Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais

Extraction of Profile and Context Features in Social Networks for Educational Resources Recommendation

Crystiam Kelle Pereira

Programa de Pós-graduação em
Ciência da Computação
Universidade Federal de Juiz de
Fora (UFJF)
crystiamkelle@gmail.com

Fernanda Campos

Programa de Pós-graduação em
Ciência da Computação
Universidade Federal de Juiz de Fora
(UFJF)
fernanda.campos@ufjf.edu.br

Victor Ströele

Programa de Pós-graduação em
Ciência da Computação
Universidade Federal de Juiz de
Fora (UFJF)
victor.stroele@ice.ufjf.br

Regina Braga

Programa de Pós-graduação em
Ciência da Computação
Universidade Federal de Juiz de
Fora (UFJF)
regina.braga@ufjf.edu.br

José Maria N. David

Programa de Pós-graduação em
Ciência da Computação
Universidade Federal de Juiz de Fora
(UFJF)
jose.david@ufjf.edu.br

Resumo

Este trabalho apresenta uma proposta que faz recomendações educacionais individualizadas, usando informações geradas espontaneamente nas Redes Sociais. O objetivo é identificar características do perfil e do contexto do usuário através do uso desses sistemas e da sua interação social e então, fazer recomendações educacionais aderentes ao seu perfil e contexto. A avaliação da proposta foi feita a partir do desenvolvimento de um protótipo, de uma prova de conceito e um estudo de caso. Os resultados da avaliação mostraram a viabilidade e uma aceitação relevante por parte dos usuários no sentido de extrair informações sobre os seus interesses educacionais, geradas automaticamente da rede social Facebook, enriquecê-las, encontrar interesses implícitos e usar essas informações para recomendar recursos educacionais. Foi verificada também a possibilidade da recomendação de pessoas, permitindo a formação de uma rede de interesses em torno de um determinado tema, indicando aos usuários bons parceiros para estudo e pesquisa.

Palavras-Chave: *Sistemas de Recomendação, Redes Sociais, Perfil do Usuário.*

Abstract

This paper presents a proposal that generates individualized educational recommendations, using information spontaneously generated in Social Networks. The main goal is to identify characteristic of user's profile and context through the use of these systems and his/her social interaction and then make consistent educational recommendations. The proposal evaluation was made from the development of a prototype, a proof of concept and a case study. The evaluation results showed the viability and relevant users' acceptance in order to extract information about their educational interests, automatically generated from the Facebook social network, enrich these information, find implicit interests and use these information to recommend educational resources. It was also validated the possibility of people recommendation, enabling the establishment of interest network, based on a specific subject, showing good partners to study and research.

Keywords: *Recommendation System, Social Network, User Profile*

1 Introdução

Nos últimos anos nota-se um aumento na busca por informações através da Internet. Essa busca tem ocorrido tanto através de cursos formais, nas modalidades à distância e semipresenciais, quanto na procura informal por conteúdos de interesse. Existem inúmeros recursos educacionais distribuídos em diferentes repositórios, que abordam um conjunto amplo de assuntos e que possuem objetivos educacionais distintos. Em virtude disso, a escolha adequada desses recursos é um desafio para os educadores e para os próprios alunos.

Nesse cenário, um sistema de recomendação (SR) tem papel importante para auxiliar educadores e alunos a encontrarem recursos educacionais relevantes e pertinentes aos seus perfis e ao contexto em que estão inseridos. Segundo [1], um SR pode ser “qualquer sistema que produza recomendações individualizadas ou que tenha o efeito de guiar o usuário de forma personalizada para objetos do seu interesse ou que lhes sejam úteis dentre diversas opções possíveis”.

Para que se torne possível gerar recomendações personalizadas, observa-se a necessidade de informações que ajudem a definir o perfil do usuário e auxiliem na identificação de suas necessidades e interesses. Por outro lado, observam-se alguns sistemas que capturam, espontaneamente, uma grande quantidade de dados dos usuários, bem como a interação entre eles, evidenciando, muitas vezes, suas preferências e interesses. Esse é o caso, por exemplo, das Redes Sociais. Segundo [2], o site de rede social mais popular entre estudantes universitários é o Facebook¹. De acordo com uma pesquisa realizada pela QuacquarelliSymonds [3] entre 2012 e 2013, com 918 entrevistados em 26 países distribuídos pela Europa, Ásia, América Latina, América do Norte e África, o uso do Facebook entre os estudantes varia entre 78% e 96% nessas regiões, sendo que a América Latina é a região que mais utiliza o Facebook.

Assim, neste trabalho, foram extraídas e exploradas as informações disponíveis nas redes sociais, bem como o interesse dos seus usuários em estarem atualizados através do seu uso. Por meio da extração de informações disponíveis nas redes sociais, busca-se identificar e inferir características, preferências e interesses educacionais dos usuários, usando técnicas de Extração de Texto e Web Semântica. O objetivo principal é a identificação de características do perfil e do contexto do usuário a partir de informações geradas, espontaneamente, através do uso das redes sociais, para auxiliar no processo de seleção e recomendação de recursos educacionais adequados a esse perfil. Além disso, busca-se explorar o longo tempo despendido

pelos estudantes em redes sociais, levando até eles orientações e recursos educacionais.

Este trabalho pretende ainda contribuir com o projeto BROAD [4][5][6][7] que engloba pesquisas relacionadas a investigação e adoção de tecnologias de software, tais como ontologias, serviços web semânticos, agentes e workflow para construir uma arquitetura para a busca personalizada por objetos de aprendizagem, bem como para a sua composição em conteúdos educacionais. Este trabalho pretende evoluir a proposta do projeto BROAD, acrescentando características de perfil e do contexto educacional dos usuários, através da aquisição automática de informações disponibilizadas por usuários de redes sociais, possibilitando assim, a recomendação de recursos educacionais personalizados.

O artigo está organizado da seguinte forma: na seção 2 são apresentados alguns trabalhos relacionados, na seção 3 é dada uma visão geral da arquitetura de recomendação, nomeada BROAD-RSI, já na seção 4 são apresentados o protótipo desenvolvido e a avaliação do modelo de recomendação feita através de uma prova de conceito e um estudo de caso. Por fim, na seção 5 são feitas as considerações finais.

2 Trabalhos relacionados

Em [8] é apresentada uma proposta que busca modelar o contexto a partir de mensagens de redes sociais, e usar esse contexto para melhorar a relevância na recuperação de documentos na Web, tendo como objetivo apoiar a aprendizagem colaborativa. Em [9], os autores apresentam uma proposta de recomendação baseada na avaliação de conteúdo em uma rede social educacional, que foi criada no escopo do trabalho apresentado. Os alunos compartilham uns com os outros as suas classificações para recursos didáticos, sendo considerados o estilo de aprendizagem e o nível de conhecimento para atribuir pesos diferentes às recomendações.

Em [10], os autores apresentam uma abordagem para melhorar o entendimento e conhecimento dos interesses e preferências dos usuários a partir de informações disponíveis em sistemas colaborativos, mais especificamente redes sociais. A proposta é extrair automaticamente um perfil do usuário, sendo para isso consideradas as postagens dos usuários nas redes sociais. Esse trabalho não tem como objetivo principal a recomendação, no entanto menciona essa possibilidade.

Em [11] é apresentada uma técnica de recomendação baseada na construção implícita de perfis de usuários através da observação dos valores de metadados dos conteúdos que esses usuários acessam. A técnica de recomendação proposta utiliza a filtragem colaborativa combinada com

¹www.facebook.com

uma técnica de agrupamento de usuários por interesses similares. Em [12] é apresentado o UbiGroup, um modelo de recomendação ubíqua de conteúdo educacional para grupos de alunos. Esse modelo leva em conta o perfil dos alunos e o contexto onde eles estão inseridos. O UbiGroup faz uso de agentes de software para realizar o processo de recomendação. Feita a associação são realizados filtros considerando o perfil e o contexto do aluno, que são definidos manualmente.

O sistema de recomendação BROAD-RSI traz aspectos inovadores em relação aos trabalhos aqui relacionados [9][10][12] no sentido de extrair e considerar informações geradas espontaneamente nas redes sociais, mais especificamente no Facebook, já que esses sistemas obtêm informações acerca do usuário através da aplicação manual de formulários.

Na proposta apresentada em [8] as redes sociais já haviam sido exploradas para extração automática de contexto a partir das discussões em grupos, sendo esse usado para recuperação de documentos na Web. A proposta deste trabalho traz avanços em relação a [8] ao tornar possível gerar recomendações individuais, especificamente educacionais, coerentes com os interesses e as preferências de qualquer usuário na rede social, mesmo que este não esteja inserido em um grupo com discussões no âmbito educacional.

A possibilidade de envio das recomendações para a rede social do usuário permite que ele possa usar parte do tempo que passa interagindo para ter acesso a conteúdos educacionais e que possa, ainda, compartilhar tais conteúdos, recomendá-los para outros usuários, enfim, utilizar todos os recursos intrínsecos das redes sociais. Tais possibilidades trazem progressos em relação aos trabalhos [11][12] uma vez que esses estão inseridos dentro de sistemas acadêmicos, tais como ambientes virtuais de aprendizagem, restringindo assim o acesso às recomendações para usuários que estejam de alguma forma inseridos em um contexto acadêmico.

3 BROAD-RSI

O projeto BROAD-RSI (Sistema de Recomendação Baseado nas Interações Sociais)[6][13][14], descrito no escopo deste artigo, busca gerar recomendações educacionais individualizadas, usando informações geradas espontaneamente nas Redes Sociais.

O processo simplificado de extração das informações do usuário e do seu uso na recomendação, conforme está sendo adotado pelo BROAD-RSI, pode ser visualizado na figura 1. Usando a rede social Facebook, um sistema independente da arquitetura do BROAD-RSI, o usuário inicia o processo realizando o seu acesso através da sua conta nessa rede social. Logo após, ainda através da API (*Application Programming Interface*) do Facebook, o usuário é

questionado sobre a autorização para que suas informações sejam usadas pelo BROAD-RSI. Diante da autorização dada pelo usuário, o primeiro módulo da arquitetura é responsável pela extração das informações, que incluem dados de perfil, lista de amigos, páginas curtidas, postagens e grupos. Na etapa seguinte, as informações são filtradas, buscando aquelas que definem interesses educacionais, preferências sobre dias e horários de acesso, idiomas, mídias e, ainda, os dados de perfil, tais como nome, e-mail, idade e gênero.

Uma vez filtradas as informações necessárias, os interesses identificados passam por uma etapa de enriquecimento, usando algumas técnicas de Extração de Informação para identificar interesses relacionados. Após a etapa de enriquecimento dos interesses é feita a representação semântica desses e também do perfil e do contexto do usuário. Por fim, o perfil e contexto são usados para determinar os parâmetros de busca e relevância dos recursos educacionais e pessoas com interesses semelhantes. A etapa posterior à busca inclui a priorização dos recursos/pessoas encontrados e envio dessas recomendações para o usuário através de um aplicativo na rede social.

Figura 1 - Fluxo de Recomendação

3.1 Arquitetura do BROAD-RSI

A arquitetura do BROAD-RSI, mostrada na figura 2, pode ser dividida em 4 módulos: Extração de Informação; Enriquecimento do Perfil; Representação Semântica e Recomendação.

Os principais objetivos da arquitetura são: (1) especificar uma estratégia para extrair informações acerca do interesse educacional e informações básicas de perfil dos usuários, usando para isso a rede social Facebook; (2) determinar itens de perfil educacional do usuário a partir das

informações extraídas; e enriquecer o perfil extraído, buscando tópicos de interesses implícitos; (3) gerar a representação semântica desse perfil e (4) considerar as informações extraídas para sugerir recursos educacionais apropriados.

Figura 2 - Arquitetura do BROAD-RSI

3.1.1 Extração de Informações

O módulo de extração de informações é responsável por identificar as informações disponíveis através do perfil e das interações do usuário, selecionar quais são relevantes para a recomendação e filtrar tópicos relacionados aos seus interesses educacionais.

Através de sua API, o Facebook disponibiliza informações referentes às conexões de cada usuário, informações públicas do seu perfil (nome, idade, cidade de nascimento e residência, dados educacionais, dados profissionais, idiomas falados), seus interesses e preferências (livros, músicas, filmes, jogos, páginas, entre outros) e ainda os grupos dos quais ele participa.

A definição de quais informações seriam consideradas no BROAD-RSI foi feita em duas etapas.

Na primeira, buscou-se identificar quais eram as informações relevantes para Sistemas de Recomendação educacional, ou seja, quais elementos do perfil e do contexto do usuário seriam úteis para gerar recomendações relevantes. A investigação dessa questão foi feita através de um mapeamento sistemático [14], cujo objetivo foi identificar os elementos contextuais encontrados em trabalhos relacionados a Sistemas de Recomendação educacionais sensíveis ao contexto e fazer uma análise quantitativa das categorias de contexto, além de uma lista abrangente de informações em cada categoria contextual.

Na segunda etapa foi feito um filtro no conjunto de informações identificadas na primeira etapa, considerando

apenas aquelas possíveis de serem extraídas da rede social Facebook. Ao final das duas etapas, foram definidos os elementos contextuais indicados na tabela 1. As informações foram categorizadas de acordo com o framework definido em [15], conforme detalhado em [14].

Categoria	Subcategoria	Elementos de Contexto
Tempo	Horário	Horário de acessos aos sistemas
	Intervalo de tempo	Dia da semana e período do dia com maior número de acesso
Atividades	Tópicos	Disciplinas cursadas
Usuário	Informações básicas	Nome Qualificação Cidade onde reside Histórico educacional Idade
	Conhecimento	Conhecimento de Idiomas Competências educacionais
	Preferências Interesses	Interesses gerais e educacionais Preferência (mídia, comunicação, música, notícias, filmes)
Relacionamentos Pessoais		Interesses comuns entre amigos Similaridade entre pessoas Grupos de discussão nos quais o usuário participa

Tabela 1 - Elementos contextuais usados no BROAD-RSI

Depois de extraídas as informações citadas anteriormente, faz-se necessário filtrá-las, buscando quais, dentre os diversos interesses do usuário, podem ser consideradas de cunho Educacional. Para isso, são consideradas: as áreas de interesse definidas no Perfil Educacional do usuário e as páginas curtidas, cuja categoria esteja no âmbito Educacional, tais como, páginas referentes a faculdades, escolas, linhas de pesquisa, entre outras. Tais categorias são definidas para cada página no Facebook no momento de criação da página. É extraído também o nível educacional do usuário (Ensino Médio, Graduação, Pós-Graduação).

Apesar das informações educacionais não serem obrigatórias, elas estão presentes no perfil de muitos usuários. Segundo um estudo feito no escopo deste trabalho com uma amostra de 251 usuários, 92 deles possuíam, em seu perfil público, informações sobre áreas educacionais de interesse, ou seja, 36,65%. Os participantes do estudo não necessariamente estavam inseridos em um ambiente acadêmico. Eles eram amigos na rede social Facebook de alguns alunos voluntários do mestrado em Ciência da Computação da Universidade Federal de Juiz de Fora. Considerando o amplo uso do Facebook, o quantitativo de pessoas beneficiadas por uma recomendação educacional pode ser significativo.

Além das informações relacionadas aos interesses educacionais dos usuários, são capturadas também aquelas relativas às suas preferências em relação aos dias, horários e períodos (manhã, tarde e noite) de acesso à rede social. Pretende-se, com isso, enviar recomendações educacionais nos horários mais oportunos de acordo com o perfil de acesso do usuário.

São extraídas, ainda, as preferências por tipos diferentes de mídias, a partir dos quantitativos dos diferentes tipos de recursos compartilhados pelo usuário no Facebook, tais como Livros, Músicas, Vídeos, Imagens e Jogos. Essa preferência ajuda a priorizar recursos de acordo com o seu formato, não sendo usada, portanto, para eliminar uma recomendação, mas apenas para dar mais importância e prioridade a um subconjunto delas.

Para acessar as informações do usuário foi usada a API fornecida pelo Facebook para desenvolvedores de aplicativos, chamada GraphAPI, na versão 2.0. Essa API é uma plataforma para a criação de aplicativos e está disponível para os membros dessa rede social. Ela permite aos aplicativos usar as conexões sociais e informações de perfil para o desenvolvimento de aplicações próprias com diferentes propósitos. Todos os aplicativos desenvolvidos estão sujeitos a aceitação dos termos de privacidade. A GraphAPI usa o protocolo RESTful [16] e as respostas são dadas em formato XML. A exploração da GraphAPI foi feita através

da biblioteca chamada RestFB [17]. A escolha dessa biblioteca foi feita por ela ser escrita em Java, possuir ampla documentação, contendo exemplos de utilização, ter licença *Open Source* e ser muito utilizada por aplicativos desenvolvidos em Java que utilizam a API do Facebook.

3.1.2 Enriquecimento de Perfil

O enriquecimento dá-se com maior ênfase nos tópicos de interesse, ou seja, o enriquecimento é feito a partir das páginas que o usuário demonstrou ter interesse ou de grupos que ele participa. Para realizar o enriquecimento é usado o texto descritivo de cada uma das páginas, onde são identificados o idioma, as entidades, os conceitos e as palavras-chave explicitamente ou implicitamente relacionados ao texto.

Para a identificação desses conceitos, palavras-chave, entidades e idioma foi usada a API AlchemyAPI². AlchemyAPI é uma ferramenta que oferece um conjunto de funcionalidades sobre o texto fornecido, como análise de sentimentos, extração de tags, detecção de idioma, extração de entidades, categorização de texto, extração de conceitos relacionados, extração de palavras-chave, entre outros. A maioria destas funcionalidades usa métodos estatísticos de processamento de linguagem natural [18] e algoritmos de aprendizagem de máquina [19].

A escolha da AlchemyAPI foi feita apoiada em resultados apresentados em [20] e [21] nos quais a AlchemyAPI apresentou bons resultados em relação a outras APIs, tais como F-EXT³, Ltasks⁴ e Boilerpipeline⁵. Além disso, a AlchemyAPI possui suporte para diferentes idiomas, fator que foi determinante para a sua escolha, já que, por se tratar de dados extraídos de redes sociais, resultantes da colaboração de pessoas de diferentes nacionalidades, é possível que páginas escritas em diferentes idiomas sejam encontradas.

A detecção de idioma é usada para determinar quais demais funções podem ser aplicadas ao texto, já que algumas delas são restritas a um conjunto limitado de idiomas. A identificação do idioma é relevante ainda para descobrir idiomas de domínio do usuário, caso essa informação não esteja explicitada no seu perfil.

Um exemplo de enriquecimento de perfil, onde são identificados temas que não estavam explicitamente declarados no perfil do usuário pode ser visto na figura 3. Um usuário X curtiu uma página sobre o tema “DBPEDIA” no Facebook. As informações das páginas são apresentadas em formato de texto, ou seja, sem uma estrutura definida que permita a identificação de tópicos de interesses relacionados ao tema “DBPEDIA”. Quando o texto é submetido aos métodos de extração de informação da AlchemyAPI,

²<http://www.alchemyapi.com/>

³ Site: <http://ltasks.com>

⁴ Site: <http://www.opencalais.com>

⁵ Site: <https://code.google.com/p/boilerpipeline>

citados anteriormente, é possível identificar no texto alguns temas, como por exemplo: “Semantic Web”, “Wikipedia”, “Query Language”, “RDF”, “Linked Data”. Dessa

forma, é possível encontrar outros temas implícitos, relacionados ao tema DBPEDIA e, assim, conhecer melhor possíveis interesses do usuário X, enriquecendo o seu perfil.

LANGUAGE: English

DBpedia is the Semantic Data Web interface to Wikipedia. It exposes Wikipedia as a database and provides a live demonstration of power of Structured Web Data (aka, RDF-based Linked Data). DBpedia includes the following: 1. REST-based API (SPARQL) 2. Query Language (SPARQL) 3. Ability to serialize results in a myriad of formats - XML, JSON, RDF, N3 4. Full Text Search Interface 5. Lots of examples!

Click here to learn more about concepts.

Visual JSON API

Entities	Concept	Relevance	Linked Data
Keywords	Semantic Web	0.944142	dbpedia freebase yago website
Taxonomy			
Concepts			
Document Sentiment	World Wide Web	0.657546	dbpedia freebase
Targeted Sentiment			
Relations			
Language			
Title			
Author			
Text			
Feeds			

Entity	Relevance	Sentiment	Type	Subtypes	Linked Data
Wikipedia	0.880065	neutral	Company		
Query Language	0.576089	neutral	FieldTerminology		

Keyword	Relevance	Sentiment
RDF-based Linked Data	0.93056	neutral
Semantic Data Web	0.879067	neutral

Figura 3 - Exemplo de enriquecimento de perfil

3.1.3 Representação Semântica

O módulo de representação semântica é responsável por criar uma ontologia do perfil do usuário e dos interesses encontrados, associando as informações extraídas às classes e propriedades pertinentes. Além disso, esse módulo é responsável pelo armazenamento das informações em uma base de dados semântica.

A representação semântica do perfil e do contexto do usuário se torna importante para futuras inferências, associações e buscas semânticas que podem auxiliar tanto no enriquecimento do perfil do usuário, quanto no enriquecimento dos tópicos de interesses.

Para a representação do perfil e do contexto do usuário foram usados os vocabulários FOAF [22] e SIOC [23], com algumas extensões necessárias.

Os interesses educacionais, bem como a extensão dos conceitos associados a eles, também dão origem a uma base semântica de interesses educacionais extraídos do Facebook, exemplificada na figura 4. Essa representação semântica foi implementada como parte BROAD-RSI e é gerada automaticamente como consequência do seu uso.

Todos os interesses identificados e aqueles inferidos na etapa de enriquecimento de perfil são relacionados através

de propriedades semânticas, de tal forma que quando um usuário expressa interesse em alguma área é feita também uma consulta na base do BROAD-RSI para buscar interesses que em outro momento tenham sido relacionados e possam, então, enriquecer os interesses do usuário.

A geração e armazenamento desta base semântica são feitas usando Jena, Sesame e a linguagem SPARQL para consultas.

Ainda usando o exemplo do interesse do usuário X por DBPEDIA, a etapa de representação semântica insere na base semântica do BROAD-RSI, através de propriedades semânticas, que o tema “DBPEDIA” está relacionado aos temas “Semantic Web”, “Wikipedia”, “Query Language”, “RDF”, “Linked Data”. Dessa forma, outros usuários que façam uso do BROAD-RSI futuramente e demonstrem, por exemplo, interesse no tema “Linked Data” terão o seu interesse enriquecido com o tema “DBPEDIA”, mesmo que esse não esteja explicitamente identificado. Esse enriquecimento é possível realizando uma busca semântica na base gerada pelo próprio uso do BROAD-RSI.

Para a representação semântica dos recursos educacionais foi usada a ontologia criada para o padrão OBAA, que [24] engloba também o padrão LOM [25].

Figura 4—Exemplos de Ontologia de Áreas de Interesses Educacionais

3.1.4 Recomendação

O módulo de recomendação é responsável por realizar buscas no repositório de recursos educacionais, estabelecendo a relação entre o perfil do usuário e as características dos recursos educacionais armazenados nessas fontes, considerando as informações extraídas das redes sociais, conforme mostrado na tabela 2.

Nesse módulo é definida também a prioridade de recomendação a partir da aderência dos recursos educacionais ao perfil e ao contexto do usuário. Além da recomendação de recursos educacionais, é responsabilidade deste módulo recomendar pessoas que possuem interesses semelhantes aos do usuário.

O primeiro passo da recomendação é buscar recursos educacionais relacionados aos interesses capturados do usuário e aos interesses inferidos a partir da etapa de enriquecimento semântico. A busca é feita nos metadados General.Title, General.Keyword e General.Description. Ela retorna um conjunto de recursos que, então, é processado visando determinar a prioridade das recomendações, ou seja, quais dos recursos são mais ou menos aderentes ao perfil do usuário.

Considerando a dinamicidade das interações nas redes sociais, foi adotado no processo de recomendação uma priorização por interesses recentes, já que acredita-se que tão logo o usuário tenha expressado interesse por algum tema, seja mais relevante lhe enviar recomendações a respeito deste tema. Sendo assim, os interesses mais recentemente manifestados através das redes sociais são considerados

mais relevantes no envio de recomendações.

A diversificação das recomendações enviadas é outra questão que é tratada no processo de recomendação. É desejável, nesta proposta, que o usuário receba recomendações sobre os diversos assuntos pelos quais ele tem interesse, não ficando a recomendação restrita a um ou outro interesse.

Para aumentar a eficiência e precisão da busca na base de metadados, é usada uma busca por índice usando o Lucene⁶. O Lucene é uma biblioteca *Open Source* para indexação e consulta de textos. Essa biblioteca trabalha em duas etapas: (1) indexação, onde os dados originais são indexados gerando uma estrutura de dados relacionados para pesquisa baseada em palavras-chave; (2) pesquisa, que é feita no índice criado, por palavras digitadas pelo usuário, gerando resultados ordenados pela similaridade do texto com a consulta. A similaridade do documento é usada na priorização da recomendação.

A geração do índice no Lucene foi feita da seguinte forma: para cada registro do banco de dados, que contém os valores dos metadados, é gerado um *Document Lucene* distinto, composto por cinco campos (*Fields*): identificador, título, descrição, palavras-chave e um campo (*Field*) não indexado que armazena o endereço de acesso ao documento.

Após a filtragem é feita a etapa de determinação da aderência do objeto de aprendizagem ao perfil e contexto do usuário, onde é verificado se as características desses objetos são aderentes aos itens de perfil e de contexto do usuário. Os valores V (verdadeiro) e F (falso) são atribuídos a partir da acareação entre a característica do usuário

⁶<http://lucene.apache.org/core/>

e o valor do metadado do objeto de aprendizagem. Por exemplo, se o objeto de aprendizagem é escrito em Português e o usuário possui domínio do idioma Português será atribuído valor V para a relação estabelecida e, assim sucessivamente para os outros metadados.

Metadados LOM	Informações extraídas
General.Title	Interesses identificados e interesses inferidos
General.Description	
General.Keyword	
General.Language	Idiomas falados
Technical.Format	Preferência por mídias(Visuais/Verbais)
Educational.Context	Grau de Formação
Educational.TypicalAgeRange	Idade

Tabela 2 - Relação entre metadados e informações extraídas

Depois de calculada a aderência dos recursos educacionais é aplicado o grau de relevância do documento a partir

do índice de similaridade resultante da etapa de filtragem usando o Lucene. Por fim, os recursos são ordenados de acordo com o seu grau de aderência e relevância do documento em relação à busca, respectivamente.

Após realizadas todas as etapas das abordagens apresentadas, o resultado é um conjunto de recursos educacionais (objetos de aprendizagem, links, vídeos, textos) aderentes e priorizados de acordo com o perfil e o contexto do usuário. As recomendações são armazenadas em uma base de dados semântica, onde poderão ser acessadas futuramente para melhoria da própria abordagem de recomendação usada pelo BROAD-RSI.

A entrega da recomendação pode ser feita ao usuário de duas formas: (1) através da solicitação do usuário dentro do aplicativo do BROAD-RSI e (2) através do envio automático da recomendação na rede social do usuário.

Uma visão geral do processo de recomendação baseado em repositórios de objetos de aprendizagem é mostrada na figura 5.

Figura 5 - Visão geral do processo de recomendação do BROAD-RSI

3.1.4.1 Recomendação de Pessoas

O módulo de recomendação também é responsável pela recomendação de pessoas. Ela consiste em sugerir a um determinado usuário outras pessoas que possuem interesses similares aos seus. Para isso, são considerados todos os interesses do usuário para o cálculo de similaridade, sendo que quanto mais interesses comuns os usuários têm, mais similares eles são. Por exemplo, caso o usuário A peça uma recomendação de pessoas que possuem interesse na área de Gerenciamento de Projetos, serão apresentados a ele outras pessoas que possuem tal interesse e, para cada uma dessas pessoas, será mostrado um valor numérico que determina o grau de similaridade com o usuário. Para este cálculo são considerados todos os interesses comuns, conforme mostra a equação 1: seja A o usuário que solicitou a

recomendação e B o usuário recomendado, temos que:

$$Graude\text{similaridade} = \frac{N^{\circ} de\text{interesses} comuns\ entre\ A\ e\ B}{N^{\circ} de\text{interesses} de\ A} \quad (1)$$

A rede de usuários com interesses comuns é formada considerando todos os usuários do BROAD-RSI, funcionando da seguinte forma: caso o usuário A queira ver a rede de interesses em torno do tema “Engenharia de Software” será feita uma busca semântica por todos os usuários com interesse nesse tema, que já são amigos dele na rede social Facebook. Logo após, são consultados os demais usuários com interesse em Engenharia de Software que não estão diretamente ligados ao usuário A. Tendo todos os usuários com o mesmo interesse reconhecidos, são identificadas as conexões entre eles de tal forma que seja possível mostrar ao usuário A como ele pode alcançar todas as pessoas que possuem interesse em Engenharia de

Software a partir dos seus contatos na rede social.

A implementação da recomendação de pessoas foi feita com o apoio de um conjunto de ferramentas de software chamadas Prefuse, que é uma arquitetura para a visualização de grafos e árvores desenvolvida na linguagem de programação Java. Ela possui um conjunto de recursos para modelagem de dados, visualização e interação.

Na Figura 6 pode ser vista a formação de uma Rede de Interesse em torno do tema Ciência da Computação. O usuário representado pelo rótulo (A) é aquele para o qual serão feitas as recomendações. Os nós representados pelo rótulo (B) são os amigos diretamente ligados a esse usuário, portanto, pessoas que possuem o mesmo interesse e já fazem parte do círculo de amizades dele. Já os nós representados pelo rótulo (C), são usuários que possuem o mesmo interesse e não estão ligados diretamente ao usuário. Através das conexões mostradas é possível indicar os amigos que possuem interesses comuns e, ainda, quais são os amigos que podem conectá-lo a outras pessoas, que ainda não fazem parte do seu círculo de amizade.

Figura 6 - Rede de pessoas com um determinado interesse comum

4 Metodologia de avaliação

Para permitir a avaliação da proposta do BROAD-RSI foi implementado um Protótipo, mostrado na figura 7. Através desse protótipo é possível o usuário fazer login, usando a sua conta do Facebook, dar permissão de uso às suas informações e, então, serem extraídas aquelas já citadas anteriormente. Para o desenvolvimento do protótipo,

foi utilizada a linguagem de programação Java e ainda as seguintes tecnologias: Tomcat, Jena, Sparql, API Graph através da biblioteca RestFB, AlchemyAPI e MySQL.

A avaliação desta proposta foi feita através de uma prova de conceito e um estudo de caso. A realização da prova de conceito se justifica pela necessidade de demonstrar a viabilidade técnica do modelo proposto, dos conceitos e das tecnologias envolvidas no projeto. Apesar de não apresentar o formalismo de um estudo experimental, a utilização da prova de conceito neste domínio contribuiu também para a avaliação preliminar da proposta, o ajuste do instrumento de avaliação, a amenização de dificuldades técnicas pontuais e o refinamento dos passos que foram, posteriormente, aplicados em um estudo de caso.

Na prova de conceito dois usuários foram convidados a avaliar as recomendações geradas para eles, após as etapas de extração de informação e enriquecimento do perfil.

A partir da avaliação feita pelos dois usuários foi possível perceber que a identificação de interesses educacionais obteve um bom percentual de aceitação, atingindo 75% de aceitação para o primeiro usuário e 64% para o segundo, mostrando a viabilidade de extração das informações educacionais a partir das interações dos usuários no Facebook.

Foi possível perceber também que o uso dessas informações para realizar recomendações educacionais aos usuários foi bem aceita, gerando recomendações relevantes para os usuários. O primeiro usuário indicou aprovar 73,9% das recomendações oferecidas, enquanto o segundo usuário aceitou 77% delas.

A primeira etapa da avaliação foi importante para verificar a viabilidade técnica da proposta e dar os primeiros indícios acerca de sua aceitação. Através da sua realização foi formulado o questionário usado como instrumento de avaliação no estudo de caso. A partir do uso do protótipo, durante a realização da prova de conceito, foi possível ajustar alguns parâmetros para a avaliação, solucionar problemas técnicos e perceber algumas questões que poderiam ser melhor avaliadas em um estudo de caso.

Através da realização da prova de conceito percebeu-se, ainda, a necessidade de adoção de uma métrica para avaliar o sistema de recomendação de forma que tal avaliação pudesse ser futuramente comparada a outros sistemas de recomendação.

Assim sendo, a realização da prova de conceito motivou o planejamento e execução de um estudo de caso visando aumentar a capacidade de observação e a confiança na viabilidade e eficácia da proposta do BROAD-RSI.

BROAD-RSI
Recomendação Educacional baseada em interações sociais

Objetivo: O BROAD-RSI é ajudar a encontrar recursos educacionais apropriados aos seus interesses e às suas preferências, dando-lhe uma recomendação.

Facebook
Você está conectado com 12 outros friends via BROAD-RSI.
Ver as recomendações que temos para você

Característica	Valor	Avaliação
Nome	[Redacted]	
Email	[Redacted]	
Idade	31	
Cidade	[Redacted]	
Gênero	[Redacted]	
Dia de maior acesso		
Horário de maior acesso	11	
Nível Educacional	Pós Graduação	
Mãe	nek-lap	
Idioma	Portuguese	
Idioma	Espanhol	

Tema	Links Externos	Facebook	Avaliação
Data_mining	http://www.analytics.com	[Facebook icon]	[Recommend] [Compare]
Data_mining	http://csl.cmu.edu/courses/2013/09/01/reading/reading/lecture_10/lec10.pdf	[Facebook icon]	[Recommend] [Compare]
Data_mining	http://www.analytics.com	[Facebook icon]	[Recommend] [Compare]
Data_mining	http://www.analytics.com	[Facebook icon]	[Recommend] [Compare]
Data_mining	http://www.analytics.com	[Facebook icon]	[Recommend] [Compare]
Data_mining	http://www.analytics.com	[Facebook icon]	[Recommend] [Compare]
Collaborative_filtering	http://www.dss.org/	[Facebook icon]	[Recommend] [Compare]
Collaborative_filtering	http://pandas.ameba.net.edu/projects.html	[Facebook icon]	[Recommend] [Compare]

Interesses Educacionais

Categoria	Interesses Educacionais	Repositório	Recomendar	Personas	Interesses	Avaliação
Education	Data_mining	[Icon]	[Icon]	[Icon]	[Icon]	
Education	E-learning	[Icon]	[Icon]	[Icon]	[Icon]	
Education	Computer_science	[Icon]	[Icon]	[Icon]	[Icon]	

Interesses Diversos

Categoria	Interesses Diversos	Repositório	Recomendar	Personas	Interesses	Avaliação
Community	DATA_ANALYTICS_AND_LS	[Icon]	[Icon]	[Icon]	[Icon]	
Interest	DATA_MINING	[Icon]	[Icon]	[Icon]	[Icon]	
Education	DEVMEDIA	[Icon]	[Icon]	[Icon]	[Icon]	
Interest	E-LEARNING	[Icon]	[Icon]	[Icon]	[Icon]	

Interesses Gerais

Interesses Gerais	/science/computer_science/artificial_intelligence
Interesses Gerais	/technology_and_computing
Interesses Gerais	/technology_and_computing/software

Figura 7 - Tela do protótipo do BROAD-RSI.

4.1 Estudo de caso

De acordo com [26], a vantagem de um estudo de caso é que ele é mais fácil de planejar e mais realista, mas a desvantagem é que os resultados dificilmente podem ser generalizados e são mais difíceis de interpretar.

No caso desta pesquisa a melhor forma de avaliação é um estudo de caso por se tratar de uma pesquisa empírica realizada em um contexto real, que deseja investigar a aceitação por parte dos usuários do processo de recomendação proposto através do BROAD-RSI.

O estudo de caso conduzido nesta pesquisa tem caráter descritivo. A pesquisa descritiva foi empregada na observação dos seguintes fenômenos:

- Aceitação do usuário em relação às características de perfil e de contexto identificadas através de suas redes sociais;
- Aceitação dos usuários em relação às recomendações educacionais feitas de acordo com as características de perfil e de contexto extraídas.

4.1.1 Objetivos do estudo de caso

Os objetivos deste estudo de caso foram definidos de acordo com a abordagem Goal/Question/Metric (GQM) [27]. Este estudo de caso possui dois objetivos principais relacionados à extração das informações a partir da interação dos usuários em suas redes sociais e às recomendações oferecidas ao usuário a partir das suas características de perfil e de contexto, a citar:

- Analisar a extração de características de perfil e de contexto do usuário para o propósito de avaliar*

a extração de informações sobre o perfil e os interesses educacionais do usuário **com relação à precisão da extração do ponto de vista do usuário no contexto** de usuários da rede social Facebook, que possuem interesses relacionados à área de Computação e expressam tais interesses através das suas interações nesta rede social.

- Analisar as recomendações educacionais através do BROAD-RSI para o propósito de avaliar as recomendações **com relação à relevância das recomendações do ponto de vista do usuário no contexto** de recomendações geradas a partir do BROAD-RSI para usuários da rede social Facebook, que possuem interesses relacionados à área de Computação e expressam tais interesses através das suas interações na rede social.*

4.1.2 Questões e métricas

As questões que se deseja investigar com este estudo de caso são:

- **Questão 1:** Se extrairmos dados, gerados espontaneamente, a partir da interação dos usuários em redes sociais, então será possível identificar características de perfil e de contexto do usuário? (Q1)
- **Questão 2:** Se usarmos as características de perfil e de contexto extraídas da interação do usuário em redes sociais, então será possível realizar recomendações educacionais relevantes a esses usuários? (Q2)

Em [28] é apresentada uma discussão sobre os métodos de avaliação de recomendações, sendo eles resumidos em

três principais métricas mais comumente utilizadas em Sistemas de Recomendação: *Precision*, *Recall*, *False Positive Rate*.

No caso da avaliação feita neste trabalho foi usada somente a métrica *Precision*, porque as outras métricas dependem de variáveis que não são possíveis de serem calculadas na proposta do BROAD-RSI.

Essa métrica não representa um valor percentual. Segundo [28], quanto mais próximo de 1 o resultado da métrica *Precision*, melhor o Sistema de Recomendação.

Segundo [29], a métrica *Precision* também é usada com frequência para avaliar Sistemas de Extração de Informação, sendo definida como a quantidade de informações corretamente extraídas sobre todas as informações extraídas. Sendo assim, a mesma métrica será usada para avaliar a extração de características de perfil, os interesses extraídos e os interesses descobertos pela etapa de enriquecimento.

4.1.3 Seleção dos indivíduos

O estudo de caso foi realizado com 14 participantes, estando eles distribuídos entre estudantes do mestrado em Ciência da Computação da Universidade Federal de Juiz de Fora e profissionais da área de Computação (analistas de sistema e professores). Um requisito para a seleção dos indivíduos foi possuir uma conta na rede social Facebook e expressar interesses através das suas interações nessa rede social.

Vale ressaltar que a escolha dos indivíduos levou em consideração a área da Computação, por que o repositório de recursos educacionais usado como fonte para as recomendações do protótipo contém catalogados recursos específicos para essa área.

4.1.4 Coleta dos dados

Os dados usados para a realização desta pesquisa foram coletados usando método direto, através de entrevistas semiestruturadas [26]. Perguntas fechadas foram usadas na avaliação feita pelos usuários para cada item de perfil, cada interesse extraído e inferido, para cada recomendação gerada pelo BROAD-RSI e, ainda, após o uso do BROAD-RSI para verificar alguns pontos da sua experiência.

Para realizar a coleta dos dados foi implementado no protótipo do BROAD-RSI, uma funcionalidade onde os usuários podem avaliar as características de perfil, os interesses e as recomendações.

4.1.5 Resultados

Os participantes do estudo de caso foram convidados a avaliar os itens de perfil extraídos através das suas interações nas redes sociais. Os itens avaliados foram idiomas falados, preferências por mídia, grau de formação, dia e

horário de maior acesso. Busca-se através das análises feitas a seguir obter indícios a respeito da seguinte questão apresentada neste trabalho: se extrairmos dados, gerados espontaneamente, a partir da interação dos usuários em redes sociais, então será possível identificar características de perfil e de contexto do usuário? (Q1). Aplicando a métrica *Precision* para cada um dos elementos tem-se os valores mostrados na tabela 3.

	Idiomas	Preferência por mídias	Grau de formação	Dia de maior acesso	Horário de maior acesso
Concordo Plenamente	1,00	0,25	1,00	0,31	0,54
Concordo Parcialmente	0,00	0,50	0,00	0,46	0,08
Não concordo, nem discordo	0,00	0,00	0,00	0,15	0,23
Discordo Parcialmente	0,00	0,17	0,00	0,00	0,08
Discordo Plenamente	0,00	0,08	0,00	0,08	0,08

Tabela 3 - Avaliação dos usuários sobre as características de perfil e de contexto extraídos

Os participantes do estudo de caso avaliaram também cada interesse educacional extraído, respondendo à seguinte questão: *este tema de interesse identificado através das suas interações na rede social Facebook representa de fato um interesse educacional para você?* Foi explicado aos participantes que interesse educacional deve ser interpretado como qualquer assunto para o qual ele tem interesse em obter conhecimento, em aprender algo.

Os 14 participantes avaliaram 255 interesses extraídos, ressaltando que há repetição de interesses somente para usuários distintos. Na tabela 5 são mostrados os resultados das avaliações feitas por eles, já com aplicação da métrica *Precision*.

Os participantes do estudo de caso também foram convidados a avaliar cada um dos interesses enriquecidos através de técnicas de Extração de Informação, totalizando 295 temas, respondendo à seguinte questão: *os interesses descobertos representam um dos seus reais interesses educacionais?*

Vale esclarecer que os interesses enriquecidos são distintos dos interesses extraídos, já que uma vez que o interesse tenha sido extraído, o mesmo termo ou tema não aparece na lista de interesses enriquecidos.

Na tabela 4 são apresentados os resultados da avaliação dos usuários, seguindo a métrica *Precision* para os interesses extraídos e para os interesses descobertos.

	Interesses extraídos	Interesses descobertos
Concordo Plenamente	0,47	0,50
Concordo Parcialmente	0,22	0,16
Não concordo, nem discordo	0,10	0,16
Discordo Parcialmente	0,04	0,06
Discordo Plenamente	0,18	0,13

Tabela 4 - Avaliação dos usuários sobre os interesses extraídos e enriquecidos.

Analisando os resultados obtidos com a realização do estudo de caso, foi possível perceber que a extração das características de perfil dos usuários foi bem avaliada pelos participantes, sendo que o horário de maior acesso foi a característica que obteve menor aceitação, atingindo índice de 0,62 de aceitação (parcial + total). Essa característica apresentou índice de 0,23 de neutralidade por parte dos usuários, sendo percebido durante o estudo de caso que boa parte dos usuários não conseguiu analisar esse item por não terem clara percepção do horário que interagem no Facebook.

Em relação aos interesses extraídos também houve uma boa aceitação chegando a atingir o índice de 0,69 contra 0,22 de rejeição, sendo que quanto mais próximo de 1 o resultado da métrica *Precision*, melhor o sistema de recomendação. Esse dado traz indícios de que a seleção das categorias das páginas foi satisfatória para filtrar os interesses educacionais dos usuários.

Os interesses descobertos pela técnica de extração de informação também foram bem avaliados pelos usuários atingindo o índice de 0,66.

Apesar da boa aceitação observada na prova de conceito e corroborada no estudo de caso em relação aos interesses extraídos e enriquecidos, foi considerada no BROAD-RSI a possibilidade de confirmação dos interesses extraídos e enriquecidos, pois deseja-se obter um conjunto de interesses que, de fato, represente interesses do usuário, descartando assim recomendações sobre assuntos que não sejam do seu interesse e, consequentemente, minimizando um possível desinteresse do usuário em relação às recomendações e ao uso do BROAD-RSI.

De forma geral, foi possível perceber que existem informações relevantes nas redes sociais que podem auxiliar na tarefa de identificar características de perfil e de contexto dos usuários, reforçando os indícios de que a primeira questão pode ser respondida com uma afirmação positiva.

A avaliação das recomendações foi feita solicitando aos participantes do estudo de caso que avaliassem cada

uma das recomendações em relação à sua relevância, respondendo à seguinte pergunta: *considerando que você está buscando recursos educacionais a respeito deste tema, esta recomendação é relevante para você?* Busca-se através das análises feitas a seguir obter indícios a respeito da seguinte questão apresentada neste trabalho: se usarmos as características de perfil e de contexto extraídas da interação do usuário em redes sociais, então será possível realizar recomendações educacionais relevantes a esses usuários? (Q2).

O repositório local de recursos educacionais usado pelo BROAD-RSI tinha, no momento da avaliação, 60 recursos educacionais relacionados às áreas de engenharia de software, redes de computadores, computação gráfica e banco de dados. Dentre os recursos educacionais podemos citar, videoaulas, apostilas, artigos científicos, listas de exercícios, apresentações e palestras.

O resultado das avaliações das recomendações, aplicando a métrica *Precision*, é apresentada na tabela 5. Foram analisadas 84 recomendações.

	Resultado da avaliação
Concordo Plenamente	0,39
Concordo Parcialmente	0,19
Não concordo, nem discordo	0,14
Discordo Parcialmente	0,08
Discordo Plenamente	0,19

Tabela 5 - Resultados da avaliação das recomendações

As recomendações oferecidas aos usuários obtiveram índice de 0,58 de aceitação contra 0,27 de rejeição.

Acredita-se que um dos fatores que poderiam contribuir para o aumento do índice de aceitação das recomendações do repositório é a diversificação e aumento do número de recursos cadastrados no repositório para atender os diferentes interesses extraídos e enriquecidos.

Considerando os índices de aceitação, há indícios de que a segunda questão pode ser respondida também com uma afirmação positiva, já que a partir das características de perfil e de contexto extraídas da interação do usuário em redes sociais foram geradas recomendações educacionais a esses usuários e eles avaliaram a maior parte delas como relevantes.

4.1.6 Ameaças a validade

As ameaças à validade deste estudo serão tratadas em relação a: (1) validade do constructo, que é a confiança que se tem de que as medidas que estão sendo estudadas realmente representam aquilo que o pesquisador tem em mente e o que é investigado de acordo com as questões de pesquisa; (2) validade interna, que é a confiança que se tem

de que o efeito observado é realmente devido à manipulação feita, e não a outros fatores; (3) validade externa que é a confiança que se tem que o efeito observável é generalizável [26].

Entende-se que uma ameaça em relação ao constructo é aplicação do questionário tendo uma situação hipotética, onde o usuário foi estimulado a analisar as recomendações supondo que ele estava, naquele momento, necessitando de uma recomendação para aquele tema. Sabe-se que uma análise mais minuciosa precisaria ser feita para reconhecer a utilidade de um recurso educacional para o usuário ao longo do tempo, uma vez que fatores externos, tais como o uso do recurso, a atividade que o usuário está realizando no momento da recomendação, a qualidade do recurso, entre outros, podem exercer influência no quão relevante uma recomendação é para o usuário.

Em relação à validade interna deste estudo, uma ameaça é a influência que a qualidade dos recursos educacionais exerce sobre o interesse do usuário, ou seja, é possível que a qualidade do material desperte maior ou menor interesse sobre o usuário, fazendo com que ele avalie a recomendação sob essa ótica. Outra ameaça é a catalogação do recurso educacional, já que se a catalogação não tiver sido feita de modo eficiente pode prejudicar a recuperação do recurso ou o cálculo de aderência ao perfil do usuário. Em relação a essas duas ameaças foi adotado o critério de usar apenas recursos produzidos e catalogados por professores com domínio nas suas áreas de atuação, sendo que tais materiais já são usados com frequência em disciplinas do curso de Ciência da Computação.

Em relação à validade externa, as principais ameaças que poderiam impedir a generalização dos resultados obtidos para qualquer usuário da rede social Facebook são a restrição da área dos recursos educacionais catalogados, conforme já mencionado anteriormente e a possibilidade de que os usuários não tenham na sua rede social nenhum tipo de interesse educacional que possa ser capturado, ou seja, que ele não curta nenhuma página relacionada a assuntos do seu interesse ou não participe de nenhum grupo com foco em assuntos do seu interesse ou ainda que não tenha nenhuma conexão com páginas a respeito de interesses educacionais. Em relação à primeira ameaça poderiam ser usadas outras fontes de recursos educacionais para buscar atender todos os diferentes tipos de interesses e também foi feita uma restrição do público alvo, buscando somente pessoas com interesse em áreas da Computação. Porém, ainda deve-se considerar a necessidade de uma ampliação das áreas de interesse dos recursos catalogados no repositório do BROAD ou o uso de outros repositórios para gerar recomendações. Em relação à segunda ameaça, o contexto de aplicação do estudo de caso já foi limitado a participantes que expressam seus interesses educacionais através das suas interações na rede social Facebook, já que esse é um requisito básico para que o modelo proposto seja

eficaz.

5 Considerações finais

O sistema de recomendação apresentado neste trabalho traz aspectos inovadores no sentido de extrair e considerar informações geradas espontaneamente nas redes sociais, mais especificamente no Facebook, bem como gerar recomendações individuais, especificamente educacionais, coerentes com os interesses e as preferências identificados a partir da análise das informações oriundas desse sistema.

O desenvolvimento da arquitetura do BROAD-RSI, do protótipo, da prova de conceito e do estudo de caso mostraram a viabilidade técnica e a aceitação dos usuários em relação à extração das informações, seu enriquecimento e das recomendações feitas.

A geração automática de recomendações, considerando as características individuais dos usuários pode auxiliá-los na tarefa de encontrar recursos educacionais relevantes e compatíveis com o seu perfil, podendo assim despertar maior interesse a respeito de um determinado conteúdo.

O envio da recomendação diretamente na rede social do usuário contribui para a facilidade de acesso aos recursos educacionais e permite que o usuário use todos os recursos de interação fornecidos por esses ambientes.

Outra contribuição é a possibilidade da recomendação de pessoas, possibilitando a formação de uma rede de interesses em torno de um determinado tema. O cálculo de similaridade entre os interesses das pessoas aproxima de um determinado usuário todas as pessoas com as quais ele possui interesses comuns, informando a ele quais dentre elas têm maior similaridade, indicando assim bons parceiros para estudo e pesquisa, por exemplo.

A proposta apresentada neste trabalho é bem ampla, contemplando diferentes áreas de conhecimento e, por isso, entende-se que a mesma pode ser refinada e melhorada em cada uma das etapas, buscando aumentar os índices de aceitação tanto da extração das características de perfil e contexto, quanto das recomendações.

A API usada neste trabalho como apoio na extração de informações, a *AlchemyAPI*, apresentou bons resultados, porém possui limitações, principalmente em relação aos idiomas dos textos analisados. Alguns métodos não estão disponíveis para todos os idiomas, inclusive para o Português. Por isso, observou-se que os resultados apresentados são mais satisfatórios quando os textos estão no idioma Inglês.

O número de recursos educacionais disponíveis no repositório do BROAD-RSI é uma limitação para a geração de recomendações relevantes para os usuários. Acredita-se que se o repositório dispusesse de mais recursos e se esses

abordassem um conjunto maior de temas, as recomendações seriam melhor avaliadas pelos usuários.

O presente trabalho se limitou a buscar recursos educacionais que contivessem termos exatos aos interesses extraídos e enriquecidos, não tendo explorado todas as possibilidades das técnicas de recuperação de informação [30], como por exemplo buscando por partes do nome ou a combinação entre diferentes interesses.

5.1 Trabalhos futuros

Em relação à Extração de Informação seria importante ampliar o conjunto de informações usadas na extração, seja através da inclusão de postagens individuais em redes sociais, postagens em grupo e compartilhamentos; seja através da inclusão, junto com as redes sociais, de outros sistemas que possam complementar essa extração, tais como ambientes virtuais de aprendizagem, sistemas acadêmicos e sistemas de apoio a MOOCS. No caso de inclusão de outros sistemas seria muito interessante também avaliar a qualidade do perfil extraído nos diferentes sistemas.

A ampliação da proposta para outras redes sociais também é uma sugestão de trabalhos futuros. A proposta foi projetada para sofrer poucas adaptações em caso de expansões para outras redes, sendo necessário, no entanto, considerar as peculiaridades das outras redes sociais, como por exemplo, as permissões e as formas de extração de informações de cada ambiente.

Por fim, a própria abordagem de recomendação usada neste trabalho pode ser refinada e melhorada, considerando, por exemplo, as recomendações feitas dentro da rede de pessoas que possuem interesses comuns, as avaliações das recomendações realizadas pelos usuários do BROAD-RSI, a interação do usuário dentro do BROAD-RSI, através de logs ou mineração de dados, por exemplo.

A categorização eficaz dos grupos nos quais o usuário participa nas redes sociais e a consideração dos interesses extraídos deles na recomendação também é um possível trabalho futuro, bem como a adaptação da proposta para a recomendação para grupos.

Agradecimentos: à FAPEMIG, CNPq, UFJF e Capes pelo apoio financeiro ao projeto.

Referências

- [1] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Model. User-adapt. Interact.*, vol. 12, no. 4, pp. 331–370, 2002.
- [2] R. Junco, "The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement," *Comput. Educ.*, vol. 58, no. 1, pp. 162–171, Jan. 2012.
- [3] M. Vanoozzi and L. Bridgestork, "Students' Online Usage Global Market Trends Report," 2013.
- [4] T. Nery, F. Campos, R. Braga, N. Santos, E. Mattos, K. Engineering, and C. S. Program, "BROAD Project: Semantic Search and Application of Learning Objects," *IEEE Technol. Eng. Educ.*, vol. 7, no. 3, 2012.
- [5] F. Campos, R. Braga, T. Nery, and N. Santos, "Rede de Ontologias: apoio semântico a linha de produtos de objetos de aprendizagem," *An. do Simpósio Bras. Informática na Educ.*, vol. 23, no. Sbie, pp. 26–30, 2012.
- [6] C. K. Pereira, F. Campos, V. Ströele, M. N. David, and R. Braga, "Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais," *XXV SBIE 2014*, vol. 25, no. Cbie, pp. 506–515, 2014.
- [7] P. Rezende, F. Campos, R. Braga, and J. M. N. David, "Broad-rs: uma arquitetura para recomendação de objetos de aprendizagem sensível ao contexto usando agentes," *An. do X Congr. Bras. Ensino Super. a Distância*, pp. 11–13, 2013.
- [8] E. Fritzen, S. W. M. Siqueira, and L. C. V. de Andrade, "Recuperação Contextual de Informação na Web para Apoiar Aprendizagem Colaborativa em Redes Sociais," *XXIII An. do Simpósio Bras. Informática na Educ.*, vol. 23, no. Simpósio Brasileiro de Informática na Educação, pp. 26–30, 2012.
- [9] P. Dwivedi and K. K. Bharadwaj, "E-Learning Recommender System for Learners in Online Social Networks Through Association Retrieval," *Proc. CUBE Int. Inf. Technol. Conf. - CUBE '12*, p. 676, 2012.
- [10] G. Fernandes and S. Siqueira, "Building users profiles dynamically through analysis of messages in collaborative systems," pp. 75–82, 2013.

- [11] M. F. R. Casagrande, G. Kozima, and R. Willrich, "Técnica de Recomendação Baseada em Metadados para Repositórios Digitais Voltados ao Ensino," *XXIV Simpósio Bras. Informática na Educ. (SBIE 2013)*, vol. 24, pp. 677–686, Nov. 2013.
- [12] L. G. A. Ferreira, J. L. V. Barbosa, and J. C. Gluz, "Um Modelo de Recomendação Ubíqua de Conteúdo para Grupos de Aprendizes," *XXIV Simpósio Bras. Informática na Educ. (SBIE 2013)*, vol. 24, no. Cbie, pp. 697–706, Nov. 2013.
- [13] C. K. Pereira, F. Campos, V. Ströele, J. M. N. David, and R. Braga, "Extracting User Profile and Context from Social Networks and Virtual Learning Environments to Recommend Educational Resources," *Proc. 13th Int. Conf. WWW/Internet 2014*, pp. 171–178, 2014.
- [14] C. K. Pereira, F. Campos, R. Braga, V. Ströele, and J. M. N. David, "Elementos de Contexto em Sistemas de Recomendação no Domínio Educacional: um Mapeamento Sistemático," *XIX Conferência Int. sobre Informática na Educ.*, vol. 19, p. 10, 2014.
- [15] K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic, S. Member, and E. Duval, "Context-Aware Recommender Systems for Learning: A Survey and Future Challenges," vol. 5, no. 4, pp. 318–335, 2012.
- [16] Oracle, "The Java EE 6 Tutorial," Redwood City, 2013.
- [17] M. Allen, "RestFB - A Lightweight Java Facebook Graph API and Old REST API Client," *Transmogrify, LLC*. [Online]. Available: <http://restfb.com/>.
- [18] G. G. Chowdhury, "Natural language processing," *Annu. Rev. Inf. Sci. Technol.*, vol. 37, pp. 51–89, 2003.
- [19] F. Sebastiani, "Machine Learning in Automated Text Categorization," *ACM Comput. Surv.*, vol. 34, no. 1, pp. 1–47, 2002.
- [20] G. Fernandes, "Captura de perfis dinâmicos de usuários a partir da análise de mensagens em redes sociais," Universidade Federal do Estado do Rio de Janeiro, 2013.
- [21] T. Kovacic, "Evaluating Web Content Extraction Algorithms," UNIVERSITY OF LJUBLJANA, 2012.
- [22] D. Brickley and L. Miller, "FOAF Vocabulary Specification 0.99," 2014. [Online]. Available: http://xmlns.com/foaf/spec/#term_Document. [Accessed: 07-Jan-2015].
- [23] J. Breslin, S. Decker, A. Harth, and U. Bojars, "SIOC: an approach to connect web-based communities," *Int. J. Web Based Communities - IJWBC*, vol. 2, no. 2, pp. 133–142, 2006.
- [24] R. M. Vicari, J. C. Gluz, L. Passerino, E. Santos, T. Primo, L. Rossi, A. Bordignon, P. Behar, R. Filho, and V. Roesler, "The OBAA Proposal for Learning Objects Supported by Agents," *Syntax Semant.*, no. Aamas, pp. 9–12, 2010.
- [25] E. Duval and W. Hodgins, "A LOM Research Agenda," pp. 1–17, 2003.
- [26] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Springer Berlin Heidelberg, 2012.
- [27] V. R. Basili, "Software modeling and measurement: the Goal/Question/Metric paradigm," *Quality*, p. 24, 1992.
- [28] G. Shani and A. Gunawardana, "Evaluating recommendation systems," *Recomm. Syst. Handb.*, pp. 257–298, 2011.
- [29] A. C. Álvarez, "Extração de informação de artigos científicos: uma abordagem baseada em indução de regras de etiquetagem," USP - Universidade de São Paulo, 2007.
- [30] C. D. Manning and P. Raghavan, *An Introduction to Information Retrieval*. Cambridge University Press, 2009.