

# RECOMENDAÇÃO ASSISTIDA POR COMPUTADOR PARA ORIENTAÇÃO DE TRABALHOS DE CONCLUSÃO DE CURSO BASEADA NO PERFIL INFORMACIONAL E NA CAPACIDADE DE ORIENTAÇÃO DOCENTE

*Douglas de Lima Feitosa\**  
*Vitor Rafael Morais Silva\*\**  
*Fábio Paraguaçu\*\*\**  
*Roberta Vilhena Vieira Lopes\*\*\*\**

## RESUMO

Os Sistemas de Combinação Social são sistemas de recomendação capazes de sugerir pessoas às outras, ao invés de recomendar produtos ou serviços. Esses sistemas são aplicados em diversos contextos, inclusive na Educação a Distância (EAD), modalidade educacional que necessita dispor de um conjunto de ferramentas computacionais, no sentido de viabilizar o processo de ensino aprendizagem. A identificação de possíveis orientadores e a sobrecarga de trabalho são algumas das dificuldades existentes neste processo. Pesquisas nesta área investem na idéia de trabalhar os interesses de alunos, professores e gestores da educação. Para tanto, este trabalho utilizou uma abordagem quali-quantitativa, de maneira que os dados foram coletados em um Ambiente Virtual de Aprendizagem (AVA) e através da aplicação de um questionário estruturado aos orientadores em potencial. Desta forma, este trabalho tem o objetivo de descrever a aplicação de duas técnicas de recomendação, aplicadas através de um Modelo Informacional de Combinação Social que auxilia o Gestor de EAD no processo de alocação de orientadores de trabalhos monográficos. Serão apresentados três softwares, desenvolvidos por meio de algoritmos genéticos, que foram aplicadas no âmbito do Curso de Bacharelado em Sistemas de Informação à Distância, ofertado em conjunto pela Universidade Federal de Alagoas e Universidade Aberta do Brasil. Os resultados obtidos apontam que os métodos de filtragens propostos podem ser utilizados para auxiliar os gestores de EAD na redução dos casos de incompatibilidades de perfis de alunos e orientadores, bem como na redução da sobrecarga de trabalho dos docentes.

## Palavras-Chave:

Combinação Social. Sistemas de Recomendação. Modelo Informacional. Educação a Distância. Ambientes Virtuais de Aprendizagem.

\* Mestre em Modelagem Computacional de Conhecimento pela Universidade Federal de Alagoas. Doutorando em Administração de Empresas na Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas (Brasil).  
E-mail: douglas-feitosa@uol.com.br.

\*\* Aluno no Curso de Bacharelado em Ciência da Computação da Universidade Federal de Alagoas. Universidade Federal de Alagoas (Brasil).  
E-mail: vrafael1812@gmail.com.

\*\*\* Doutor em Informatique pela Université d'Aix-Marseille III. Professor adjunto II da Universidade Federal de Alagoas (Brasil).  
E-mail: fabioparagua2000@gmail.com.

\*\*\*\* Doutora em Ciências da Computação pela Universidade Federal de Pernambuco. Professora Adjunta I da Universidade Federal de Alagoas (Brasil).  
E-mail: rv21@hotmail.com.

## I INTRODUÇÃO

Em virtude da grande quantidade de informações e da disponibilidade facilitada das mesmas, proporcionadas pelo acesso à internet, as pessoas se deparam com uma grande diversidade de opções. São inúmeros os casos onde um indivíduo possui pouca ou

quase nenhuma experiência pessoal para realizar escolhas dentre as várias possibilidades que lhe são apresentadas.

No sentido de diminuir as dúvidas e necessidades que temos frente à escolha entre alternativas, geralmente confiamos nas recomendações que são passadas por outras pessoas, as quais podem chegar de forma direta

ou através de textos de recomendação, opiniões de revisores de filmes e livros, impressos de jornais, dentre outros (CAZELLA; NUNES; REATEGUI, 2010).

Terveen e Hill (2001) descrevem recomendação como um recurso que ajuda usuários a fazer escolhas em um universo de alternativas, agindo como uma espécie de filtro de informações. Em geral, uma recomendação é baseada nas preferências/características do produto, serviço ou pessoa a ser recomendada, considerando também as características daquele que requisitou a recomendação.

Neste contexto, os Sistemas de Recomendação são ferramentas que auxiliam no aumento da capacidade e eficácia deste processo de indicação, já bastante consolidados nas relações sociais. Em um sistema típico, as pessoas fornecem recomendações como entradas que o sistema agrega e direciona para os indivíduos considerados potenciais interessados neste tipo de recomendação (CAZELLA; NUNES; REATEGUI, 2010).

Em meio ao processo de recomendação deste tipo de sistema, os Perfis de Usuários representam o interesse do usuário com relação a vários assuntos em um momento particular. Cada termo que um perfil de usuário expressa é, num certo grau, característica de um usuário particular incluindo todas as informações diretamente solicitadas a ele e aprendidas, de maneira implícita, durante sua interação na web (CARREIRA et al, 2004). Fisicamente, o perfil do usuário pode ser visto como uma base de dados onde a informação sobre o usuário, incluindo seus interesses e preferências, é armazenada e pode ser dinamicamente mantida (ROUSSEAU et al, 2004; POO; CHING; GOH, 2003).

Na *web* encontram-se inúmeros tipos de perfis de usuário com diferentes graus de complexidade. Eles são desenvolvidos no contexto do *e-commerce*, *e-learning* e *e-community*, por exemplo.

No que concerne à Educação a Distância (EAD), modalidade de ensino mediada por Ambientes Virtuais de Aprendizagem (AVA), são geradas quantidades massivas de dados através das interações realizadas por seus usuários (WITTEN; FRANK, 2005). Para converter estes dados em sugestões úteis, são utilizadas diversas abordagens de filtragens de dados, à medida que são definidos os dados mais relevantes para construção da recomendação em questão.

No Brasil, a EAD apresenta-se em processo de forte crescimento. Segundo o Anuário Brasileiro de Educação Aberta e a Distância (ABED), em 2008, havia mais de 2,5 milhões de alunos matriculados em 1.181 cursos, disponibilizados em 257 instituições credenciadas.

Em virtude da relevância da EAD para o sistema educacional, percebe-se que é escassa a quantidade de softwares que auxiliam os gestores de EAD no que se refere à aplicação de recursos humanos e materiais, tornando necessário, portanto, o desenvolvimento de modelos informacionais que sirvam de auxílio ao planejamento das práticas pedagógicas nos cursos de ensino superior.

No que concerne às diversas possibilidades pedagógicas proporcionadas pelos AVA, existe uma que representa uma importante etapa nos cursos de graduação: a elaboração de Trabalhos de Conclusão de Curso. Primo e Silva (2007) explicam que durante a elaboração de trabalhos monográficos, por meio de AVA, podem ocorrer problemas que provocam sobrecarga de trabalho e estresse para alunos e orientadores. Dentre os problemas encontrados, estão, principalmente, o número excessivo de alunos e a incompatibilidade de perfis de orientadores e orientandos, que podem ocasionar atrasos na entrega dos trabalhos e, conseqüentemente, no cronograma dos cursos.

Assim sendo, através do presente trabalho, buscou-se responder a seguinte questão: "É viável desenvolver técnicas de recomendação, baseadas no perfil informacional e na capacidade de orientação docente, para recomendação de orientadores de trabalhos monográficos?".

O presente artigo seguirá a seguinte organização: Na seção 2, serão abordados os métodos utilizados para realização da pesquisa; Na seção 3, serão apresentados os resultados; Na seção 4, será apresentada uma breve discussão sobre os resultados da pesquisa; Na seção 5, serão apresentadas as conclusões do trabalho.

## 2 METODOLOGIA

O presente trabalho trata do desenvolvimento de 2 (duas) técnicas de recomendação, validadas através de um modelo informacional de combinação social voltado a auxiliar o gestor de EAD no processo de

contratação de orientadores. Para tanto, buscou-se identificar quais as variáveis que podem afetar a escolha de um orientador voltado às práticas de construção monográfica em um curso superior em EAD.

Já em posse dos principais conceitos teóricos, necessários ao embasamento do trabalho, foi efetuada a escolha do curso de bacharelado em Sistemas de Informação, da Universidade Federal de Alagoas, para que as técnicas de filtragens baseadas no desempenho dos alunos e na capacidade laboral dos professores fossem validadas. Para tanto, foram selecionados 10 (dez) orientadores em potencial, que cumprem com os requisitos demandados pela coordenação do curso estudado, tendo eles já participado de ao menos 1 (um) módulo de capacitação, promovido pelo setor responsável pela EAD na instituição.

Além disso, foram selecionados os 20 (vinte) últimos alunos concluintes a acessar o AVA, sendo 5 (cinco) de cada Pólo onde o curso é ofertado, representando, portanto, algo em torno de 50% do universo dos alunos concluintes do curso em 2011.

Após a coleta dos dados, foram desenvolvidos softwares que utilizam modelos de algoritmos genéticos, conhecidos pela sua aplicação na resolução de problemas de alta complexidade e que contemplam o uso das duas técnicas de recomendação propostas, sendo o terceiro software uma síntese destas: a Filtragem Baseada em Desempenho e a Filtragem Baseada em Capacidade, desenvolvidas para a resolução do problema apresentado neste trabalho.

Para tanto, foram utilizadas 2 (duas) bases de dados convertidas em tabelas, sendo a primeira referente aos históricos escolares dos alunos; e a segunda referente aos questionários aplicados para os orientadores em potencial.

Um processo de mineração de dados foi realizado por meio das técnicas de recomendação supreferidas, utilizadas no sentido de gerar sugestões de orientadores, à medida que identificam padrões similares de perfis e consideram os dados referentes ao desempenho dos alunos nos AVA, as áreas de domínio dos orientadores e as questões relacionadas à capacidade de captação de alunos orientandos.

Após o desenvolvimento das três diferentes versões do software (uma abordando a incompatibilidade de perfis, uma abordando a

sobrecarga de alunos e uma terceira abordando os dois aspectos de maneira simultânea) foi realizada então uma comparação destas, identificando qual seria o algoritmo genético que melhor atende ao modelo computacional utilizado para validação das técnicas de recomendação propostas e, conseqüentemente, se houve minimização do problema tratado neste trabalho.

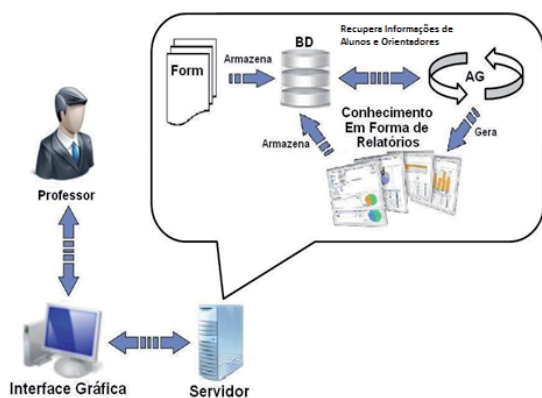
## 2.1 O modelo para validação das filtragens

A Figura 1 apresenta o modelo informacional de combinação social aplicado para validação das técnicas de recomendação propostas. O Modelo segue a seqüência abaixo:

- 1) Inicialmente são fornecidas à ferramenta duas bases de dados: uma contendo os registros históricos das notas dos alunos e outra com os questionários a serem respondidos pelos professores;
- 2) Para tanto, num primeiro momento, o professor interage com a interface do sistema, informando sua chave de acesso (*login* e senha);
- 3) Após a entrada das informações, o sistema apresenta o questionário estruturado, elaborado no sentido de extrair o máximo de informações úteis dos professores, no que concerne à formação de seus perfis e estipulação da capacidade máxima de orientandos;
- 4) Logo após, o professor insere algumas informações importantes como nome, regime de trabalho (dedicação exclusiva ou parcial) e o número atual de orientandos, seguidos de suas respostas, de maneira que cada resposta dada pelo professor representará um valor VERDADEIRO ou FALSO;
- 5) Essas informações serão armazenadas na base de dados, representando o seu perfil e a sua capacidade máxima de orientação de alunos;
- 6) Em seguida, o sistema acessa a base de dados do Histórico de Alunos e extrai os perfis dos mesmos. Um determinado intervalo de notas, definido pelo próprio gestor, indica se o aluno tem afinidade com a referida disciplina. Caso contrário, que o aluno não tem afinidade. O gestor define ainda o número de alunos que cada

- orientador em regime de dedicação parcial e em regime de dedicação exclusiva pode ter;
- 7) As bases de dados dos perfis dos professores e alunos serão convertidas em tabelas binárias com o mesmo número de colunas. Em paralelo, será construída uma base de dados contendo a capacidade máxima de alunos e o número atual de alunos de cada professor;
  - 8) Estas tabelas serão usadas pela função de adaptação do algoritmo genético, que construirá planos de alocação de perfis de orientadores para cada perfil de aluno existente. Os planos gerados serão exibidos em uma tela para que os gestores do processo escolham o que melhor se enquadra às suas necessidades.

**Figura 1 - O Modelo Proposto**



Fonte: Os Autores

## 2.2 Filtragem baseada em desempenho

No geral, as técnicas de recomendação utilizadas em AVA se restringem à sugestão de conteúdos, portanto, dificilmente utilizam o fator “desempenho do usuário” para realizar as recomendações.

Para este trabalho, pensou-se em um tipo de filtragem diferenciado que levasse em conta esse fator e que denominamos “Filtragem Baseada em Desempenho”. A ideia desse método é oferecer um novo tipo de técnica de filtragem de informações aos desenvolvedores que buscam resolver problemas onde o desempenho é considerado um fator relevante.

A técnica de Filtragem Baseada em Desempenho consiste na mineração dos dados úteis que estão relacionados ao desempenho dos usuários de AVA. Os logs que indicam o tempo de resolução de atividades e os registros de notas disponibilizados no AVA são considerados relevantes à definição dos perfis, com base no desempenho dos usuários.

Para validação do método, foram utilizados as informações referentes aos registros históricos de notas dos estudantes do curso de bacharelado em Sistemas de Informação – UFAL.

## 2.3 Filtragem baseada em capacidade

A técnica de Filtragem Baseada em Capacidade leva em consideração as informações relacionadas à capacidade do orientador em captar novos alunos. Desta forma, são levadas em conta as informações disponíveis no AVA sobre o regime de trabalho no qual o orientador trabalha e sobre sua situação atual em relação ao número de orientandos.

No que diz respeito ao regime de trabalho, são consideradas as informações referentes à carga horária do professor, de maneira a estabelecer qual seria a capacidade máxima de alunos desejável para ele.

Quanto às informações referentes ao número atual de alunos captados pelo orientador, estas são utilizadas no sentido de melhor distribuir os alunos e evitar o problema da sobrecarga de trabalho.

Para tanto, houve a necessidade de que fosse criada uma base de dados que contempla a capacidade máxima de alunos e o número atual de alunos em orientação para cada orientador.

## 2.4 Variações do algoritmo genético de holland para o problema

Os modelos de algoritmos genéticos procuram utilizar o princípio evolucionário encontrado na natureza para encontrar soluções para problemas computacionais, tomando como base para isto, o modelo natural da sobrevivência do mais apto. Por esta razão, são categorizados dentro do contexto da inteligência artificial como algoritmos de busca cega que se utilizam ainda de uma abordagem de metáforas entre os

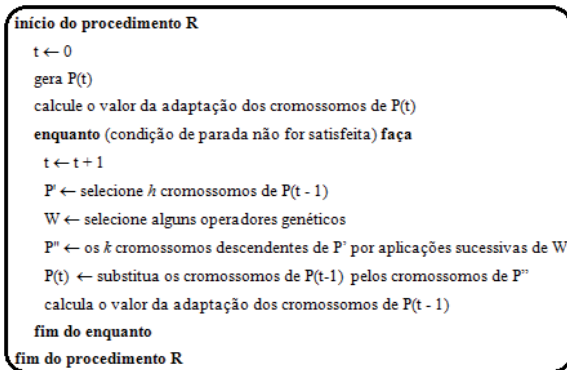


conceitos da biologia e dos algoritmos de busca (CARVALHO, 2008).

Dada a natureza do problema desta pesquisa, optou-se pelo uso do algoritmo genético de Holland, tendo em vista a complexidade encontrada em questões relacionadas à recomendação de agentes humanos em grandes sistemas educacionais.

Holland propôs dois algoritmos genéticos, denominados de  $R_1$  e  $R_d$  (HOLLAND, 1995). No algoritmo  $R_1$ , em cada iteração do algoritmo será gerado um novo indivíduo que substituirá um dos indivíduos da população atual. Enquanto que, no algoritmo  $R_d$ , em cada iteração do algoritmo serão gerados  $m$  descendentes, onde  $m$  é o tamanho da população atual, os quais irão substituir todos os indivíduos da população atual. Em seguida, Holland propôs o algoritmo R que simula o comportamento desses dois algoritmos dependendo dos valores dos parâmetros  $h$  e  $k$ . Os indivíduos da população em todos esses algoritmos são representados por cadeias binárias, denominadas de cromossomo.

**Figura 2 - Funcionamento do Algoritmo R**



Fonte: Holland (1995)

Para o problema de planejamento de orientadores de monografias, baseado no perfil do professor e do aluno e no número máximo de orientandos, convencionou-se que o cromossomo é um vetor com dimensão igual ao número de alunos e que o conteúdo das suas células é um número natural pertencente ao intervalo [1,número de perfis de professores].

Por exemplo, considere a Tabela 1, referente aos perfis dos orientadores, e a Tabela 2, referente aos perfis de alunos.

O Cromossomo  $c_1=[1,3,2,1,3]$  representa respectivamente:

- 1) A alocação do orientador cujo perfil está na 1ª linha da Tabela 1 para o aluno que está na 1ª linha da Tabela 2;
- 2) A alocação do orientador cujo perfil está na 3ª linha da Tabela 1 para o aluno que está na 2ª linha da Tabela 2;
- 3) A alocação do orientador cujo perfil está na 2ª linha da Tabela 1 para o aluno que está na 3ª linha da Tabela 2;
- 4) A alocação do orientador cujo perfil está na 1ª linha da Tabela 1 para o aluno que está na 4ª linha da Tabela 2; e
- 5) A alocação do orientador cujo perfil está na 3ª linha da Tabela 1 para o aluno que está na 5ª linha da Tabela 2.

**Tabela 1 - Representação dos Perfis dos Orientadores**

Orientadores	Disciplinas				
	A	B	C	D	E
Orientador 1	1	1	0	0	1
Orientador 2	0	1	0	1	1
Orientador 3	1	0	1	1	1

Fonte: Os Autores

**Tabela 2 - Representação dos Perfis dos Alunos**

Alunos	Disciplinas				
	A	B	C	D	E
Aluno 1	1	0	1	0	0
Aluno 2	0	0	1	1	0
Aluno 3	1	1	1	0	1
Aluno 4	1	0	0	0	0
Aluno 5	1	1	1	1	1

Fonte: Os Autores

As variáveis  $t$ ,  $h$ ,  $k$ ,  $P(t)$ ,  $P'$ ,  $P''$  e  $W$  do algoritmo R representam o número de gerações, o número de cromossomos selecionados para gerar descendentes, o número de descendentes que devem ser gerados, a população na geração  $t$ , o reservatório dos cromossomos que irão gerar novos cromossomos, o reservatório dos novos cromossomos gerados e um conjunto de operadores genéticos que serão aplicados aos cromossomos do reservatório  $P'$  para gerar os cromossomos do reservatório  $P''$ , os quais

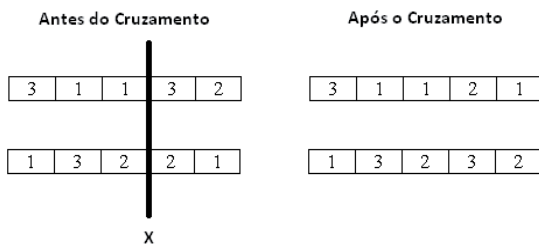
farão parte da próxima população  $P(t+1)$ , respectivamente. O número de gerações é incrementado em cada execução do laço **faça-enquanto** para simular a contagem do tempo genealógico. A população atual  $P(t)$  é formada por um conjunto de  $m$  cromossomos, os quais representam alguns dos possíveis resultados para o sistema adaptativo complexo em análise.

A adaptação de um cromossomo é dada pela função objetiva  $\hat{o}$  do sistema adaptativo em análise, se a função objetiva for de maximização, então a função de adaptação  $f$  do algoritmo genético será igual a  $\hat{o}$ , caso contrário, a função de adaptação  $f$  do algoritmo genético será igual a  $-\hat{o}$ . Essa convenção se faz necessária porque o algoritmo genético, por definição, estará sempre em busca do cromossomo mais adaptado.

A condição de parada do algoritmo genético pode ser o menor valor de adaptação considerado satisfatório, um número de iterações determinado, ou qualquer combinação dessas duas condições. Para este trabalho, será encontrado pelo menos um cromossomo com adaptação igual ao número de alunos.

Considere as Tabelas 1 e 2 para as ilustrações dos operadores genéticos trabalhados por Holland, em seu algoritmo genético, que foram o cruzamento de um ponto de corte, inversão e mutação. O operador de cruzamento de um ponto de corte tem por objetivo combinar a cadeia de dois cromossomos da seguinte maneira: seja  $a_1a_2...a_m$  e  $b_1b_2...b_m$  dois cromossomos de  $P(t)$ , selecione aleatoriamente um número  $x$ , pertencente ao conjunto  $\{1, 2, \dots, m\}$  e então constrói-se dois novos cromossomos  $a_1...a_xb_{x+1}...b_m$  e  $b_1...b_xa_{x+1}...a_m$ . A Figura 3 ilustra um cruzamento, sendo que  $x = 3$ .

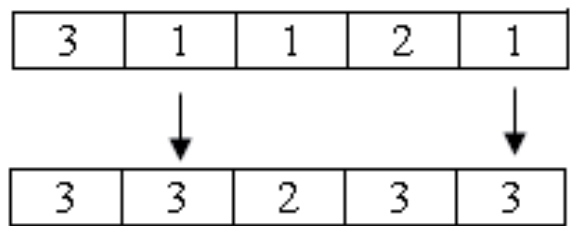
Figura 3 - Exemplo de Cruzamento



Fonte: Os Autores

O operador de mutação tem por objetivo alterar a cadeia de um cromossomo da seguinte maneira: seja  $a_1a_2...a_m$  um cromossomo de  $P(t)$ , selecione aleatoriamente  $n$  números  $x_1, x_2, \dots, x_n$  com  $n < m$ , pertencentes ao conjunto  $\{1, 2, \dots, m\}$  e então constrói-se um novo cromossomo  $a_1...a_{x_1-1}a_{x_1}a_{x_1+1}...a_{x_2-1}a_{x_2}a_{x_2+1}...a_m$ . Os caracteres em negrito das posições  $x_i$  dessa cadeia representam o complemento do caractere que ocupava essa mesma posição no cromossomo  $a_1a_2...a_m$  de  $P(t)$ . A Figura 4 ilustra uma mutação considerando  $n = 2$ ,  $x_1 = 2$  e  $x_2 = 5$ .

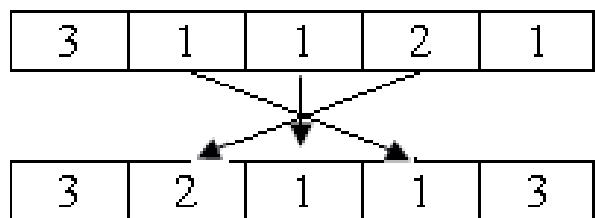
Figura 4 - Exemplo de Mutação



Fonte: Os Autores

O operador de inversão, tem por objetivo alterar a ordem dos caracteres de uma sub-cadeia de um cromossomo da seguinte maneira: seja  $a_1a_2...a_m$  um cromossomo de  $P(t)$ , selecione aleatoriamente dois números  $x_1$  e  $x_2$ , pertencentes ao conjunto  $\{1, 2, \dots, m\}$  e então constrói-se um novo cromossomo  $a_1...a_{x_1-1}a_{x_2}a_{x_2-1}...a_{x_1+1}a_{x_1}a_{x_2+1}...a_m$ . A Figura 5 ilustra uma inversão considerando que  $x_1 = 2$  e  $x_2 = 4$ .

Figura 5 - Exemplo de Inversão



Fonte: Os Autores

Apesar de apresentarem funcionalidades diferentes, as variações do Algoritmo Genético de Holland utilizadas nos softwares, descritos nas próximas seções, possuem a mesma condição de parada que é a de que deverá ser encontrado pelo menos um cromossomo com adaptação igual ao número de alunos.

### 3 RESULTADOS

No sentido de identificar a técnica de recomendação que melhor atenderia ao modelo proposto no trabalho, foram desenvolvidos 3 (três) sistemas de combinação social, sendo 1 (um) que trabalha a minimização das incompatibilidades de perfis de orientadores e alunos; 1 (um) que trabalha a minimização da sobrecarga dos orientadores; e 1 (um) que trabalha ambos os aspectos.

Apesar de apresentarem funcionalidades diferentes, as variações do Algoritmo Genético de Holland utilizadas nos softwares, descritos abaixo, possuem a mesma condição de parada que é a de que deverá ser encontrado pelo menos um cromossomo com adaptação igual ao número de alunos.

#### 3.1 Aplicações do algoritmo genético de Holland

O Algoritmo Genético de Holland, descrito na seção anterior, foi aplicado no contexto do Curso de Bacharelado em Sistemas de Informação, ofertado pela Universidade Federal de Alagoas em parceria com a Universidade Aberta do Brasil, em quatro cidades do estado de Alagoas (Maceió, Santana do Ipanema, Olho d'Água das Flores e Maragogi). Para elaboração da tabela com os perfis dos alunos, foram selecionados os 20 (vinte) últimos alunos a acessar o AVA do curso, sendo 5 (cinco) de cada cidade. Enquanto que para a elaboração da tabela com os perfis dos orientadores, foram selecionados 10 (dez) candidatos em potencial, que já elaboraram e/ou ministraram disciplinas no curso. Um questionário, que abrange 22 disciplinas já ministradas, foi aplicado.

**Figura 6 - Modelo de Questionário para Extração de Perfis de Orientadores**

Questionário – Orientação de Monografias		
Curso de Sistemas de Informação (EAD)		
<b>Orientador:</b>		
<b>Regime de Trabalho: Dedicção Exclusiva ( ) ou Parcial ( )</b>		
Disciplinas	Possui conhecimentos significativos sobre a disciplina e pode orientar os alunos acerca de temas relacionados (Por favor, assinale sua resposta com um "X")	
	Sim	Não
<b>Disciplina A</b>		
<b>Disciplina B</b>		
<b>Disciplina C</b>		
<b>Disciplina (...)</b>		
<b>Disciplina V</b>		

Fonte: Os Autores

Deste modo, obteve-se as tabelas 3, 4 e 5, que foram utilizadas como base de dados dos sistemas tratados neste trabalho.

**Tabela 3 - Tabela de Perfis dos Alunos**

Alunos	Disciplinas													V
	A	B	C	D	E	F	G	H	I	J	K	(...)		
Aluno 1	1	0	1	0	1	0	1	1	0	1	1	-	1	
Aluno 2	1	1	1	1	0	1	1	0	0	0	0	-	0	
Aluno 3	1	1	0	0	0	1	1	1	0	0	0	-	0	
Aluno 4	1	1	1	0	1	1	0	1	0	0	0	-	1	
Aluno 5	1	1	0	0	0	1	1	1	0	0	1	-	1	
Aluno 6	1	1	0	0	0	1	0	0	0	1	0	-	1	
Aluno 7	1	1	1	0	1	1	1	1	0	1	1	-	0	
Aluno 8	0	0	1	0	0	1	1	1	0	0	1	-	0	
Aluno 9	0	1	1	1	0	1	1	0	0	0	0	-	1	
Aluno 10	1	1	1	0	0	1	1	1	0	0	0	-	1	
Aluno 11	1	1	1	1	0	0	0	0	0	0	0	-	0	
Aluno 12	0	1	0	1	1	0	1	1	0	0	0	-	0	
Aluno 13	1	1	1	0	1	1	0	0	0	1	0	-	0	
Aluno 14	1	1	1	1	1	1	1	1	0	1	1	-	1	
Aluno 15	1	1	1	1	0	0	1	0	0	0	0	-	0	
Aluno 16	1	1	0	1	0	0	1	1	0	0	0	-	1	
Aluno 17	1	1	0	1	0	1	1	1	0	1	0	-	1	
Aluno 18	1	1	1	1	0	1	1	1	0	0	0	-	1	
Aluno 19	1	1	0	0	1	1	1	1	0	0	0	-	1	
Aluno 20	1	1	0	1	0	0	1	0	0	0	0	-	1	

Fonte: Dados da Pesquisa

**Tabela 4 - Tabela de Perfis dos Orientadores**

Orientadores	Disciplinas													
	A	B	C	D	E	F	G	H	I	J	K	(...)	V	
Orientador 1	1	0	0	0	0	0	0	0	0	0	0	0	-	0
Orientador 2	1	0	0	0	0	0	0	1	1	0	0	-	0	0
Orientador 3	1	1	0	1	1	0	0	1	1	0	0	-	0	0
Orientador 4	1	0	1	0	0	0	0	0	0	0	0	-	0	0
Orientador 5	1	1	1	1	1	0	1	0	0	0	0	-	0	0
Orientador 6	0	0	0	0	0	0	1	1	0	0	0	-	0	0
Orientador 7	0	1	0	1	1	0	0	1	1	0	0	-	0	0
Orientador 8	1	1	0	1	1	0	0	1	1	0	0	-	1	0
Orientador 9	0	0	1	0	0	0	0	0	0	0	0	-	0	0
Orientador 10	1	1	0	0	1	0	0	1	1	0	0	-	1	0

Fonte: Dados da Pesquisa

A Tabela 5 indica a capacidade máxima de orientação de alunos e o número atual de alunos orientandos. Para evidenciar o funcionamento dos aplicativos 2 e 3, convencionou-se que, dos 10 (dez) orientadores, 5 (cinco) são professores em regime de dedicação exclusiva (Capacidade Máxima de Alunos = 4) e 5 (cinco) são professores em regime de dedicação parcial (Capacidade Máxima de Alunos = 2). Como, neste caso, os orientadores não haviam sido distribuídos, todos possuem o número atual de alunos = 0. No entanto, o modelo em questão pode ser aplicado em casos onde já ocorreu a distribuição de orientadores.

**Tabela 5 - Tabela dos Parâmetros de Distribuição de Alunos**

Orientadores	Capacidade Máxima de Alunos	Núm. Atual de Alunos Orientandos
O1	4	0
O2	4	0
O3	2	0
O4	4	0
O5	2	0
O6	2	0
O7	2	0
O8	4	0
O9	2	0
O10	4	0

Fonte: Os Autores

### 3.2 Resultados do software I – minimização das incompatibilidades de perfis

O Software 1 utiliza a técnica de Filtragem Baseada em Desempenho. A função de adaptação verificará se existe ao menos 2 (duas) coincidências entre os perfis dos professores e dos alunos, e é definida pelo sistema abaixo:

$$adapt(C, i) = \begin{cases} adapt(C, i + 1), se (compatibilidade > 1) \wedge (i < n^{\circ} \text{ de alunos}) \\ adapt(C, i + 0), se (compatibilidade \leq 1) \wedge (i < n^{\circ} \text{ de alunos}) \\ 1, se (compatibilidade > 1) \wedge (i = n^{\circ} \text{ de alunos}) \\ 0, se (compatibilidade \leq 1) \wedge (i = n^{\circ} \text{ de alunos}) \end{cases}$$

Onde:

C = cromossomo verificado

i = índice verificado e  $1 \leq i \leq n^{\circ} \text{ de alunos}$

$1 \leq i \leq n^{\circ} \text{ de alunos}$

$$compatibilidade = \{P_{prof} [C(i)] \cap P_{aluno} (i)\}$$

$P_{prof}$  = perfil de professor sugerido e  $1 \leq P_{prof} \leq n^{\circ}$  de professores

$P_{aluno}$  = perfil de professor sugerido e  $1 \leq P_{aluno} \leq n^{\circ}$  de alunos

Definida a variação do AG de Holland, iniciou-se o software, que começou o processo de correlação de perfis e sugestão de planos de alocação de orientadores. No primeiro momento, o cromossomo inicial foi gerado aleatoriamente, com adaptação = 12.

**Figura 7 - Cromossomo Inicial Utilizado no Software I**

3	2	1	2	7	9	6	7	1	5	10	1	4	10	2	8	1	4	1	6
---	---	---	---	---	---	---	---	---	---	----	---	---	----	---	---	---	---	---	---

Fonte: Os Autores

Como já citado na seção anterior, para que a condição de parada do software seja satisfeita, um cromossomo com adaptação igual ao número de alunos deverá ser encontrado, logo *Adaptação* = 20. Desta maneira, o algoritmo executou as tarefas de cruzamento e mutação, considerando o *Ponto de Corte = Média da Adaptação da População Atual*, prosseguindo até a 8ª geração de cromossomos, onde foi encontrado o cromossomo mais adaptado que, por sua vez, representa o plano de alocação de orientadores.



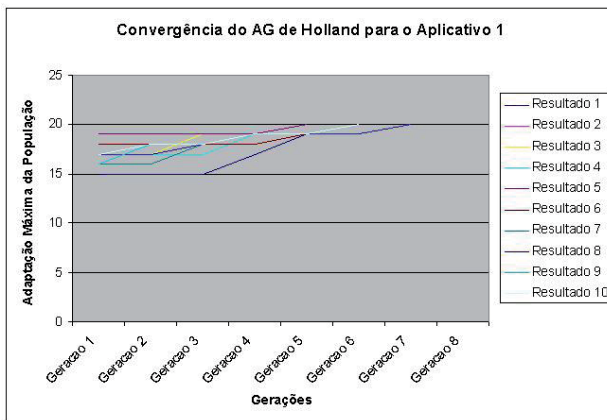
**Figura 8 - Plano de Alocação de Orientadores Sugerido no Software 1**

7	3	8	2	3	5	10	5	4	8	10	6	10	5	3	8	5	4	6	6
---	---	---	---	---	---	----	---	---	---	----	---	----	---	---	---	---	---	---	---

Fonte: Os Autores

Para verificar a convergência do algoritmo, o software foi executado 10 vezes, resultando no gráfico de convergência que demonstra a evolução do algoritmo através das novas populações geradas.

**Gráfico I - Convergência do Algoritmo de Holland no Software 1**



Fonte: Dados da Pesquisa

### 3.3 Resultados do software 2 - minimização da sobrecarga dos orientadores

O Software 2 utiliza a técnica de Filtragem Baseada em Capacidade. A função de adaptação verificará se o número atual de alunos atende à capacidade máxima de alunos estabelecida para os professores, e é definida pelo sistema a seguir:

$$adapt(c = \{G_1, G_2, \dots, G_n\}, i) = \sum^n orienta(G_i)$$

Onde:

$$c = \{G_1, G_2, \dots, G_n\}$$

$G$  = Perfil do professor

$i$  e  $n$  são tais que

$$\{\forall i, n \in \mathbb{N} \mid (n = \#c) \wedge (1 \leq i \leq n)\}$$

$$\{\forall i, n \in \mathbb{N} \mid (n = \#c) \wedge (1 \leq i \leq n)\}$$

$\#c$  = cardinalidade do conjunto  $c$

$$orienta(G_i) = \begin{cases} orienta(G_{i+1}) + 1, & \text{se } (orientados(G_i) < maxOrientados) \wedge (i < n^\circ \text{ de alunos}) \\ orienta(G_{i+1}) + 0, & \text{se } (orientados(G_i) = maxOrientados) \wedge (i < n^\circ \text{ de alunos}) \\ 1, & \text{se } (orientados(G_i) < maxOrientados) \wedge (i = n^\circ \text{ de alunos}) \\ 0, & \text{se } (orientados(G_i) = maxOrientados) \wedge (i = n^\circ \text{ de alunos}) \end{cases}$$

$orientados(G) = n^\circ$  atual de orientandos =

$maxOrientados = n^\circ$  máximo de alunos permitidos por professor

O Software 2 foi iniciado, obedecendo as regras pré-estabelecidas para a alocação de orientadores, de maneira que apenas a capacidade máxima de alunos orientandos são consideradas. Deste modo, o cromossomo inicial foi gerado aleatoriamente com adaptação = 17.

**Figura 9 - Cromossomo Inicial Utilizado no Software 2**

8	1	1	9	7	5	2	7	10	3	3	2	4	8	2	2	7	2	3	4
---	---	---	---	---	---	---	---	----	---	---	---	---	---	---	---	---	---	---	---

Fonte: Os Autores

A condição de parada do sistema foi a mesma utilizada para o Software 1. Logo, para que a condição seja satisfeita um cromossomo com adaptação igual ao número de alunos deverá ser encontrado. Portanto, o algoritmo executou as tarefas de cruzamento e mutação, também considerando o *Ponto de Corte = Média da Adaptação da População Atual*, prosseguindo até a 5ª geração de cromossomos, onde foi encontrado o cromossomo mais adaptado que representa o plano de alocação de orientadores neste caso.

**Figura 10 - Plano de Alocação de Orientadores Sugerido no Software 2**

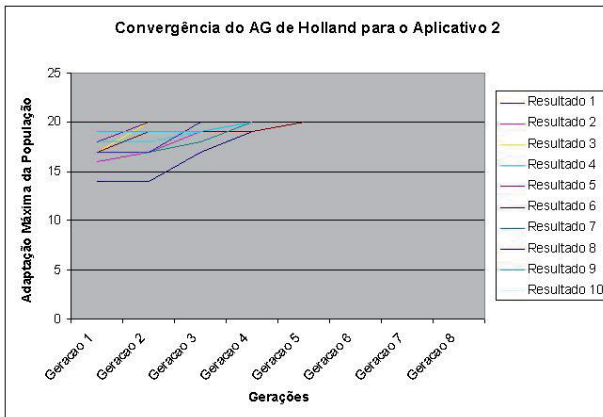
8	1	5	9	10	7	4	3	10	3	6	1	4	8	4	2	1	2	5	10
---	---	---	---	----	---	---	---	----	---	---	---	---	---	---	---	---	---	---	----

Fonte: Os Autores

Seguindo o mesmo procedimento aplicado para o Software 1, o sistema foi executado 10

vezes no sentido de verificar a convergência do algoritmo. Esse procedimento, que resultou no gráfico de convergência abaixo, demonstra a evolução do algoritmo através das novas populações geradas.

**Gráfico 2 - Convergência do Algoritmo de Holland no Software 2**



Fonte: Dados da Pesquisa

### 3.4 Resultados do software 3 – síntese dos softwares 1 e 2

O Software 3 utiliza uma técnica de Filtragem Híbrida, que representa uma síntese das filtragens utilizadas nos softwares 1 e 2. A função de adaptação verificará se o professor em questão tem capacidade (número máximo de alunos orientados) para orientar um dado aluno  $e$ , caso a resposta seja positiva, a função prossegue com a verificação da compatibilidade dos perfis do aluno e do orientador:

$$\text{adapt}(C, i) = \begin{cases} \text{adapt}(C, i + 1) + 1, & \text{se } ((\text{orienta}(G_i) \times \text{compatibilidade}) \geq 2) \wedge (i < n^{\circ} \text{ de alunos}) \\ \text{adapt}(C, i + 1) + 0, & \text{se } ((\text{orienta}(G_i) \times \text{compatibilidade}) < 2) \wedge (i < n^{\circ} \text{ de alunos}) \\ 1, & \text{se } ((\text{orienta}(G_i) \times \text{compatibilidade}) \geq 2) \wedge (i = n^{\circ} \text{ de alunos}) \\ 0, & \text{se } ((\text{orienta}(G_i) \times \text{compatibilidade}) < 2) \wedge (i = n^{\circ} \text{ de alunos}) \end{cases}$$

O Software 3 foi iniciado, sintetizando as regras pré-estabelecidas para a alocação de orientadores nos softwares 1 e 2, de maneira que o cromossomo inicial foi gerado aleatoriamente com adaptação = 13.

**Figura 11 - Cromossomo Inicial Utilizado no Software 3**

3 4 2 7 6 1 1 6 10 4 9 8 2 6 1 6 6 2 7 10

Fonte: Os Autores

Respeitando a mesma condição de parada dos softwares anteriores, o algoritmo executou as tarefas de cruzamento e mutação, também considerando o *Ponto de Corte = Média da Adaptação da População Atual*, prosseguindo até a 9ª geração de cromossomos, onde foi encontrado o cromossomo mais adaptado que representa o plano de alocação de orientadores sugerido.

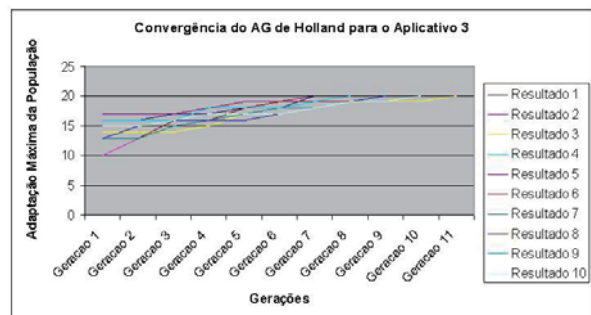
**Figura 12 - Plano de Alocação de Orientadores Sugerido no Software 3**

3 10 3 10 4 8 9 6 9 4 4 8 8 5 4 7 6 2 2 10

Fonte: Os Autores

Assim sendo, o software 3 também foi executado 10 vezes no sentido de verificar o quanto o algoritmo é convergente. O procedimento resultou no gráfico de convergência abaixo, demonstrando a evolução do algoritmo das novas populações geradas.

**Gráfico 3 - Convergência do Algoritmo de Holland no Software 3**



Fonte: Dados da Pesquisa

## **4 DISCUSSÃO**

Após propor a aplicação das variações do Algoritmo Genético de Holland ao problema abordado neste trabalho, pode-se destacar 2 (dois) aspectos: o primeiro é do ponto de vista metodológico, quando é utilizada uma metáfora da Teoria da Origem das Espécies de Darwin para resolver um problema considerado complexo. Na forma como metaforicamente foi definido o cromossomo, isto é, o ingrediente do problema de planejamento de orientadores de monografias em AVA, que faça o papel metafórico do cromossomo biológico. De forma análoga, também foram definidos os ingredientes que podem ser chamados de adaptação, cruzamento e mutação. Por similaridade com os seus correspondentes biológicos, pôde-se abstrair desse modo de escolha, um procedimento heurístico, baseado nas técnicas de recomendação propostas, para alocação de orientadores no contexto da EAD. Um segundo aspecto a ser considerado é que, o uso do Algoritmo Genético de Holland não permite a interferência do usuário no ambiente, limitando sua experiência e conhecimentos quanto à aplicação do método ao problema.

Quanto às técnicas de filtragem de informação utilizadas neste trabalho, percebe-se que estas se distinguem daquelas aplicadas na maior parte dos sistemas de recomendação voltados à EAD, à medida que permitem a sugestão de planos de alocação que minimizam as incompatibilidades de perfis de alunos e orientadores e a sobrecarga de alunos orientandos. Enquanto isso, os sistemas de recomendações convencionais voltados à EAD focam apenas a recomendação de materiais didáticos.

Em Feitosa et al (2011) foi apresentado o Software 1, que sugere planos de alocação levando em conta a compatibilidade dos perfis dos alunos e dos orientadores, mas não leva em consideração uma distribuição de alunos equivalente para os orientadores. Em termos práticos isso poderia gerar problemas no sentido de que determinados orientadores ficassem sem alunos, enquanto que outros estariam sobrecarregados.

O Software 2 sugere planos de alocação levando em conta apenas a capacidade máxima de captação de alunos orientandos, considerando

o regime de trabalho dos docentes. O trabalho de distribuição de orientadores foi feito com sucesso. Todas as recomendações atenderam aos limites de trabalho dos orientadores. No entanto, esse sistema não levou em conta a compatibilidade dos perfis. Na prática, isso poderia gerar problemas, à medida que o professor poderia apresentar interesses divergentes daqueles apresentados pelo aluno, tanto em termos de práticas metodológicas, como em termos de áreas afins de conhecimento.

Entretanto, o Software 3, ao sintetizar as regras utilizadas nos softwares 1 e 2, constituiu-se no sistema final para validação das filtragens propostas neste trabalho. Apesar de ter um aumento razoável no seu tempo de processamento, o sistema oferece ao Gestor de EAD um plano que o auxilia na minimização das dificuldades relacionadas à incompatibilidade de perfis e sobrecarga de trabalho docente. Se o Software 3 não resolve o problema do planejamento de orientadores, ao menos facilita a sua resolução, já que o gestor de EAD passa a dispor de um plano construído de alocação de orientadores, seja para acatá-lo ou para que sejam feitos ajustes ao plano que considera ideal.

## **5 CONCLUSÕES**

Os resultados apresentados pela pesquisa são considerados satisfatórios. Todas as versões dos sistemas de recomendação desenvolvidos podem ser aplicadas em escala bem maior do que a amostra utilizada neste trabalho. Além disso, os três sistemas permitem a flexibilização de seus parâmetros, dando autonomia ao gestor educacional e garantindo que as recomendações reflitam os interesses dos responsáveis pelo processo de planejamento de docentes.

Apesar dos bons resultados, entende-se que as técnicas de filtragem baseada em desempenho e filtragem baseada em capacidade, aliadas a outras técnicas de recomendação mencionadas em Cazella, Nunes e Reategui (2010), poderiam gerar informações ainda mais consistentes, principalmente se forem levados em consideração os aspectos psicológicos relacionados aos agentes humanos.

Isso se deve ao fato, principalmente, de que o ser humano é um ser complexo e devido às suas singularidades torna-se difícil a formalização de um perfil baseado apenas no

questo desempenho, visto que o desempenho do aluno reflete uma série de outras variáveis.

Por outro lado, as filtragens de informações aplicadas neste trabalho constituem-se em mais duas alternativas, em termos de técnicas de recomendação aplicadas a AVA, que podem ser utilizadas pelos desenvolvedores de sistemas computacionais, nos casos onde os fatores “desempenho” e “capacidade de orientação” forem relevantes.

Através da proposta do modelo desenvolvido espera-se que, a partir de agora, os gestores de EAD tenham, à sua disposição, soluções automatizadas que os auxiliem na minimização de problemas como a má aplicação de recursos na gestão educacional e as dificuldades enfrentadas no processo de construção monográfica.

Para trabalhos futuros, pretende-se aprofundar os estudos relacionados às técnicas de recomendação baseadas no desempenho dos

alunos e na capacidade laboral dos docentes, como também buscar-se-á a minimização de problemas relacionados às interações entre professores-tutores e alunos.

Dentre as possibilidades de trabalhos voltados para Educação, há o interesse de que seja trabalhada a alocação de professores-tutores para as disciplinas realizadas em AVA. Existe ainda a possibilidade de que sejam desenvolvidas novas técnicas de filtragem, que utilizem os dados gerados nos AVA pela interação entre professores-tutores e alunos.

Do ponto de vista da Inteligência Artificial, estuda-se a possibilidade da aplicação de algoritmos mais convergentes e que permitam uma maior interferência do desenvolvedor em relação ao ambiente, dentre eles o algoritmo genético baseado em tipos abstratos de dados por este permitir identificar mudanças de perfil (Professor e/ou Aluno) durante o processo de ensino-aprendizagem.

## **COMPUTER-ASSISTED RECOMMENDATION FOR MONOGRAPHS ADVISORY BASED ON INFORMATIONAL PROFILE AND TEACHERS' ADVISORY CAPACITY**

### **Abstract**

Social Combination Systems are recommendation systems able to suggest people, instead of recommending products or services. These systems are applied in various contexts, including Distance Learning (DL), modality of education that needs to have a set of computational tools in order to facilitate the teaching-learning process. The identification of potential advisors and work overload are some of the difficulties in this process. Researches in this area invest in the idea of working interests of students, teachers and education managers. Therefore, this study used a qualitative and quantitative approach, so that the data were collected in a Virtual Learning Environment (VLE) and through the application of a structured questionnaire to potential advisors. Thus, this paper aims to describe the application of two recommendation techniques, implemented through a Informational Model of Social Combination that assists the Educational Manager in the process of monographs advisors allocation. Three systems developed through genetic algorithms will be presented, which were applied in the Bachelor's course of Information Systems, offered jointly by the Federal University of Alagoas and the Open University of Brazil. These results indicate that the proposed filtering methods can be used to help educational managers in order to reduce cases of incompatibility of profiles of students and advisors as well as reducing the workload of teachers.

### **Keywords:**

Social Combination. Recommendation Systems. Informational Model. Distance Learning. Virtual Learning Environments.

---

Artigo recebido em 23/07/2011 e aceito para publicação em 11/12/2011

---

## REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO (ABED). **Informações do anuário estatístico de educação aberta e a distância**. Disponível em: [http://www2.abed.org.br/visualizaDocumento.asp?Documento\\_ID=555](http://www2.abed.org.br/visualizaDocumento.asp?Documento_ID=555). Acesso em 15 dez. 2010.

CARREIRA, R.; CRATO, J.; GONÇALVES, D.; JORGE, J. A. Evaluating adaptive user profiles for news classification. 9th International Conference on Intelligent User Interfaces, 2004, [S.l.]. **Proceedings...** [S.l.; s.n.], 2004.

CARVALHO, L. F. B. S. **Um jogo de damas evolutivo**. [S.l.; s.n.], 2008.

CAZELLA, S. C.; NUNES, M. A. S. N.; REATEGUI, E. **A ciência do palpite**: estado da arte em sistemas de recomendação. PucRIO, 2010. Disponível em: <http://www.dcomp.ufs.br/~gutanunes/hp/publications/JAI4.pdf>. Acesso em: 20 jan. 2011.

POO, D.; CHING, B.; GOH, J. A hybrid approach for user profiling. 36th Annual Hawaii International Conference on System Sciences, 2003, [S.l.]. **Proceedings...** [S.l.; s. n.], 2003.

PRIMO, L.; SILVA, C. R. O. **A prática de orientação a distância na elaboração de tccs**. 2007. Disponível em: <http://www.abed.org.br/congresso2007/tc/424200725707pm.pdf>. Acesso em: 10 fev. 2011.

ROUSSEAU, B.; BROWN, P.; MALONE, P.; FOSTER, P.; MENDIS, V. Personalised resource discovery searching over multiple repository types: using user and information provider profiling. CEIS, 2004, [S.l.]. **Proceedings ...** [S.l.; s.n.], 2004.

TERVEEN, L.; HILL, W. **Human-computer collaboration in recommender Systems**. [S. l.]: Addison Wesley, 2001.

WITTEN, I. H.; FRANK, E. **Data mining**: practical machine learning tools and techniques. [S.l.]: Elsevier, 2005.

FEITOSA, D. D. L.; SILVA, V. R. M. E.; LOPES, R. V. V.; PARAGUAÇU, F. D. D. C. An application of holland's genetic algorithm for the problem of planning advisors in virtual learning environments. 13th iee joint international computer science and information technology conference, 2011, [S.l.]. **Proceedings...**[S.l.; s.n.], 2011.