SWEETS: um Sistema de Recomendação de Especialistas aplicado a uma plataforma de Gestão de Conhecimento[a]

Edeilson M. Silva [1]

Ricardo A. Costa [1 2]

Lucas R. B. Schmitz [2]

Silvio R. L. Meira [1 2]

Resumo: As organizações, com o intuito de aumentarem o seu grau de competitividade no mercado, vêm a cada instante buscando novas formas de evoluir a produtividade e a qualidade dos produtos desenvolvidos, além da diminuição de custos. Para que tais objetivos possam ser alcançados é primordial explorar ao máximo o potencial de seus colaboradores e os possíveis relacionamentos que esses colaboradores têm uns com os outros, ou seja, encontrar e partilhar conhecimento tácito. Como o conhecimento tático está na mente das pessoas, é difícil de ser formalizado e documentado, por isso, o ideal seria identificar e recomendar a pessoa que detém o conhecimento.

Diante disso, o presente artigo apresenta o Sistema de Recomendação de Especialistas SWEETS e a sua implantação no ambiente a.m.i.g.o.s., uma plataforma de gestão de conhecimento baseada em conceitos voltados às redes sociais. O SWEETS foi desenvolvido em duas versões, 1.0 e 2.0. A versão 1.0, de forma pró-ativa, aproxima pessoas com especialidades em comum, ora pelos seus conhecimentos (perfil de escrita), ora pelos seus interesses (perfil de leitura). Já a versão 2.0 do SWEETS não atua de forma pró-ativa, ou seja, é necessário que haja a requisição de um usuário especialista em determinada área, e é baseada em *folksonomia* para extração de uma ontologia, fundamental para identificar as especialidades das pessoas de forma mais eficaz. Esta ontologia é refletida pela co-ocorrência das *tags* (conceitos) em relação aos itens (instâncias) e é independente de domínio, sendo a principal contribuição desse trabalho.

A implantação do SWEETS no a.m.i.g.o.s. visa trazer benefícios como: minimizar o problema de comunicação na corporação, prover um incentivo ao conhecimento social e partilhar conhecimento; proporcionando, assim, à empresa, a utilização mais eficaz dos conhecimentos de seus colaboradores.

---

[a] Este trabalho foi apoiado pelo Instituto Nacional de Ciências e Tecnologias para Engenharia de Software (INES – www.ines.org.br), financiado pelo CNPq e FACEPE.

[1] Centro de Informática, UFPE, Caixa Postal 7851 – Recife, PE - Brasil
{ems,rac,srlm@cin.ufpe.br}

[2] C.E.S.A.R - Centro de Estudos e Sistemas Avançados do Recife, Rua Bione, 220 – Bairro do Recife - 50.030-390 - Recife, PE - Brasil
{ricardo.araujo,lucas.schmitz,silvio@cesar.org.br}

Abstract: The organizations aim at increasing industrial competitiveness. In this context, they have been searching for new ways to improve their productivity, the quality of their products, and costs reduction. To achieve these goals, it is essential to explore the collaborators' potentials and the relationship among them, thus finding and sharing tacit knowledge. Since the tacit knowledge is embedded in the people's mind, it is therefore hard to be formalized and documented. This way identifying and recommending the person that retains the mentioned knowledge is the best option. In this context, this work presents the Specialist Recommendation System (SWEETS) and its application into the a.m.i.g.o.s environment, a knowledge management platform based on the social network concept. The SWEETS system has two versions, 1.0 and 2.0. The 1.0 version, puts people with specialties in common together proactively, either by their knowledge (according to the writing profile), or by their interest (according to the reading profile). On the other hand, the SWEETS second version does not act in a proactive way. Thereby, it is necessary to request for a specialist in a specific field, and it uses folksonomy to extract an ontology, which is essential to identify people's skills effectively. This ontology is reflected by the tags (concept) co-occurrence relating them to items (instances). In addition, such ontology is independent of domain, being the main contribution of this work. Applying the SWEETS system into the a.m.i.g.o.s. environment aims at bringing some benefits, such as: minimize the communication problem in the corporation, provide an encouragement of social knowledge and knowledge sharing; Therefore, a better usage of the collaborators knowledge may be expected.

1 Introdução

De acordo com estudo apurado pela IBM[b], os CEOs das organizações espalhados pelo mundo estão interessados em estabelecer um clima e cultura que ofereçam suporte que ajude as suas empresas a inovar, incluindo: desenvolvimento de novos produtos, serviços e mercados; a criação de novos modelos empresariais; e a melhoria de operações existentes. Para tanto, os CEOs têm que explorar ao máximo o potencial das suas empresas. Esse potencial envolve o conhecimento inerente aos seus funcionários (colaboradores) e os relacionamentos que esses funcionários têm uns com os outros, ou seja, a capacidade de encontrar e partilhar conhecimento tácito. Assim, prover rápido acesso a essa classificação de conhecimento é questão primordial para as empresas que estão constantemente preocupadas em evitar esforços duplicados e em inovar.

[b] IBM Global Business Services, Global CEO Study 2006, disponível em: <http://www.ibm.com/bcs/ceostudy>. Acesso em 1508/2008

De acordo com estimativa feita por Duhon [7], 50 a 90% do conhecimento corporativo está na mente dos seus empregados – conhecimento tácito, por isso, é difícil de ser formalizado e documentado [17]. Em decorrência de tal dificuldade, uma solução apropriada seria recomendar uma pessoa que detenha o conhecimento relacionado ao assunto.

Além disso, uma prática comum e bem conhecida entre os seres-humanos é a procura por informações entre pessoas de grupos que se relacionam. Isso porque as pessoas tendem a dar uma maior credibilidade às informações advindas das suas redes de contatos – colegas e amigos [2], [16], [19]. Segundo a análise que consta em [11], os colaboradores de uma organização obtém 50 a 75% de suas informações diretamente de outras pessoas.

Embora as redes pessoais possam atuar como um caminho para se obter respostas rápidas, às vezes elas não são suficientes para alcançar diretamente quem tem informações necessárias sobre uma determinada área. Assim, essas redes têm como característica o alcance limitado. As pessoas em uma rede pessoal podem atuar como intermediadores, facilitando assim o contato com pessoas ainda desconhecidas [8], e com isso, prover uma maior interatividade, comunicação e colaboração entre as pessoas.

Em uma ferramenta com contexto social, se alguém recebe uma determinada requisição de informação, é mais provável esta pessoa responder à requisição se esta vier de uma pessoa amiga ao invés de uma pessoa desconhecida. Assim, segundo Ehrlich [8], um sistema que identifique especialistas poderia refletir o contexto social em que as pessoas estão encaixadas.

Ainda, de acordo com Ehrlich [8], há algum tempo, pesquisadores da área argumentam que todo sistema de rede social deveria ser adicionado de tecnologias capazes de localizar especialistas, assim a procura por pessoas certas para sanar um determinado problema poderia ser bem mais eficiente.

Diante deste cenário, o presente artigo tem como objetivo apresentar um sistema de recomendação de especialistas e a sua implantação em uma plataforma de gestão de conhecimento a.m.i.g.o.s. (AMbiente para Integração de Grupos e Organizações Sociais), atualmente em uso no C.E.S.A.R. – Centro de Estudos e Sistemas Avançados do Recife [4]. Para atingir tal objetivo, este trabalho está organizado da seguinte forma: (seção 2) apresenta conceitos inerentes a Redes Sociais; (seção 3) apresenta particularidades da área de Gestão de Conhecimento; (seção 4) apresenta alguns sistemas de recomendação de especialistas disponíveis na literatura; (seção 5) apresenta o sistema de recomendação de especialista desenvolvido em ambas as versões, 1.0 e 2.0; (seção 6) apresenta o estudo de caso utilizado; (seção 7 e 8) apresenta os experimentos e resultados obtidos e, por fim, (seção 9) apresenta as conclusões.

2 Redes Sociais

A teoria de redes sociais aborda os relacionamentos sociais como nós e ligações entre estes nós. Cada nó representa um ator na rede social, e as ligações representam algum tipo de ligação existente entre estes atores. Existem diversas formas de ligação entre os atores, cada uma representando contextos distintos acerca da rede social [22].

Desde meados da década de 90, as redes sociais baseadas na Web vêm sendo desenvolvidas, apresentando um crescimento rápido tanto no número de redes quanto em seus escopos. Segundo Golbeck [12], estas redes podem ser vistas como grandes repositórios de dados que armazenam informações sobre cada um de seus usuários.

Em [12], Golbeck afirma ainda que uma comunidade online só pode ser considerada uma rede social na Web se: (1) ela é acessível na Web, sendo requerido para isto apenas um navegador; (2) usuários podem definir seus relacionamentos com os demais membros de forma explícita; (3) o sistema possui mecanismos nativos para a criação destas conexões; (4) os relacionamentos são visíveis e navegáveis.

Em [9], Erickson e Kellogg começaram a trabalhar no desenvolvimento de um ambiente multiusuário que permitiria a comunicação e colaboração em grupos, onde conhecimento comunitário poderia ser criado. Eles afirmavam que o uso de redes sociais poderia se tornar um mecanismo eficiente para compartilhar e disseminar o conhecimento individual, o que o credencia como uma abordagem interessante para iniciativas de gestão de conhecimento (GC). De acordo com Staab [21], estes ambientes podem reter informações relevantes sobre seus usuários, bem como produzidas por estes, provendo os primeiros passos para o gerenciamento e disseminação do conhecimento.

3 Gestão de Conhecimento

A gestão do conhecimento (GC) em uma empresa de software, de acordo com Choi e Lee [4], é uma oportunidade para criar uma linguagem comum compartilhada pelos desenvolvedores de software, de forma que eles possam interagir, trabalhar e compartilhar conhecimentos e experiências. Nos últimos tempos, a área de GC vem recebendo uma atenção especial por parte de organizações em busca de diferenciais competitivos [3].

Em [17], Nonaka caracteriza o conhecimento em dois tipos, o tácito e o explícito. O ultimo é composto basicamente pelo conhecimento que pode ser documentado e distribuído, já o conhecimento tácito reside na mente humana, sem seu comportamento e sua percepção, o que o torna difícil sua formalização e distribuição. GC pode ser visto como um processo para criação, coleta, transferência e aplicação do conhecimento.

Todo programa de GC precisa balancear o tipo de conhecimento no qual está focando. Baseado em seu foco, os programas podem ser classificados em um dos quatro estilos definidos em [4] e apresentados abaixo:

- Passivo: O conhecimento não é gerenciamento de forma sistemática. Existe pouco interesse em GC;
- Orientado a sistema: Enfatiza a codificação e reuso do conhecimento. Como consequência, aumenta a capacidade de codificação através do uso de TI, diminuindo então a complexidade no acesso e uso do conhecimento;
- Orientado a pessoas: Enfatiza a aquisição e compartilhamento do conhecimento tácito e a experiência interpessoal. O conhecimento é originado através de redes sociais informais e seu significado não pode ser simplesmente obtido de uma base de dados ou repositório;
- Dinâmico: Explora ambos os tipos de conhecimento de forma dinâmica, similar a uma organização de comunicação intensiva. Possuem forte dependência do conhecimento cultural.

A redução na perda do Capital Intelectual referente à saída de empregados das organizações, a redução do custo para o desenvolvimento de novos produtos, e o aumento da produtividade por tornar o conhecimento mais facilmente acessível a todos são alguns dos diversos benefícios da adoção de uma estratégia de GC. Em [4], Choi e Lee realizaram um estudo empírico com o intuito de validar os quatro estilos de GC propostos. O trabalho analisou o efeito e o custo de cada método em relação à performance organizacional. Foram aplicados questionários para 100 organizações coreanas. Os resultados deste trabalho indicaram que o estilo dinâmico, segundo os autores, é o mais efetivo, porém de longe o mais custoso dos quatro – apesar de não apresentarem valores concretos que justifiquem tal conclusão, enquanto o estilo passivo resultou numa performance significativamente mais baixa que os demais. Não foram identificadas diferenças significativas na performance organizacional das empresas que utilizam o estilo orientado a sistemas ou o orientado a pessoas.

4 Sistemas de Recomendação de Especialistas

Um Sistema de Recomendação de Especialistas é uma alternativa para prover acesso ao conhecimento implícito/tácito – conhecimento que está na mente das pessoas. Neste sentido, a literatura apresenta algumas iniciativas que buscam prover acesso ao conhecimento tácito das pessoas. Dentre elas, podem ser mencionadas *ReferralWeb* [13], ERS [23], TABUMA [20], ICARE [18] e o *SmallBlue* [14].

ReferralWeb [13] é um sistema de recomendação de especialistas que combina conceitos inerentes a Redes Sociais e Filtragem Colaborativa para prover recomendações personalizadas, oferecendo prioridade aos especialistas que estão mais próximos do usuário, ou seja, aqueles com distância social menor. Os relacionamentos existentes entre as pessoas foram extraídos a partir de *logs* de e-mails, pois acreditavam na hipótese de que os e-mails seriam uma rica fonte de extração de relacionamentos sociais (rede social). Porém, tal característica pode ser considerada um problema, pois levantam importantes preocupações de privacidade da informação.

O ERS *(Expert Recommendation System*) [23] utiliza métodos de recuperação de informação para retornar pessoas e/ou organizações com forte relevância para uma palavra-chave ou documento, ou seja, utiliza base de documentos para encontrar especialistas com relevância entre o tópico consultado e a pessoa. A análise dos documentos associados aos usuários como premissa básica para inferir se um usuário é ou não especialista em uma área pode ser um problema, pois a eficiência das recomendações estaria diretamente relacionada à quantidade e à qualidade desses documentos.

Da mesma forma que o ERS, o TABUMA [20] é um SRE que explora a capacidade dos documentos de texto para refletir os interesses de um usuário. Com isso, é utilizado um conjunto de documentos de textos que são associados com o trabalho do usuário para a geração do seu perfil. O fato de essa ferramenta poder receber como entrada dos usuários quaisquer tipos de documentos de textos, que indicam as habilidades e experiências dos usuários, a torna bastante flexível. Tal característica é um aspecto positivo. Contudo, essa liberdade ao usuário pode ser dada como negativa, caso esse usuário submeta documentos que não reflitam suas especialidades.

O ICARE [18] é um Sistema de Recomendação de Especialistas Sensível a Contexto e utiliza ontologia de domínio para realizar as recomendações. Assim, o ICARE, a partir da entrada de palavras-chaves pelos usuários, recomenda especialistas considerando informações tanto do usuário que faz a requisição, quanto dos especialistas que são recomendados. Dessa forma, promove recomendações personalizadas, pois estas mudam de acordo com o usuário e o instante de tempo em que a solicitação é feita. Ou seja, são recomendados os especialistas mais adequados a oferecer ajuda em um determinado instante de tempo. Dentre outras informações, são consideradas: a disponibilidade do especialista, o cargo que o especialista ocupa na organização, a distância social entre o usuário alvo e o especialista e a reputação do especialista entre o conjunto de pessoas que se relacionam com ele.

O *SmallBlue* [14] tem como objetivo encontrar especialistas, comunidades e redes sociais em grandes companhias, através de técnicas de mineração de dados, recuperação de informação e análises de redes sociais. A rede social no *SmallBlue* é extraída a partir de mensagens de e-mails (similar ao *ReferralWeb*) e de mensagens instantâneas, o que implica dizer que o usuário possui controle sobre o conteúdo utilizado. Conforme já mencionado, há

um problema de privacidade das pessoas em usar essas informações. Entretanto, o uso de informações desse gênero pode ser importante se os seguintes fatores forem considerados: o uso de e-mails ser uma atividade comum; novos e-mails serem gerados constantemente; e haver facilidade de uso, pois as pessoas já usam e-email no seu dia-a-dia, não havendo nenhum trabalho adicional requerido para o usuário, além de dar permissão para utilizar seus dados.

Conforme supracitado, um especialista pode ser identificado pelas mais variadas formas: autoria de artigos; conteúdos de e-mails e mensagens instantâneas, entre outros. Além disso, as recomendações desses especialistas podem possuir vantagens significativas se atuarem de forma personalizada, como são os casos dos sistemas *ReferralWeb, ICARE* e *SmallBlue*. Um exemplo de uma recomendação personalizada é a proximidade social do usuário para com o especialista, assim o usuário pode escolher um especialista que esteja mais acessível a ele.

Normalmente, os SREs utilizam técnicas de Recuperação da Informação aliadas a ontologias para identificar os especialistas de domínio. Essas ontologias, que representam um domínio em particular, são pré-definidas, ou seja, a identificação de especialistas é limitada a um domínio em restrito. Contrário a isso, a ferramenta SWEETS, na versão 2.0, apresentada na seção 5, e aplicada no a.m.i.g.o.s (*AMbiente para Integração de Grupos e Organizações Sociais*), identifica especialistas de domínios nos mais variados contextos, pois a ontologia utilizada na referida ferramenta surge à medida que as interações (colaboração) entre os usuários acontecem.

5 SWEETS

O SWEETS é um sistema de recomendação de especialistas disponível em duas versões (1.0 e 2.0). Peculiaridades inerentes a essas versões são apresentadas respectivamente nas seções 5.1 e 5.2.

5.1 SWEETS 1.0

SWEETS 1.0 é um sistema pró-ativo para Rede Social que recomenda pessoas que possuem conhecimento relacionando a um determinado assunto (um especialista). O SWEETS 1.0 utiliza filtragem baseada em conteúdo para prover as recomendações. Essas recomendações visam aproximar as pessoas que possuem interesse em comum e não são realizadas explicitamente a partir da entrada de um usuário.

À medida que um usuário interage com outros usuários, ou seja, produzindo conhecimento, o SWEETS 1.0 gera um perfil definido como *perfil de escrita*. Da mesma forma, os usuários podem ler conteúdos produzidos por outros usuários. Estes conteúdos

podem ser qualificados como relevantes ou não, a partir da atribuição de uma nota, que varia entre 1 e 5. Os conteúdos relevantes são aqueles classificados pelos usuários com nota maior ou igual a 4. Estes conteúdos, lidos e qualificados positivamente pelos usuários, geram um segundo perfil, definido como *perfil de leitura*.

Com o perfil de escrita é possível aproximar usuários com especialidades em comum, enquanto que com o perfil de leitura é possível aproximar pessoas que detenham conhecimento relacionado a conteúdos que o usuário esteja interessado. Ambas as formas de recomendações são realizadas a partir dos interesses do usuário, ora por conhecimento escrito, ora por conhecimento lido. Desta forma, não é preciso que o usuário forneça explicitamente o assunto ao qual deseja encontrar um especialista.

Para representar ambos os perfis de leitura e escrita, foi utilizado o Modelo Algébrico intitulado Modelo Espaço Vetorial (VSM). O VSM representa os termos/palavras mais relevantes associados a cada conteúdo do usuário como um vetor, sendo as freqüências/pesos, com que os termos ocorrem, as coordenadas deste vetor [1]. São considerados todos os termos relevantes após a exclusão de *stopwords*, ou seja, termos de categorias gramaticais como artigos, preposições e interjeições. De acordo com [1], estes pesos permitem o "casamento" parcial entre os perfis dos usuários, representados por seus conteúdos, e são usados para calcular o grau de similaridade entre os perfis. O grau de similaridade é calculado através da medida do co-seno do ângulo entre os vetores representados pelo VSM, o que o faz variar entre 0 e 1. Assim, quando maior a similaridade entre perfis, menor será o ângulo entre os vetores, o que implica em um maior valor do co-seno [1].

O SWEETS 1.0 recomenda apenas pessoas com interesses em comum, seja através da escrita ou da leitura além de não oferecer liberdade ao usuário para pesquisar especialistas por um assunto qualquer. Este problema, unido às falhas existentes em análise pura de textos planos motivou a evolução do SWEETS 1.0 para o SWEETS 2.0.

5.2 SWEETS 2.0

SWEETS 2.0 é um sistema de recomendação de especialistas não pró-ativo que pode ser implantado em qualquer ambiente computacional que possua informações associadas aos usuários e que utilize os conceitos inerentes a *folksnomia*.

Folksonomia, segundo Mika [15] é "uma renovação lingüística para a categorização colaborativa, a partir do uso livre de palavras-chaves". Ou seja, é um mecanismo de *tagging* social em que as pessoas colaboram para a sua criação. Esta colaboração é possível a partir de livres descrições de objetos compartilhados realizadas pelos usuários.

Para Mika [15], *folksonomia* aparece como uma boa alternativa para o surgimento de ontologias simples (*lightweight ontology*) e criação de metadados. Uma das vantagens do uso de *folksonomia* para a extração de simples ontologias é que esta pode abranger vários

domínios, diferentemente de uma ontologia de domínio que representa conhecimento relacionado a um domínio fechado. Assim, o uso de ontologias simples surgidas a partir de uma *folksonomia* pode ser aplicado em vários contextos.

Da mesma forma que em [15], para fazer modelos de redes de *folksonomia* em um nível abstrato, pode ser utilizado um sistema de grafo tripartido (três partes), em que o conjunto de vértices é definido em 3 (três) conjuntos disjuntos em que, cada um desses conjuntos correspondem, respectivamente, aos atores (usuários), conceitos (*tag*s) e as instâncias anotadas (por exemplo, documentos, sites, imagens, etc.). Dessa forma, um sistema de *tagging* social permite que usuários etiquetem objetos com conceito, criando assim associações ternárias entre usuário, o conceito e o objeto. Assim, uma *folksonomia* (T) é definida por um conjunto de anotações. A partir disso, segundo Mika [15], é possível utilizar o tradicional modelo bipartido de ontologia (conceitos e instâncias), fundamental para a construção do SWEETS. A Figura 1 apresenta a arquitetura do SWEETS.

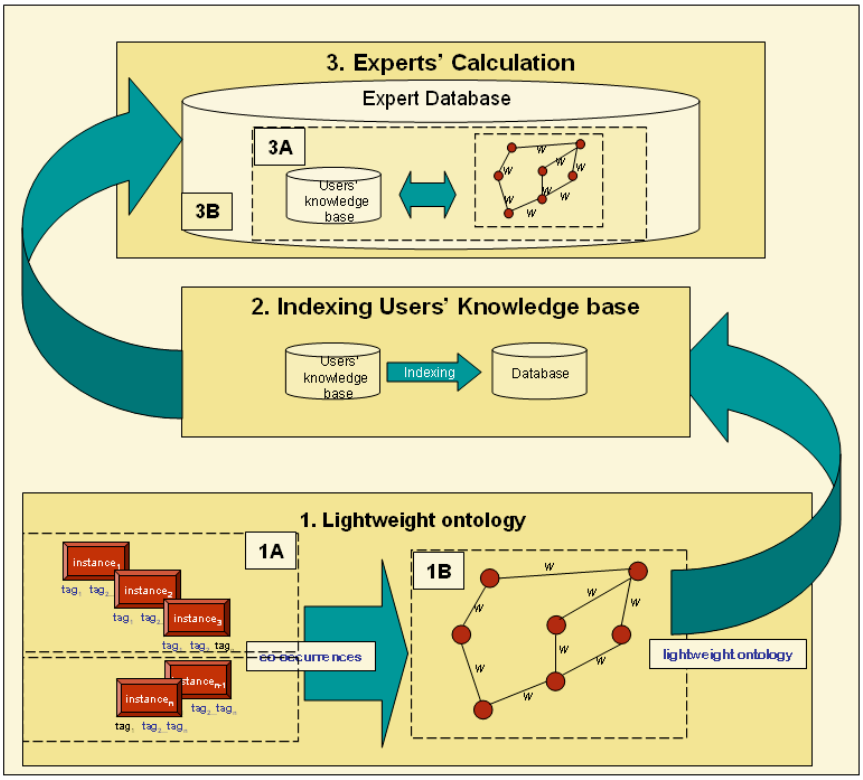

Figura 1. Arquitetura do SWEETS

A arquitetura do SWEETS está dividida em três camadas. A primeira (1) cria uma simples ontologia O_{ci}, que é refletida pela co-ocorrência das *tag*s (conceitos) em relação aos itens (instâncias). Os relacionamentos entre os termos/conceitos da ontologia O_{ci}, são ponderados pela quantidade de instâncias (I) que são etiquetadas (*tagged*) com ambos os termos, ou seja, o número de vezes que os termos co-ocorrem em instâncias diferentes. Este é um método básico de mineração de texto, em que os termos são geralmente associados pela sua co-ocorrência em documentos [6], [10], [15].

Uma vez que a ontologia O_{ci}, foi criada, na segunda camada (2), é realizada a extração de cada conceito que possuem relacionamentos maior que 1, seus relacionamentos e respectivos pesos. Sabe-se que essa quantidade mínima deve ser maior que 1, porém, não há uma definição exata do valor ideal. Nesta versão inicial do projeto, é adotado o mínimo de 4 relacionamentos – isso será explicado com mais detalhes na seção 6. É importante ressaltar

que este número mínimo de relacionamentos é configurável e, por isso, pode ser modificado a qualquer instante.

Após isso, tais informações são representadas em um Modelo Espaço Vetorial $\overrightarrow{Wo_{t_i}} = \left[Wo_{t_1},\ Wo_{t_2},\ \ldots\ , Wo_{t_{n-1}},\ Wo_{t_n}\right]$, em que $\overrightarrow{Wo}_t$ representa o vetor de pesos do conceito chave e seus relacionamentos. O peso do conceito chave possui maior relevância que qualquer outro peso do relacionamento, por isso, este peso é determinado por $\max\left(\overrightarrow{Wo}\right)+1$. Com isso, para cada conceito e seus relacionamentos, haverá uma representação vetorial, formando assim um conjunto de vetores.

Após esse processo, é realizada a indexação da base de conhecimento de cada usuário. Essa indexação é executada utilizando todo o conhecimento produzido pelos usuários no ambiente ao qual o SWEETS está sendo implantado. Com isso, para cada vetor $\overrightarrow{Wo}_t$ haverá um vetor $\overrightarrow{Fu_{t_i}} = \left[Fu_{t_1},\ Fu_{t_2},\ \ldots\ Fu_{t_{m-1}},\ Fu_{t_m}\right]$ equivalente, com a diferença que $\overrightarrow{Fu}_t$ representa a freqüência de cada termo em relação ao usuário e $\overrightarrow{Wo}_t$ representa os pesos dos termos na O_{ci}. É importante ressaltar que n, o tamanho do vetor $\overrightarrow{Wo}_t$ e m, tamanho do vetor $\overrightarrow{Fu}_t$ são iguais, sendo $Fu_{t_m} \geq 0$ e $Wo_{t_m} \geq 0$.

E, por fim, na terceira camada da Figura 1 é realizado o cálculo do grau de especialidade dos usuários em relação aos conceitos. Para isso, são utilizadas as representações vetoriais de $\overrightarrow{Fu}_t$ e $\overrightarrow{Wo}_t$. O cálculo das especialidades pode ser realizado por algoritmos de análise de similaridade, por exemplo, *co-seno*, *coeficiente de Pearson e Jaccard* [1].

O presente trabalho adotou o algoritmo do *co-seno*, conforme mostra a Equação 1.

$$GEuc = \frac{\sum_{i=1}^{n}\left(Wo_{t_i} \cdot Fu_{t_i}\right)}{\sqrt{\sum_{i=1}^{n}\left(Wo_{t_i}\right)^2} \cdot \sqrt{\sum_{i=1}^{n}\left(Fu_{t_i}\right)^2}} \qquad (1)$$

A função *GEuc* representa o grau de especialidade do usuário em relação a um conceito e seu valor, variando entre 0 e 1. O valor 0 representa a não especialidade entre o

usuário e o conceito, enquanto o 1 representa a total especialidade do usuário em relação ao conceito. De acordo com Baeza [1], esta função é inversamente relacionada ao ângulo entre $\overrightarrow{Fu}_t$ e $\overrightarrow{Wo}_t$, pois quanto menor é o ângulo entre $\overrightarrow{Fu}_t$ e $\overrightarrow{Wo}_t$, maior o valor do *co-seno*, e maior é a correlação entre $\overrightarrow{Fu}_t$ e $\overrightarrow{Wo}_t$, sendo que esse valor independe do tamanho do vetor.

6 Estudo de Caso

Acrônimo de Ambiente Multimídia para Integração de Grupos e Organizações Sociais, o a.m.i.g.o.s [5] tem por objetivo prover a infra-estrutura necessária para a criação de redes sociais virtuais para os mais diversos fins. Dentre estes fins, pode-se destacar o seu uso para estimular a criação e o compartilhamento do conhecimento pelos seus membros, podendo estes estarem relacionados a uma organização social.

O a.m.i.g.o.s. foi construído com base nas definições de redes sociais na Web apresentadas na seção 2, em que é permitida a criação explícita das redes sociais através dos usuários e seus contatos. Cada contato é explicitamente adicionado por cada usuário, mesmo que dentro de uma mesma organização, e este relacionamento é navegável por qualquer outro membro da rede social.

Nas próximas seções são apresentadas algumas das principais funcionalidades com suas características e possíveis usos. Essas funcionalidades principais são aquelas que permitem a construção do conhecimento pertinente a cada usuário, além das que oferecem subsídios para a criação da *folksonomia* – premissas básicas para implantar a ferramenta SWEETS 2.0.

6.1 Perfis

Cada usuário possui um perfil no a.m.i.g.o.s. Este perfil consiste de um conjunto de dados preenchidos na forma de cadastro, que definem algumas propriedades simples do usuário, como local de residência, idiomas que possui conhecimento, endereço de e-mail, identificadores de aplicações de mensagem instantânea (*Windows Live Messenger*, *Skype*, *Google Talk*, dentre outros), e uma descrição de suas áreas de interesse.

Porém a parte mais relevante do perfil não é preenchida pelo usuário, e sim inferida pelo sistema. Isto inclui o índice de atividade do usuário dentro do ambiente, que é calculado à medida que o mesmo participa de atividades de produção ou consumo do conhecimento existente na rede social; o conjunto de assuntos sobre os quais o usuário possui conhecimento, que foi inferido através da identificação dos termos de maior relevância

postados pelo usuário nas diversas atividades realizadas dentro do ambiente; e o conjunto de respostas às perguntas lançadas aos usuários.

Esta abordagem pretende cobrir uma ampla gama de possibilidades de identificação do usuário no sistema, particularmente em termos de práticas emergentes e distribuídas ao longo do tempo.

6.2 Histórias

Histórias são destinadas ao registro, compilação e apresentação de conhecimentos emergentes entre os participantes da rede. Construídas de forma gradual, através de contribuições espontâneas ou induzidas, qualquer usuário do sistema pode inserir no ambiente suas próprias histórias de sucesso ou dilemas, à medida que as considere relevantes para o objetivo da rede social.

Cada história pode incluir objetos dos mais diversos tipos (arquivos texto, apresentações, arquivos de áudio, arquivos de vídeo), de forma que o conhecimento depositado possa ser enriquecido por outros recursos. De modo similar, uma história também pode estar relacionada a outras histórias, permitindo que os usuários criem histórias maiores, compostas por várias pequenas histórias.

Adicionalmente as histórias podem estar associadas a uma ou mais comunidades, o que indica que, apesar do autor ser um usuário em específico, o conhecimento construído encontra-se de alguma forma relacionado a estas comunidades. Cada usuário do sistema poderá, adicionalmente, atuar como um revisor do conteúdo inserido por seus pares, avaliando qualitativamente as contribuições disponibilizadas neste ambiente. Esta avaliação pode ser realizada de uma das duas formas:

- Adição de comentários que contribuam para a evolução da história, criando-se assim uma história mais rica, com mais participantes e novos conhecimentos. À medida que a história for acrescida de comentários, é criado um diálogo associado ao conhecimento em construção;
- Atribuição de uma nota, variando de uma (1) a cinco (5) estrelas, às histórias que lê. Permitindo que este conhecimento, expresso através de histórias, possa ser apresentado através de um ranking que indique as mais relevantes para os membros daquela rede social.

6.3 Relacionamentos

O a.m.i.g.o.s dá suporte a praticamente todos os mecanismos de relacionamentos existentes nas atuais redes sociais. Nele cada usuário pode adicionar a sua lista de contatos

qualquer outro usuário também membro da rede social. Esta lista de contatos pode ser agrupada em grupos, facilitando a organização dos contatos pelo seu usuário.

6.4 Comunidades Virtuais

Comunidades podem ser vistas como agregações de pessoas com objetivos em comum. O a.m.i.g.o.s dá suporte à criação de manutenção de comunidades (Figura 2) por parte de seus usuários, podendo estes convidarem membros de sua lista de contatos a participar das discussões ou atividades a serem realizadas no âmbito da comunidade.

Cada comunidade possui uma série de mecanismos para a criação e compartilhamento do conhecimento. O principal mecanismo de criação e compartilhamento do conhecimento é o fórum de discussão, onde os membros da comunidade podem iniciar discussões sobre os mais diversos assuntos. E nestas discussões, todos os outros mecanismos de criação e compartilhamento de conhecimento podem ser utilizados (ex: as histórias, os objetos, a folksonomia).

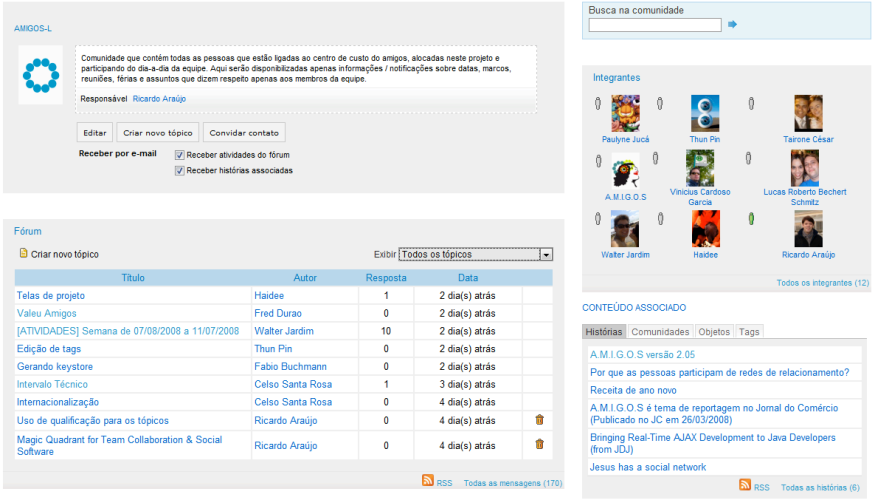

Figura 2. Comunidade no a.m.i.g.o.s.

Um segundo mecanismo de compartilhamento do conhecimento é a associação de histórias à comunidade. Esta associação pode ser realizada por qualquer membro da comunidade ao criar uma história no sistema. Caso deseje-se, é possível até mesmo que a história seja visível apenas pelos membros das comunidades relacionadas. Um terceiro mecanismo é a associação de objetos à comunidade, que, assim como as histórias, podem ser

associados pelo proprietário do objeto a qualquer uma das comunidades às quais este pertence.

Adicionalmente, uma comunidade pode possuir uma série de outras comunidades relacionadas ou afins, permitindo que comunidades que possuam focos semelhantes ou alguma intersecção de interesses possam facilmente navegar entre si.

6.5 Objetos

O a.m.i.g.o.s permite a adição de conhecimento através do conceito de objetos. Um objeto pode ser visto como qualquer meio eletrônico pelo qual o conhecimento pode estar formalizado, ou que pode ser utilizado para a construção de novos conhecimentos. Desta forma, qualquer arquivo pode ser armazenado e disponibilizado no sistema, como, por exemplo, documentos, artigos, apresentações, planilhas, vídeos, áudios e *URLs* externas. Todos os objetos possuem controle básico de versões, a ser definido pelo usuário, e permissões de acesso. Para permitir uma maior colaboração por parte dos usuários sobre um determinado objeto, todo e qualquer objeto pode receber comentários dos usuários, adicionando a possibilidade do surgimento de diálogos acerca do conhecimento existente em um objeto.

6.6 Folksonomia

Para permitir uma classificação do conhecimento armazenado em seu ambiente, facilitando assim a descoberta de informações por parte de seus membros, o a.m.i.g.o.s possui um mecanismo de *folksonomia* [15]. Através de *folksonomia* os usuários podem classificar, de forma social e colaborativa, o conteúdo disponível no ambiente, como, por exemplo, comunidades, histórias, objetos, comentários e discussões em fóruns. Para isto precisam apenas adicionar palavras-chave que identifiquem o assunto sendo tratado pelo conteúdo classificado. Adicionalmente, o sistema permite a visualização de todos os marcadores criados pelos usuários através de uma *tagcloud*. Desta forma, o usuário pode acessar rapidamente todo e qualquer conteúdo associado a um determinado marcador.

7 Experimento Preliminar

Antes que o SWEETS fosse implementado e implantado no ambiente do a.m.i.g.o.s., foi realizada uma pesquisa junto a seus usuários cadastrados que faziam parte do c.e.s.a.r.. O principal objetivo dessa pesquisa era verificar/constatar a real necessidade de implantação de um Sistema de Recomendação de Especialistas de Domínio nesse ambiente.

O questionário foi disponibilizado no a.m.i.g.o.s. e recomendado explicitamente a 100 usuários, distribuídos da seguinte forma: 40% eram Programadores, 40% Analistas de Sistemas e 20% Gerente de Projetos. Deste grupo, apenas 30 responderam prontamente o questionário, dos quais 50% eram programadores, 43,33% eram Analistas de Sistemas e somente 6,66% eram Gerentes de Projetos. Não há uma razão que justifique este percentual, já que este era um questionário totalmente aberto, ou seja, não direcionado a um grupo em específico.

Um dos objetivos da aplicação do questionário foi levantar, quantitativamente, informações de quais posturas os usuários costumam ter para a resolução de um determinado problema e como essas pessoas mostram-se interessadas em ajudar um colega na resolução de um problema. Para entender como os usuários se portam diante de um problema, foi feita a seguinte pergunta: "Quando você se depara com uma dificuldade/dúvida sobre algo relacionado ao seu trabalho, que postura toma para resolver o problema?". A Figura 3 mostra os resultados obtidos.

Dos resultados obtidos, 26,67% pessoas responderam que procuram contato direto com outras pessoas na tentativa de sanar seu problema, sejam essas pessoas amigas ou seu chefe, enquanto 56,67% tentariam resolver o problema através de pesquisas pela web.

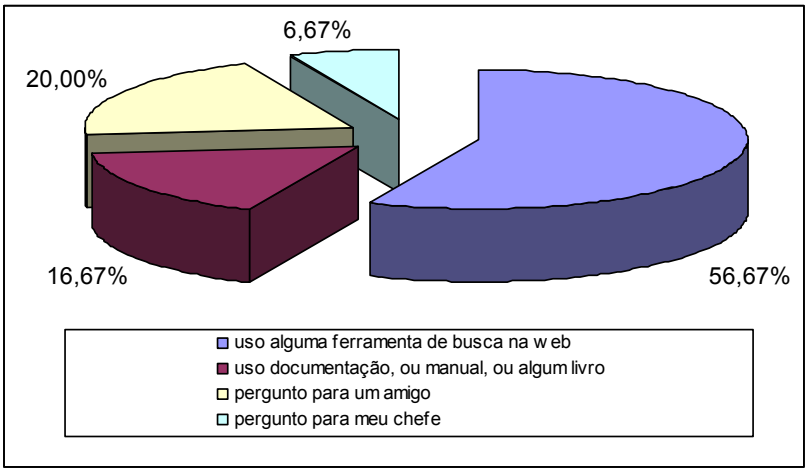

Figura 3. Postura dos usuários para resolver um problema.

Além desse questionamento, foi levantada uma questão com o objetivo de identificar a disponibilidade/interesse das pessoas em ajudar uns aos outros. Com isso, constatou-se que 53,33% das pessoas sempre ajudaram seus colegas nas últimas 5 vezes em que foram

procurados e apenas 10% dos entrevistados não ajudaram de forma alguma, ou porque estavam ocupados ou porque não tinham um contato prévio com a pessoa que detinha a dúvida.

Diante do levantamento dos dados previamente apresentados, pode-se chegar a conclusão de que seria interessante disponibilizar um ambiente em que as pessoas pudessem colaborar entre si.

Um incentivo para que essa colaboração surgisse seria a identificação de especialidades entre os usuários e serviria para oferecer subsídios para solução de um determinado problema. Estes foram alguns dos motivos que justificaram a implantação de um sistema que permita a identificação de especialistas no a.m.i.g.o.s.

A identificação de especialistas no a.m.i.g.o.s. possui dependência direta com a produção de conhecimento por parte dos usuários e, por isso, essa foi uma das questões abordadas no questionário. Como resultado dessa questão constatou-se que apenas 20% dos usuários produzem conhecimento constantemente, 33,3% produzem conhecimento entre 1 a 3 vezes por semana e a maioria das pessoas, representando 46,66% das pessoas raramente produziam conhecimento no a.m.i.g.o.s., pois, de uma forma geral, acham que mecanismo como lista de discussão resolvem muito bem o problema de comunicação que eles têm.

Como forma de sanar este problema, ou seja, proporcionar um incentivo às pessoas produzirem conhecimento no a.m.i.g.o.s. foi levantada a seguinte questão: "Caso houvesse um Sistema de Recomendação de Especialistas de Domínio implantado no a.m.i.g.o.s., você se sentiria mais motivado a produzir conhecimento com uma maior freqüência no ambiente?" 53,33% das pessoas responderam que sim, 36,67% responderam que talvez, pois primeiro deveriam utilizar a ferramenta para dar sua opinião e somente 10% das pessoas responderam que não. A grande maioria dos usuários que afirmaram que se sentiriam mais motivados a produzir conhecimento no a.m.i.g.o.s. alega que, com a implantação de um Sistema de Recomendação de Especialistas, uma maior geração de conhecimento seria intuitiva e natural, pois os colaboradores teriam interesse em crescer na empresa e, esta, com certeza, seria uma excelente oportunidade para os colaboradores projetarem suas especialidades. Essa possibilidade de promover uma maior colaboração entre as pessoas é importante tanto para os colaboradores quanto para a própria empresa, no caso o c.e.s.a.r., pois as habilidades dos colaboradores poderiam ser melhor exploradas, aumentando assim a eficiência e eficácia no nível de produção.

8 Experimento e Resultado da Implantação do SWEETS no a.m.i.g.o.s.

O sistema de Recomendação de Especialista SWEETS, em ambas as versões 1.0 e 2.0, foi desenvolvido e integrado no ambiente a.m.i.g.o.s., uma plataforma de Gestão de Conhecimento utilizada no C.E.S.A.R.. Esta integração junto ao a.m.i.g.o.s. tinha os seguintes objetivos:

1. utilizar particularidades de Redes Sociais em um sistema que identifique especialistas de domínio;
2. identificar especialistas de domínio;
3. promover melhorias na colaboração entre os usuários da plataforma a.m.i.g.o.s.;
4. incentivar uma maior produção de conhecimento dos usuários do a.m.i,g.o.s.;
5. agilizar o processo de resolução de um determinado problema;
6. e alocar ou re-alocar pessoas em projetos com uma maior eficácia.

A grande maioria desses objetivos, com exceção dos objetivos 1 e 2, só poderão ser alcançados com a análise dos resultados gerados pelo SWEETS durante um período de tempo mais longo, por exemplo, 4, 5 ou 6 meses. De posse das informações produzidas neste período, poderão ser realizadas análises mais concretas a partir da comparação dessas informações com informações geradas em períodos anteriores.

A análise da qualidade das recomendações geradas pelo SWEETS foi realizada de duas maneiras distintas. A primeira forma de análise foi aplicada em ambas as versões, 1.0 e 2.0, a partir da utilização de um conjunto restrito de 18 pessoas (colaboradores do C.E.S.A.R.). Este conjunto restrito de pessoas se deve à dificuldade de encontrar pessoas disponíveis para realizar tal análise, porém, acredita-se que este número seja o suficiente para analisar o nível de satisfação das recomendações geradas. Enquanto a segunda análise foi aplicada somente na versão 2.0, pois nesta versão há como identificar explicitamente as especialidades dos usuários. O objetivo de tal análise era verificar o nível de conhecimento ou interesse dos usuários em assuntos dos quais eles tinham sido detectados como especialistas.

8.1 Resultados do SWEETS 1.0

Em um primeiro momento foi implantada no a.m.i.g.o.s. a versão 1.0 do SWEETS, que não usa uma entrada explícita do usuário para gerar as recomendações. Ou seja, as recomendações são geradas de acordo com o conhecimento produzido (escrita) e conhecimento adquirido (leitura). A construção de conhecimento do usuário é possível a partir da criação de histórias, comunidades ou tópicos de comunidades, objetos ou

comentários sobre qualquer um desses ítens. A forma como esse conhecimento é produzido, determinando assim as especialidades dos usuários e como as recomendações funcionam são mostrados na Figura 4. De acordo com a pesquisa qualitativa realizada entre os usuários do a.m.i.g.o.s., dentre os 83,33% dos usuários que receberam recomendações, 86,67% desses usuários ficaram insatisfeitos. Provavelmente, o universo de usuários que receberam recomendações foi limitado a 83,33%, pois o restante, 16,67%, não produziu conhecimento no a.m.i.g.o.s. o suficiente para receberem as recomendações.

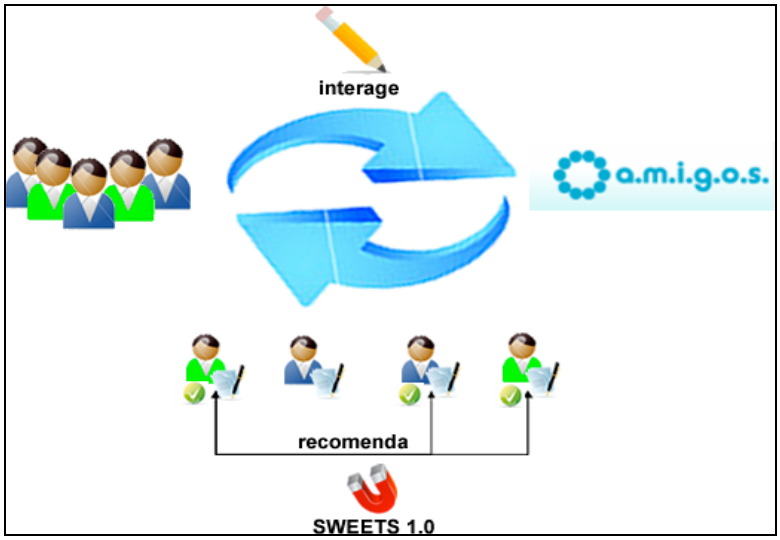

Figura 4. Determinação de especialidades no a.m.i.g.o.s. para serem usados no SWEETS 1.0.

Na grande maioria dos casos, as insatisfações eram por dois motivos: não permitir a entrada do usuário para solicitar um especialista de domínio e os especialistas recomendados eram sempre aqueles que detinham conhecimento similares ao próprio usuário. A não possibilidade ao usuário de entrar explicitamente com uma solicitação de especialistas de domínio força a aplicação a recomendar de forma pró-ativa, assim as recomendações são geradas a qualquer instante de tempo. Dessa forma, o sistema de recomendação atua da seguinte maneira: um usuário X recebe recomendações de outros usuários que possuem interesses comuns aos seus. Assim, são aproximadas pessoas que detenham conhecimentos similares (Figura 4). Tal característica proporcionou uma grande insatisfação pelos usuários, pois foi gerado um número demasiado de recomendações desnecessárias.

Para minimizar a limitação das recomendações de especialistas que detinham conhecimentos similares, foi inserida uma nova categoria de determinação de conhecimento. Com esta nova categoria, os usuários poderiam classificar positivamente conteúdos que leram e gostaram, assim, um novo conhecimento é construído: o conhecimento do interesse dos usuários. A Figura 5 mostra o funcionamento das recomendações com base na leitura dos usuários.

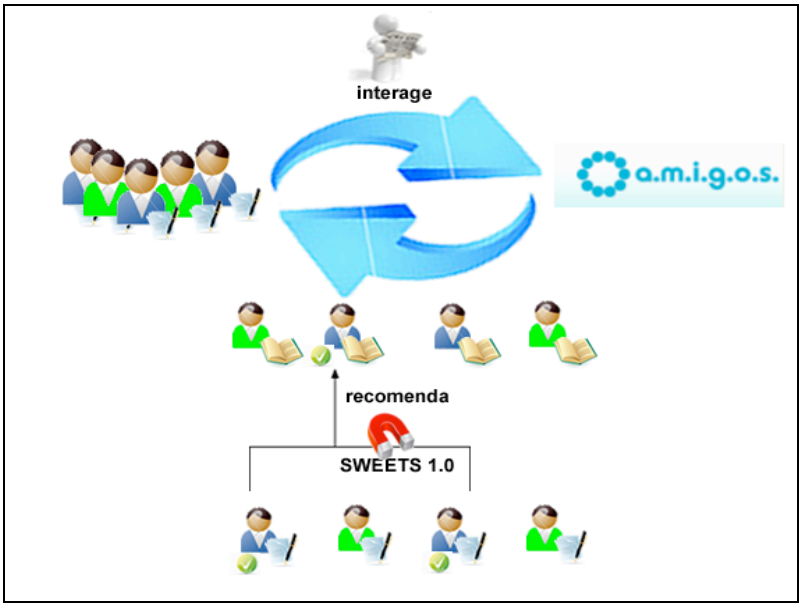

Figura 5. Recomendações de especialistas com base em conteúdos de interesses do usuário.

Uma vez que os usuários classificam os conteúdos de seus interesses, um novo perfil para cada usuário é construído: o de leitura. A partir da construção desse novo conhecimento, são recomendados especialistas que possuem conhecimento relacionado aos conteúdos de interesse desses usuários. Os conteúdos que os usuários podem classificar positivamente no a.m.i.g.o.s são: uma história, uma comunidade, ou um tópico de uma comunidade. De acordo com o que foi apresentado na seção 5.1, a classificação desses conhecimentos é realizada a partir da atribuição de uma nota, que varia de 1 a 5. Sendo que, somente conteúdos avaliados com notas 4 e 5 são considerados para construção do perfil de leitura do usuário.

Com a adoção dessa nova prática, as recomendações passaram a conter usuários com conhecimentos distintos aos dos usuários que recebem as recomendações. Como conseqüência, o nível de insatisfação dos usuários com as recomendações sofreu uma sutil

modificação, diminuindo para 77,78%. A insatisfação ainda continua alta, pois o problema que os usuários têm de não poderem requerer um especialista quando realmente necessitam, ainda persiste.

8.2 Resultados do SWEETS 2.0

Para resolver os problemas levantados pelos usuários encontrados no SWEETS 1.0, foi implantada no a.m.i.g.o.s. a versão 2.0. Nesta nova versão os usuários podem entrar explicitamente com uma requisição de especialista, a partir de uma palavra-chave. Tal particularidade resolve o problema das recomendações geradas desnecessariamente, ou seja, sem a prévia necessidade do usuário.

Na versão 2.0 do SWEETS é utilizada *folksonomia* para a criação de uma ontologia O_{ci}, que tem o objetivo de tornar as recomendações mais eficazes. A *folksonomia*, no a.m.i.g.o.s., surge à medida em que os usuários associam *tags* aos ítens (*tagged*). Esses itens são: um objeto (um documento ou site externo ao a.m.i.g.o.s.); uma história e comentários sobre esta história; uma comunidade, tópicos e comentários sobre esses tópicos. A ontologia O_{ci} é composta por relacionamentos entre conceitos (*tags*) associados a um determinado peso. Ambos são determinados pela co-ocorrência dos conceitos nos elementos utilizados para a criação da *folksonomia*. Quanto maior o peso entre os conceitos, maior a frequência que eles co-ocorrem, e maior a relação semântica entre eles. A principal vantagem desse mecanismo é a não pré-determinação da ontologia, pois ela surge à medida que os usuários interagem no a.m.i.g.o.s.. Dessa forma, a ontologia não está limitada a um domínio em particular.

A criação/atualização dessa ontologia exige um alto custo computacional, especialmente, se a *folksonomia* for extensa. É o caso da *folksonomia* do a.m.i.g.o.s. que possui em torno de 450 *tags*. Por isso, esse processo é realizado periodicamente no momento em que o a.m.i.g.o.s. não estiver ou estiver sendo pouco utilizado. Essa periodicidade é a cada 15 dias, e pode ser determinada pelo administrador do ambiente. A Figura 6 mostra o resultado da ontologia gerada a partir *folksonomia* do a.m.i.g.o.s..

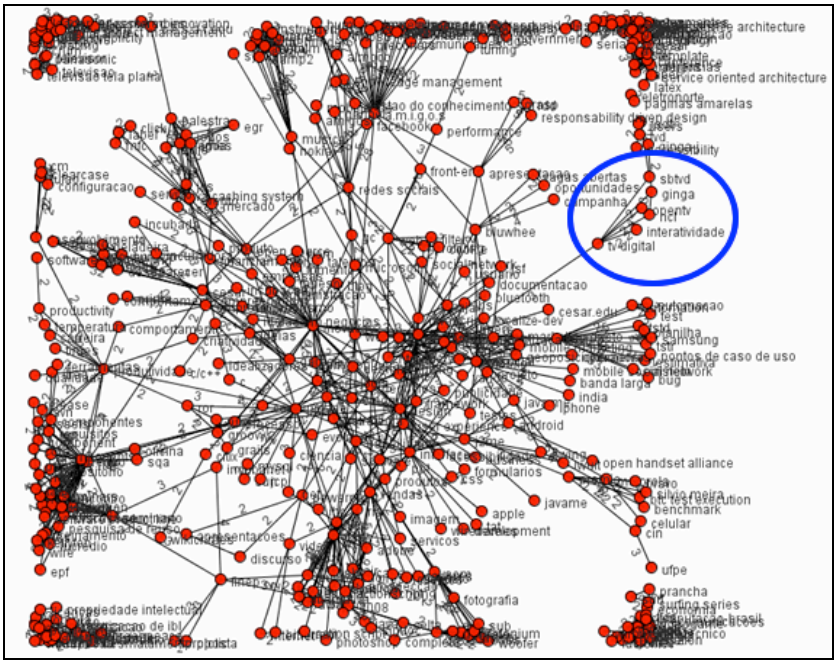

Figura 6. Ontologia O_{ci} que surgiu no a.m.i.g.o.s.

A ontologia O_{ci}, emergente do a.m.i.g.o.s., possui um total de 901 relacionamentos, por isso, o emaranhado de informações mostrado na Figura 6 não está claro. Apesar disso, é possível notar, na parte destacada com uma elipse, evidências semânticas na ontologia O_{ci} enfatizadas pelos relacionamentos dos termos *sbtvd*, interatividade, *ginga-j* e *opentv* ao termo *tv digital*.

Surgiram na ontologia conceitos que têm significados idênticos representados em relacionamentos distintos, por exemplo, "redes sociais" e "rede social". Tal problema era esperado, já que os conceitos (*tags*) associados aos itens são descritos livremente pelos usuários. Além disso, pode haver erros de digitação. A livre descrição é uma característica de *folksonomia*. Esses problemas comprometem o enriquecimento da ontologia, logo, torna a identificação dos especialistas menos eficaz.

Um dos caminhos para minimizar esses problemas é realizar um trabalho de conscientização dos usuários no enriquecimento da ontologia, ou seja, fazer com que os usuários se comprometam a "alimentar" a *folksonomia* adequadamente (de forma coerente). Esse trabalho de conscientização pode ser feito se for passada aos usuários a idéia de que esse enriquecimento é fundamental tanto para eles, quanto para a empresa. Pois, se o sistema

puder detectar corretamente as suas habilidades, seus potenciais poderão ser melhor explorados, logo, os seus níveis de satisfação e produtividade na empresa tendem a aumentar. Com isso, ganha tanto o colaborador quanto a empresa.

O passo seguinte para inferir as especialidades dos usuários foi indexar a base de conhecimento pertinente a cada usuário. Para isso, são utilizadas as informações textuais associadas aos usuários, ou as informações pré-processadas dos perfis desses usuários, que contém os termos mais relevantes e a frequência com que eles ocorrem. A diferença das informações textuais dos usuários e dos seus perfis é que os perfis contém informações adicionais dos conteúdos dos objetos (.doc ou .rtf, por exemplo), enquanto as informações textuais possuem somente o conhecimento postado pelo usuário.

Conforme apresentado na seção 5, para esta indexação, são considerados os conceitos/termos da ontologia O_{ci} e seus respectivos relacionamentos. É importante ressaltar que não são considerados todos os conceitos da ontologia, pois, a utilização desses conceitos está limitada à quantidade de relacionamentos que possuem. A quantidade mínima de relacionamentos deve ser maior que 1, Para este experimento foram considerados os conceitos que possuem no mínimo 4 relacionamentos. A justificativa para tal escolha é: quando o experimento foi executado com um mínimo de 3 relacionamentos foram identificados 387 especialistas, de um total de 916 usuários - uma quantidade irreal, já que aproximadamente 50% desses usuários são usuários assíduos do a.m.i.g.o.s.. Em um segundo instante, o experimento foi executado considerando os conceitos que possuem no mínimo 5 relacionamentos e, como resultado, obteve-se pouco mais de 60 especialistas, ou seja, especialistas em diversas áreas não foram englobados no resultado final. Com isso, chegou-se a um valor intermediário, 4.

A partir da análise da ontologia e de seus relacionamentos, acredita-se que o mínimo ideal de relacionamentos para cada conceito a ser considerado seja 4. Com o uso do sistema, pode-se chegar a um valor mais adequado, por isso, tal valor é configurável.

Para cada conceito considerado, se o mesmo for simples, é utilizado o seu peso no perfil pré-processado. Contrário a isso, se o termo for composto, tal como "Redes Sociais", são consideradas obrigatoriamente as informações textuais dos usuários para o cálculo da relevância do termo, já que os perfis dos usuários não fazem referência aos termos compostos. O grau de especialidade varia entre 0 e 1. Para que uma pessoa possa ser considerada especialista é preciso que este grau de especialidade esteja acima de um limiar. O limiar adotado para o presente experimento foi 0.8. Não há uma justificativa para a escolha deste limiar, poderia ser adotado um limiar 0.7 ou 0.6, contudo, definiu-se 0.8 para aumentar o nível de precisão de aplicação, já que é um número que se aproxima de 1 (valor máximo permitido). A Figura 7 mostra um exemplo de pesquisa na interface do SWEETS 2.0 no a.m.i.g.o.s..

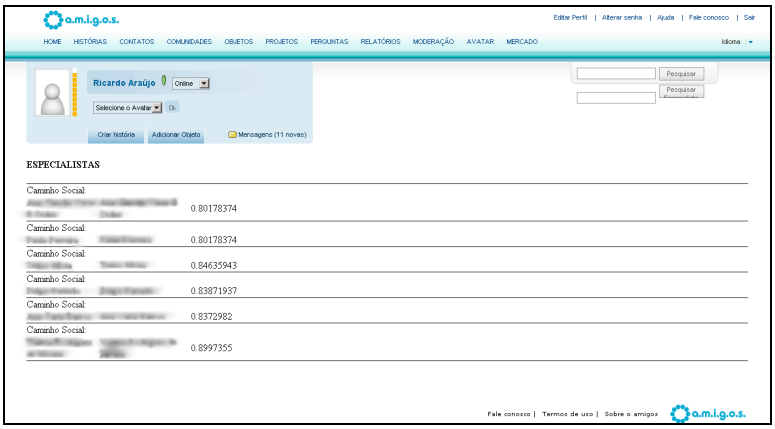

Figura 7. Interface do SWEETS 2.0 no a.m.i.g.o.s.

Na Figura 7 é mostrado um exemplo de pesquisa de especialistas em "requisitos" e a lista de especialistas retornada. Os nomes dos especialistas estão borrados, pois essas são informações confidenciais do C.E.S.A.R.. A lista de especialistas recomendada, por padrão, é ordenada de forma decrescente de acordo com o grau de especialidade. Além disso, apesar de ainda não estar disponível na interface, os especialistas podem ser ordenados por sua disponibilidade (online ou ocupado) ou pela distância social em relação ao usuário que solicitou a recomendação. Para cada especialista recomendado, é possível visualizar o número de pessoas intermediárias entre o usuário e o referido especialista.

Além do número, é possível visualizar o caminho completo da rede social de uma pessoa a outra. Cada nome que compõe o caminho social pode ser um link para o perfil de cada usuário.

8.2.1 Análise da qualidade das recomendações do SWEETS 2.0

O SWEETS 2.0 identificou 121 especialistas em variados assuntos. Diante da identificação dessas especialidades, foi elaborado e aplicado um questionário personalizado para aproximadamente 50% desses usuários com o objetivo de avaliar o nível de conhecimento ou interesse dos usuários nos assuntos aos quais foram identificados como especialistas. Desse universo de 50% de pessoas, aproximadamente 75% delas responderam o questionário.

As informações adquiridas com a aplicação desse questionário foram primordiais para avaliar o nível de qualidade das recomendações geradas pelo SWEETS 2.0. É importante ressaltar que, para evitar uma avaliação tendenciosa, não foi informado aos usuários que

esses teriam sido identificados pelo SWEETS 2.0 como possíveis especialistas nos assuntos apresentados em cada questionário.

Para a análise de qualidade das recomendações foi utilizada a métrica de precisão, que é medida pela razão entre os ítens relevantes e a quantidade total de ítens. Ou seja, a quantidade de especialidades detectadas corretamente pelo total de especialidades detectadas. Com isso, chegou-se a uma precisão de aproximadamente 57,26%.

O percentual de especialidades erradas ficou em 23,08% [esta margem de erro poderia ser menor se a folksonomia estivesse mais rica], enquanto o restante, 19,66%, é um percentual possivelmente justificável, pois este se refere às especialidades em que os usuários têm interesse, porém, não possuem conhecimento. Isso implica dizer que, possivelmente os usuários costumam produzir conhecimento sobre esses assuntos, uma das premissas para que as especialidades possam ser detectadas.

Outra avaliação realizada foi a utilização do sistema por um conjunto de 18 pessoas (as mesmas pessoas que avaliaram a versão 1.0 do SWEETS) e, diferentemente da versão 1.0, houve um nível de satisfação por parte de todos os usuários - na versão 2.0 -, no que tange à possibilidade de se permitir a entrada de uma especialidade pelo usuário para requerer especialistas, evitando assim recomendações desnecessárias. Durante a realização da pesquisa houve um colaborador que pesquisou por um especialista em "Redes de Computadores" e não obteve nenhum especialista recomendado. Isso ocorreu, pois a área informada não constava na ontologia O_{ci} emergente da *folksonomia*, logo, não há como identificar possíveis especialistas.

9 Conclusões

Este trabalho mostrou a ferramenta SWEETS, que tem por objetivo recomendar especialistas de domínio e a sua implantação na rede social a.m.i.g.o.s., uma plataforma de gestão de conhecimento utilizada pelo C.E.S.A.R.. Antes que a SWEETS fosse desenvolvida foi realizada uma pesquisa entre os colaboradores do C.E.S.A.R. e usuários da plataforma a.m.i.g.o.s. para verificar a real necessidade de implantação de um Sistema de Recomendação de Especialistas de Domínio nesta plataforma. De acordo com 53,33% das pessoas, uma ferramenta como a SWEETS seria tão importante que eles se sentiriam até com uma maior motivação para produzir conhecimento no ambiente, já que apenas 20% dos usuários produzem conhecimento assiduamente no ambiente. A construção do conhecimento no ambiente é premissa básica para identificar as especialidades das pessoas. A ferramenta SWEETS foi desenvolvida em duas versões, 1.0 e 2.0.

A versão 1.0 não possibilita uma requisição explícita dos usuários para recomendar especialistas, pois utiliza os perfis dos usuários para promover as recomendações. Assim, são

recomendados especialistas com perfis similares. Diante de uma pesquisa junto aos usuários, tal característica é uma desvantagem dessa versão, pois, além de não permitir que o usuário forneça explicitamente o assunto ao qual queira um especialista, sempre são recomendadas pessoas que detenham o mesmo conhecimento do usuário que recebe a recomendação.

Para minimizar o problema de aproximar (recomendar) somente pessoas com perfis similares, foi implantada à versão 1.0 uma nova forma de gerar os perfis das pessoas. Esta nova abordagem tinha o objetivo de identificar conhecimentos de interesse do usuário e não apenas conhecimento produzido pelo usuário. O conhecimento de interesse de uma pessoa é identificado a partir da do seu perfil de leitura, que é classificado a partir de informações que as pessoas leram e gostaram.

O SWEETS 1.0 utiliza somente técnicas de Recuperação de Informação para analisar textos planos e realizar as recomendações, ou seja, não existe nenhuma análise semântica dos termos. Com isso, o SWEETS 1.0 tende a uma menor eficácia, pois conhecimentos similares podem ser expressados utilizando termos diferentes. Com o intuito de adicionar semântica às recomendações e promover assim melhorias na qualidade dessas recomendações, foi desenvolvida e implantada no a.m.i.g.o.s. a versão 2.0 do SWEETS, que permite a entrada explícita para requisição de especialistas e usa ontologia como forma de representação de conhecimento. Diferente ICARE [18] - apresentado na seção 4 - essa ontologia não é prédefinida, pois emerge a partir de uma *folksonomia*, isso torna a identificação de especialistas livre de domínio - principal vantagem dessa abordagem.

O uso de *folksonomia* para o surgimento de ontologia há limitações voltadas à forma como a *folksonomia* é enriquecida, pois compromete diretamente a qualidade da ontologia O_{ci} gerada, logo, compromete a qualidade das recomendações. Isto é, as entradas dos usuários podem conter erros de digitação ou podem ser inconsistentes. Ou seja, ora o usuário associa a *tag* "Rede Social" a um item, ora associa uma *tag* diferente a outro item, mas que expressam o mesmo significado, por exemplo, "Redes Sociais".

Com a pesquisa realizada entre o conjunto de usuários que usou o SWEETS 2.0, observou-se que houve uma melhora na satisfação com a forma em que as recomendações são geradas, pois nesta versão não são geradas recomendações desnecessárias, o usuário só recebe a recomendação de um especialista quando ele a solicita.

Na versão 2.0, como as especialidades são detectadas explicitamente, pode ser realizada uma segunda análise que verifica o nível de precisão nas especialidades identificadas para os usuários. Essa precisão foi de 57,26% enquanto o percentual de erro ficou em 23,08%.

Já os 19,66% restantes podem ser justificados, pois, apesar dos usuários não serem especialistas nesses assuntos, manifestam interesse sobre eles; logo, há a probabilidade desses usuários terem postado informações no ambiente referentes a esses assuntos.

Para minimizar os problemas previamente apresentados, proporcionando assim efetivas melhorias na qualidade das recomendações, como trabalho futuro pretende-se aplicar um algoritmo de *stemming* (redução do termo ao seu radical) para gerar a ontologia O_{ci} e, com isso, espera-se diminuir a quantidade de redundâncias dos conceitos representados na ontologia e aumentar o peso semântico entre esses conceitos. Além disso, pretende-se, ainda, enriquecer a ontologia O_{ci} utilizando um thesaurus como, por exemplo, o *thesauro WordNet* [10]. Para cada termo da ontologia O_{ci} seria verificada sua existência no *thesaurus*, respectivos sinônimos, hierarquias e relacionamentos entre demais termos e, caso exista, todas essas informações seriam agregadas à ontologia O_{ci}.

Além disso, pretende-se utilizar técnicas/metodologias de aprendizagem de máquina para proferir recomendações de *tags* aos usuários, sempre que esses usuários estiverem descrevendo os itens da *folksonomia*.

Agradecimentos

Este trabalho foi apoiado pelo Instituto Nacional de Ciência e Tecnologia para Engenharia de Software (INES – www.ines.org.br), financiado pelo CNPq e FACEPE, processos 573964/2008-4 e APQ-1037-1.03/08.

Referências

[1] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, Boston, MA, May 1999.

[2] Stephen P. Borgatti and Rob Cross. A relational view of information seeking and learning in social networks. *Manage. Sci.*, 49(4):432–445, 2003.

[3] Ranjit Bose. Knowledge management metrics. *In Industrial Management & Data Systems*, 104(6):457–468, 2004.

[4] Byounggu Choi and Heeseok Lee. An empirical investigation of km styles and their effect on corporate performance. *Inf. Manage.*, 40(5):403–417, 2003.

[5] Ricardo A. Costa, Robson Y. S. Oliveira, Edeilson M. Silva, and Silvio R. L. Meira. A.m.i.g.o.s: knowledge management and social networks. In *SIGDOC '08: Proceedings of the 26th annual ACM international conference on Design of communication*, pages 235–242, New York, NY, USA, 2008. ACM.

[6] Douglass R. Cutting, David R. Karger, Jan O. Pedersen, and John W. Tukey. Scatter/gather: a cluster-based approach to browsing large document collections. In

SIGIR'92: Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval, pages 318–329, New York, NY, USA, 1992. ACM Press.

[7] Bryant Duhon. It's all in our heads. *Inform*, 12(8):8–13, 1998.

[8] Kate Ehrlich and N. Sadat Shami. Searching for expertise. In *CHI'08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1093–1096, New York, NY, USA, 2008. ACM.

[9] Thomas Erickson and Wendy A. Kellogg. Knowledge communities: Online environments for supporting knowledge management and its social context. In *Ackerman MA, Pipek V and Wolf W (eds) Beyond*, pages 299–325. MIT Press, 2003.

[10] Ronen Feldman, Ido Dagan, and Haym Hirsh. Mining text using keyword distributions. *J. Intell. Inf. Syst.*, 10(3):281–300, 1998.

[11] Gartner. The knowledge worker investment paradox. Julho 2002.

[12] Jennifer Ann Golbeck. *Computing and applying trust in web-based social networks*. PhD thesis, College Park, MD, USA, 2005. Chair-Hendler James.

[13] Henry Kautz, Bart Selman, and Mehul Shah. ReferralWeb: Combining social networks and collaborative filtering.

[14] Ching-Yung Lin, Kate Ehrlich, Vicky Griffiths-Fisher, and Christopher Desforges. Smallblue: People mining for expertise search. *Multimedia, IEEE*, 15(1):78–84, 2008.

[15] Peter Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):5–15, March 2007.

[16] Bonnie Nardi. It's not what you know, it's who you know: Work in the information age. *First Monday*.

[17] Ikujiro Nonaka and Hirotaka Takeuchi. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, May 1995.

[18] Helô Petry, Patrícia Tedesco, V. Vieira, and A. C. Salgado. Icare: A context-sensitive expert recommendation system. In *ICAI'08: Proceedings of the Workshop on Recommender Systems*, pages 53–58.

[19] Gabriele Plickert, Rochelle R. Cote, and Barry Wellman. It's not who you know, it's how you know them: Who exchanges what with whom? *Social Networks*, 29(3):405–429, July 2007.

[20] Tim Reichling, Kai Schubert, and Volker Wulf. Matching human actors based on their texts: design and evaluation of an instance of the expertfinding framework. In *GROUP'05: Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, pages 61–70, New York, NY, USA, 2005. ACM.

[21] Steffen Staab, Pedro Domingos, Peter Mika, Jennifer Golbeck, Li Ding, Tim Finin, Anupam Joshi, Andrzej Nowak, and Robin R. Vallacher. Social networks applied. *IEEE Intelligent Systems*, 20(1):80–93, January 2005.

[22] Stanley Wasserman, Katherine Faust, and Dawn Iacobucci. *Social Network Analysis: Methods and Applications (Structural Analysis in the Social Sciences)*. Cambridge University Press, November 1994.

[23] Takashi Yukawa, Kaname Kasahara, Toshiro Kita, and Tsuneaki Kato. An expert recommendation system using concept-based relevance discernment. In *ICTAI'01: Proceedings of the 13th IEEE International Conference on Tools with Artificial Intelligence*, page 257, Washington, DC, USA, 2001. IEEE Computer Society.