



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ELETRÔNICA
ENGENHARIA INDUSTRIAL ELÉTRICA -
ELETRÔNICA/TELECOMUNICAÇÕES
DISCIPLINA DE TRABALHO DE CONCLUSÃO DE CURSO 2



SINAPSE - SISTEMA INFORMATIZADO AVALIADOR DO POTENCIAL DE SINERGIA PARA EQUIPES

Relatório de Trabalho de
Conclusão de Curso 2, de-
senvolvido pelo estudante
Ricardo Esteban Rojas
Guerrero, orientado pelo
Prof. Dr. Dario Eduardo
Amaral Dergint.

CURITIBA
OUTUBRO 2013

AGRADECIMENTOS

A Deus por sempre me guiar e auxiliar nos momentos mais difíceis.

Aos meus familiares pela perseverança e esperança na vida que demonstram.

Ao Prof. Dr. Dario Eduardo Amaral Dergint por orientar, partilhar ideias e conhecimentos e pela compreensão e incomensurável paciência que teve comigo.

Ao Prof. Dr. Gustavo Benvenuto Borba e ao Prof. Dr. Miguel Antonio Sovierzoski pela participação na banca de defesa.

RESUMO

No paradigma contemporâneo, percebe-se que o trabalho em equipe é de fundamental relevância, visto a interdisciplinaridade inerente da atividade humana e o crescente emprego dessa neste mundo globalizado. Nessa visão, o sistema informatizado SINAPSE objetiva justamente a estimação do potencial de cooperação interpessoal em projetos. Para tanto, emprega-se um modelo, alicerçado na teoria de redes sociais, a considerar as múltiplas relações sociais e o dinamismo ao longo do tempo a fim de serem identificados prováveis integrantes de uma equipe. O desenvolvimento do ambiente do sistema avaliador se dá em Linguagem de Programação C# com base no framework ASP.NET. Por fim, com o suporte do SINAPSE, espera-se que as organizações possam formar as melhores equipes para as demandas requeridas.

Palavras-chave

Sistema informatizado. Engenharia colaborativa. Potencial de cooperação. Redes sociais.

ABSTRACT

In the contemporary paradigm, one realizes that teamwork is of fundamental relevance, given the inherent interdisciplinarity of human activity and the increasing use of it in this globalized world. On that regard, the computerized system SINAPSE aims precisely on the estimation of the interpersonal cooperation potential on projects. Therefore, it employs a model, based on the theory of social networks, considering the multiple social relations and the dynamism over time in order to identify possible members for a team. The estimator system's environment is developed in the C# Programming Language based on the ASP.NET framework. Finally, with the SINAPSE's support, it is expected that organizations can form the best teams for the required demands.

Keywords

Computerized system. Collaborative engineering. Cooperation potential. Social networks.

LISTA DE FIGURAS

Figura 1 - Diagrama de casos de uso do sistema	9
Figura 2 - Sociograma de Moreno ou Grafo em Estrela.....	15
Figura 3 - Comunidade, Grupo e Equipe	17
Figura 4 - Extração de informações para a construção da rede social egocêntrica	21
Figura 5 - Diagrama de classes.....	30
Figura 6 - Tela inicial do SINAPSE	31
Figura 7 - Tela ao serem adicionadas Competências.....	32
Figura 8 - Tela que exibe o <i>grid</i> de Candidatos encontrados	33
Figura 9 - Tela com exemplo de exibição de mensagem.....	33
Figura 10 - Tela ao serem adicionados Candidatos.....	34
Figura 11 - Tela ao ser pressionado o botão Formar Equipe	35
Figura 12 - Candidatos a serem analisados para conformar a Equipe	38
Figura 13 - Matriz de teste	40
Figura 14 - MEMENTO: Candidatos a serem analisados para conformar a Equipe.....	44
Figura 15 - MEMENTO: Matriz de teste	45

LISTA DE QUADROS

Quadro 1 - Interações (PAT).....	12
Quadro 2 - Matriz relacional de um pesquisador	22
Quadro 3 - Fator de Proximidade Individual de um pesquisador	25
Quadro 4 - Ausência do valor do peso atribuído à relação de coautoria	26
Quadro 5 - Rol completo dos pesos	26
Quadro 6 - GIVIC: Matriz relacional obtida manualmente.....	41
Quadro 7 - GIVIC: Rede social egocêntrica obtida manualmente	41
Quadro 8 - MEMENTO: Matriz relacional obtida manualmente	43
Quadro 9 - MEMENTO: Rede social egocêntrica obtida manualmente.....	43
Quadro 10 - Cronograma	47

SUMÁRIO

1	INTRODUÇÃO	8
1.1	JUSTIFICATIVA	8
1.2	OBJETIVO	9
1.3	ORGANIZAÇÃO DOS CAPÍTULOS	10
2	ANÁLISE DOS MODELOS	11
2.1	MODELOS REJEITADOS	11
2.1.1	<i>La Dynamique de la Confiance</i>	11
2.1.2	<i>Modelo Computacional de Percepção de Contextos de Atividade para Identificação de Comunidades</i>	12
2.2	REDES SOCIAIS	13
2.2.1	<i>Representações de Redes Sociais</i>	14
2.2.2	<i>Abordagem Egocêntrica</i>	14
2.3	CONCEITOS-CHAVE DA ANÁLISE DE REDES SOCIAIS	15
2.4	ASPECTOS SOCIAIS DA ANÁLISE DE REDES SOCIAIS	18
2.5	CONSTRUÇÃO DAS REDES EGOCÊNTRICAS	18
2.5.1	<i>Relação de Coautoria</i>	19
2.5.2	<i>Relação de Orientador-Orientado</i>	19
2.5.3	<i>Representação da Rede Social Egocêntrica</i>	20
2.6	CÁLCULO DO FATOR DE PROXIMIDADE INDIVIDUAL	21
2.6.1	<i>Contratempo</i>	25
2.7	FATOR DE PROXIMIDADE SOCIAL	26
2.8	CONCLUSÃO	27
3	DESENVOLVIMENTO	28
3.1	CONSIDERAÇÕES E AJUSTES DO MODELO	28
3.1.1	<i>Definição do Candidato mais Apto</i>	28
3.1.2	<i>Automatização do Processo de Obtenção de Informações</i>	29
3.2	ARQUITETURA	29
3.3	MODO DE OPERAÇÃO	31
3.4	MENSAGENS	38
3.5	TESTES	39
3.5.1	<i>Metodologia dos Testes</i>	39
3.5.2	<i>Teste com o GIVIC</i>	40
3.5.3	<i>Teste com o MEMENTO</i>	42
3.6	CONCLUSÃO	46
4	GESTÃO	47
4.1	CRONOGRAMA	47
4.2	CONCLUSÃO	47
5	CONCLUSÕES	49
6	REFERÊNCIAS BIBLIOGRÁFICAS	52

1 INTRODUÇÃO

Este documento apresenta o projeto de desenvolvimento de um sistema informatizado capaz de, dadas certas demandas, conformar as equipes mais aptas.

Para tal, obteve-se em mãos a dissertação de mestrado intitulada Modelo Computacional para Formação de Equipes baseado nas Redes Egocêntricas de Líderes de Projeto (REINHARDT, 2008), a qual propõe um modelo computacional, alicerçado na teoria de redes sociais para a formação de equipes de pesquisa baseado nas redes sociais egocêntricas de líderes de projeto, a considerar as múltiplas relações sociais e o dinamismo ao longo do tempo para calcular um fator de proximidade relacional utilizado para identificar integrantes pertencentes a um grupo.

Com ligeiras modificações e como pode ser visto no decorrer deste trabalho, grande parte do sistema implementado automatiza esse modelo.

1.1 *Justificativa*

Sabe-se que cada qual tem sua criatividade estimulada a partir da interação com os demais em uma equipe de pessoas de áreas e perfis variados; logo, tal equipe se torna, progressivamente, uma potência em reflexão e imaginação. A ideia, o projeto, a nova estratégia engendrada de cada encontro tende a ser mais bem sucedida quando as pessoas constituintes dessa possuem conhecimentos diversos que se complementam.

Já a incompreensão, o equívoco, a suspeita, a desconfiança, que pode haver entre pessoas de uma mesma equipe, são fatores que podem desencadear uma ruptura em qualquer conexão, ligação entre essas.

Desse modo, um sistema capaz de prevenir a reunião de pessoas incompatíveis, no contexto de um projeto, ou, ainda, de possibilitar determinar se os membros da equipe apresentarão competências complementares, fomentando um ambiente de real cooperação, torna-se demasiado útil.

Útil na conformação de equipes a terem o melhor desempenho em determinadas situações e condições, e em evitar o desperdício de recursos humanos, monetários e de tempo em projetos empresariais ou industriais ou acadêmicos ou de qualquer natureza.

1.2 Objetivo

Desenvolver sistema informatizado cujo intuito será estimar o potencial de cooperação de determinado grupo de pessoas no âmbito de um projeto com o propósito de formar uma equipe.

A Figura 1 ilustra o diagrama de casos de uso do sistema a ser implementado.

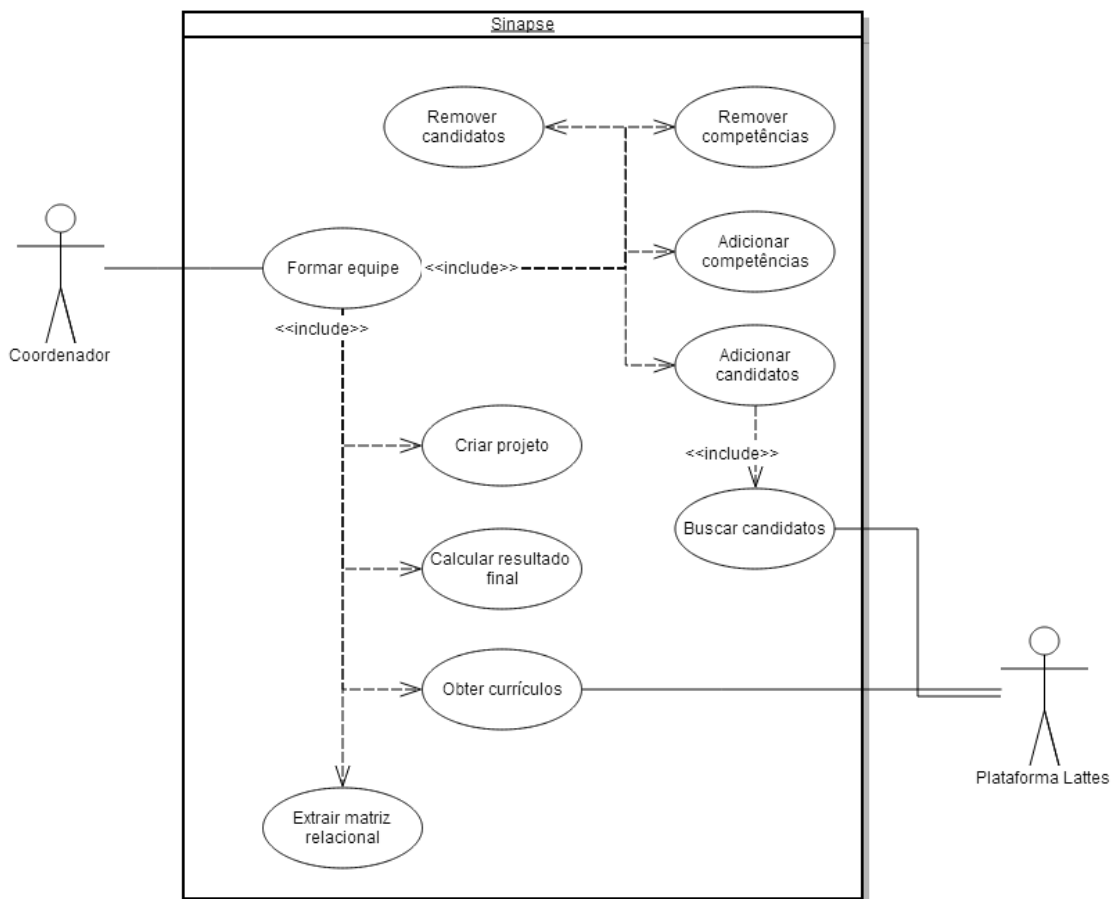


Figura 1 - Diagrama de casos de uso do sistema

Fonte: Autoria própria.

Visando o objetivo geral, os objetivos específicos que delimitam o escopo deste trabalho são:

- Examinar a fundamentação teórica do modelo computacional selecionado para servir de base ao sistema informatizado que se deseja desenvolver;

- Automatizar e ajustar esse modelo para que a abrangência de utilização do sistema implementado contemple também empresas e não tão somente ambientes acadêmicos;
- Implementar o software que permita, a um coordenador, avaliar quais são os candidatos aptos a, no contexto de um projeto, conformarem uma equipe na qual efetivamente haja cooperação e diversidade;
- Efetuar testes com o software para verificar a coerência dos resultados fornecidos pelo mesmo.

1.3 Organização dos Capítulos

Com a intenção de expor as ideias clara e efetivamente, este Relatório está dividido em cinco capítulos. Existe, na verdade, o sexto capítulo, mas esse é somente o das referências bibliográficas.

Neste capítulo, apresentam-se os objetivos deste TCC bem como sua justificativa.

Uma análise dos modelos cogitados para o desenvolvimento do sistema é apresentada no capítulo dois. Expõe-se, então, a fundamentação teórica do modelo selecionado para a concepção do sistema.

No terceiro capítulo, são encontrados os aspectos referentes ao desenvolvimento do software. Inicialmente, comenta-se acerca das considerações e ajustes pelas quais o modelo selecionado passou para constituir a base do sistema a ser implementado. Em seguida, apresenta-se a arquitetura do sistema e o modo de operação desse após sua codificação. Finalmente, realizam-se testes para averiguar se o sistema fornece resultados coerentes.

O cronograma do TCC é mostrado no quarto capítulo.

Por fim, as conclusões são mostradas no capítulo cinco.

2 ANÁLISE DOS MODELOS

Inicialmente, na seção 2.1, são apresentados os modelos que foram examinados, mas rejeitados para o desenvolvimento do SINAPSE.

Já a partir da seção 2.2, são expostos os conceitos e as ideias que respaldam e culminam no modelo empregado e automatizado pelo SINAPSE. Tal modelo é o da dissertação de Reinhardt (2008); muitíssimo, portanto, do que é apresentado no decorrer do capítulo é extraído dessa dissertação.

2.1 Modelos Rejeitados

2.1.1 La Dynamique de la Confiance

A priori, leu-se o livro *La Dynamique de la Confiance* (LE CARDINAL; GUYONNET; POUZOULLIC, 1997) com a intenção de se utilizar o método ali proposto como base do sistema a ser implementado. Tal método, a partir do Dilema do Prisioneiro, propõe uma descrição existencial de situações de interação sob uma forma acessível à linguagem usual através da expressão dos temores de ser traído, dos incentivos da cooperação e das tentações de trair. Eis por que esse método se denomina PAT-Miroir, já que promove a indicação de temores (*Peurs*), incentivos (*Attraits*) e tentações (*Tentations*) defronte de um "espelho" (*Miroir*). Assim, membros de um grupo de trabalho se reúnem e participam das fases resumidas a seguir:

- Definição do projeto, a qual engloba Reformulação do projeto (nova exposição, pelo grupo de trabalho, dos objetivos a serem atendidos), Criação de lista de agentes (pessoas que podem intervir de algum modo no desenrolar do projeto) e Agrupamento de tais agentes em atores (agentes que possuem funções, papéis e pontos de vista parecidos);
- Preenchimento do quadro de interações (temores, incentivos e tentações) no qual se encontram, em linhas e colunas, os atores anteriormente definidos. O Quadro 1 exemplifica esse de interações (PAT);

	Ator 1	Ator 2	Ator 3
Ator 1			
Ator 2			
Ator 3			

Quadro 1 - Interações (PAT)

Fonte: Adaptado de Le Cardinal, Guyonnet e Pouzoullic (1997, p. 219).

- Avaliação de itens (enunciados a tratarem de um temor, de um incentivo ou de uma tentação);
- Classificação desses itens;
- Agrupamento dos itens em temas incontornáveis;
- Elaboração de preconizações, as quais são recomendações visando diminuir os temores, a limitar as tentações e a favorecer, incrementar, efetivar os incentivos dos atores com relação ao projeto analisado.

Sem embargo, apesar desse método se mostrar bastante efetivo em estabelecer uma direção a seguir dentro do emaranhado de dilemas concernentes a projetos complexos, o sistema informatizado a advir do PAT-Miroir seria basicamente um ambiente a promover a atuação recíproca dos membros do grupo, o ateliê de criatividade das fases anteriormente mencionadas, ou seja, ao fim e ao cabo, não se teria um sistema automatizado, como era a intenção deste TCC, a ser utilizado por tão somente um coordenador a querer montar uma equipe.

Portanto, o PAT-Miroir foi deixado de lado na problemática deste TCC.

2.1.2 Modelo Computacional de Percepção de Contextos de Atividade para Identificação de Comunidades

Depois, teve-se acesso à dissertação de mestrado intitulada Modelo Computacional de Percepção de Contextos de Atividade para Identificação de Comunidades (BORTOLI, 2006). Tal dissertação apresenta um modelo conceitual de percepção que permite capturar e representar o contexto de atividade dos indivíduos e colocá-los em correlação para identificar comunidades. Esse modelo de percepção é baseado na análise de conteúdo dos artefatos textuais produzidos ou utilizados pelos indivíduos, permitindo-lhes perceberem outros que realizam ou realizaram atividades em contextos similares, proporcionando assim a identificação de comunidades potenciais. Nesse modelo experimental, são utilizadas técnicas de recuperação da informação (cálculo do TF/IDF, remoção de *stop words*, *stemming*).

Esse propõe ainda a análise da evolução das comunidades ao longo do tempo, além de propor métricas para avaliar o desempenho de algoritmos de identificação de comunidades, denominadas desempenho local e global.

No entanto, também apartou-se desse modelo, pois, enquanto esse monitora e analisa diversos artefatos textuais a fim de capturar contextos de atividade, os quais nada mais são do que vetores de termos relevantes acompanhados de seus valores de TF/IDF (que medem a relevância de tais termos), é o coordenador, ao empregar o SINAPSE, quem define os termos relevantes a serem varridos de um tipo específico de artefato textual: mais detalhes no Capítulo 3. Neste caso, tampouco se faz necessária a remoção de *stop words*. Por fim, a definição de comunidade, como pode ser visto na seção 2.3, difere da de equipe.

2.2 Redes Sociais

As redes sociais podem ser definidas como conjuntos de indivíduos e das relações que mantêm juntos esses indivíduos. Tais indivíduos podem ser elementos ou subconjuntos, como pesquisadores, grupos de pesquisa, universidades ou países.

As redes sociais vêm sendo amplamente utilizadas como ferramenta quantitativa em ciência social para a compreensão de como indivíduos ou organizações se relacionam. O desenvolvimento no campo de redes sociais se iniciou aproximadamente em 1930, com diversos grupos trabalhando de forma independente em diferentes áreas. Reinhardt (2008) menciona que, por volta da década de 30, existiam duas linhas de pesquisa para redes sociais. A primeira foi desenvolvida pelo grupo da Universidade de Harvard com o objetivo de encontrar subgrupos de pessoas em grandes grupos. A abordagem sociocêntrica foi derivada dessa linha de pesquisa. A segunda linha de pesquisa foi desenvolvida pelo grupo de antropologistas da Universidade de Manchester e abriu caminho para o estudo de comunidades partindo do ponto de vista do indivíduo, chegando ao desenvolvimento da abordagem egocêntrica.

Combinando informações relacionais com os atributos dos indivíduos em um panorama social, foi possível enriquecer a percepção social em relação à abordagem sociológica tradicional. As duas principais abordagens para o estudo de redes sociais são as já mencionadas sociocêntrica e egocêntrica. A abordagem

sociocêntrica é usada para estudar as redes sociais onde os limites são claramente especificados através da quantificação das relações entre os participantes desse grupo definido. O foco dessa abordagem é a medição de padrões de interação e como esses padrões surgem. Já a abordagem egocêntrica ganhou popularidade por focar os indivíduos, os grupos e as comunidades. Nessa abordagem são estudadas as relações do ponto de vista do indivíduo, sendo possível avaliar os aspectos da rede social individual.

2.2.1 Representações de Redes Sociais

As relações das redes sociais podem ser representadas de diversas maneiras, como gráficos, matrizes ou grafos de relacionamento. Os gráficos podem ser usados na análise de uma determinada relação em particular (como o número de ligações de cada indivíduo da rede social). As matrizes representam os relacionamentos entre os indivíduos através da valoração de cada relação. Um caso particular de matrizes são as chamadas matrizes relacionais, em que os indivíduos são dispostos em linhas e colunas e o valor da relação social entre dois indivíduos compõe os elementos da matriz. Essas matrizes são indicadas na análise de padrões de interação, equivalência estrutural e análise de blocos. Os grafos ou sociogramas permitem representar relações dirigidas ou não, sendo comum a utilização de grafos valorados, estruturas matemáticas básicas utilizadas para visualização das redes sociais. Os grafos são indicados na análise de centralidade, prestígio, intermediação, coesão de subgrupos e métodos de análise de díades e tríades dentro das redes sociais.

2.2.2 Abordagem Egocêntrica

Uma rede egocêntrica consiste na rede social focada em um determinado indivíduo (chamado de *ego*), nos demais indivíduos com os quais possui ligações (chamados de *alters*) e na medição das ligações sociais entre o *ego* e os *alters* e das ligações entre os *alters*.

Apesar de muitas análises de redes sociais buscarem um completo exame da rede social em larga escala (em que todas as conexões entre os membros da rede são visíveis), a abordagem egocêntrica busca uma visão mais restrita. Usando técnicas de análise de redes sociais egocêntricas, é possível entender como os

indivíduos se relacionam, os papéis individuais e os tipos de interações que se desenvolvem com os demais em sua volta.

Na dissertação de Reinhardt (2008), apenas as ligações diretas entre o indivíduo central (*ego*) e os membros de sua rede de contato (*alters*) são consideradas. Tal simplificação da abordagem egocêntrica reduz a complexidade computacional do modelo proposto. Esse tipo de rede egocêntrica simplificada é conhecido como grafo em estrela, ou sociograma de Moreno, que é indicado para a identificação de líderes ou indivíduos isolados. O sociograma de Moreno é ilustrado pela Figura 2, onde os diversos indivíduos da rede egocêntrica estão ligados ao indivíduo central por uma determinada relação social representada pelas linhas.

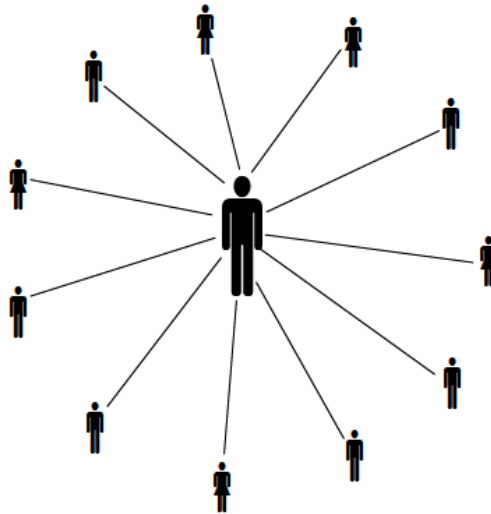


Figura 2 - Sociograma de Moreno ou Grafo em Estrela

Fonte: Reinhardt (2008, p. 22).

2.3 Conceitos-Chave da Análise de Redes Sociais

A seguir, são definidos os conceitos-chave da análise de redes sociais utilizados na dissertação de Reinhardt (2008).

Ator: A análise de redes sociais está baseada em entender as ligações entre as entidades sociais e as implicações dessas ligações. Essas entidades sociais são chamadas de atores. Atores são indivíduos discretos agrupados ou unidades de coletividades sociais. Exemplos de atores são os pesquisadores de um grupo de pesquisa, os grupos de pesquisa de uma universidade e as universidades no

sistema científico mundial. O uso do termo ator não implica que esta entidade possua vontade própria ou a habilidade de atuar. Muitas aplicações de redes sociais focam a coletividade de atores que são todos do mesmo tipo, como os pesquisadores de um centro de pesquisa. Pode-se chamar esta coletividade de monomodal. No entanto, alguns métodos permitem o estudo de atores conceitualmente de diferentes tipos, ou de diferentes grupos, como a participação dos pesquisadores em congressos científicos. Tanto os pesquisadores como os congressos são atores de tipos diferentes em uma mesma rede social. Pode-se chamar essa de rede social de bi-modal.

Ligações Sociais: Os atores são ligados a outros atores por ligações sociais. Existem diversos tipos de ligações sociais. A definição de ligação social é a ligação estabelecida entre um par de atores. Alguns dos exemplos de ligações mais comumente utilizados na análise de redes sociais são: avaliação de um ator por outro (a exemplo, expressando amizade, respeito ou credibilidade); relações formais (como autoridade e hierarquia); relacionamentos biológicos (como descendência, gênero); interações comportamentais (comunicar, enviar mensagens ou coautoria); associação ou afiliação (como participação em eventos sociais ou pertencer a um clube social); transferência de recursos materiais (como transações econômicas, empréstimos etc.); movimentações entre locais ou sociais (como migrações e modificação da posição ou grupo social); conexões físicas (como estradas, pontes, linhas telefônicas ligando dois pontos).

Díade: É o relacionamento que estabelece uma ligação entre dois atores. Tal ligação é uma propriedade de um par e não de apenas um ator em particular. A análise da rede social entre pares utiliza as díades como unidade de análise. A análise das díades consiste no par de atores envolvidos e na probabilidade de existir uma ligação entre eles, dessa relação ser recíproca ou não, ou se tipos específicos de múltiplos relacionamentos tendem a ocorrer em conjunto ou não. Na dissertação de Reinhardt (2008), a construção da rede egocêntrica de pesquisadores é realizada atribuindo um valor para mensurar as diversas díades formadas entre o líder de projeto (*ego*) e os demais pesquisadores (*alters*).

Comunidade: Pode ser definida como o conjunto de atores que partilham algo, mas não necessariamente conhecem uns aos outros ou interagem entre si. Por exemplo, a comunidade científica brasileira é composta por pesquisadores das mais

diversas áreas que possuem em comum o trabalho de promover o desenvolvimento científico e tecnológico.

Grupo: Pode ser caracterizado como um conjunto de atores que conhecem uns aos outros. Entretanto, não há necessariamente uma cooperação entre os elementos do grupo, sendo as interações caracterizadas como fracas, pois não existem necessariamente tarefas conjuntas com um objetivo em comum. O grupo de pesquisa do CPGEI é um exemplo, pois nesse grupo há diversos pesquisadores trabalhando em diferentes projetos de pesquisa.

Equipe: É um conjunto de atores que se conhecem mutuamente e cooperam para atingir um objetivo em comum, podendo haver o compartilhamento de algum artefato no qual estão trabalhando. Uma equipe de pesquisa trabalha em conjunto para atingir os objetivos do projeto de pesquisa e compartilham recursos como laboratórios, softwares, rede de computadores, documentos, livros e outros artefatos.

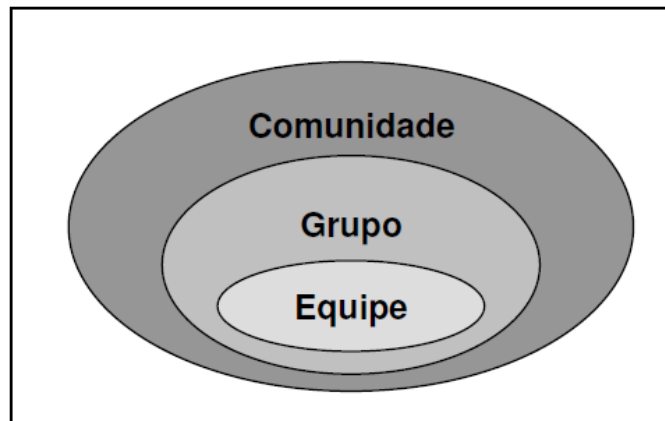


Figura 3 - Comunidade, Grupo e Equipe

Fonte: Reinhardt (2008, p. 26).

Relação: A coleção de ligações sociais de um tipo específico entre os membros de um grupo é chamada de relação. A exemplo, o conjunto formado pelos professores orientadores e seus alunos orientados nos cursos de pós-graduação de mestrado e doutorado, e o conjunto de coautorias entre os pesquisadores em diversos artigos científicos publicados. É importante perceber que uma relação se refere a um conjunto de ligações sociais, que possuem um determinado tipo de medição, entre um par de um conjunto específico de atores.

2.4 Aspectos Sociais da Análise de Redes Sociais

Nesta seção, são apresentados os aspectos sociais da análise de redes sociais considerados relevantes para o modelo de Reinhardt (2008).

Familiaridade: Define o nível de interação que dois atores tiveram entre si no passado. Uma grande familiaridade reduz as incertezas sobre como o outro ator irá se comportar no futuro. Pesquisadores que tiveram a experiência de trabalhar conjuntamente no passado devem estar familiarizados com o estilo de trabalho, hábitos e preferências uns dos outros, e, em virtude disso, deverá ser mais confortável trabalhar novamente em conjunto no futuro, estando a familiaridade relacionada com o grau de convivência dos pesquisadores no meio acadêmico.

Proximidade Geográfica: É definida como a distância física entre dois atores. Quando um par de atores se localiza geograficamente perto um do outro, eles possuem maior oportunidade para se encontrar, realizar comunicações espontâneas e interagir informalmente. Em contraste, quando esses atores estão longe um do outro, como no caso de pesquisadores que atuam em diferentes universidades, eles possuem poucas chances de realizar um trabalho conjunto e terão mais dificuldade em manter uma rotina de comunicações e coordenação para suportar uma relação de trabalho cooperativo do que dois atores que trabalham no mesmo espaço físico.

Homofilia: Define-se como a similaridade entre dois indivíduos. Os atores preferem se relacionar com outros atores que são similares a eles em características como idade, religião, nível de educação etc. No contexto de um grupo de pesquisa, a homofilia em relação às ideias e experiências proporciona a formação de equipes orientadas a objetivos comuns. Em um centro de pesquisa, dois pesquisadores que compartilham a mesma área de pesquisa frequentemente possuem experiências e objetivos similares. Por outro lado, pesquisadores em diferentes áreas de pesquisa possuem poucas experiências em comum, o que reduz a probabilidade de um trabalho em conjunto no futuro.

2.5 Construção das Redes Egocêntricas

A fim de construir as redes egocêntricas, definem-se quais relações sociais podem ser evidenciadas e como calcular a intensidade dessas ligações sociais entre os indivíduos. A escolha dessas relações depende dos aspectos sociais

relacionados com a necessidade de aplicação do modelo computacional proposto e da disponibilidade de documentos onde tais relações possam ser evidenciadas. No modelo de Reinhardt (2008), admite-se que, para os pesquisadores engajados em um projeto de pesquisa, as ligações sociais podem ser mensuradas pelo nível de esforço para o compartilhamento intelectual. Com isso, infere-se que pesquisadores com ligações fortes trabalham e publicam juntos, enquanto ligações fracas ocorrem com pesquisadores que trabalham em uma mesma área, porém não necessariamente trabalham ou publicam juntos.

Para o modelo de Reinhardt (2008), foram escolhidas as relações sociais que possuem aspectos sociais que aumentam a probabilidade de dois pesquisadores trabalharem juntos, como a familiaridade, a proximidade geográfica e a homofilia. As relações, escolhidas por causa desses aspectos e da facilidade de serem evidenciadas nos documentos da Plataforma Lattes, foram as de coautoria em trabalhos científicos e as relações de orientador-orientado.

2.5.1 Relação de Coautoria

Define-se como coautoria qualquer forma de ligação entre dois pesquisadores para a elaboração e divulgação de trabalhos científicos, como a publicação de artigos científicos, periódicos, livros, coparticipação em eventos científicos ou em outras atividades culturais. Essa relação é valorada de acordo com a frequência de ocorrência e o período de tempo em que ocorreu, sendo tais medições proporcionais ao resultado do trabalho científico realizado em conjunto, estando assim ligada ao aspecto social de familiaridade. Tais medições são utilizadas para identificar padrões em uma rede social entre os autores e coautores, proporcionando entendimento de como um autor trabalha e lida com seus colegas. Esses comportamentos são considerados como um fator que influencia o processo de formação de uma equipe.

2.5.2 Relação de Orientador-Orientado

É considerada na dissertação de Reinhardt (2008) como uma das mais coesas relações no meio acadêmico por possuir a capacidade de promover novas relações, como a relação de coautoria, e possuir duração de um a cinco anos. Esse tipo de relação está relacionado aos aspectos sociais de familiaridade, homofilia e proximidade.

Podem-se distinguir quatro tipos de orientações: a de iniciação científica, mestrado, doutorado e pós-doutorado. A relação de orientação de iniciação científica é a mais fraca dentre as relações de orientação, ocorrendo com os alunos de graduação que buscam conhecer o meio científico e colaboram com pequenas pesquisas e tarefas rotineiras da equipe de pesquisa. Essa relação na maioria dos casos dura apenas um ano e possui pouca ou nenhuma produção científica. A orientação de mestrado já possui maior intensidade, podendo durar de dois a três anos, sendo o mestrando responsável por construir um trabalho de pesquisa em uma determinada área do conhecimento. A relação de orientação de doutorado é a mais intensa, podendo durar até cinco anos, e proporciona a produção de um trabalho inédito de pesquisa, com diversos artigos relacionados. A relação de pós-doutorado dura em média um ano e busca o aperfeiçoamento em determinada área inovadora do conhecimento, relacionada à área de pesquisa do pós-doutorando. Pelo seu nível de especialização, classifica-se essa relação de orientação em conjunto com a relação de doutorado.

No modelo de Reinhardt (2008), as relações de orientador-orientado são caracterizadas como variáveis binárias, uma vez que se verifica apenas a existência dessa relação no decorrer dos anos. Utilizam-se coeficientes de correlação para diferenciar as relações de orientador-orientado de acordo com o tipo e época de ocorrência, obtendo-se, dessa forma, um valor proporcional à intensidade das relações. Tais coeficientes de correlação são empíricos.

2.5.3 Representação da Rede Social Egocêntrica

A construção da rede social egocêntrica de um pesquisador é realizada através do conjunto dos pesquisadores pertencentes ao mesmo grupo de projeto e das relações entre esses e o pesquisador analisado. Conforme mostra a Figura 4, o modelo computacional teórico e proposto por Reinhardt (2008) possui as seguintes estruturas para a extração de informações para a construção da rede social egocêntrica dos pesquisadores:

- Uma máquina de busca na Plataforma Lattes para obter as informações relacionais sobre o pesquisador a ser representado, os grupos de projeto e os membros desses grupos;
- Um mecanismo de extração de informação para identificar informações relacionais sobre a rede social dos pesquisadores;

- Um método para armazenar o valor das ligações sociais evidenciadas e calcular a composição destas relações para construir a matriz relacional da rede egocêntrica do pesquisador representado.

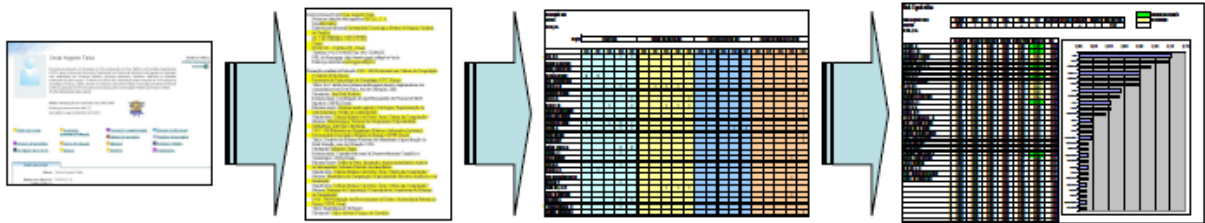


Figura 4 - Extração de informações para a construção da rede social egocêntrica

Fonte: Adaptado de Reinhardt (2008, p. 48).

2.6 Cálculo do Fator de Proximidade Individual

No âmbito do processo de identificação dos pesquisadores que serão indicados para a formação da equipe, o fator de proximidade individual está relacionado com o ator analisado e cada indivíduo de sua rede egocêntrica, e é valorado pela média ponderada entre as relações medidas, a qual é normalizada pelo número de integrantes da rede egocêntrica. Cada peso é proporcional à contribuição de cada relação para o fator de proximidade entre os indivíduos, considerando o período em que ocorreu a relação e os fatores sociais propostos de familiaridade, proximidade geográfica e homofilia.

Ao ser calculada a média ponderada das relações de orientador-orientado e coautoria no decorrer do tempo, utiliza-se esse valor para construir uma composição entre as múltiplas relações de coautoria (métrica) e orientador-orientado (binárias). A normalização desses valores ao final das médias produz um valor proporcional à relação entre os dois atores analisados, porém somente pode ser utilizado como referência para o ponto de vista do ator pesquisado em relação aos atores com que se relaciona, proporcionando uma noção qualitativa da relação.

É considerado, para a coleta de informação, o período de cinco anos por ser o valor médio para conclusão de um trabalho de doutorado. A contribuição de relações anteriores a esse período foi considerada desprezível em relação ao aspecto de familiaridade, utilizado no trabalho da dissertação de Reinhardt (2008) como um aspecto contribuinte para o fator de proximidade entre os pesquisadores.

Depois de serem obtidas as informações sobre cada relação escolhida, constrói-se uma matriz relacional com os valores de cada relação em função do tipo, da intensidade e do ano de ocorrência. O Quadro 2 mostra um exemplo dessa matriz em que, na primeira coluna, estão os pesquisadores que pertencem à rede social egocêntrica do pesquisador analisado, na segunda coluna, encontra-se a relação de coautoria nos cinco anos, e, nas demais colunas, para o mesmo período, repetem-se as relações de orientador-orientado.

Cesar Augusto Tacla DOUTOR TACLA, C. A.																					
relação:	CO-AUTORIA					ORIENTADOR-MESTRANDO					ORIENTADOR-DOUTORANDO					ORIENTADOR INICIAÇÃO CIENTÍFICA					
	2003	2004	2005	2006	2007	2003	2004	2005	2006	2007	2003	2004	2005	2006	2007	2003	2004	2005	2006	2007	
AVILA, B. C.				1	1																
Banaszewski, Roni Fábio										1											
Baptista, Rafael de Pauli																			1		
BARTHÈS, Jean-paul	6	3																			
Betiz, Fiori Angelo Rosot																				1	
BORTOLI, E. M.				1			1	1	1												
Brancaleone, Fernando																				1	
BRITO, R. C.					1					1											
CALIARI, F. M.			1		1			1	1	1											
Camargo, Fernando Guedes																				1	
Correa, Edgar Fernando																				1	
ENEMBRECK, Fabricio	1		1	3	1																
FREDDO, A.				1	1									1	1						
Gavon, Terezinha																				1	
Gazzoni, Joacir																				1	
GIMENEZ-LUGO, G.A.				2	2																
HUBNER, J. F.				1																	
Lazzari, Rômulo de																				1	
LUGO, G. G.				1																	
PARAISO, E. C.		1																			
Paulo, Marco Antônio Sales de																				1	
RAMOS, M. P.				1																	
REINHARDT, J. C. V.								1	1	1											
Santos, Nei Roberto dos																				1	
SCALABRIN, E.				1	1																
WOZNAK-GIMENEZ, A. B.				1																	
Yamamoto, Anderson Seiji																				1	

Quadro 2 - Matriz relacional de um pesquisador

Fonte: Adaptado de Reinhardt (2008, p. 48).

Para cada indivíduo (m), as relações de coautoria são mensuradas através da frequência de ocorrência, e é realizada uma média ponderada atribuindo empiricamente um peso a cada ano de ocorrência dessas relações. Assim, a relação de coautoria pode ser expressa por:

$$R_{ca_m} = \frac{\sum_{j=1}^K \left(\sum_{i=1}^n rca_i \cdot \beta_j \right)}{\sum_{j=1}^K \beta_j}$$

em que:

- R_{ca} é o valor da relação de coautoria do indivíduo m ;
- rca_i é a variável binária de ocorrência da coautoria i ;
- n é o número máximo de coautorias realizadas em um mesmo período;
- β_j é o peso atribuído ao período j , com $j = 1, \dots, K$;
- K é o número total de períodos analisados.

Para cada indivíduo (m), as relações de orientador-orientado são mensuradas através da média ponderada entre os anos de ocorrência e os pesos atribuídos a cada ano. Essas relações podem ser expressas por:

$$R_{dr_m} = \frac{\sum_{j=1}^K (rdr_j \cdot \theta_j)}{\sum_{j=1}^K \theta_j}$$

$$R_{ms_m} = \frac{\sum_{j=1}^K (rms_j \cdot \gamma_j)}{\sum_{j=1}^K \gamma_j}$$

$$R_{ic_m} = \frac{\sum_{j=1}^K (ric_j \cdot \delta_j)}{\sum_{j=1}^K \delta_j}$$

em que:

- R_{dr} , R_{ms} , R_{ic} são, respectivamente, os valores das relações de orientador-orientado para doutorandos, mestrandos e estudantes com bolsas de iniciação científica do indivíduo m ;

- rdr_j , rms_j e ric_j são as variáveis binárias de ocorrência das relações escolhidas de orientador-orientado, sendo $rdr_j = rms_j = ric_j = 1$ caso exista a relação, e $rdr_j = rms_j = ric_j = 0$ caso contrário;
- θ_j , γ_j e δ_j são os pesos atribuídos para cada uma das relações no período j , com $j = 1, \dots, K$;
- K é o número total de anos analisados.

O valor do fator de proximidade individual bruto é dado pela média ponderada entre essas relações. Cada peso é proporcional à contribuição de cada relação para os aspectos sociais de familiaridade, proximidade geográfica e homofilia. O valor do fator de proximidade individual bruto entre os atores pode ser expresso por:

$$P_m = \frac{(Rdr_m \cdot \alpha_{dr}) + (Rms_m \cdot \alpha_{ms}) + (Ric_m \cdot \alpha_{ic}) + (Rca_m \cdot \alpha_{ca})}{(\alpha_{dr} + \alpha_{ms} + \alpha_{ic} + \alpha_{ca})}$$

em que:

- P_m é o fator de proximidade individual bruto, entre o indivíduo m e os elementos de sua rede egocêntrica;
- α_{ca} , α_{ic} , α_{ms} e α_{dr} são os pesos atribuídos a cada uma das relações escolhidas.

Como a importância de cada indivíduo em uma rede egocêntrica depende da visão subjetiva, o valor do fator de proximidade individual bruto deve ser normalizado para transparecer a visão individual de cada indivíduo analisado. Assim, o fator de proximidade individual pode ser expresso por:

$$|P_m| = \frac{P_m}{\sum_{i=1}^N P_i}$$

em que:

- $|P_m|$ é o fator de proximidade individual do pesquisador m ;
- P_i é o fator de proximidade individual bruto do ator i ;

- *N* é o número máximo de agentes na rede egocêntrica.

Com os valores do fator de proximidade individual calculados, é possível a construção de um quadro com os valores da relação de proximidade da rede social egocêntrica. O Quadro 3 mostra os valores do fator de proximidade individual do exemplo:

Rede Egocêntrica											
Cesar Augusto Tacla DOUTOR TACLA, C. A.	relação:	2003	2004	2005	2006	2007	co-autoria	mestrado	doutorado	iniciação	soma Méd POND.
	peso:	1	2	3	4	5	2	3	4	1	2,947
	co-autoria	mestrado	doutorado	iniciação	Média POND	Média Arit	Dist				
FREDDO, A.	0,600	0,000	0,600	0,000	0,360	0,122	0,878				
CALIARI, F. M.	0,533	0,800	0,000	0,000	0,347	0,118	0,882				
ENEMBRECK, Fabrício	1,400	0,000	0,000	0,000	0,280	0,095	0,905				
REINHARDT, J. C. V.	0,000	0,800	0,000	0,000	0,240	0,081	0,919				
GIMENEZ-LUGO, G. A.	1,200	0,000	0,000	0,000	0,240	0,081	0,919				
BORTOLI, E. M.	0,267	0,600	0,000	0,000	0,233	0,079	0,921				
BRITO, R. C.	0,333	0,333	0,000	0,000	0,167	0,057	0,943				
BARTHES, Jean-paul	0,800	0,000	0,000	0,000	0,160	0,054	0,946				
AVILA, B. C.	0,600	0,000	0,000	0,000	0,120	0,041	0,959				
SCALABRIN, E.	0,600	0,000	0,000	0,000	0,120	0,041	0,959				
Banaszewski, Roni Fábio	0,000	0,333	0,000	0,000	0,100	0,034	0,966				
HUBNER, J. F.	0,267	0,000	0,000	0,000	0,053	0,018	0,982				
LUGO, G. G.	0,267	0,000	0,000	0,000	0,053	0,018	0,982				
RAMOS, M. P.	0,267	0,000	0,000	0,000	0,053	0,018	0,982				
WOZNAK-GIMENEZ, A. B.	0,267	0,000	0,000	0,000	0,053	0,018	0,982				
Bettez, Fiori Angelo Rosot	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Brancaleone, Fernando	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Camargo, Fernando Guedes	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Correa, Edgar Fernando	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Gavron, Terezinha	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Gazzoni, Joacr	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Yamamoto, Anderson Seiti	0,000	0,000	0,000	0,333	0,033	0,011	0,989				
Baptista, Rafael de Pauli	0,000	0,000	0,000	0,267	0,027	0,009	0,991				
Lazzari, Rômulo de	0,000	0,000	0,000	0,267	0,027	0,009	0,991				
PARAISQ, E. C.	0,133	0,000	0,000	0,000	0,027	0,009	0,991				
Paulo, Marco Antônio Sales de	0,000	0,000	0,000	0,267	0,027	0,009	0,991				
Santos, Nel Roberto dos	0,000	0,000	0,000	0,267	0,027	0,009	0,991				

Quadro 3 - Fator de Proximidade Individual de um pesquisador

Fonte: Adaptado de Reinhardt (2008, p. 48).

2.6.1 Contratempo

Durante a análise desse fator, houve um sério contratempo que foi a ausência do valor do peso atribuído à relação de coautoria na Tabela 2 da dissertação (REINHARDT, 2008, p. 60), como pode ser verificado no Quadro 4.

relação:	2003	2004	2005	2006	2007	mestrado	doutorado	iniciação	soma Méd Pond.
peso:	1	2	3	4	5	3	4	1	2,947

Quadro 4 - Ausência do valor do peso atribuído à relação de coautoria

Fonte: Adaptado de Reinhardt (2008, p. 60).

Sem esse valor, não era claro o modo como se chegava aos valores do Fator de Proximidade Individual dessa tabela.

Afortunadamente, percebeu-se que a Figura 6 dessa dissertação (REINHARDT, 2008, p. 48) possui uma excelente resolução e, fazendo-se um zoom in de cerca de 800% nessa, encontra-se o quadro correto de relação-peso que foi de fato utilizado para o modelo:

relação:	2003	2004	2005	2006	2007	co-autoria	mestrado	doutorado	iniciação	soma Méd Pond.
peso:	1	2	3	4	5	2	3	4	1	2,947

Quadro 5 - Rol completo dos pesos

Fonte: Adaptado de Reinhardt (2008, p. 48).

Dessa forma, pôde-se por fim compreender efetivamente como se determinava o Fator de Proximidade Individual.

2.7 Fator de Proximidade Social

Na seção 2.2.2, comentou-se acerca da rede egocêntrica simplificada, conhecida como grafo em estrela ou sociograma de Moreno, empregada no modelo proposto por Reinhardt (2008) a fim de reduzir a complexidade computacional desse modelo. No entanto, a perda de informação causada por tal simplificação foi minimizada na dissertação de Reinhardt (2008) devido à sobreposição das diversas redes egocêntricas e o uso da cadeia de recomendações no processo de identificação de indivíduos por intermédio do uso do Fator de Proximidade Social. Mais detalhes relacionados a esse fator se encontram na dissertação (REINHARDT, 2008, p. 61).

Em um primeiro momento, a informação sobre tal fator nessa dissertação foi considerada insuficiente para uma boa compreensão desse. Assim, tentou-se contatar o autor da dissertação em três ocasiões via e-mail, mas não se obteve

qualquer resposta da parte dele. Entretanto, relendo-se atentamente os conceitos embasando esse fator, compreendeu-se finalmente o mecanismo desse. Mesmo assim, o Fator de Proximidade Social foi descartado no cerne do SINAPSE.

Esse fator é obtido pela média ponderada entre os Fatores de Proximidade Individuais dos pesquisadores-candidatos em relação ao líder de pesquisa e aos demais pesquisadores que compõem o grupo de pesquisa. Os pesos dessa ponderação seriam definidos empiricamente de acordo com o papel que cada indivíduo desempenha no grupo. E aí se encontra o motivo do descarte: Quem definiria esses pesos? O coordenador? Baseado em quê? Mesmo na dissertação, não se encontram esses pesos em nenhum momento na seção que trata do Método de Cálculo do Fator de Proximidade Social a fim de que sirvam de referência à determinação desses.

Desse modo, não se empregou, no SINAPSE, tal fator na identificação de candidatos para a conformação de uma equipe.

2.8 Conclusão

Antes de se chegar ao modelo escolhido para servir de base à concepção do SINAPSE, cogitaram-se outros dois modelos também com essa finalidade. Após análise, contudo, ambos foram considerados inapropriados e, na seção 2.1, justificou-se o porquê.

Examinou-se, então, a fundamentação teórica do modelo computacional selecionado e proposto por Reinhardt (2008), o qual é capaz de identificar indivíduos, com a intenção de formar uma equipe de pesquisa, através da representação da rede egocêntrica do líder de projeto e demais integrantes da rede social.

A noção de proximidade na rede social egocêntrica é dada pelo fator de proximidade individual, obtido a partir do cálculo das médias ponderadas das relações em função do tipo (coautoria, relações orientador-orientado) e do período em que ocorreram.

Do modelo, entretanto, o fator de proximidade social é rejeitado e a razão disso se encontra na seção 2.7.

Por fim, o início do próximo capítulo trata dos devidos ajustes pelos quais teve de passar o modelo computacional teórico para constituir o cerne do SINAPSE.

3 DESENVOLVIMENTO

Este capítulo apresenta o desenvolvimento e funcionamento do SINAPSE.

3.1 *Considerações e Ajustes do Modelo*

Como visto no capítulo 2, o modelo computacional proposto por Reinhardt (2008) tem como entrada o grupo de pesquisadores que serão analisados e as informações disponíveis sobre os mesmos na Plataforma Lattes. Um grupo desses é facilmente consultado via site do Diretório dos Grupos de Pesquisa no Brasil; tal diretório contém, para cada grupo, a área de pesquisa, a lista de nomes dos integrantes e o papel que cada um desempenha nesse. De posse de um grupo, determinam-se os parâmetros da rede social egocêntrica do líder do grupo que possam ser utilizados na valoração de relações sociais para orientar a indicação dos demais pesquisadores desse grupo com o intuito de formação de uma equipe de pesquisa.

O SINAPSE, sem embargo, objetiva não somente ser um software utilizado em instituições educacionais, mas também em empresas. Sendo assim, um Coordenador, responsável por um grupo de pessoas de algum setor, usaria o SINAPSE a fim de, dentre os possíveis Candidatos desse grupo, conformar a Equipe referente a um Projeto. Perceba-se que o termo Candidato é, então, empregado em detrimento de pesquisador, pois se expande o escopo de atuação do sistema.

Naturalmente, o período de cinco anos de 2003 a 2007, considerado quando da construção de uma matriz relacional, foi atualizado para o de 2008 a 2012.

Ademais, no decorrer deste capítulo, o Fator de Proximidade Individual é renomeado/simplificado para Fator de Proximidade. Tal Fator, por valorar quão "próximos" são dois indivíduos com relação aos aspectos sociais de familiaridade, proximidade geográfica e homofilia, é tido como o Potencial de Sinergia entre esses mesmos indivíduos.

3.1.1 **Definição do Candidato mais Apto**

Enquanto a primeira etapa de funcionamento do modelo de Reinhardt (2008) se baseia em um *ego* já definido como sendo o líder do grupo de pesquisa analisado, o SINAPSE requer do Coordenador a inserção das Competências necessárias ao âmbito do Projeto para, de posse dos Currículos Lattes de cada

Candidato suscetível a integrar a Equipe, quantificar a ocorrência dessas Competências nesses Currículos.

Dessa forma, o Candidato, cujo Currículo tenha a maior quantidade de ocorrências, é considerado o Candidato mais apto, o de melhor perfil, e, por conseguinte, o ego da rede social egocêntrica a ser obtida.

3.1.2 Automatização do Processo de Obtenção de Informações

De acordo a Reinhardt (2008), a obtenção e processamento de informações relacionais sem a intervenção de um usuário é um dos desafios na análise de redes sociais. A análise de grandes redes sociais é limitada por aspectos relativos à fadiga gerada pela construção manual de longas listas de indivíduos e das ligações existentes entre esses.

Nesse sentido, o SINAPSE trata de ser original ao eliminar essa fadiga. Não tanto por auferir a rede social egocêntrica a partir da matriz relacional do Candidato mais apto, visto que essa etapa de cálculos não representa um grande desafio computacional, mas sim por automatizar a extração de informações da Plataforma Lattes a fim de elaborar tal matriz relacional.

Assim, desenvolveu-se o SINAPSE de forma a simular, basicamente, uma pessoa visitando o site Buscar Currículo Lattes com a intenção de conseguir os dados da Plataforma Lattes necessários à determinação do Candidato mais apto e à posterior criação da matriz relacional desse mais apto concernente aos demais Candidatos suscetíveis à formação da Equipe. Mais detalhes na seção 3.3.

3.2 Arquitetura

O SINAPSE é uma aplicação web de arquitetura cliente-servidor. Foi desenvolvida em Linguagem de Programação C# empregando-se o ambiente de desenvolvimento integrado (IDE na sigla em inglês) Microsoft Visual Studio Express 2012 para a Web (com base no framework ASP.NET 4.5) e se baseia em Formulários Web ASP.NET.

O diagrama de classes da Figura 5 mostra as classes que representam as entidades do sistema:

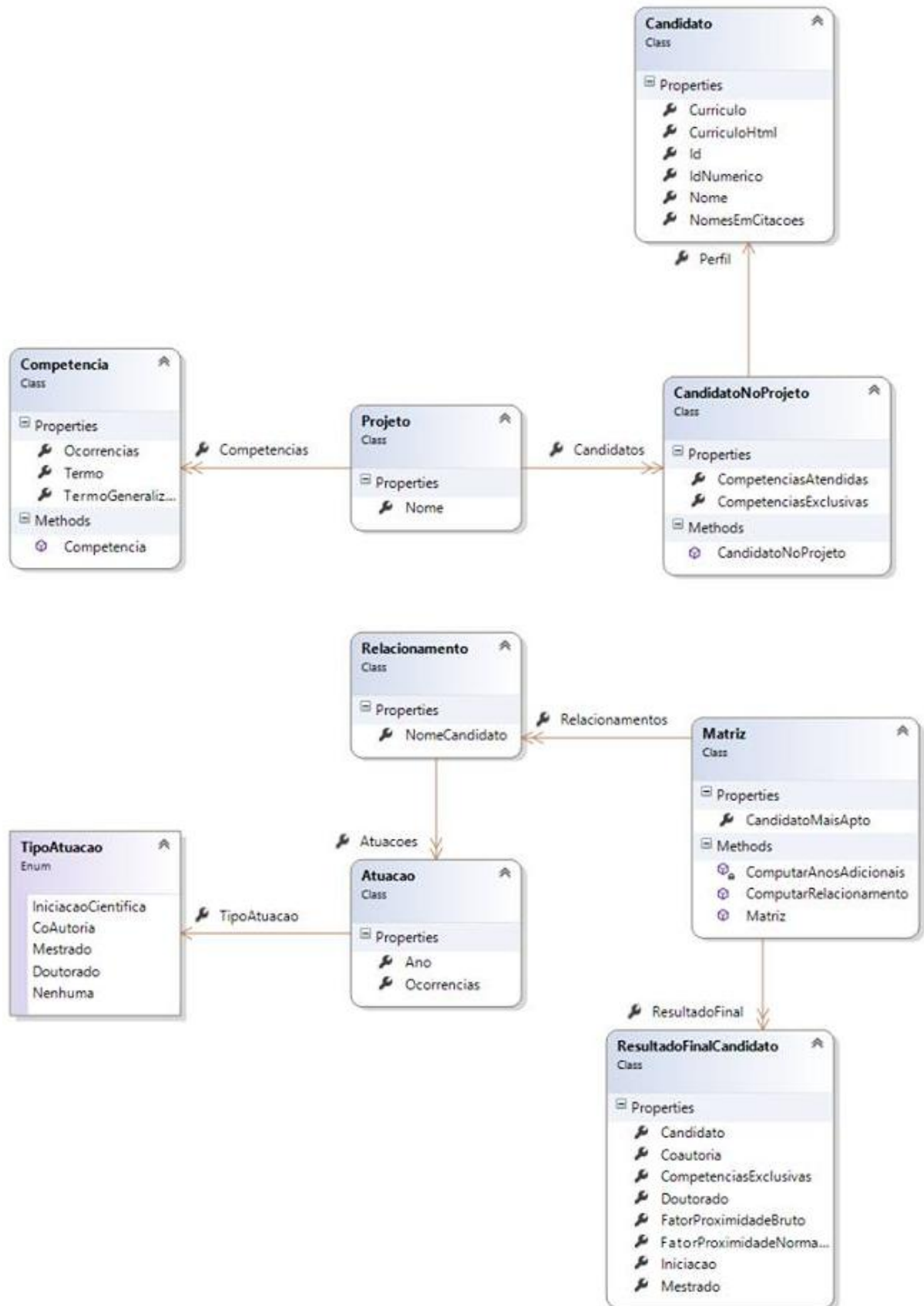


Figura 5 - Diagrama de classes

Fonte: Autoria própria.

A classe **Equipe** (da página **Equipe.aspx**) teria o papel de Camada de Apresentação (*User Interface*), de Camada de Negócios (*Business Logic Layer*) e de Camada de Acesso a Dados (*Data Access Layer*) do SINAPSE.

A Plataforma Lattes, por fim, seria o banco de dados empregado pelo sistema.

3.3 Modo de Operação

Nesta seção, apresenta-se o modo como o SINAPSE funciona, explicitando-se os principais *scripts*, métodos, classes e eventos do código desenvolvido que permitem tal funcionamento.

A Figura 6 mostra a tela inicial do SINAPSE:

A imagem mostra a interface de usuário do SINAPSE. No topo, há o logotipo "Sinapse" em um círculo. Abaixo, há um formulário com os seguintes elementos:

- Nome do Projeto:** Um campo de texto vazio.
- Competências Necessárias:** Um campo de texto vazio com um botão "Adicionar" ao lado.
- Busca de Candidatos:** Um campo de texto vazio com um botão "Pesquisar" ao lado.
- Formar Equipe:** Um botão destacado.
- Competências Seleccionadas:** Uma área vazia com uma linha horizontal de separação.
- Candidatos:** Uma área vazia com uma linha horizontal de separação.

No rodapé da interface, há o texto "© 2013 - Sinapse".

Figura 6 - Tela inicial do SINAPSE

Fonte: Autoria própria.

O Coordenador dá um nome ao Projeto, digita as Competências e, para cada uma dessas, clica no botão Adicionar. Esse botão executa uma chamada de um *javascript adicionarCompetencia*, o qual, empregando o componente *jQuery Tokeninput* (licença GPL ou MIT) desenvolvido por Smith (2011), adiciona cada Competência à lista Competências Seleccionadas. Essas Competências serão mais tarde lidas e enviadas ao servidor para processamento.

The screenshot shows the 'Sinapse' web application interface. At the top left, the logo 'Sinapse' is displayed. Below it, there is a form with several sections:

- Nome do Projeto:** A text input field containing the value 'GIVIC'.
- Competências Necessárias:** A text input field containing 'Geomática' and a button labeled 'Adicionar'.
- Competências Selecionadas:** A list area showing 'Engenharia Biomédica' with a close button (x).
- Busca de Candidatos:** A text input field and a button labeled 'Pesquisar'.
- Candidatos:** A section header for the candidate list, which is currently empty.
- Formar Equipe:** A button located at the bottom left of the form.

Figura 7 - Tela ao serem adicionadas Competências

Fonte: Autoria própria.

Em seguida, o Coordenador digita o nome de um Candidato e clica no botão Pesquisar. Executa-se, assim, o método **BuscarCandidatos**, o qual, empregando a classe **WebRequest**, faz uma requisição HTTP a simular uma pessoa buscando o Candidato em questão utilizando o site Buscar Currículo Lattes da Plataforma Lattes. O retorno dessa requisição HTTP é um objeto com diversos dados relativos a essa, inclusive o código-fonte que o navegador exibiria. Neste caso, converte-se esse código em uma *string* e, com uso do *parser Html Agility Pack* (licença Ms-PL) desenvolvido por Mourier, DarthObiwan e Jessynoo (2012), reconverte-se essa em um objeto a conter o HTML estruturado (similar a XML) para fácil manipulação via LINQ (que é um meio de rodar uma *query* em uma estrutura de dados). Obtém-se, então, uma lista dos Candidatos para a *grid* de resultados, como mostra a Figura 8.

Competências Necessárias

Competências Selecionadas

Busca de Candidatos

Candidatos

Adicionar à Equipe	Nome	Formação
<input type="checkbox"/>	Elmeri Esteves Homem	Graduação em Ciências Econômicas pela Universidade de Caxias do Sul, Brasil(2011)
<input type="checkbox"/>	Leyza Elmeri Baldo Dorini	Doutorado em Ciência da Computação pela Universidade Estadual de Campinas, Brasil(2009)
Candidatos encontrados: 2		

Figura 8 - Tela que exibe o *grid* de Candidatos encontrados

Fonte: Autoria própria.

Como no código-fonte os dados dos Candidatos estão contidos em elementos HTML ``, a busca efetuada no objeto resultado do *parsing* usa esse padrão. O resultado é uma lista, a qual é retornada pelo método **BuscarCandidatos**. Se o resultado não for nulo, o método **BindGrid** é executado para que esse seja exibido no *grid*. No entanto, se for nulo, a mensagem "Nenhum candidato encontrado com esse nome." é exibida.

Nenhum candidato encontrado com esse nome.

Nome do Projeto

Competências Necessárias

Competências Selecionadas

Busca de Candidatos

Candidatos

Figura 9 - Tela com exemplo de exibição de mensagem

Fonte: Autoria própria.

Se o resultado da pesquisa exibe o Candidato desejado, o Coordenador clica na caixa de seleção à esquerda para adicionar o Candidato à lista Candidatos. No evento **onclick** dessa caixa de seleção, a função **adicionarCandidato** em *javascript* é chamada. Essa utiliza o mesmo *Tokeninput* citado acima (e utilizado para a lista de Competências Seleccionadas) com o fim de preencher a lista Candidatos. Quando essa adição à lista acontece, além do nome ser armazenado e exibido, guarda-se também o ID (*identifier*) do Currículo Lattes; este, porém, não é exibido e será posteriormente enviado ao servidor.

The screenshot displays a web interface for candidate selection. It features a search bar labeled 'Busca de Candidatos' with a 'Pesquisar' button. Below the search bar is a 'Formar Equipe' button. To the right, under the 'Candidatos' section, two candidate names are listed: 'Leyza Elmeri Baldo Dorini' and 'Luís Alberto Lucas'. Below this is a table with columns for 'Adicionar à Equipe', 'Nome', and 'Formação'. The table shows one candidate, 'Luís Alberto Lucas', with a checked checkbox in the 'Adicionar à Equipe' column. The 'Formação' column contains the text: 'Doutorado em Engenharia Elétrica e Informática Industrial pela Universidade Tecnológica Federal do Paraná, Brasil(2009)'. At the bottom of the table, it indicates 'Candidatos encontrados: 1'.

Adicionar à Equipe	Nome	Formação
<input checked="" type="checkbox"/>	Luís Alberto Lucas	Doutorado em Engenharia Elétrica e Informática Industrial pela Universidade Tecnológica Federal do Paraná, Brasil(2009)
Candidatos encontrados: 1		

Figura 10 - Tela ao serem adicionados Candidatos

Fonte: Autoria própria.

Finalmente, clica-se no botão Formar Equipe, que enviará todos os dados ao servidor para processamento e posterior retorno do resultado final, com a lista dos Candidatos a serem analisados para conformar a Equipe, como mostra a Figura 11.

The screenshot shows a web form with the following sections:

- Nome do Projeto:** A text input field containing "GIVIC".
- Competências Necessárias:** An empty text input field with an "Adicionar" button next to it.
- Competências Seleccionadas:** A list of selected competencies, each in a button with an 'x' icon: "Engenharia Biomédica", "Geomática", "Modelagem, controle, otimização e supervisão de processos industriais", "Processamento e análise de imagens", and "Classificadores fuzzy".
- Busca de Candidatos:** An empty text input field with a "Pesquisar" button next to it.
- Candidatos:** A list of selected candidate names, each in a button with an 'x' icon: "Leyza Elmeri Baldo Dorini", "Luís Alberto Lucas", "Myriam Regattieri De Biase da Silva Delgado", "Rodrigo Minetto", "Tania Mezzadri Centeno", "Allan Palmerio Vieira", "André Donato Zanon", "Diego de Faria do Nascimento", "Guilherme da Silva Violada", "Lucas Chociay", "Neiton Arlindo Soares", and "Thayse Marques Solis".
- Formar Equipe:** A button located at the bottom left of the form.

Figura 11 - Tela ao ser pressionado o botão Formar Equipe

Fonte: Autoria própria.

Essa parte começa no cliente: antes de os dados serem enviados ao servidor, esses são tratados no cliente. O botão Formar Equipe chama a função *javascript* **FillJson**, que formata o que vai ser enviado ao servidor (Competências e Candidatos) para que seja corretamente recebido e tratado. Já no servidor, quando os dados são recebidos, instancia-se um objeto da classe **Projeto** através da chamada ao método **CriarProjeto**. Por meio da classe **JavaScriptSerializer**, do .Net, os dados que vêm do cliente são transferidos para listas de objetos das classes **Candidato** e **Competencia**. Em outros pontos do sistema e neste momento, quando a lista de Competências é criada, o método **RemoverDiacriticos** é executado ao final de uma *string*. Dessa forma, os diacríticos são removidos para que as diversas comparações entre *strings* funcionem corretamente. É feito, então, o uso do componente de *stemming* **PTStemmer** (licença GNU Lesser GPL) desenvolvido por Oliveira (2010), que possui a classe **OrengoStemmer**. Através do método **GetWordStem**, obtém-se a raiz de palavras em diversas variações. Por exemplo, a palavra Robótica seria transformada em "robo", o que permite que variações da palavra Robótica sejam encontradas nos Currículos. Essas palavras "estemizadas" são armazenadas na propriedade **TermoGeneralizado** do objeto **competencia**, que contém a variação original do termo.

É hora de recuperar o Currículo de cada Candidato junto ao Lattes. Em um processo semelhante à busca de Candidatos, um HTTP *request* é feito, através da classe **WebRequest**, a partir do método **ObterCurrículoEmHtml**, o qual recebe como parâmetro o ID do Currículo do Candidato e retorna o código-fonte do Currículo já no formato HTML do *Html Agility Pack*. Transforma-se também esse HTML em uma *string* armazenada na propriedade **Currículo**. O motivo para essas duas conversões é o fato de certas buscas, como a de termos, por exemplo, ser facilmente executada sobre uma *string*. Já na busca de nomes de Candidatos e outros elementos do Currículo, esses são mais facilmente encontrados em um objeto bem estruturado, conforme o *Html Agility Pack* fornece. Logo em seguida, os termos ou Competências contidos no Currículo de cada Candidato são contabilizados. Também o método **ExtrairNomesEmCitacoes** é invocado para armazenar, no perfil do Candidato, as variações de nomes que pode haver para ele em "Nome em citações bibliográficas" do seu Currículo. Finalizado esse processo, o objeto **projeto** é retornado.

O retorno do método **CriarProjeto** é então passado como parâmetro para o próximo método a ser invocado: **ExtrairMatriz**. Tal método começa escolhendo o melhor perfil, o Candidato mais apto, com base na contagem de ocorrências das Competências necessárias ao Projeto. Instancia-se, então, um objeto da classe **Matriz**, que conterá os relacionamentos entre Candidatos, o resultado final e proverá, também, métodos que realizarão cálculos conforme as regras de negócio.

Depois, as Produções que o Candidato mais apto (*ego*) e o outro Candidato (*alter*) efetuaram juntos são obtidas por meio de uma consulta LINQ e com base no elemento HTML que contém um atributo com valor **ProducoesCientificas**. O valor e o atributo são obtidos através da observação do código-fonte do Currículo Lattes. Quando um relacionamento de Coautoria é encontrado, o método **ComputarRelacionamento** do objeto **matriz** é chamado e armazena uma indicação de que no ano X o *alter* trabalhou com o *ego*. Para encontrar esse relacionamento, é verificada, inicialmente, a existência de um link para o Lattes do *alter* em questão (através do seu "Endereço para acessar este CV" contido no link) no Lattes do *ego*. Caso não haja link, uma segunda tentativa é feita: de posse do nome do *alter* e das variações desse nome em "Nome em citações bibliográficas", buscam-se esses no Lattes do *ego*.

A seguir, um processo muito semelhante é realizado para computar relacionamentos Orientador-Orientado em que o *ego* orientou ou esteja orientando o *alter*. As diferenças principais são em relação aos valores buscados em atributos, que variam de acordo ao que se procura (Iniciação Científica, Mestrado ou Doutorado) e a como a contagem de anos para cada relacionamento Orientador-Orientado é feita, com base no fato dessa estar em andamento ou não.

Após todos os relacionamentos terem sido registrados, é feita uma busca via LINQ para encontrar Candidatos que sejam detentores exclusivos de uma Competência a fim de que, por mais que o Potencial de Sinergia do *alter* seja baixo, esse *alter* seja tomado em consideração por poder contribuir com diversidade à Equipe. Levantados todos esses dados, o método **CalcularResultadoFinal** é chamado. Esse método varre os relacionamentos já armazenados no objeto **matriz** e, para cada um, efetua os cálculos pertinentes, conforme as regras de negócio (cálculos da seção 2.6 para se determinar o Fator de Proximidade). Ao fim dos cálculos, obtém-se uma lista de objetos do tipo **ResultadoFinalCandidato**. Pronta a lista, essa é complementada com a Competência exclusiva que um Candidato possa ter. Por fim, coloca-se o nome do Candidato mais apto na lista final e os resultados são exibidos ao Coordenador em uma janela *pop-up*. Se o Coordenador quiser, ele pode imprimir essa lista, como mostra a Figura 12.

Eis a lista, ordenada por potencial de cooperação com relação àquele mais apto de acordo às competências escolhidas, dos candidatos disponíveis à conformação da equipe referente ao projeto GIVIC:

Melhor perfil: Tania Mezzadri Centeno

Candidato	Potencial de Sinergia	Competência Exclusiva
Luís Alberto Lucas	34,88%	Classificadores fuzzy
Myriam Regattieri De Biase da Silva Delgado	25,58%	-
Allan Palmerio Vieira	16,28%	-
Leyza Elmeri Baldo Dorini	11,63%	-
Guilherme da Silva Violada	5,81%	-
Lucas Chociay	5,81%	-
Rodrigo Minetto	0,00%	-
André Donato Zanon	0,00%	-
Diego de Faria do Nascimento	0,00%	-
Neiton Arlindo Soares	0,00%	-
Thayse Marques Solis	0,00%	-

Imprimir

Figura 12 - Candidatos a serem analisados para conformar a Equipe

Fonte: Autoria própria.

Com relação ao resultado do exemplo utilizado nesta seção, não bastasse o fato de haver obtido o maior Potencial de Sinergia com relação àquela que foi definida como a de melhor perfil, o Candidato Luís Alberto Lucas é o único a ter a Competência Classificadores Fuzzy. Caso ele tivesse um Potencial de Sinergia de 0% (como o Candidato Rodrigo Minetto), ainda assim o Coordenador deveria tê-lo em conta ao decidir a conformação da Equipe para promover a diversidade nessa.

Pois bem, pela avaliação do SINAPSE, caso o Coordenador desejasse formar uma Equipe, referente ao Projeto GIVIC, de apenas cinco integrantes (com relação aos 12 disponíveis), esses seriam a Tania Mezzadri Centeno, o Luís Alberto Lucas, a Myriam Regattieri De Biase da Silva Delgado, o Allan Palmerio Vieira e a Leyza Elmeri Baldo Dorini.

3.4 Mensagens

Além da mensagem "Nenhum candidato encontrado com esse nome." já mencionada, as demais que o SINAPSE emite são estas:

- "Digite o nome do candidato.": quando o Coordenador pressiona o botão Pesquisar sem haver digitado o nome de um Candidato;
- "Não foi possível contatar a Plataforma Lattes. O site pode estar indisponível. Aguarde alguns minutos e tente novamente.": quando o SINAPSE não consegue acessar o site Buscar Currículo Lattes;
- "Para formar uma equipe são necessários, ao menos, três candidatos e uma competência.": quando o Coordenador pressiona o botão Formar Equipe sem ter seguido as condições que a mensagem acusa;
- "Não há relações de coautoria ou de orientador-orientado entre o candidato de melhor perfil e os demais candidatos.";
- "Nenhum candidato atende às competências indicadas para o projeto.": nenhum Candidato possui um Currículo Lattes com qualquer uma das Competências que o Projeto requer.

3.5 Testes

A fim de verificar se o SINAPSE fornece resultados coerentes, realizaram-se alguns testes com esse intuito.

3.5.1 Metodologia dos Testes

Encontrava-se, por meio do site do Diretório dos Grupos de Pesquisa no Brasil, como já mencionado na seção 3.1, um grupo de pesquisa da UTFPR como o GIVIC, utilizado no exemplo da seção 3.3. A partir de um grupo desses, tal qual foi construída a matriz do Quadro 2 da seção 2.6, fazia-se manualmente, via Excel, a matriz relacional do pesquisador líder do grupo. Em seguida, construía-se um quadro com os valores da relação de proximidade da rede social egocêntrica tal qual o Quadro 3 da seção 2.6.

Depois, em uma sessão do SINAPSE, o Nome do Projeto era o do grupo de pesquisa, as Competências Necessárias eram, normalmente, as linhas de pesquisa desse grupo, e os Candidatos, geralmente, os recursos humanos. Por causa desse procedimento, o pesquisador líder do grupo acabaria sendo, muito provavelmente, o Candidato de melhor perfil.

industriais; Processamento e análise de imagens; Classificadores fuzzy;

- Candidatos: Leyza Elmeri Baldo Dorini; Rodrigo Minetto; Luís Alberto Lucas; Tania Mezzadri Centeno; Myriam Regattieri De Biase da Silva Delgado; Allan Palmerio Vieira; Lucas Chociay; André Donato Zanon; Diego de Faria do Nascimento; Neiton Arlindo Soares; Guilherme da Silva Violada; Thayse Marques Solis.

Assim, como dito na seção anterior, faz-se manualmente, via Excel, a matriz relacional do pesquisador líder do grupo (Tania Mezzadri Centeno):

Tania Mezzadri Centeno	COAUTORIA				ORIENTADOR-MESTRANDO					ORIENTADOR-DOUTORANDO					ORIENTADOR INICIAÇÃO CIENTÍFICA							
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012		
Leyza Elmeri Baldo Dorini					1																	
Luís Alberto Lucas	2	1			1						1	1										
Myriam Regattieri De B. da S. Delgado	2	2			1																	
Rodrigo Minetto																						
Allan Palmerio Vieira					2																1	
André Donato Zanon																						
Diego de Faria do Nascimento																						
Guilherme da Silva Violada																					1	
Lucas Chociay																					1	
Neiton Arlindo Soares																						
Thayse Marques Solis																					1	1

Quadro 6 - GIVIC: Matriz relacional obtida manualmente

Fonte: Autoria própria.

Depois, constrói-se um quadro com os valores da relação de proximidade da rede social egocêntrica:

Tania Mezzadri Centeno	coautoria	mestrado	doutorado	iniciação	P	P						
Leyza Elmeri Baldo Dorini	0,333333333	0	0	0	0,066666667	0,095238095						
Luís Alberto Lucas	0,6	0	0,2	0	0,2	0,285714286						
Myriam Regattieri De B. da S. Delgado	0,733333333	0	0	0	0,146666667	0,20952381						
Rodrigo Minetto	0	0	0	0	0	0						
Allan Palmerio Vieira	0,666666667	0	0	0,266666667	0,16	0,228571429						
André Donato Zanon	0	0	0	0	0	0						
Diego de Faria do Nascimento	0	0	0	0	0	0						
Guilherme da Silva Violada	0	0	0	0,333333333	0,033333333	0,047619048						
Lucas Chociay	0	0	0	0,333333333	0,033333333	0,047619048						
Neiton Arlindo Soares	0	0	0	0	0	0						
Thayse Marques Solis	0	0	0	0,6	0,06	0,085714286						
	relação:	2008	2009	2010	2011	2012	coautoria	mestrado	doutor.	iniciação	sigma	Pi
	peso:	1	2	3	4	5	2	3	4	1	0,7	

Quadro 7 - GIVIC: Rede social egocêntrica obtida manualmente

Fonte: Autoria própria.

Ao serem comparados os valores de $|P|$ (Fator de Proximidade normalizado ou Potencial de Sinergia) do Quadro 7 com os da lista resultante da Figura 12 do exemplo da seção 3.3, percebe-se a divergência desses.

Por meio da matriz de teste da Figura 13, nota-se que, enquanto nessa há somente uma ocorrência de Coautoria, no ano de 2012, entre o Allan Palmerio Vieira e a Candidata de melhor perfil (Tania Mezzadri Centeno), na matriz relacional do Quadro 6 há duas. Já para a Thayse Marques Solis, na matriz de teste, não são computadas as ocorrências de Orientação de Iniciação Científica dos anos de 2011 e de 2012, as quais sim aparecem na matriz relacional.

Examinando-se o porquê dessas discrepâncias, percebe-se que, no Lattes da Tania, com relação ao Allan, na ocorrência de Coautoria não encontrada pelo SINAPSE, ele está como "VIEIRA, ALLAN", diferindo de "Allan Palmerio Vieira" (Nome) e de "VIEIRA, A. P." (Nome em citações bibliográficas), como consta no Lattes dele. No caso da Thayse, seu sobrenome está erroneamente grafado no Lattes da Tania: está como Sollis (duas letras "l") em vez de Solis (uma letra "l").

No entanto, essas divergências, decorrentes de falha humana, relacionadas a esses dois Candidatos, apesar de exporem uma limitação do SINAPSE, poderiam ser facilmente retificadas caso links para seus Lattes fossem inseridos nessas ocorrências no Lattes da Tania.

Para os demais Candidatos, os dados da matriz relacional feita manualmente coincidem com os da matriz de teste.

Em suma, o resultado deste teste foi considerado coerente no tocante ao modo como o SINAPSE opera.

3.5.3 Teste com o MEMENTO

Dados de entrada:

- Nome do Projeto: MEMENTO;
- Competências Necessárias: Ontologias; Redes sociais; Simulação; Sistemas Colaborativos; Sistemas multi-agentes;
- Candidatos: Ademir Roberto Freddo; Gustavo Alberto Giménez Lugo; Cesar Augusto Tacla; Jomi Fred Hübner; Ana Luiza Kaori Waki; Josiane Michalak Hauagge Dall Agnol; Emmerson Santa Rita da Silva; Maurício Chicalski; Guilherme Lincon Barczynsyn; Raphael Jardim Lopes; Josiane Melchiori Pinheiro Ferreira; Roni Fabio Banaszewski.

Faz-se manualmente, então, a matriz relacional do pesquisador líder do grupo (Cesar Augusto Tacla):

Cesar Augusto Tacla	COAUTORIA					ORIENTADOR-MESTRANDO					ORIENTADOR-DOUTORANDO					ORIENTADOR INICIAÇÃO CIENTÍFICA				
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012
	Ademir Roberto Freddo	2	4	2	2	2						1	1	1						
Gustavo Alberto Giménez Lugo	1	1			1															
Jomi Fred Hübner																				
Ana Luiza Kaori Waki																				1
Emmerson Santa Rita da Silva													1	1	1					
Guilherme Lincon Barczynsyn																				1
Josiane Melchiori Pinheiro Ferreira					2									1	1					
Josiane M. Hauagge Dall Agnol			1	1								1	1	1	1					
Maurício Chicalski																	1			1
Raphael Jardim Lopes																				1
Roni Fabio Banaszewski		2	3	2	7	1	1					1	1	1	1					

Quadro 8 - MEMENTO: Matriz relacional obtida manualmente

Fonte: Autoria própria.

Constrói-se, em seguida, um quadro com os valores da relação de proximidade da rede social egocêntrica:

Cesar Augusto Tacla	COAUTORIA					MESTRADO					DOUTORADO					INICIAÇÃO					P					P					
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012	
Ademir Roberto Freddo	1,6					0					0,4					0					0,48					0,156182213					
Gustavo Alberto Giménez Lugo	0,533333333					0					0					0					0,106666667					0,034707158					
Jomi Fred Hübner	0					0					0					0					0					0					
Ana Luiza Kaori Waki	0					0					0					0,333333333					0,033333333					0,010845987					
Emmerson Santa Rita da Silva	0					0					0,8					0					0,32					0,104121475					
Guilherme Lincon Barczynsyn	0					0					0					0,333333333					0,033333333					0,010845987					
Josiane Melchiori Pinheiro Ferreira	0,666666667					0					0,6					0					0,373333333					0,121475054					
Josiane Michalak Hauagge Dall Agnol	0,466666667					0					0,933333333					0					0,466666667					0,151843818					
Maurício Chicalski	0					0					0					0,466666667					0,046666667					0,015184382					
Raphael Jardim Lopes	0					0					0					0,333333333					0,033333333					0,010845987					
Roni Fabio Banaszewski	3,733333333					0,2					0,933333333					0					1,18					0,383947939					
	relação:	2008	2009	2010	2011	2012	coautoria	mestrado	doutor.	iniciação	sigma_Pi																				
	peso:	1	2	3	4	5	2	3	4	1	3,073333																				

Quadro 9 - MEMENTO: Rede social egocêntrica obtida manualmente

Fonte: Autoria própria.

Depois, obtém-se a lista resultante da sessão do SINAPSE, como mostra a Figura 14:

Eis a lista, ordenada por potencial de cooperação com relação àquele mais apto de acordo às competências escolhidas, dos candidatos disponíveis à conformação da equipe referente ao projeto MEMENTO:

Melhor perfil: Cesar Augusto Tacla

Candidato	Potencial de Sinergia	Competência Exclusiva
Roni Fabio Banaszewski	40,14%	-
Ademir Roberto Freddo	16,33%	-
Josiane Michalak Hauagge Dall Agnol	15,87%	-
Emmerson Santa Rita da Silva	10,88%	-
Josiane Melchiori Pinheiro Ferreira	8,16%	-
Gustavo Alberto Giménez Lugo	3,63%	-
Maurício Chicalski	1,59%	-
Guilherme Lincon Barczyszyn	1,13%	-
Raphael Jardim Lopes	1,13%	-
Ana Luiza Kaori Waki	1,13%	-
Jomi Fred Hübner	0,00%	-

Imprimir

Figura 14 - MEMENTO: Candidatos a serem analisados para conformar a Equipe

Fonte: Autoria própria.

Comparando-se os valores de $|P|$ (Fator de Proximidade normalizado ou Potencial de Sinergia) do Quadro 9 com os da lista resultante da Figura 14, nota-se a discordância desses.

Dessa forma, extrai-se a matriz de teste da Figura 15:

Melhor perfil: Cesar Augusto Tacla

Nome	Tipo	Ano	Ocorrencias
Ademir Roberto Freddo	CoAutoria	2008	2
Ademir Roberto Freddo	CoAutoria	2009	4
Ademir Roberto Freddo	CoAutoria	2010	2
Ademir Roberto Freddo	CoAutoria	2011	2
Ademir Roberto Freddo	Doutorado	2008	1
Ademir Roberto Freddo	Doutorado	2009	1
Ademir Roberto Freddo	Doutorado	2010	1
Ana Luiza Kaori Waki	IniciacaoCientifica	2012	1
Emmerson Santa Rita da Silva	Doutorado	2010	1
Emmerson Santa Rita da Silva	Doutorado	2011	1
Emmerson Santa Rita da Silva	Doutorado	2012	1
Guilherme Lincon Barczyszyn	IniciacaoCientifica	2012	1
Gustavo Alberto Giménez Lugo	CoAutoria	2008	1
Gustavo Alberto Giménez Lugo	CoAutoria	2009	1
Gustavo Alberto Giménez Lugo	CoAutoria	2012	1
Josiane Melchiori Pinheiro Ferreira	Doutorado	2011	1
Josiane Melchiori Pinheiro Ferreira	Doutorado	2012	1
Josiane Michalak Hauagge Dall Agnol	CoAutoria	2010	1
Josiane Michalak Hauagge Dall Agnol	CoAutoria	2011	1
Josiane Michalak Hauagge Dall Agnol	Doutorado	2009	1

Figura 15 - MEMENTO: Matriz de teste

Fonte: Autoria própria.

Graças à matriz de teste, percebe-se que, enquanto nessa não consta nenhuma ocorrência de Coautoria, no ano de 2012, entre a Josiane Melchiori Pinheiro Ferreira e o Candidato de melhor perfil (Cesar Augusto Tacla), na matriz relacional do Quadro 8 há duas.

Analisando-se o motivo dessa discrepância, nota-se que, no Lattes do Tacla, com relação à Josiane, nas duas ocorrências de Coautoria não encontradas pelo SINAPSE, ela está como "Ferreira, J. M. P.", diferindo de "Josiane Melchiori Pinheiro Ferreira" (Nome) e de "FERREIRA, Josiane Melchiori Pinheiro" ou de "Ferreira, Josiane M. P." (Nomes em citações bibliográficas), como consta no Lattes dela.

Assim como no Teste com o GIVIC, entretanto, essa divergência, decorrente novamente de falha humana, poderia ser facilmente corrigida se links para o Lattes da Josiane fossem inseridos nessas ocorrências no Lattes do Tacla.

Para os demais Candidatos, os dados da matriz relacional feita manualmente coincidem com os da matriz de teste.

O resultado deste teste, portanto, também foi considerado coerente com relação ao modo como o SINAPSE opera.

3.6 Conclusão

Mediante o exemplo da seção 3.3, verificou-se que o SINAPSE cumpre o objetivo de permitir a um Coordenador avaliar quais são os Candidatos capazes de, no contexto de um Projeto, conformarem uma Equipe na qual realmente haja cooperação e diversidade.

Ocorre, todavia, a limitação, citada durante os testes e decorrente da inexistência de um link ou da falta de correspondência relacionada aos nomes, que não permite uma correta determinação do Potencial de Sinergia entre o Candidato afetado por essa e o de melhor perfil. A fim de contornar essa situação, haveria a possibilidade de tentar implementar algum modo de indagar o Coordenador quando da eventualidade de nomes muito similares. Outra forma seria a de exigir sempre ou a inclusão de um link ou a correta escrita do nome no momento de preencher um Currículo Lattes.

Ademais, pela seção 3.4, sabe-se que o SINAPSE informa o Coordenador quando não consegue contatar a Plataforma Lattes. Isso, contudo, significa que o sistema ficará inoperante enquanto essa estiver inacessível ou, ainda, se a Plataforma Lattes estiver com algum problema, esse também afetará o SINAPSE.

Finalmente, para que uma empresa pudesse usufruir completamente da funcionalidade do SINAPSE, essa teria de possuir uma base de dados tão rica de informações quanto a Plataforma Lattes. Ficaria a questão, entretanto, de analisar como ficariam as relações de Coautoria e de Orientador-Orientado no interior dessa empresa. Quiçá pudessem ser, respectivamente, algo como relações de Trabalho/Projeto Conjunto e de Mentor-Trainee.

4 GESTÃO

4.1 Cronograma

O Quadro 10 mostra o cronograma do TCC, levando em consideração o exposto nos demais capítulos deste Relatório.

O total de horas efetivo e disposto nesse encontra-se disperso em oito meses, de fevereiro a setembro deste ano. A razão de tanto espalhamento se deve a uma série de transtornos de ordem pessoal que acometeram o autor deste TCC.

Atividade	Tempo (h)
Leitura/Compreensão do livro <i>La Dynamique de la Confiance</i>	30
Leitura/Compreensão da dissertação <i>Modelo Computacional de Percepção de Contextos de Atividade para Identificação de Comunidades</i>	10
Leitura/Compreensão da dissertação <i>Modelo Computacional para Formação de Equipes baseado nas Redes Egocêntricas de Líderes de Projeto</i>	15
Análise de requisitos	10
Pesquisa acerca do parser <i>Html Agility Pack</i> e dos componentes <i>jQuery Tokeninput</i> e <i>PTStemmer</i>	12
Desenvolvimento da Interface Gráfica	15
Codificação	390
Testes e ajustes	40
Total	522

Quadro 10 - Cronograma

Fonte: Autoria própria.

4.2 Conclusão

Desafortunadamente, não houve condições de se definir antecipadamente um cronograma para que, neste capítulo, pudesse ser feita uma comparação entre o planejado com o de fato executado.

Todavia, apesar de todas as atividades do Quadro 10 terem sido cumpridas, percebe-se que, caso um bom planejamento houvesse ocorrido, a dispersão

mencionada poderia ter sido mitigada. Quando com problemas, contudo, põe-se de lado, frequentemente, a racionalidade.

5 CONCLUSÕES

O objetivo geral deste TCC era o de desenvolver um sistema informatizado visando estimar o potencial de cooperação de determinado grupo de pessoas no âmbito de um projeto com o propósito de formar uma equipe. Para tanto, foram analisados três modelos com o intuito de constituir o cerne desse sistema.

O primeiro a ser examinado foi o método denominado PAT-Miroir (LE CARDINAL; GUYONNET; POUZOULLIC, 1997), o qual foi deixado de lado, pois o sistema informatizado a derivar-se desse seria basicamente um ambiente a promover uma atuação recíproca dos membros do grupo, isto é, ao fim e ao cabo, não se teria um sistema automatizado a ser utilizado por tão somente um coordenador a querer montar uma equipe.

O seguinte foi o Modelo Computacional de Percepção de Contextos de Atividade para Identificação de Comunidades (BORTOLI, 2006), do qual também se apartou, já que, enquanto esse monitora e analisa diversos artefatos textuais a fim de capturar contextos de atividade, que nada mais são do que vetores de termos relevantes acompanhados de seus valores de TF/IDF (os quais medem a relevância de tais termos), seria o coordenador, ao empregar o SINAPSE, quem definiria os termos relevantes a serem varridos de um Currículo Lattes de um candidato. Por último, tal modelo, como o próprio nome desse dá a entender, tem por objetivo identificar comunidades e não a conformação de equipes, intento do SINAPSE.

Já o terceiro, intitulado Modelo Computacional para Formação de Equipes baseado nas Redes Egocêntricas de Líderes de Projeto (REINHARDT, 2008), promove a análise de redes sociais com o fim de determinar os parâmetros da rede social egocêntrica do líder de projeto que possam ser usados na valoração de relações sociais para orientar a indicação de pesquisadores com o objetivo de formação de uma equipe de pesquisa. Para isso, determina a rede social egocêntrica dos pesquisadores levando em consideração os aspectos das múltiplas relações analisadas (coautoria, relações orientador-orientado) e aspectos temporais (período de ocorrência de determinada relação). Tendo em vista o objetivo geral do TCC, este foi o modelo selecionado para servir de base ao sistema que se desejava implementar.

Assim, em posse desse terceiro modelo, definiram-se os objetivos específicos a delimitarem o escopo do TCC.

Durante o exame da fundamentação teórica do modelo computacional selecionado, chegou-se à determinação do Fator de Proximidade Individual, obtido a partir do cálculo das médias ponderadas das relações em função do tipo (coautoria, relações orientador-orientado) e do período em que ocorreram. Houve um contratempo na análise desse fator, mas tal obstáculo foi contornado ao se encontrar o quadro correto de relação-peso utilizado na obtenção do fator.

Sem embargo, como o SINAPSE objetivava não somente ser um software utilizado em instituições educacionais, mas também em empresas, o termo candidato passou a ser empregado em detrimento de pesquisador, já que se expandia o âmbito de atuação do sistema. Ainda nesse ajuste ao SINAPSE, o Fator de Proximidade Individual foi renomeado para Fator de Proximidade e, por valorar quão "próximos" são dois indivíduos com relação aos aspectos sociais de familiaridade, proximidade geográfica e homofilia, foi usado como o Potencial de Sinergia que esses mesmos indivíduos teriam. No modelo, então, mais adaptações foram implementadas no que tange à definição do candidato mais apto e à automatização do processo de obtenção de informações e, por fim, procedeu-se à análise de requisitos e à codificação do sistema propriamente dito.

Finalmente, verificou-se que o objetivo geral e os objetivos específicos haviam sido atingidos, visto que o SINAPSE permite a um coordenador, de fato, avaliar quais são os candidatos capazes de, no contexto de um projeto, conformarem uma equipe na qual realmente haja cooperação e diversidade.

Mesmo havendo uma limitação, citada durante os testes e decorrente da inexistência de um link ou da falta de correspondência relacionada aos nomes, que não permitiria uma correta determinação do Potencial de Sinergia entre o candidato afetado por essa e o de melhor perfil, tal limitação poderia ser contornada através da implementação de algum modo de indagar o coordenador quando da eventualidade de nomes muito similares. Poderiam ser investigadas outras possibilidades de detecção desses casos, mas essas precisariam ser suficientemente seguras e tão eficientes quanto questionar o coordenador. Outro modo seria o de exigir sempre ou a inclusão de um link ou a correta escrita do nome no momento de preencher um Currículo Lattes.

Ademais, se não consegue contatar a Plataforma Lattes, o SINAPSE informa o Coordenador. Isso, contudo, significa que o sistema ficará inoperante enquanto

essa estiver inacessível ou, ainda, quando a Plataforma Lattes estiver com algum problema, esse também afetará o SINAPSE.

Aliás, para que uma empresa pudesse usufruir completamente da funcionalidade do SINAPSE, essa teria de possuir uma base de dados tão rica de informações quanto a Plataforma Lattes. Ficaria a questão, entretanto, de analisar como ficariam as relações de coautoria e de orientador-orientado no interior dessa empresa. Sugere-se, respectivamente, algo como relações de trabalho/projeto conjunto e de mentor-trainee.

No intuito de promover melhorias no SINAPSE, além da forma de contornar a limitação supracitada, poderiam ser implementadas medidas de segurança que impedissem ataques sobre vulnerabilidades conhecidas em sistemas baseados na Web. Além disso, para tarifar o serviço provido pelo SINAPSE, se a cobrança fosse por uso, bastaria alterar o código para armazenar as estatísticas desse uso. Com relação a backup, esse já é feito normalmente e diariamente por empresas que hospedam sistemas Web. No estágio atual do SINAPSE, não seria necessário, pois não se mantém um banco de dados e uma ou mais cópias do código são suficientes.

Por fim, como resultados alcançados deste TCC, ademais de atingir os objetivos propostos, citam-se: a possibilidade de prosseguir no campo de avaliação do potencial de cooperação interpessoal em projetos e de sistemas informatizados capazes dessa avaliação; a capacidade de redução do desperdício de dinheiro e de tempo com equipes que não teriam coesão e não alcançariam sequer os resultados desejados em projeções; e a melhora na competitividade das empresas, pois o SINAPSE permitiria a conformação de equipes que maximizariam lucros possíveis no tocante a projetos.

6 REFERÊNCIAS BIBLIOGRÁFICAS

REINHARDT, Joel Carlos Vieira. **Modelo Computacional para Formação de Equipes baseado nas Redes Egocêntricas de Líderes de Projeto**. 2008. 94 f. Dissertação (Mestrado em Informática Industrial) – Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná, Curitiba, 2008.

LE CARDINAL, Gilles; GUYONNET, Jean-François; POUZOULLIC, Bruno. **La Dynamique de la Confiance**. Paris: Dunod, 1997.

BORTOLI, Eliane Maria de. **Modelo Computacional de Percepção de Contextos de Atividade para Identificação de Comunidades**. 2006. 163 f. Dissertação (Mestrado em Informática Industrial) – Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná, Curitiba, 2006.

Diretório dos Grupos de Pesquisa no Brasil. Disponível em: <<http://dgp.cnpq.br/buscaoperacional/>>. Acesso em: 29 out. 2013.

Buscar Currículo Lattes. Disponível em: <<http://buscatextual.cnpq.br/buscatextual/busca.do>>. Acesso em: 29 out. 2013.

SMITH, James. **jQuery Tokeninput**. Version 1.6.0. San Francisco: 2011. Disponível em: <<http://loopj.com/jquery-tokeninput/>>. Acesso em: 29 out. 2013.

MOURIER, Simon; DARTHOBIVAN; JESSYNOO. **Html Agility Pack**. Version 1.4.6. France: 2012. Disponível em: <<http://htmlagilitypack.codeplex.com/>>. Acesso em: 29 out. 2013.

OLIVEIRA, Pedro. **PTStemmer: A Stemming toolkit for the Portuguese language**. Version 2.0. San Francisco: 2010. Disponível em: <<https://code.google.com/p/ptstemmer/>>. Acesso em: 29 out. 2013.