário cadastrado no sistema, pode avaliar os documentos disponíveis, cadastrando notas de 1 até 5 para cada documento, cada nota lançada no sistema será utilizada para receber recomendação;
- Receber Recomendações: o usuário cadastrado no sistema receberá as recomendações conforme o seu perfil;
- Gerar Recomendações: o sistema *SysPaperAdvice* gera a recomendação conforme o perfil do usuário, utilizando a FBC e a FC, conforme detalha a próxima seção.

### 4.3 Algoritmo de Recomendação

Conforme descrito anteriormente, o *SysPaperAdvice* utiliza a FH, combinando as técnicas de FBC e FC. Para realizar a FBC, implementamos um algoritmo com os seguintes passos:

1. Após realizar o cadastro no SR, durante o primeiro acesso, o usuário receberá uma lista de 21 artigos, para avaliar. Esta lista será gerada de forma aleatória e cada artigo deverá ser avaliado com uma nota entre 1 e 5, sendo 5 a melhor;
2. A partir das avaliações realizadas, por meio da correlação de *Pearson*, o SR irá calcular a similaridade entre o novo usuário e os demais usuários cadastrados no sistema;
3. Com os resultados da similaridade calculados, o usuário será agrupado junto aos usuários com maior similaridade, sendo que o grupo de usuários similares será formado com os usuários cuja similaridade seja igual ou superior a 0.75, para formar a sua vizinhança;
4. Com base nos artigos que foram avaliados em comum por todos os usuários da vizinhança, será aplicado o cálculo da média ponderada, para calcular a predição do usuário com relação aos artigos existentes no SR;
5. Caso a nota do resultado da predição seja igual ou superior a 3, o artigo fará parte da lista de recomendações geradas pelo SR, mostrando em ordem decrescente do maior para o menor, todos os documentos serão mostrados na tela.

Este algoritmo será reexecutado toda vez que o usuário acessar, novamente, a opção *Recomendação FC* no menu do *SysPaperAdvice* ou realizar uma nova avaliação.

A Figura 11 apresenta o código-fonte desenvolvido, adaptado de Guimarães [2014], para o cálculo da correlação de *Pearson* e a Figura 12 apresenta o código-fonte para realizar o cálculo da predição (média ponderada), adaptado de Gonçalves et al. [2016], ambos desenvolvidos em linguagem de programação PHP.

```

<?php
    class Pearson {
        private $list = array();
        public function __construct($list=array()) {
            if(is_array($list)){
                $this->setList($list);
            }else{echo "O argumento passado no construtor não é um Array";
            }
        }
        private function setList($list){
            $this->list = $list;
            $this->calcula();
        }
        public function calcula(){//Calcula a soma dos produtos entre as notas
            for ($i = 0 ; $i < sizeof($this->list[0]); $i++){
                echo "\n" ;
                $nota1 = $this->list[0][$i];
                $nota2 = $this->list[1][$i];
                $valor1 += $nota1*$nota2;
            };
            $valor2 = array_sum($this->list[0]); //Soma as notas de cada usuário entre si
            $valor3 = array_sum($this->list[1]);
            for ($p = 0; $p < sizeof($this->list[0]); $p++){//Calcula a soma dos quadrados das notas
                $valor4 += pow($this->list[0][$p],2);
                $valor5 += pow($this->list[1][$p],2);
            };
            $n = count($this->list[0]);//Calculo do COEFICIENTE DE PEARSON
            $coel = ($n * $valor1 - $valor2 * $valor3);
            $coel2 = sqrt((($n * $valor4 - pow($valor2,2)) * ($n * $valor5 - pow($valor3,2))) / $coel);
            $coeficiente = $coel / $coel2;
            $valor = $coeficiente;
            $_SESSION['valor'] = $valor; //grava uma sessão com o nome de numero e com valor
            return $coeficiente;//Retorna o resultado do cálculo
        }
    }
?>

```

Figura 11 – Código-fonte do cálculo da correlação de Pearson. Fonte: Adaptado de Guimarães [2014]

```

if(isset($list2) && is_array($list2)) {
    $geral['atual']['usuario'] = $_COOKIE["iduser"];
    $upload = array();
    $total_users = 1 + count($list2);
    $complemento = $_COOKIE["iduser"].','.join(',',$list2);
    $sql5 = "SELECT upload_idupload FROM avaliacao_fc WHERE usuario_idusuario IN ($complemento)
    GROUP BY upload_idupload HAVING Count(*) >= $total_users";
    $query = $conexao->query($sql5);
    while($valores_users = $query->fetch_array()){
        $upload[] = $valores_users[0];
    }
    $total_upload_iguais = count($upload);
    $upload = join(',',$upload);
    foreach($geral as $key => $user){
        $sql6 = "SELECT nota FROM avaliacao_fc WHERE upload_idupload IN ($upload) AND usuario_idusuario = ".$user['usuario'];
        $query = $conexao->query($sql6);
        $total_upload = $query->fetch_array();
        $geral[$key]['media_geral'] = $total_upload[0]/$total_upload_iguais;
    }
    $sql_check_upload = "SELECT upload_idupload FROM avaliacao_fc AS a WHERE
    upload_idupload NOT IN (
        SELECT upload_idupload FROM avaliacao_fc AS b
        WHERE b.usuario_idusuario = ".$_COOKIE["iduser"]."
        ) AND usuario_idusuario IN (".join(',',$list2).")";
    ";
    $query = $conexao->query($sql_check_upload);
    while($linha = $query->fetch_array()){
        $resultado = checkFormula($geral, $linha[0], $conexao);
        $media_final = number_format($resultado, 2, '.', '');
        if($media_final >= 3 ){
            $aprovados[] = array(
                'id_upload' => $linha[0],
                'media_final' => $media_final);
        }
    }
}

```

Figura 12– Código-fonte do cálculo da Predição Fonte: Adaptado de Gonçalves et al. [2016].

Para a FBC onde utilizamos o Índice de *Jaccarde* implementamos um algoritmo com os seguintes passos:

1. Ao se cadastrar no sistema o usuário passa por um formulário de Áreas de interesse, onde é permitido que sejam inseridas até quatro palavras, que são armazenadas no banco de dados do sistema;
2. Depois de realizado o cadastro e inseridas palavras nos campos de Área de Interesse o sistema consulta os documentos cadastrados são convertidos do formato PDF para o formato texto, por meio da ferramenta *pdftotext*;
3. Por meio do índice de *Jaccard*, é calculado o índice de similaridade entre as palavras de interesse do usuário e os termos existentes no documento;
4. Se o valor do índice calculado for maior que 0, o documento será recomendado ao usuário.

Este algoritmo será reexecutado toda vez que o usuário acessar, novamente, a opção *Recomendação FBC* no menu do *SysPaperAdvice* ou realizar a troca de uma das áreas de interesse. A Figura 13 apresenta um trecho do código-fonte implementado para o cálculo da similaridade utilizando o índice de *Jaccard*.

```
<?php //Parte 1 ($ni) //Código Implementado Para Texto Auxiliar
$areaInteresse = "Jogo Na"; // 2 áreas de interesse
$palavrasinteresse = explode(" ", $areaInteresse); //Coloca em um Array
$ni = count($palavrasinteresse); //Conta o Array e coloca o total na variavel $ni
//Parte2 ($nj)
//Coloca no result todas as palavras
$result="em jogo futbol na sexta feira"; //Seta de onde vai começar a pesquisar as palavras -Jogo
$posresumo=strpos($result,strtolower('jogo'));//Coloca as palavras 'jogo','futbol','na','sexta','feira' no Resumo
$resumo=substr ($result, $posresumo, +26 );//Explode coloca as palavras em um Array Resumo1
$resumol = explode(" ", $resumo); //Conta o Array e coloca o total na variavel $nj
$nj = count($resumol);
//Parte3 ($nij)
$nij=termoscomum($palavrasinteresse, $resumol); //Chama a função JACCARD
$jac= jaccard($ni, $nj, $nij); //Mostra a Similaridade
echo $jac;
function termoscomum($palavrasinteresse, $resumol){
    $cont = 0;
    foreach ($palavrasinteresse as $key => $value) {
        if ($palavrasinteresse[$key] != "") {
            foreach ($resumol as $key2 => $value2) {
                if (strtoupper($palavrasinteresse[$key]) == strtoupper($resumol[$key2])) {
                    $cont++;
                }
            }
        }
    }
    return $cont;
}
function jaccard($ni, $nj, $nij){
    $jaccard = ($nij/($ni+$nj)-$nij));
    return $jaccard;
}
?>
```

Figura 13 – Código-fonte para o cálculo do índice de *Jaccard*. Fonte: Dos Autores.

## 4.4 Tecnologias Empregadas

O *SysPaperAdvice* foi desenvolvido utilizando a linguagem de programação PHP, além de recursos de HTML (*Hypertext Markup Language*) e CSS (*Cascading Style Sheets*), para modelar a interface; o Sistema Gerenciador de Banco de Dados *MySQL* para armazenamento das informações. Para auxiliar na recomendação foi utilizada a ferramenta *pdftotext*, para converter os arquivos *PDF* (*Portable Document Format*).

## 4.5 Demonstração do Protótipo Implementado

Nesta seção apresentamos as principais telas do protótipo desenvolvido - *SysPaperAdvice*, destacando suas principais funcionalidades. Na Figura 14 temos a tela inicial de cadastro de usuário.



Figura 14 – Tela Inicial de Cadastro do Usuário. Fonte: Dos autores

Na tela inicial de Cadastro do Usuário, conforme a Figura 14, temos, no menu, as opções *Home*, *Artigos* e *Trabalhos* onde é possível visualizar os documentos existentes no sistema, sem recomendação ou opção de avaliar. Ainda temos a opção de *Entrar*, permitindo ao usuário fazer o *login*, caso já seja cadastrado no sistema.

Na Figura 15 mostramos a tela para cadastro das Áreas de Interesse do Usuário. Nesta mesma tela temos a opção para realizar o *upload* de documentos para o sistema, que depois serão recomendados e avaliados. Se o usuário optar por definir suas áreas de interesse, quando clicar no botão *Salvar* o sistema processa e direciona o usuário para o menu *Recomendação FBC*. Caso o usuário clique no botão *Pular*, sem dados inseridos pelo usuário, o sistema direciona-o para a tela de *Artigos e Trabalhos* onde devem ser realizadas as avaliações de uma lista de 21 artigos, para permitir a criação do perfil do usuário, que será utilizado na FC.

**Figura 15 – Tela Para Criar o Perfil do Usuário. Fonte: Dos autores**

O sistema processa os dados salvos nas Áreas de Interesse e mostra na tela as Recomendações, conforme ilustra a Figura 16. Nessa figura, ao lado de cada item temos uma *listbox* para inserção de notas, caso o usuário opte por avaliar o artigo sugerido pelo sistema.

**Figura 16 – Tela Recomendação FBC. Fonte: Dos autores**

Na tela de Avaliação de Artigos, apresentada na Figura 17, são visualizados todos os artigos existentes no sistema e, ao lado de cada um deles existe um botão *Salvar Nota*. Cada vez que o usuário selecionar uma nota e clicar neste botão, o sistema verifica a existência de recomendação.



**Figura 17 – Tela Avaliação de Artigos. Fonte: Dos autores**

Após fazer a avaliação o usuário é direcionado para a tela apresentada na Figura 18. Nesta tela são apresentados os trabalhos conforme avaliação de um grupo de usuários vizinhos (FC) e, ao lado de cada artigo, existe uma opção para avaliar o documento.



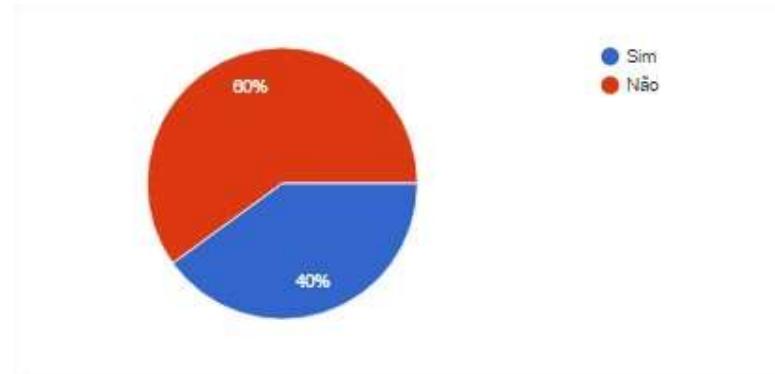
**Figura 18 – Tela Recomendação FC. Fonte: Dos autores**

#### 4.6. Testes e Validação

Ao longo da implementação vários testes foram realizados, verificando todos os requisitos apontados para o desenvolvimento do sistema. Após finalizar os testes, foram cadastrados no sistema 21 documentos para permitir a validação do sistema. Os documentos cadastrados são os Trabalhos de Graduação em Sistemas de Informação (TGSIs) desenvolvidos nos anos de 2014 e 2015.

Para validação do *SysPaperAdvice* 10 usuários utilizaram o sistema, sendo professores e alunos do curso de Sistemas de Informação. Após a utilização os usuários responderam a um questionário, apresentado no Anexo 1.

Com relação à pergunta “*Você já tinha utilizado algum Sistema de Recomendação anteriormente?*”, conforme os resultados obtidos verificamos que (60%) dos usuários não utilizaram SR, como mostra o gráfico da Figura 19.

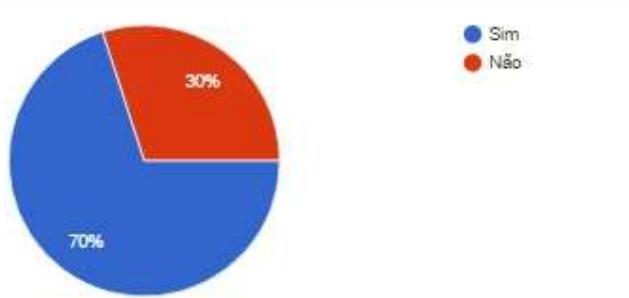


**Figura 19 – Pergunta: Você já tinha utilizado algum Sistema de Recomendação anteriormente?. Fonte: Dos autores.**

Com a pergunta “*Com relação ao Sistema de Recomendação de Artigos e Trabalhos Científicos, você acredita que as recomendações recebidas foram adequadas?*” Buscamos identificar se o sistema fez a recomendação adequadamente, todos os usuários que avaliaram o sistema responderam positivamente.

O mesmo resultado foi apresentado para a pergunta “*Caso o Sistema de Recomendação de Artigos e Trabalhos Científicos já estivesse disponível para acesso via web você utilizaria?*”, ou seja, todos os usuários responderam de forma positiva.

Na última pergunta do questionário: “*Você tem alguma sugestão de aprimoramento para o Sistema de Recomendação de Artigos e Trabalhos Científicos?*”, verificamos que (30%) dos usuários sugeriram aprimoramentos como mostra o gráfico da Figura 20. No Quadro 2 apresentamos as alterações sugeridas pelos usuários que realizaram a validação.



**Figura 20 – Pergunta: Você tem alguma sugestão de aprimoramento para o Sistema de Recomendação de Artigos e Trabalhos Científicos?. Fonte: Dos autores.**

**Quadro 2: Sugestões da Validação**

Deixar apenas um botão salvar. Na Avaliação dos artigos mostrar a nota dos que já foram avaliados, possibilitando o usuário reavaliar.
Além de gravar as áreas de Interesse no perfil do usuário, poderia haver uma forma de filtrar os artigos na própria lista de artigos, facilitando assim sua procura, além disso, uma informação que acho importante é o ano em que o artigo foi desenvolvido, poderia haver um filtro por ano, um filtro Autor e um outro Orientador.
Mostrar recomendações em um só lugar

Consideramos que os resultados da validação foram satisfatórios. Considerando o número de usuários que utilizaram o sistema observamos que futuros ajustes se fazem necessários, para assim atender as sugestões do usuário e aperfeiçoar o sistema implementado.

## 5. Considerações Finais

Acreditamos que os objetivos do trabalho foram alcançados, tendo sido implementado um protótipo de SR, aplicando as técnicas de FC e de FBC. A aplicação destas duas técnicas teve o objetivo de melhorar os resultados das recomendações, minimizando o problema da esparsidade existente na FC. As principais dificuldades encontradas envolveram o entendimento das técnicas necessárias para implementar as recomendações, em especial o cálculo da correlação de *Pearson*. Além disso, foram enfrentados problemas na conversão dos documentos *PDF*, o que prejudicou a consulta da FBC por não permitir a conversão para texto.

Na validação do sistema obtivemos resultados positivos, sendo que o sistema efetuou recomendação para todos os usuários que utilizaram o protótipo. Algumas observações foram notadas e serão apresentadas como trabalhos futuros, podemos aprimorar o protótipo implementado, reformulando algumas funcionalidades. Por exemplo, na avaliação dos artigos poderia existir apenas um botão *Salvar Nota* no final da lista de artigos, ao invés de um botão para cada avaliação. Além disso, as informações necessárias para calcular o índice de similaridade de *Jaccard* poderiam ser armazenadas no banco de dados, evitando que a contagem dos termos dos documentos precisasse ser realizada toda vez que o usuário avalia um artigo. Outra sugestão é pesquisar e implementar uma biblioteca para conversão dos documentos *PDF*, pois foi um dos problemas enfrentados.

## Referências

- Avila, C. M. O. (2006). Desenvolvimento de um Sistema de Recomendação de Artigos Científicos e Avaliações de Métodos de Extração de Palavras-Chave: UCPel. Disponível em:<http://ppginf.uepel.tche.br/TI-arquivos/2006/ChristianoAvila/PPGINF-UCPel-TI-2006-2-01.pdf>. Acesso em 27 de novembro de 2016.
- Barbosa, C. E. M. (2014). Estudo de Técnicas de Filtragem Híbrida em Sistemas de Recomendação de Produtos: UFPE, Disponível em: <http://www.cin.ufpe.br/~tg/2013-2/cemb.pdf>. Acesso em 15 de novembro de 2016.

- Barcellos, C. D., Musa, D. L., Brandão, A. L., and Warpechowski, M. (2007). Sistema de Recomendação Acadêmico para Apoio a Aprendizagem: CINTED-UFRGS. Disponível em: <http://www.cinted.ufrgs.br/ciclo10/artigos/3fDaniela.pdf>. Acesso em 27 de novembro de 2016
- Barth, F. J. (2010). Modelando o Perfil do Usuário para a Construção de Sistemas de Recomendação: Um Estudo Teórico e Estado da Arte: Revista de Sistemas de Informação da FSMA, nº6. Disponível em: [http://www.fsma.edu.br/si/edicao6/FSMA\\_SI\\_2010\\_2\\_Tutorial\\_1.pdf](http://www.fsma.edu.br/si/edicao6/FSMA_SI_2010_2_Tutorial_1.pdf). Acesso em 27 de novembro de 2016
- Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments: California State University, Disponível em: <http://josquin.cti.depaul.edu/~rburke/pubs/burke-umuai02.pdf>. Acesso em 15 de novembro de 2016.
- Cazella, S. C., Drumm, J. V., and Barbosa, J. L. V. (2010). Um Serviço Para Recomendação de Artigos Científicos Baseado em Filtragem de Conteúdo Aplicado a Dispositivos Móveis: CINTED-UFRGS. Disponível em: <http://www.seer.ufrgs.br/renote/article/download/18057/10645>. Acesso em 27 de novembro de 2016
- Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. (1992). Using Collaborative Filtering to Weave an Information Tapestry: New York – USA: Commun. ACM.
- Gonçalves, F.; Silveira, S. R.; Pertile, S. Sistema de Recomendação para Séries de TV. Frederico Westphalen: UFSM, 2016.
- Guimarães, R. (2014) Coeficiente de correlação de Pearson. Disponível em: <<http://blog.dehbora.com.br/?p=21>>. Acesso em 15 de novembro de 2016.
- Herlocker, J. L. (2000). Understanding and Improving Automated Collaborative Filtering Systems: Minnesota – USA: University of Minnesota.
- Herlocker, J. L., Konstan, J. A., and Riedl, J. (2000). Proceeding of the 2000 ACM Conference on Computer Supported Cooperative Work: New York – USA: Communications of the ACM.
- Jesus, R. P., Escobar, M. (2011). Desenvolvimento de Um Sistema de Recomendação de Eventos Com Uso de Geolocalização: Ulbra.
- Linhares, B. L., and Fernandes, A. M. R. (2012). Utilização de Técnicas de Sistemas de Recomendação para Aprimoramento de E-commerce do Tipo B2b: UNIVALI.
- Marchi, K. R. C. (2008). Sistemas de Recomendação: Uma Abordagem Geral: UEM. Disponível em: <http://kessia.blogs.unipar.br/files/2008/07/sistemas-de-recomendacao.pdf>. Acesso em 15 de novembro de 2016.
- Nascimento, C. (2011). Um Arcabouço Para Busca e Recomendação de Artigos Científicos: UFMG.
- Oliveira, L. G. (2007). Proposta de Uma Estrutura Metodológica para Implementação de Sistema de Recomendação: UFSC.
- Reategui, E. B., and Cazella, S. C. (2005). Sistema de Recomendação: XXV Congresso da Sociedade Brasileira de Computação – UNISINOS.

Resnick, P., and Varian, H. (1997). Recommender Systems. New York – USA: Communications of the ACM.

RodriguesJunior, M. B. R. (2016). Análise e Desenvolvimento de Modelos de Usuários Baseado em Múltiplos Domínios para Sistemas de Recomendação: UFBA.

Santos dos Mello, M. C. (2014). Sistema de Recomendação de Produtos Online Baseado em Palavras-Chave: UNESP

## ANEXO 1

### Sistema de Recomendação de Artigos e Trabalhos Científicos

*SysPaperAdvice*

### QUESTIONÁRIO PARA VALIDAÇÃO DO PROTÓTIPO

1) Você já tinha utilizado algum Sistema de Recomendação anteriormente?

- Sim. Qual(is)?  
 Não

2) Com relação ao Sistema de Recomendação de Artigos e Trabalhos Científicos, você considera que as recomendações recebidas foram adequadas?

- Sim  
 Não

Justifique sua resposta:

3) Caso o Sistema de Recomendação de Artigos e Trabalhos Científicos já estivesse disponível para acesso via *web* você o utilizaria?

- Sim  
 Não

Justifique sua resposta:

4) Você tem alguma sugestão de aprimoramento para o Sistema de Recomendação de Artigos e Trabalhos Científicos?