Automatização do Processo de Identificação de Presença Social em Fóruns e Chats

Júlia K. Kambara da Silva[1], Hélvia P. Pinto Bastos[2], Magda Bercht[1], Leandro Krug Wives[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – CEP 91.501-970 – Porto Alegre – RS – Brasil

[2]NIE - Instituto Federal Fluminense
CEP 28.030-130 – Rua Dr. Siqueira, 273, Campos dos Goytacazes – RJ – Brasil.

{jkksilva, bercht, wives}@inf.ufrgs.br, helviabastos@yahoo.com.br

Abstract. *One of the problems faced by current distance learning courses is the high abandonment rate. In order to address this problem, studies were performed to detect the students´ social presence, that is, their involvement within the virtual learning environment. However, to evaluate the degree of social presence, it is necessary to analyze all student´s manifestations in the environment, typically via forum and chat, which is a job that requires much effort from the teacher or tutor. Thus, this paper proposes an automatic approach to find the degree of social presence of the students. To verify this, a tool was developed, and four experiments were conducted and compared with the analysis performed by an expert. Results show that the proposed approach produces results that are close to the analysis performed by the expert.*

Resumo. *Um dos problemas enfrentados pelos cursos oferecidos em Educação a Distância é a alta taxa de desistência. A fim de minimizar esse problema, trabalhos foram realizados para descobrir a presença social do aluno, ou seja, o seu envolvimento com o ambiente virtual. Contudo, para encontrar o grau de presença social, é preciso analisar todas suas manifestações no ambiente, tipicamente fórum e chat, um trabalho que requer muito esforço do professor ou tutor. Desta forma, o que este trabalho propõe é uma abordagem automática para encontrar o grau de presença social do aluno. Para verificar esta abordagem, foi desenvolvida uma ferramenta e realizados quatro experimentos, comparados com a análise feita por um perito. Os resultados mostram que a abordagem aproxima-se da análise feita pelo perito.*

1. Introdução

O censo brasileiro da Educação Superior de 2010 confirma uma tendência de crescimento de cursos na modalidade de ensino a distância, onde em mais de seis milhões de matrículas em curso superior, que 14,6% delas são cursos oferecidos à distância (CENSUP, 2010).

Um dos maiores problemas enfrentados na Educação a Distância (EaD) é a evasão. Levy (2007), em seu estudo, constata que o principal fator de evasão nesta modalidade é o

baixo nível de satisfação do aluno. Richardon e Swan (2003) mostraram que a satisfação do aluno com o curso, a aprendizagem percebida do aluno, e sua satisfação com o instrutor, está relacionada com a percepção de Presença Social (PS).

Para tentar tratar desse problema, muitas pesquisas foram feitas sobre PS. PS é o grau de sentimento, da percepção de pertencimento no grupo, na comunidade em interação, e reação ao conectar-se com outros indivíduos por recursos de comunicação mediada por computador (TU, 2002). Dessa forma, é importante para o professor conhecer os graus de PS de seus alunos no decorrer de um curso, de modo que possa, se necessário, intervir de algum modo. Entretanto, conforme Máximo (2009), a avaliação de um grande número de alunos de forma processual a partir de suas participações em conversações é ainda extremamente difícil e lenta. Assim, frequentemente, não é possível ao professor diagnosticar problemas em tempo hábil. Desse modo, como apontam Junior e Costa (2012), existe a necessidade de desenvolvimento de ferramentas que proporcionem automação no todo ou em parte de análise qualitativa de fóruns, ajudando a obter resultados satisfatórios em intervalo de tempo exíguo.

O trabalho aqui apresentado visa justamente descrever o desenvolvimento de ferramentas que auxiliem na identificação de PS em fóruns e chats de Ambientes Virtuais de Ensino e Aprendizagem (AVEAs), com base no modelo de PS proposto por Bastos (2011). Tais ferramentas envolvem a coleta de dados a partir de um AVEA, a definição de categorias e pistas de PS, a análise da distribuição de PS nos dados coletados, tendo como base categorias e etiquetas definidas, e a geração de relatórios.

A fim de avaliar as ferramentas desenvolvidas, foram feitos experimentos e os resultados obtidos foram comparados com os encontrados por um perito. Verificou-se que a abordagem implementada pelas ferramentas aproxima-se da análise feita pelo especialista humano na grande maioria dos casos.

O artigo está organizado da seguinte forma. A seção seguinte apresenta trabalhos relacionados. A seção 3 descreve a modelagem do problema de identificação de PS. A seção 4 descreve a abordagem proposta, focando no software desenvolvido (ferramentas). A seção 5 mostra e discute os experimentos realizados e seus resultados. A seção 6 relata as conclusões do estudo.

2. Trabalhos Relacionados

Como o problema a ser tratado é focado em um modelo de PS novo, proposto por Bastos (2011), não há nenhuma ferramenta que realmente realize a análise automática da PS proposta em AVEAs. O que existe são ferramentas de análise lexicométrica que podem ser usadas para, a partir de um processo semiautomatizado, dar indícios de PS de alunos de um AVEA. Cabe salientar que a Lexicometria é um procedimento para tratar estatisticamente dados qualitativos sob fundo quantitativo para a caracterização topológica e combinatória de elementos lexicais de um corpus dado e delimitado (DAMASCENO, 2008).

Dentre as ferramentas analisadas, destacam-se as seguintes: WordSmith e Eurekha. O WordSmith[1] é um software pago que pode ser usado para gerar uma lista das palavras de um texto, em ordem alfabética ou de frequência, bem como suas correlações. Isso é importante porque o processo de identificação de PS (como será descrito na seção 3) tem como base a análise de palavras e sua frequência nos textos gerados pelos alunos em um AVEA. No entanto, o WordSmith é uma ferramenta genérica, não voltada somente para a PS, além de não estar integrada com AVEAs, sendo necessário um trabalho manual muito extenso e complexo para que os graus de PS sejam identificados.

O Eurekha[2], que está disponível na modalidade freeware para fins educativos, também gera uma lista de todas as palavras de um texto e suas respectivas ocorrências. Seu objetivo principal é agrupar documentos por similaridade de palavras, não sendo, portanto, totalmente adequado para gerar relatórios lexicométricos. Ainda, como concluem Bastos et al. (2010), a detecção de presença social pela análise automática no Eurekha mostrou-se insatisfatória devido à impossibilidade do programa em processar as diferentes estratégias discursivas, tais como processamento de palavras compostas, e a não localização da palavra no parágrafo ou na oração, assim como os termos que a antecedem e/ou sucedem. Além disso, o software não permite a definição de categorias lexicográficas, o que é fundamental para o processo de identificação de PS, conforme será descrito na seção seguinte.

Diante disso, optou-se por desenvolver um novo software, na verdade um conjunto de ferramentas capazes de realizar todas as etapas do processo de identificação de PS, facilitando o trabalho do professor ou tutor. Tal processo e ferramentas são descritos nas seções seguintes.

3. Descrição do Problema

A fim de fornecer informação o grau de envolvimento do aluno ao professor ou tutor, Bastos (2011) fez um estudo sobre Presença Social (PS), onde PS pode ser compreendida como a manifestação verbal e percepção da afetividade e interatividade dos sujeitos em relação ao ambiente virtual de ensino e aprendizagem. Dessa forma, PS é diferente do registro de *login* de alunos no ambiente virtual, pois ele só indica se o aluno entrou ou não no ambiente e quais ferramentas ou elementos acessou e em que ordem. Com base no grau de PS, o professor pode, se achar necessário, intervir de maneira a prevenir uma eventual desistência de um aluno.

Segundo Bastos (2011), a ocorrência de um indicador, que pode ser uma expressão textual ou a utilização de um recurso gráfico, em um corpus, seja em fóruns ou chats, gera o valor de PS do aluno. Em Bastos et al. (2011), tais indicadores são chamados de pistas, e foram usadas para detectar noções de PS e pertencimento. Com base nisso, Bastos et al. (2011) mostraram que as mensagens escritas em fóruns e chats fornecem indicativos importantes para o estudo de PS em AVEA, fornecendo envolvimento dos sujeitos no ambiente de aprendizagem.

[1] http://www.lexically.net/wordsmith/
[2] http://www.inf.ufrgs.br/~wives/wiki/doku.php?id=eurekha/

Como ilustra a Tabela 1, o modelo de Bastos (2011) é dividido em classes, subclasses e indicadores, que são formalizados por meio de pistas textuais. Tais pistas são palavras ou expressões da linguagem escrita que possuem relação com o contexto de cada indicador. Nesse sentido, com base em uma análise extensa do corpus textual de alguns cursos oferecidos em um AVEA específico (detalhes podem ser encontrados no documento de Bastos), Bastos definiu um conjunto de pistas para cada categoria. Cabe salientar, no entanto, que tais pistas podem mudar, não só pelo fato pelo fato do trabalho de Bastos ter sido focado em uma amostra muito pequena da população brasileira, mas também pelo fato da linguagem ser dinâmica e também com base no perfil de cada curso sendo analisado.

Tabela 1: Trecho do Modelo de Presença Social proposto por Bastos (2011)

CLASSE	SUBCLASSE	INDICADORES
AFETIVIDADE	AF1. Emoção	AF1a. Interjeições e expressões interjetivas
		AF1b. Onomatopéias
		AF1c. Emoticons, gifs e imagens
		AF1d. Pontuação repetida
		AF1e. Letras Repetidas
		AF1f. Realce
	AF2. Humor	AF2a. Expressão de humor
	AF3. Autorrevelação	AF3a. Vulnerabilidade
		AF3b. Vida pessoal
:	:	:
FORÇA	FO1. Intensidade	FO1a. Advérbios e expressões adverbiais
		FO1b. Adjetivos Comparativos
	FO2. Quantidade	FO2a. Pron. Indefinidos
	FO3. Sufixos	FO3a. Aumentativos
		FO3b. Diminutivos
	FO4. Valoração	FO4a.Termos valorativos e superlativos
	FO5. Figuras de linguagem	FO5a. Metáforas, analogias
	FO6. Ênfase	FO5a. Reforço

Sendo assim, o problema consiste em: (a) permitir ao professor ou tutor que defina, cadastre e altere categorias de PS e as pistas sociais relacionadas; (b) colete os dados relacionados com a análise desejada; e (c) gere relatórios dos graus de PS com base em extratos especificados (p.ex., por aluno, por turma, por curso, por fórum, etc.). Com base em tais requisitos, uma abordagem foi definida, envolvendo um conjunto de ferramentas e artefatos de software, seguindo uma arquitetura específica. Tal abordagem, assim como a arquitetura e seus componentes, é descrita na próxima seção.

4. Abordagem Proposta

A fim de que a ferramenta possa ser utilizada em qualquer AVEA, optou-se por uma abordagem mais flexível, onde os dados a serem processados devem ser convertidos para um formato-específico, independente de AVEA. Assim sendo, desde que os AVEAs exportem para o formato definido ou que sejam construídos softwares de conversão, o processo pode ser aplicado.

A abordagem proposta para resolver o problema de identificação dos valores de PS para um conjunto de alunos é composta, portanto, dos seguintes passos: aquisição do corpus

a ser analisado; conversão desse corpus para um formato padrão; definição de conjunto de classes, subclasses e indicadores de presença social (pistas textuais); e geração de relatórios de PS. A Figura 1 dá uma visão geral do processo e das ferramentas desenvolvidas, assim como os seus resultados. Como ilustrado na Figura 1, três módulos foram desenvolvidos para a obtenção da presença social de uma determinada turma, tendo o AVEA Moodle como base, a saber:

- Conversor HTML para XML: converte arquivos de fóruns/chats para o formato específico, em XML, especificado para a ferramenta de análise de PS (maiores detalhes sobre tal formato são dados na seção 4.1);
- Construtor de Categorias: permite ao professor criar e editar categorias, subcategorias de presença social, além de especificar suas respectivas pistas textuais. A ferramenta cria um arquivo de formato específico contendo as classes, subclasses e indicadores de presença social, que é lido pelo Analisador;
- Analisador de PS: com base no arquivo de categorias, analisa o arquivo de fóruns/chats e retorna a presença social dos alunos.

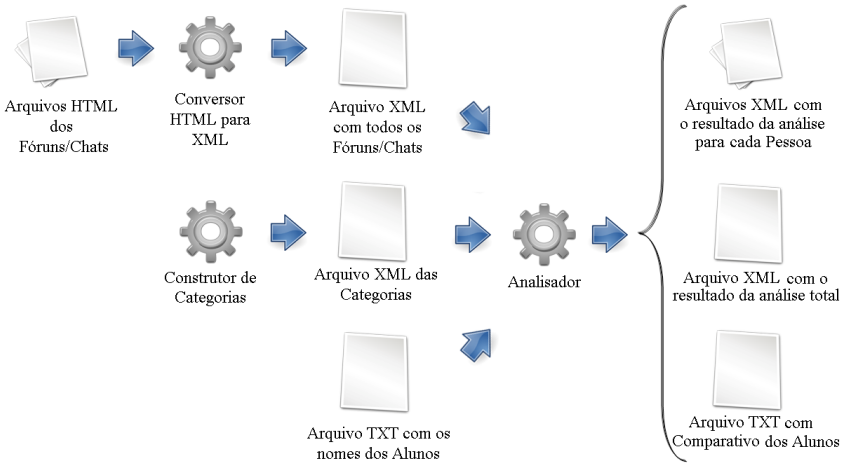

Figura 1: Abordagem proposta

Tais módulos são descritos a seguir. No entanto, é importante salientar que o conversor de HTML para XML foi elaborado porque nos experimentos iniciais realizados verificou-se ser conveniente salvar os arquivos correspondentes a um fórum ou chat diretamente em formato HTML a partir do navegador Web. Assim, o professor ou tutor pode salvar diversos arquivos correspondentes aos fóruns, chats, de um curso, transformá-los e combiná-los em um único arquivo contendo todos os dados a serem analisados. No futuro espera-se desenvolver um módulo integrado ao AVEA que permita a exportação dos conteúdos textuais, selecionados, diretamente para o formato da ferramenta.

4.1. Conversor HTML para XML

Após a coleta dos arquivos HTML a serem analisados, esses passam por um módulo que os converte para um arquivo no formato XML, ilustrado na **Figura 2** (esquerda), filtrando eventuais *tags* HTML que não são de interesse. Dessa forma, além da abordagem não ser

exclusiva de um ambiente virtual específico, ela também não diferencia fórum de chats, ambos podem ser convertidos para este mesmo formato. Tal formato visa guardar todas as informações pertinentes ao fórum, pois um fórum possui vários tópicos, tópicos possuem várias postagens, e postagens são formadas por título, autor, data/hora e conteúdo. Da mesma forma, no caso de chats, cada tópico representa uma conversa, que pode ter várias postagens, postagens representam o comentário de alguém na conversa, além de possuírem autor, data/hora e conteúdo. A diferença é que no chat uma postagem não possui título, porque representa um comentário no bate-papo.

```
<topicos>
    <topico_n>
        <postagem_m>
            <titulo> ... </titulo>
            <autor> ... </autor>
            <data> ... </data>
            <conteudo> ... </conteudo>
        </postagem_m>
        <postagem_m+1>
            ...
        </postagem_m+1>
    </topico_n>
    <topico_n+1>
        ...
    </topico_n+1>
</topicos>
```

```
<classes algoritmo="x">
    <classe algoritmo="x">
        <classe algoritmo="x">
            <pista_textual> ... </pista_textual>
            <pista_textual> ... </pista_textual>
        </classe>
    </classe>
    <classe algoritmo="x">
        <pista_textual> ... </pista_textual>
        <pista_textual> ... </pista_textual>
        <pista_textual> ... </pista_textual>
    </classe>
</classes>
```

Figura 2: Formato padrão para Fóruns/Chats (esquerda) e formato padrão para categorias (direita)

4.2. Construtor de Categorias

O módulo permite criar e organizar classes, subclasses e indicadores de PS. Como os indicadores são, na realidade, expressões léxicas a serem procuradas no texto, optou-se por deixar a busca mais inteligente, oferecendo tipos de busca diversificados (algoritmos de busca, na verdade). Detalhes sobre tais tipos ou padrões são dados a seguir. Com isto pretende-se flexibilizar a análise, aproximando-se um pouco mais da análise feita por um especialista humano. Este módulo também gera um arquivo XML, ilustrado na **Figura 2** (direita).

Tabela 2: Alguns Tipos de Buscas Oferecidos

TIPO DE BUSCA	DESCRIÇÃO
Expressão Simples	Procura no conteúdo de toda a postagem, fórum ou chat, a expressão. Uma expressão pode ser uma única palavra ou um conjunto de palavras.
Expressão Simples no Inicio da Frase	Procura a expressão apenas no inicio das frases.
Expressão Simples Início da Postagem	Procura a expressão no inicio de cada postagem.
Expressão Regular	A expressão cadastrada é vista como uma expressão regular, que será executada pelo Analisador.
Uso do Recurso Responder	Verifica se a postagem utilizou o recurso “Responder”. Este recurso só é utilizável em fóruns.
Expressões em Maiúsculo	Verifica se há expressões escritas em maiúsculos. Se houver várias palavras sequencias escritas em maiúsculo, só irá contar como uma única ocorrência.
Nomes Próprios	Verifica se há menção de algum nome próprio.

Conforme ilustrado na **Figura 2** (direita), o arquivo de categorias armazena um conjunto de classes e, para cada uma, um conjunto de pistas textuais. Uma classe também pode ter um conjunto de classes aninhadas. Entretanto a última classe filha tem que ter pelo

menos uma pista textual associada. Por motivos de simplificação, cada classe possui como atributo o algoritmo que será utilizado na busca da pista textual. Foram criados dezenove tipos/algoritmos de busca diferentes. A título de ilustração, a Tabela 2 apresenta alguns deles.

O construtor de categorias possui três áreas (ver Figura 3): a da esquerda, que mostra a árvore de classes e suas pistas textuais; a central, que permite editar as pistas textuais da categoria selecionada; e a da direita, que permite alternar o tipo de análise que é feito no arquivo em busca das pistas textuais.

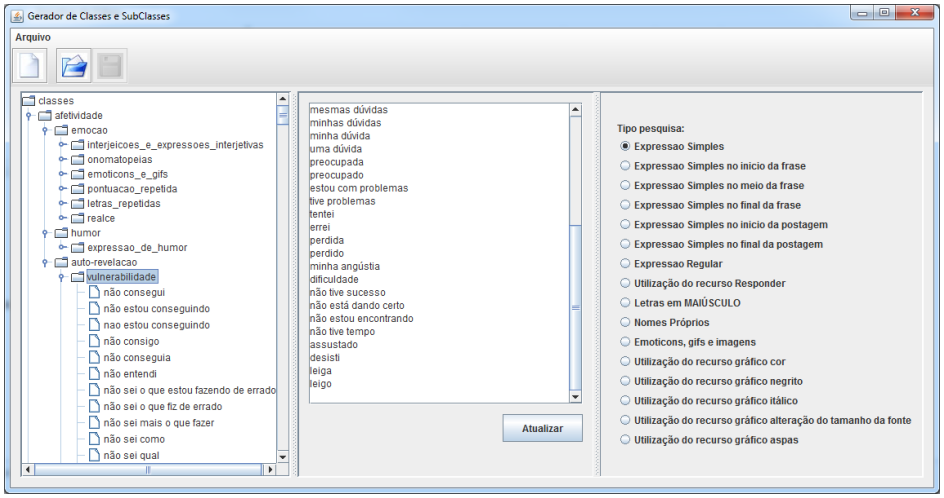

Figura 3: Construtor de categorias

4.3. Analisador

Este módulo analisa o arquivo de postagens tendo como base o arquivo de categorias definidas pelo professor ou tutor. Ele também recebe como entrada um arquivo contendo os nomes dos alunos cuja análise deve ser feita. Tal informação mostrou-se importante, pois se verificou que não só alunos fazem postagens nos fóruns/chats, mas sim também professores, tutores e monitores e deve-se excluí-los da análise de presença social. Assim, o Analisar retorna três tipos de arquivos: um conjunto de arquivos com a presença social detalhada de cada aluno; um arquivo com a presença social da turma; e um arquivo com todos os alunos e seus valores de presença social.

5. Experimentos

O AVEA escolhido para estudo de caso foi o Moodle[3]. Para fazer a avaliação da proposta, quatro experimentos foram realizados em dois fóruns distintos: dois em um fórum a partir de aqui denominado de Fórum A e dois no denominado Fórum B. Em ambos, um perito da área de linguística, e capacitado no modelo de PS proposto por Bastos (2011), fez uma análise manual, marcando as expressões e palavras que considerava pertencentes às respectivas classes de PS definidas. A comparação dos experimentos com a análise manual considera a análise do perito como ótima, ou seja, aqueles resultados da ferramenta diferentes dos da

[3] http://www.moodle.org.br/

análise manual são considerados incorretos. A diferença entre cada um dos dois experimentos são as seguintes:

- Experimento I: o perito analisou manualmente o Fórum A identificando as pistas que encontrava. Tais pistas foram, após, cadastradas no software. Procedeu-se então com a análise automática do fórum. Houve diversas iterações, onde o resultado da ferramenta foi usado pelo perito para refinar suas pistas. Ao final, as pistas foram aplicadas tanto no fórum A quanto no B.
- Experimento II: sem observar os fóruns, o perito cadastrou pistas textuais que acreditou descreverem melhor as classes, dessa forma, esperava-se obter um arquivo com pistas textuais mais genéricas, que pudessem ser utilizadas em qualquer fórum. Procedeu-se então com a análise automática dos fóruns A e B.

Como as classes são um conjunto de pistas textuais, pretendeu-se avaliar o impacto da escolha das pistas textuais e seus respectivos algoritmos no resultado.

O Fórum A é composto por 66 tópicos e 142 postagens. Nele, durante o experimento I, o perito fez várias iterações de análise manual intercaladas com as automáticas, iniciando com 679 marcações de pistas textuais e chegando a um total de 942 pistas identificadas na última iteração. Esse fato demonstra que o processo manual de identificação e marcação de pistas é, por si só, bastante oneroso, complexo e cansativo. Também mostra a utilidade da ferramenta na elaboração de pistas. As pistas foram cadastradas na ferramenta e aplicadas no mesmo fórum A. A intenção foi verificar se as regras de identificação e expressões cadastradas chegariam ao mesmo resultado. Na verdade, a ferramenta acabou identificando mais pistas do que a análise manual, conforme pode ser verificado na Tabela 3.

Tabela 3: Experimentos com os Fóruns A e B

CATEGORIAS	Fórum A			Fórum B		
	ANÁLISE MANUAL	ANÁLISE AUTOMÁTICA		ANÁLISE MANUAL	ANÁLISE AUTOMÁTICA	
		Exp. I	Exp. II		Exp. I	Exp. II
AFETIVIDADE	189	218	257	362	811	864
INTERATIVIDADE	250	251	276	680	649	953
COESÃO SOCIAL	392	414	534	1488	1453	2429
FORÇA	111	116	263	871	872	2384
TOTAL	942	999	1330	3401	3785	6630
Taxa de acerto	-	93,95%	58,90%	-	88,80%	5,00%

Verifica-se que o resultado do Experimento I para o fórum A obteve em torno de 94% de acerto. A diferença entre valores foi provocada porque a ferramenta não reconheceu o contexto de algumas pistas textuais, gerando marcações a mais, inválidas. Por exemplo, a ferramenta confundiu siglas com expressões em maiúsculo. Além disso, quando a pista consistia no pronome da terceira pessoa, a ferramenta não fez distinção entre aqueles que eram usados para coisas ou animais, ao contrário do perito, que não as marcou, visto que em PS só são interessantes aqueles que se referem a pessoas. Além disso, para a categoria Afetividade, por exemplo, onde uma das pistas é a mudança de cor do texto ou uso de negrito, verificou-se que a ferramenta encontrou repetidos códigos HTML de mudança de cor ou negrito para uma única frase ou parágrafo, contabilizando-os como pistas diferentes.

O Fórum B apresenta 74 tópicos e 7309 postagens. Devido ao grande número de postagens, o perito apenas realizou duas iterações de análise manual. Na Tabela 3 verifica-se que a taxa de acertos ficou em torno de 89%, mas isso só foi obtido após termos analisado os resultados da marcação manual do perito, pois se constatou que, pela quantidade de mensagens, o perito algumas vezes fez marcações duplas, marcações trocadas (classes inconsistentes), marcações com códigos incorretos (com erro de ortografia e similares). Tal fato reforça o fato de que a tarefa do perito é muito onerosa e complexa.

Já o experimento II, com pistas textuais genéricas especificadas pelo perito com base na sua experiência, não obteve um resultado tão bom quanto o experimento I, conforme pode ser observado na Tabela 3, isto pode ter ocorrido visto que as pistas textuais estão muito amplas e não restringem devidamente a sua classe, o que gera um contagem maior do que a esperada. Uma análise mais profunda não foi realizada, mas deve ser feita no futuro.

Contudo, apesar disso, a ferramenta mostrou ser capaz de fornecer indícios da presença social dos alunos, podendo funcionar como um mecanismo relevante de sinalização ao professor ou tutor de que determinados alunos merecem maior atenção. Tal análise, se feita manualmente, envolve muitas mensagens em diferentes fóruns, e dificilmente um professor conseguiria analisar um volume de mensagens muito grande em uma turma, quiçá em um conjunto maior de turmas.

6. Conclusão

O problema de automatização do processo de obtenção do nível de Presença Social em AVEAs foi tratado neste artigo nos seguintes passos: obtenção dos arquivos de fóruns/chats; conversão destes arquivos para um formato específico; geração de um arquivo com as classes, subclasses e indicadores de presença social; e por fim a análise.

Com estes passos procurou-se uma abordagem que não fosse específica de um determinado ambiente virtual. Ainda, uma vez feito o arquivo de classes e indicadores de presença social, ele pode ser utilizado em qualquer fórum ou chat.

Embora o resultado seja satisfatório em parte dos experimentos, naqueles que mostraram resultados menos interessantes, fica clara a importância do perito para povoar os indicadores, não só fazendo uma análise superficial, mas no sentido de fazer um estudo das palavras mais frequentes, levando em conta o gênero número e o grau das pistas, além da voz verbal utilizada. Todas essas variações morfológicas devem ser marcadas como um indicador de classe, o que geralmente é uma tarefa tediosa, cansativa e passível de erros. Isto também demonstra uma limitação da ferramenta, que é baseada em lexemas fixos, sem dispor de um dicionário morfológico ou regras mais complexas que envolvam um complexo analisador sintático da língua natural.

Desta forma, como trabalho futuro é interessante realizar um estudo das expressões e suas variações mais utilizadas em fóruns/chats que indicam a ocorrência de uma classe de presença social, e se for necessário, ampliar o número de tipos de busca, além de incluir um analisador morfológico e gramatical a fim de ampliar as possibilidades de análise.

Referências

Bastos, H., Bercht, M., Wives, L. K. "Análise Manual e Automática de Pistas Lexicais de Presença Social em Chat Educacional". 21º Simpósio Brasileiro de Informática na Educação (21º SBIE), 2010.

Bastos, H. "Presença Plus: Um Modelo de Identificação de Presença Social como Ferramenta Auxiliar da Tutoria em Ambientes Virtuais de Ensino e Aprendizagem". Tese de Doutorado em Informática na Educação da Universidade Federal do Rio Grande do Sul, PPGIE-UFRGS, Porto Alegre, 2011.

Bastos, H., Bercht, M., Wives, L. K. "Presença Social e Pertencimento em Fóruns Educacionais: Manifestação e Percepção de Afetividade". 22º Simpósio Brasileiro de Informática na Educação (22º SBIE), p.1047-1056, 2011.

CENSUP. "Censo da Educação Superior 2010 - Divulgação dos principais resultados do Censo da Educação Superior 2010". Outubro, 2011. Disponível em : < http://download.inep.gov.br/educacao_superior/censo_superior/documentos/2010/divulgacao_censo_2010.pdf>

Damasceno, E. A. "Lexicometria, geração de descritores, construção de ontologias e ensino de línguas: implicações e perspectivas". In: Magalhães, J. S. de & Travaglia, L. C.. (Org.). Múltiplas Perspectivas em Linguísticas. 1a ed. Uberlância: EDUFU, v.01, p. 01-3037, 2008.

Junior, L., Costa, A., "Discutindo a avaliação mediada por fóruns no contexto de cursos totalmente à distância: uma experiência no Etec-Brasil". 22º Simpósio Brasileiro de Informática na Educação (22º SBIE), p.1027-1036, 2011.

Levy, Y. "Comparing Dropouts and Persistence in e-Learning Courses". Computers & Education, v.48, issue 2, p.185-204, 2007.

Máximo, L. F. "A Efetividade de Feedbacks Informatizados Sobre a Autoregulação da Aprendizagem em Cursos a Distância : Um Estudo de Caso na Área da Computação". 105 f. Tese de Doutorado em Informática na Educação, UFRGS, Porto Alegre, 2009.

Richardson, J. C., Swan, K. "Examining social presence in online courses in relation to students' perceived learning and satisfaction". Journal of Asynchronous Learning Networks, 7 (1), p.68-88, 2003.

Tu, C.-H. "The impacts of text-based CMC on online social presence". The Journal of Interactive Online Learning, 1(2), p.1-24. 2002.

Wives, L. K**. "**Um estudo sobre Agrupamento de Documentos Textuais em Processamento de Informações não Estruturadas Usando Técnicas de "Clustering"". 1999. 102 f. Dissertação de Mestrado em Ciência da Computação – Instituto de Informática, UFRGS, Porto Alegre.

Wordsmith. "Step-by-step guide to WordSmith". Disponível em: < http://www.lexically.net/wordsmith/step_by_step/index.html >.