

**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**UMA METODOLOGIA PARA ACOMPANHAMENTO E MONITORAMENTO DE
INTERVENÇÕES DE FORMAÇÃO BASEADA EM ANÁLISE E MINERAÇÃO DE REDES
SOCIAIS**

ALEKSANDRA DO SOCORRO DA SILVA

TD – 11/2016

UFPA / ITEC / PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM-PARÁ-BRASIL
2016

**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

ALEKSANDRA DO SOCORRO DA SILVA

**UMA METODOLOGIA PARA ACOMPANHAMENTO E MONITORAMENTO DE
INTERVENÇÕES DE FORMAÇÃO BASEADA EM ANÁLISE E MINERAÇÃO DE REDES
SOCIAIS**

UFPA / ITEC / PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM-PARÁ-BRASIL
2016

**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

ALEKSANDRA DO SOCORRO DA SILVA

**UMA METODOLOGIA PARA ACOMPANHAMENTO E MONITORAMENTO DE
INTERVENÇÕES DE FORMAÇÃO BASEADA EM ANÁLISE E MINERAÇÃO DE REDES
SOCIAIS**

Tese submetida à Banca Examinadora do Programa de Pós-graduação em Engenharia Elétrica da UFPA para a obtenção do grau de Doutora em Engenharia Elétrica na área de Computação Aplicada.

UFPA / ITEC / PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM-PARÁ-BRASIL
2016

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Silva, Aleksandra do Socorro, 1972-

Uma metodologia para acompanhamento e monitoramento de intervenções de formação baseada em análise e mineração de redes sociais / Aleksandra do Socorro Silva. - 2016.

Orientador: Carlos Renato Lisboa Francês;

Coorientador: João Crisóstomo Weyl Albuquerque Costa.

Tese (Doutorado) - Universidade Federal do Pará,
Instituto de Tecnologia, Programa de Pós-Graduação em
Engenharia Elétrica, Belém, 2016.

1. Redes sociais - análise. 2. Tecnologia da informação.
3. Inclusão digital - avaliação. I. Título.

CDD 22. ed. 302.30285

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**UMA METODOLOGIA PARA ACOMPANHAMENTO E MONITORAMENTO DE INTERVENÇÕES DE
FORMAÇÃO BASEADA EM ANÁLISE E MINERAÇÃO DE REDES SOCIAIS**

AUTORA: ALEKSANDRA DO SOCORRO DA SILVA

Tese de doutorado submetida à avaliação da banca examinadora aprovada pelo colegiado do Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará e julgada adequada para a obtenção do Grau de Doutora em Engenharia Elétrica, na área de Computação Aplicada.

APROVADA EM: _____ / _____ / 2016

BANCA EXAMINADORA:

Prof. Dr. Carlos Renato Lisboa Francês (Orientador – PPGGE/UFPA)

Prof. Dr. João Crisóstomo Weyl Albuquerque Costa (Coorientador – PPGGE/UFPA)

Prof. Dr. Cláudio Alex Jorge da Rocha (Membro Externo – IFPA)

Prof. Dr. Marcos César da Rocha Seruffo (Membro Externo – UFPA/Castanhal)

Prof. Dr. Maurílio de Abreu Monteiro (Membro Externo – UNIFESSPA)

Prof. Dr. Nandamudi Lankalapalli Vijaykumar (Membro Externo – INPE)

VISTO: _____

Prof. Dr. Evaldo Gonçalves Pelaes (Coordenador do PPGGE/UFPA)

DEDICATÓRIA

À minha mãe, que foi o meu alicerce maior no início de minha vida acadêmica, quando nem eu mesma esperava. Maza: mulher de raça, esse passo é também seu.

A querida tia Coca, minha tia, minha mãe, minha amiga. Carrego você comigo neste momento.

Ao meu pai, Seu Chiquinho (in memorian), dedico este passo, que ele sempre tanto almejou para seus/suas filhos/filhas: desculpa pai, por não ter sido enquanto você estava aqui, nesta vida terrena.

A minha madrinha e mestra Maria Paraense (in memorian), que me achava a mais “sabida” de todas as suas alunas, e eu nem era. Agradeço por todos os ensinamentos que me fizeram tão cedo refletir, e ser tão eu.

AGRADECIMENTOS

Este sempre foi um momento esperado em minha vida. Demorou tanto, que algumas pessoas (infelizmente) não estão do meu lado neste momento. Agradeço a Deus, por suas grandes asas sobre mim, por toda sua proteção, para estar nesse momento vivendo esse sonho.

Esse trabalho não teria sido realizado se não fosse pela minha parceira de vida, *Silvana Rossy*, amiga, irmã e companheira de estudos — “*O que vamos fazer hoje Pink? O que tentamos há anos fazer Cérebro!*” —, e assim foram os últimos anos, em nossos laboratórios de estudos e pesquisas, implantados em nossas casas e em nossas salas de trabalho. Dona de uma sabedoria, rapidez e sensibilidade invejável, digna do privilégio de poucos humanos. Impossível conseguir este feito, se não fosse por essa amiga-irmã que a vida me deu de presente há exatos 27 anos.

Agradeço a minha companheira, *Joelma*, por toda sua paciência nos últimos anos, por seu amor e dedicação, por sua vontade de continuar ao meu lado, por me entender, quando nem eu mesma me entendo: “*espero que desculpes os meus erros por favor / nas frases desta carta / que é uma prova de afeição*”.

Agradeço a *Laila*, companhia das madrugadas, pelas alegrias proporcionadas. Agradeço à *Lully*, amor incondicional dos últimos doze anos, por ter aguentado o tranco, sendo carregada para auxiliar na árdua tarefa de escrever os artigos produzidos, junto com a *Sara Linda da Silva*, netinha linda que a vida me trouxe. Ao *Bobby*, que nos acompanhou até onde foi possível e agora é estrelinha de todos os outros: os que tem amor e àqueles que enfrentam as ruas diariamente. Amores para minha vida e história toda, que marcaram esses tempos: amor de bicho e de gente, que me deram grandes ensinamentos, junto com todos os outros que encontrei, alimentei pelo caminho e me deram e ensinaram tanto amor.

A minha *mãe*, por toda a sua paciência e por todo seu afago e verdadeiro amor, pela torcida da minha *tia* amada *Coquinha*, minhas irmãs queridas: *Norma*, *Antônia*, *Edna* e *Nazaré*, por todos os bons momentos vividos juntas, pelas músicas enraizadas em mim, pelas conversas, pelas risadas e também pela presença nos momentos mais difíceis. Aos irmãos que a vida me trouxe: *Valdemar*, por suas palavras de incentivo, força, conversas, e viagens ao longo da jornada, ao *Roberto (Robert)*, pelos bons momentos vividos, curtidos e brincados juntos em um tempo tão inocente e ao mesmo tempo tão bom. Aos irmãos *Nego* e *Zuza*, por todos os bons ensinamentos que adquiri com eles, pelas músicas, pelas aulas, pelas boas conversas. Às pequenas sobrinhas *Giovana*, *Beatriz*, minha afilhada *Maria Eduarda*, que eu possa curtir mais ao lado delas. Aos não tão pequenos, mas muito amados, *Fábio*, o eterno *Fabinho* da família, sempre presente, sempre disposto a ajudar e a me escutar; à *Eliana* e *Ariane*, símbolos de amor de mãe e de filha; aos futuros médicos da família Silva: *Emília*, que com toda a semelhança física, me inspira tanta

doçura e dedicação, *Luís Henrique*, que com todo a sua força de vontade, vem superando os desafios e obstáculos, como um verdadeiro lutador e *Ana Luísa*, a futura cirurgiã, por nossas poucas conversas, mas sempre proveitosas. Aos queridos: *Elias*, *Eloísa*, *Francisco Neto* e *Felipe*, sempre no meu coração.

Agradeço especialmente ao amigo e orientador *Renato Francês*, pois confesso que nunca conheci pessoa com tanto otimismo. Desde nossas épocas de Cálculo II, tudo ia dar certo e está dando. Obrigada *Renatinho*, pela confiança, pelas motivações, pelas oportunidades dadas e principalmente por saber lidar com essa amiga aqui. Ao Prof. *João Weyl*, pela oportunidade dada de participar do Polo Norte de formação do programa Telecentros.BR, e por todas as oportunidades de boas conversas que tive em sua companhia.

Ao prof. *Vijay*, que nunca hesitou em nos ajudar, nos momentos de mais aflição. Ao Prof. *Eloi*, por toda a boa convivência que tive desde minha graduação. Aos membros da banca Prof. *Maurílio*, Prof. *Marcos Seruffo* e *Cláudio Alex*, por aceitar o convite de compartilhar e discutir sobre os nossos resultados. Ao *Alex*, que não é *Cláudio Alex*, um grande amigo e companheiro no início de minha vida acadêmica, parceiro de projetos, de boas conversas e de muitas risadas.

Ao amigo e primo *Jefrey Belém* e à amiga *Luciana Faria (Laxi)* pelos momentos de descontração e às amigas *Thaís Tavares (Tatá)* e *Eulália Mata (Eulaka)*, por suas presenças constantes em minha vida. *Eulaka*, exemplo de lealdade, determinação, força de vontade e humildade. *Tatá*, companheira de missões impossíveis. Ao casal *Gelice Dias* e *Roberto Zahluth de Carvalho*, por sua amizade e por todos os momentos de companhia nas viagens ao meu *Monte Roraima (Alter do Chão)* nos últimos cinco anos. Ao parceiro paulistano *Dalton Martins*, que tanto inspirou e nos incentivou no início dessa pesquisa — sem ele não teríamos adentrado e aprofundado na análise das redes sociais.

Aos companheiros da UFRA, do grupo *G6: Edvar, Jorge, Larissa, Sabrina, Silvana*: que sigamos sempre adiante em busca de uma universidade melhor e pesquisas úteis para a sociedade. À querida *Merilene*, por toda a presteza demonstrada durante todos esses anos de UFRA, por suas palavras de incentivo e de amizade. Às queridas *Deciôla*, e *Klissiomara*, por todas as boas conversas e convivências agradáveis no decorrer de nossos projetos na UFRA. À querida *Adelinha*, (irmã da *Silvana* na UFRA), por todo o carinho conosco nos últimos anos.

Ao PPGEE — representados pelos professores e seus funcionários, pelo voto de confiança depositado em mim.

Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes.
(Martin Luther King)

SUMÁRIO

LISTA DE FIGURAS.....	XIII
LISTA DE TABELAS	XIV
LISTA DE ABREVIATURAS.....	XV
RESUMO	XVI
ABSTRACT.....	XVII
1. INTRODUÇÃO.....	1
1.1 MOTIVAÇÕES	1
1.2 CONTEXTO	3
1.3 PROBLEMÁTICA, HIPÓTESE E OBJETIVOS	5
1.4 JUSTIFICATIVAS	6
1.4.1 <i>Links</i>	6
1.4.2 Local	7
1.4.3 Função ou Papel do Participante.....	7
1.5 PRINCIPAIS CONTRIBUIÇÕES	8
1.6 ORGANIZAÇÃO DO TRABALHO.....	9
2. TÉCNICAS PARA ANÁLISE E MINERAÇÃO DE REDES SOCIAIS	11
2.1 PROLEGÔMENOS	11
2.2 REDES	11
2.3 REDES SOCIAIS	13
2.4 ANÁLISE DE REDES SOCIAIS	14
2.4.1 Análise Visual ou Visualização das Redes Sociais.....	16
2.4.2 Métricas para análise de Redes Sociais	18
2.4.2.1 Centralidade	18
2.4.2.2 Prestígio	22
2.4.2.3 Densidade da rede	22
2.4.2.4 Coeficiente de agrupamento	23
2.4.3 Softwares para Análise de Redes Sociais	24

2.5	MINERAÇÃO DE REDES SOCIAIS.....	25
2.5.1	Identificação de Comunidades ou <i>Clusters</i>	26
2.5.2	Detecção automática de comunidades	27
2.5.2.1	Função Modularidade	28
2.5.2.2	Método Louvain.....	29
2.5.3	Medida de associação entre diferentes classificações.....	30
2.6	COMENTÁRIOS FINAIS.....	31
3.	TRABALHOS RELACIONADOS.....	33
3.1	PROLEGÔMENOS	33
3.2	INTERVENÇÕES PARA AVALIAÇÃO OU ACOMPANHAMENTO DE INDIVÍDUOS	33
3.3	INTERVENÇÕES PARA A MEDIÇÃO DA EXCLUSÃO/INCLUSÃO DIGITAL.....	38
3.4	SÍNTESE DOS TRABALHOS RELACIONADOS	40
3.5	COMENTÁRIOS FINAIS.....	42
4.	METODOLOGIA PARA ACOMPANHAMENTO E MONITORAMENTO DE INTERVENÇÕES DE FORMAÇÃO EM TICS	44
4.1	PROLEGÔMENOS	44
4.2	MODELO CONCEITUAL.....	45
4.3	ABORDAGEM PARA MONITORAMENTO E ACOMPANHAMENTO DE PROGRAMAS DE FORMAÇÃO EM LARGA ESCALA	47
4.4	METODOLOGIA PARA ANÁLISE DAS REDES SOCIAIS DOS PARTICIPANTES	49
4.4.1	Inicialização	50
4.4.2	Nível Macro	50
4.4.3	Nível Meso.....	51
4.4.4	Nível Micro.....	52
4.4.5	Índices sugeridos para análise.....	53
4.5	COMENTÁRIOS FINAIS.....	54
5.	APLICAÇÃO DA METODOLOGIA	55
5.1	PROLEGÔMENOS	55
5.2	CENÁRIO DE ESTUDO: CURSO DE FORMAÇÃO DO PROGRAMA TELECENTROS.BR	55
5.3	FONTES DE DADOS.....	57
5.3.1	Mensagens instantâneas.....	58

5.3.2	Atributos dos participantes	58
5.3.3	Avaliação sobre os Monitores.....	59
5.4	RESULTADOS DA APLICAÇÃO DA METODOLOGIA	59
5.4.1	Inicialização	59
5.4.2	Nível Macro	60
5.4.3	Nível Meso.....	61
5.4.3.1	Definição de <i>clusters</i> por atributos: <i>papel</i> e <i>polo</i>	61
5.4.3.2	Detecção Automática de Comunidades e confronto entre <i>Clusters</i>	65
5.4.3.3	Análise visual das estruturas de redes das comunidades	68
5.4.4	Nível Micro.....	70
5.5	COMENTÁRIOS FINAIS.....	72
6.	CONCLUSÕES.....	74
6.1	CONSIDERAÇÕES.....	74
6.2	CONTRIBUIÇÕES.....	75
6.3	DIFICULDADES ENCONTRADAS	76
6.4	TRABALHOS FUTUROS	77
6.5	PUBLICAÇÕES GERADAS.....	78
	REFERÊNCIAS BIBLIOGRÁFICAS.....	81

LISTA DE FIGURAS

Figura 2.1 Exemplo de grafo não orientado.....	13
Figura 2.2 Exemplo de grafo orientado.	13
Figura 2.3 Exemplo de Sociograma. Fonte: Nooy et al., (2005).	17
Figura 2.4 Exemplo de rede de pesquisadores com parceria em artigos científicos.....	19
Figura 2.5 Exemplo de rede de troca de mensagens entre orientador e estudantes.	22
Figura 2.6 Exemplo de duas possíveis configurações para a vizinhança do nó <i>a</i>	24
Figura 2.7 Grafo simples com três comunidades, identificadas pelos círculos tracejados. Fonte: Fortunato e Castellano (2009).	27
Figura 3.1 Modelo de três estágios proposto pelo ITU. Fonte: Adaptado de ITU (2014).....	40
Figura 4.1 Modelo Conceitual.	46
Figura 4.2 Abordagem para monitorar/acompanhar intervenções de formação em TICs em larga escala.....	48
Figura 4.3 Atividades da metodologia para analisar as redes sociais dos participantes.	50
Figura 5.1 Valores de densidade das redes nos diferentes momentos analisados.	62
Figura 5.2 Valores de centralidade por proximidade das redes nos diferentes momentos analisados.....	63
Figura 5.3 Rede de Tutores.....	63
Figura 5.4 Rede de Monitores.....	64
Figura 5.5 Rede de todos os participantes (tutores e monitores).	65
Figura 5.6 Rede do <i>Cluster</i> ID-3 (Método Louvain).	69
Figura 5.7 Rede do <i>Cluster</i> ID-3 (Método Louvain) após operação de redução de vértices....	69
Figura 5.8 Rede do <i>Cluster</i> ID-1 (Método Louvain).	69
Figura 5.9 Rede do <i>Cluster</i> ID-1 (Método Louvain) após operação de redução de vértices....	70

LISTA DE TABELAS

Tabela 2.1 Distribuição dos países nas categorias de 1980 e 1994 – tabela adaptada de Nooy et al. (2005).	31
Tabela 3.1 Trabalhos relacionados, aspectos avaliados e métodos utilizados.	41
Tabela 4.1 Índices extraídos das redes.	53
Tabela 5.1 Categoria, Descrição e Valores para o Atributo.	59
Tabela 5.2 Índices da Rede Telecentros.BR.	61
Tabela 5.3 Índices das Redes de Tutores e/ou de Monitores.	62
Tabela 5.4 Percentual de participantes (%) quanto ao polo nas comunidades detectadas pelo Louvain.	66
Tabela 5.5 Distribuição dos atores quanto aos papéis nas comunidades detectadas pelo Louvain.	67
Tabela 5.6 Distribuição dos Monitores com mais alto grau de centralidade por polos.	70
Tabela 5.7 Indicadores de Centralidade e Prestígio em ordem decrescente de grau de centralidade.	71

LISTA DE ABREVIATURAS

ACTIC	<i>Acreditación de competencias en tecnologías de la información y la comunicación</i>
AHP	<i>Analytic Hierarchy Process</i>
ANT	<i>Actor Network Teory</i>
ARS	Análise de Redes Sociais
CC	<i>Clustering Coefficient</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
iDCA	<i>Instant Digital Competence Assessment</i>
IDI	<i>ICT Development Index</i>
ITU	<i>International Telecommunication Union</i>
LTS	Laboratório de Tecnologias Sociais
SID	Secretaria de Inclusão Digital
SNA	Social Network Analysis
SNAM	Social Network Analysis and Mining
SNSs	Social Network Sites
TICs	Tecnologias de Informação e Comunicação
UFPA	Universidade Federal do Pará

RESUMO

Apesar dos avanços na qualidade e disponibilidade de tecnologias de informação e comunicação (TICs), o nível de acesso e habilidade no uso desses recursos ainda é desigual. Isto é particularmente evidente nos países em desenvolvimento. Para reduzir o fosso entre estes níveis de utilização das TICs, organizações (públicas e privadas) investem na expansão da infraestrutura, proporcionando acesso e formação, por meio de intervenções em TICs. Nesta tese, é mostrado que existe uma lacuna nas abordagens atuais, para monitorar e avaliar as intervenções de formação em TICs em larga escala. Assim, é proposta uma metodologia com base nos *links* entre os participantes, baseada em análise e mineração de redes sociais. Os *links* representam que os participantes estão em um estágio de apropriação das tecnologias. Esta metodologia está estruturada em diferentes níveis com propósitos específicos: (i) analisar se existe intensificação de *links*, de acordo com o tempo decorrido na formação; (ii) analisar a influência do papel ou da localidade do participante nas comunidades/clusters; (iii) identificar os participantes com maior participação na formação ou detectar variáveis associadas à baixa participação. O cenário de estudo para a aplicação da metodologia é o curso de formação de indivíduos no programa Telecentros.BR — criado no Brasil para instalar telecentros e habilitar indivíduos para o uso das TICs. Encontramos que: (i) a análise das interações nos diferentes períodos de tempo reflete os objetivos de cada fase de treinamento, com destaque para o aumento da densidade na fase em que os participantes desenvolvem e disseminam os seus projetos; (ii) análises conforme os papéis dos participantes (isto é, tutores ou monitores) revelam que as interações foram influenciadas pelo centro de formação (ou região) ao qual o participante pertence (isto é, uma comunidade contém principalmente membros da mesma região e sempre com a presença de tutores, contrariando as expectativas, que visavam a intensa colaboração dos participantes, independentemente da região geográfica); (iii) participantes da região nordeste do país estão associados com o maior grau de centralidade na plataforma de rede social; e (iv) diversos participantes que tiveram baixo grau de centralidade relataram que a internet era instável, principalmente no norte do país.

Palavras-chave: inclusão digital, formação em inclusão digital, modelos para acompanhamento e avaliação de programas de formação, análise e mineração de redes sociais.

ABSTRACT

Despite advances in the quality and availability of information and communication technologies (ICT), the level of access and skills in using these resources remains unequal. This is particularly evident in developing countries. To reduce the gap between these levels of employing ICT, organizations (both public and private) invest in the expansion of infrastructure by providing ICT access and ICT training through ICT interventions. In this thesis, it is shown that there is a gap in current approaches to monitor and evaluate large-scale ICT training interventions. Therefore, a methodology is proposed based on the links between participants, social networks analysis and mining. The links represent that participants are at the stage of technological appropriation. This methodology is structured in different levels with specific purposes: (i) to analyze whether there is an intensification of *links* according to the time in the training; (ii) to analyze the influence of the role or participant's location in communities/clusters; (iii) identify participants with higher participation in training or detect variables associated with low participation. The study scenario for the application of the methodology is the qualification of individuals in Telecentros.BR Program — created in Brazil to install telecenters and qualify individuals to use ICTs. We found that (i) the analysis of interactions in different time periods reflects the objectives of each phase of training, highlighting the increased density in the phase in which participants develop and disseminate their projects; (ii) analysis according to the roles of participants (i.e., tutors or monitors) reveals that the interactions were influenced by the training center (or region) to which the participant belongs (that is, a community contains mainly members of the same region and always with the presence of tutors, contradicting expectations which aimed for intense collaboration of the participants, regardless of the geographic region); (iii) participants of the northeast region of country are associated with the highest degree centrality in social networking platform; (iv) several participants who had low degree centrality reported that the internet was unstable, especially in the north of country.

Keywords: digital inclusion, digital inclusion training, models for monitoring and evaluating training programs, social network analysis and mining.

1. INTRODUÇÃO

Neste capítulo, são apresentadas algumas motivações para a realização desta pesquisa. Na sequência, descrevem-se o contexto; a problemática e hipótese levantada, bem como os objetivos da tese em questão; justificativas para os conceitos-chave abordados na metodologia; e principais contribuições. Por fim, apresenta-se a organização desta tese.

1.1 Motivações

O acesso à informação é um direito fundamental para a vida em sociedades democráticas, estabelecido através dos princípios e disposições da Declaração Universal dos Direitos Humanos (UNITED NATIONS, 2012), bem como dos Objetivos do Milênio (UNITED NATIONS, 2015). Para alguns pesquisadores, o acesso às tecnologias de informação e comunicação é essencial para o desenvolvimento econômico (HEEKS, 2008; SASSI e GOAIED, 2013), para a igualdade social (BABIC, et al., 2015; KENNEDY, 2014), para a melhoria nos cuidados de saúde (BABIC, et al., 2015; HAGE et al., 2013; LUCAS, 2008) e sistemas de educação (ASHRAF, et al., 2009; BABIC, et al., 2015), inovações no setor financeiro (DINIZ, 2012), melhoria dos serviços sociais e de governança (BASU, 2004; KHAN et al., 2011; WOUTERS et al., 2015), mudanças nas políticas públicas dos países (ANDRADE e URQUHART, 2012), e expansão da participação democrática (LUNA-REYES et al., 2012; UNITED NATIONS, 2001).

Tecnologias de informação e comunicação incluem computadores, Internet, software de computador, equipamentos periféricos, telefones de linha fixa e móveis, que permitem o acesso à informação e comunicação entre as entidades, tais como indivíduos, organizações e países. Entretanto, a existência de comunidades inteiras sem acesso às TICs¹ é ainda uma realidade em grande parte do mundo e no Brasil. De acordo com a União Internacional de Telecomunicações (do inglês: *International Telecommunication Union – ITU*) (ITU, 2014), 4.3 bilhões de pessoas ainda não estão *on-line* e 90 por cento vivem em países em desenvolvimento. O resultado é que estas comunidades estão afastadas dos benefícios da

¹ O termo acesso diz respeito à disponibilidade/prontidão das TICs.

globalização econômica. Esta é a situação de muitas comunidades quilombolas e indígenas que vivem às margens dos rios e em áreas rurais isoladas da Amazônia legal brasileira.

No Brasil, país de dimensões continentais, há grandes desigualdades no acesso às TICs. De acordo com dados de 2014 da Pesquisa Nacional por Amostra de Domicílios (IBGE, 2016) — pesquisa por amostragem que fornece indicadores por estado, o Distrito Federal tem a maior taxa no país de domicílios com computadores com Internet (64,09%), seguido de São Paulo (57,28), entretanto a diferença é desmedida para os estados que apresentam os menores índices que são Pará (16,78) e Maranhão (15,67). Quando se comparam as diversas regiões do país, as desigualdades são também desmesuradas: a região Sudeste possui o maior percentual (51,79%), seguida pelas regiões Sul (48,43%), Centro Oeste (42,4%), enquanto os piores índices são das regiões Nordeste (27,73%) e Norte (22,49%), representando as duas regiões com percentuais abaixo do Brasil (42,09%). Ou seja, apesar do crescimento econômico e social do Brasil nos últimos anos, as desigualdades de acesso às TICs entre os estados/ regiões geográficas ainda permanecem e representam verdadeiros desafios para se alcançar um acesso coletivo mais igualitário.

Assim, nas últimas décadas, países em desenvolvimento têm demonstrado um grande interesse na implementação de intervenções de TICs em áreas rurais/urbanas, seja para aumentar o desenvolvimento do país ou para diminuir as desigualdades internas. Intervenções de TICs podem ser representadas por cursos/treinamentos, oficinas práticas/teóricas, programas de formação continuada, projetos de desenvolvimento/inação, fornecimento de acesso às diferentes TICs, dentre outros. De fato, a crença é que as intervenções ajudem a reduzir o fosso digital (do inglês: *digital divide*²), que é um desafio a ser superado, tanto internamente (e.g. em um determinado país, região, cidade), quanto internacionalmente (e.g. entre países, continentes).

A despeito de muitos relatos sobre intervenções de TICs (ASHRAF et al., 2009; CALVANI et al., 2012; TELES e JOIA, 2011; VENKATESH e SYKES, 2013) em vários segmentos da sociedade (e.g. para idosos, estudantes, professores, jovens empresários, funcionários públicos, e moradores de áreas rurais e isoladas), sem a investigação adequada

²O termo *digital divide*, difundido nos anos 90 do século passado pelos Estados Unidos da América faz referência à diferença entre excluídos e incluídos. No Brasil, o termo usado normalmente como sinônimo é *fosso digital*, mas também se usa o termo *Exclusão Digital*, embora de fato, este último indique apenas o lado dos excluídos, em oposição ao termo *inclusão digital*.

do impacto destas iniciativas³, é difícil determinar o sucesso da intervenção. Assim, acompanhar e avaliar intervenções de TICs ou medir o fosso digital, representam atividades fundamentais, que podem ajudar gestores, pesquisadores e profissionais no processo de apoio à decisão.

1.2 Contexto

Na literatura de Inclusão Digital, é possível identificar duas linhas principais para abordar o tema Inclusão digital, a primeira, que possui trabalhos relacionados à inclusão digital como fornecimento do acesso à infraestrutura de TICs (e.g. redes, dispositivos), em que destacam-se discussões sobre custo do acesso, posse dos recursos de TICs, qualidade dos acessos ou ainda sobre suas implicações na economia (FEIJÓO et al., 2016; JORGENSON e VU, et al., 2016; ZHOU et al., 2011) ou a segunda, que possui trabalhos relacionados à oferta de formação para indivíduos que possuem pouca/ninguma habilidade para lidar com as TICs (e.g. por meio de cursos, oficinas), em que as discussões são referentes à relatos de intervenções de formação, formas de avaliar, medidas para o fosso digital (ASHRAF et al., 2009; CALVANI et al., 2012; SALINAS e SÁNCHEZ, 2009; TELES e JOIA, 2011; VENKATESH e SYKES, 2013).

Em outra linha⁴, considera-se que a preocupação principal deve ser com o nível de apropriação⁵ das TICs pelos indivíduos (BUZATO, 2008; PSCHIEDT, 2012; WILLIAMS, 2005) e não apenas a visão de que o acesso e o provimento de capacidades básicas para o uso das TICs são suficientes, mas sim que esses indivíduos possam experimentar um processo de transformação e reinvenção de significados das TICs e não serem apenas mero consumidores de uma ação que foi concebida para eles por financiadores/idealizadores, administradores/políticos. Com essa visão, Buzato (2008) argumenta que os sujeitos que utilizam um determinado espaço público de provimento de acesso, podem subverter as estratégias que foram estabelecidas para eles, por meio de usos e táticas, que podem constituir

³ Uma iniciativa é uma intervenção de qualquer natureza (e.g. curso, programa, projeto). É um termo bastante encontrado em textos brasileiros.

⁴ Esta linha harmoniza-se com os objetivos desta tese.

⁵ Nesta tese, o termo apropriação de TICs está alinhado com a visão de Kahen (1995) e diz respeito ao aumento da capacidade ou maior auto-suficiência/autonomia no uso de determinadas TICs por um indivíduo.

formas de inclusão digital não previstas e nem sequer pensadas pelos produtores (quem concebe a estratégia).

No Brasil, em 2011, foi criada a Secretaria de Inclusão Digital (SID), vinculada ao Ministério das Comunicações (BRASIL, 2016a), que passou a assumir responsabilidades relacionadas à conteúdos digitais, formação e infraestrutura para Inclusão Digital. A SID é responsável pela coordenação do Fórum de Inclusão Digital, que reúne instâncias do governo federal para o desenvolvimento de ações na área e atua em parceria com os governos estaduais e municipais, universidades, institutos federais de ensino e com organizações da sociedade civil, dentre outros. Dentre as ações acompanhadas pela SID, na linha de acesso às TICs, formação e governança, estão (BRASIL, 2016a): os telecentros, os centros de condicionamento de computadores, as cidades digitais e projetos voltados à juventude rural.

Apesar do avanço na implantação e execução dos projetos no Brasil na última década, há uma carência de estudos que reflitam a preocupação com a efetividade desses programas, como demonstrado em Mori (2011). Uma instigante conclusão é realizada em seu estudo, a respeito da análise desses programas, citada no trecho a seguir:

[...] contudo, o que se percebe sobre o conjunto de estudos sobre telecentros é que eles não têm conseguido dar conta de aspectos relacionados à análise de programas e projetos públicos e privados de larga escala, enquanto elementos de uma política pública recente e das especificidades da implantação dessas iniciativas no Brasil. (p. 25).

Na literatura, é possível encontrar investigações de sucesso que avaliam o impacto das intervenções de TIC para superar a exclusão digital, no nível micro, ou seja, aquelas que são realizadas com foco nos beneficiários das ações em seu contexto local (ADAM e WOOD, 1999; ASHRAF et al., 2008; ASHRAF, et al., 2009; AVGEROU, 2008; CHANG et al., 2012; TELES e JOIA, 2011; VENKATESH e SYKES, 2013). Estas investigações (ASHRAF et al. 2008; ASHRAF, et al., 2009; CHANG et al., 2012; TELES e JOIA, 2011), frequentemente são realizadas por meio de questionários e/ou entrevistas e aplicadas a um pequeno número de respondentes, favorecendo uma visão detalhada do impacto das intervenções de TICs dentro de um contexto social particular. Tal abordagem, que é dispendiosa em termos de recursos físicos e humanos, e que necessitaria de um longo tempo para ser concluída, quando a escala dos participantes cresce, é inadequada para investigar intervenções em um extenso território geográfico. Outrossim, as avaliações em larga escala são muitas vezes padronizadas e não

consideram as diferentes condições de acesso ou os níveis socioeconômicos e culturais dos participantes. Em suma, existe uma lacuna de estratégias, modelos ou *frameworks*, necessários para avaliar o impacto e monitoramento de intervenções de TICs em larga escala.

1.3 Problemática, Hipótese e Objetivos

A problemática geral deste trabalho é realizar o acompanhamento/monitoramento de intervenções de formação — em larga escala, que tenham como objetivo prover habilidades/conhecimentos básicos de TICs para indivíduos. Essa problemática remete a uma questão bastante geral: *como realizar o acompanhamento/monitoramento de uma intervenção de formação em TICs em larga escala?*

Assume-se aqui, no tratamento da questão supracitada, a hipótese de que o acompanhamento/monitoramento deve privilegiar as conexões/*links*⁶ estabelecidos entre indivíduos ou com outras entidades, como aspecto-chave, uma vez que esses *links* podem representar apropriação de recursos úteis de TICs, evidenciando mais autonomia e pró-atividade desses indivíduos. Nesse sentido, assume-se especificamente, a hipótese de que a definição do acompanhamento/monitoramento baseado nos *links* estabelecidos entre os indivíduos, contribui para a análise da apropriação de TICs dos indivíduos em uma intervenção de formação.

Assim, o objetivo geral desta tese é desenvolver uma metodologia que permita realizar o acompanhamento da intervenção, baseada nos *links* estabelecidos pelos indivíduos na formação.

Esse objetivo remete a outros objetivos subjacentes, quando são considerados os seguintes aspectos sobre a intervenção: (i) ocorre em um período de tempo definido; (ii) abrange locais diferentes, com diferentes aspectos demográficos, socioeconômicos, de infraestrutura de telecomunicações; (iii) possui participantes com diferentes funções/papéis; e (iv) possui participantes com diferentes graus de conexão.

Assim, é possível destacar como principais objetivos específicos desta tese, os que seguem:

⁶ Nesse contexto, as conexões ou links são estabelecidos através de uma ferramenta computacional/software.

- Analisar a intensidade das conexões, de acordo com o tempo decorrido na formação;
- Definir *clusters*⁷ diferentes de participantes classificados por *papel*;
- Definir *clusters* diferentes de participantes classificados por *local*;
- Selecionar algoritmo/método computacional que execute rapidamente e produza *clusters* de boa qualidade, passíveis de serem analisados;
- Utilizar algoritmo computacional para realizar simulações de descoberta de comunidades;
- Analisar *clusters* encontrados pelo algoritmo computacional e *clusters* definidos pelo atributo *local* quanto à localidades dos participantes. A intenção é verificar se nos grupos há homogeneidades dos participantes quanto ao local ou se há alguma heterogeneidade, com a localidade dos participantes sendo bem distribuída;
- Analisar a distribuição dos papéis dos participantes nos *clusters* encontrados pelo algoritmo computacional;
- Identificar participantes que possuam alto grau de conexão;
- Identificar as variáveis associadas com a baixa conexão dos participantes;
- Aplicar a metodologia desenvolvida em um domínio real, mais especificamente em um programa de inclusão digital em larga escala, conduzido em diferentes regiões do Brasil.

1.4 Justificativas

Os conceitos de *links*, papel/função e local do participante são considerados conceitos-chave no desenvolvimento da metodologia, e por isso são tratados a seguir.

1.4.1 *Links*

Os *links* estabelecidos significam que os indivíduos estão interagindo entre si, e adicionalmente podem indicar que os indivíduos estão se apropriando de um importante recurso útil de TIC, que refletem a autonomia e pró-atividade desses indivíduos. Por isso, representam a base da metodologia aqui apresentada. Isso corrobora com argumentos que

⁷ Um *cluster* é formado por um conjunto de participantes que possuem propriedades comuns ou que são mais fortemente relacionados. Nesta tese, o termo *cluster* será usado como sinônimo de comunidades.

defendem que as tecnologias não devem ser analisadas apenas quanto ao número de acessos, mas também nas formas pelas quais os usuários criativamente absorvem tais tecnologias (BEREND, 2013; SRINIVASAN, 2013).

Neste contexto, a área denominada Análise de Redes Sociais (ARS) (do inglês: *Social Network Analysis – SNA*), que proporciona uma metodologia para analisar as relações sociais existentes entre as entidades em uma rede social, demonstra ser apropriada para a metodologia apresentada nesta tese.

1.4.2 Local

Localização é um dos principais fatores do fosso digital em muitos países. Por exemplo, os cidadãos chineses que residem em áreas urbanas são muito mais propensos a ter acesso à Internet do que aqueles que vivem em áreas rurais; há uma taxa de penetração de Internet de 62,8% nas áreas urbanas da China, mas apenas 28,8% nas áreas rurais da China (CNNIC, 2015). O fosso urbano-rural também ocorre na posse de telefones (fixo ou móvel). Por exemplo, de acordo com o censo populacional e habitacional realizado na Índia em 2011, 82% dos domicílios urbanos indianos têm acesso a um telefone comparados com 54% das famílias rurais (ITU, 2014). No Brasil, a região Sudeste possui o maior percentual de computadores com Internet (51,79%), contra 22,49% da região Norte (IBGE, 2016). Cada local possui diferentes aspectos demográficos, socioeconômicos, de infraestrutura de telecomunicações. Assim, em uma intervenção de formação em TICs, quando são comparados indivíduos situados em diferentes locais, diferentes aspectos podem ser observados, refletindo essas diferenças. Dessa forma, é considerado um conceito-chave na metodologia desenvolvida nesta tese, podendo representar uma região ou cidade de um país, por exemplo.

1.4.3 Função ou Papel do Participante

Certeau (1994), classifica os produtores como grupos ou classes de indivíduos capazes de criar, impor e manter espaços disciplinares e prescrever formas de ordenamento e representação social. Para o autor, estariam entre os produtores os governos, as corporações de negócios, os poderes judiciários, as instâncias de planejamento urbano, as elites acadêmicas, os fornecedores de serviços públicos, as corporações profissionais, dentre outros. Os consumidores seriam todos aqueles constrangidos a viver (supostamente) de acordo com

as formas de ordenamento e disciplina prescritas pelos *produtores* e por meio dos produtos (físicos e simbólicos) que lhes são impostos. Nessa visão, Buzato (2008) argumenta que os sujeitos que utilizam um determinado espaço público de provimento de acesso, por exemplo, pode ser visto tanto como produtor, quanto como consumidor, pois os produtores desse lugar (e.g. idealizadores, financiadores, administradores, políticos) podem até criar uma estratégia de inclusão que pode ser traduzida na forma de um sistema disciplinar, mas esses consumidores (que são os indivíduos atendidos por um projeto) podem subverter essa estratégia por meio de usos e táticas, que podem constituir formas de inclusão digital não idealizadas e até mesmo não desejadas pelos produtores, mas que têm potencial para a transformação das relações de hegemonia-subalternidade.

Na intenção de analisar as relações de interação, o papel ou função do participante, por representar diferentes posições hierárquicas em uma intervenção de formação em TICs, também é um conceito-chave embutido na metodologia desenvolvida nesta tese.

1.5 Principais Contribuições

No que concerne ao desenvolvimento da metodologia, são consideradas as seguintes contribuições:

- Proposta de uma metodologia de acompanhamento/monitoramento de intervenções de formação em TICs;
- Proposta de uma metodologia multinível que permite diferentes níveis de agregação entre os participantes, possibilitando analisar: o nível da rede completa (nível macro), o nível de comunidades/*cluster* (nível meso), e o nível de um determinado participante (nível micro);
- Proposta de uma metodologia que considera como base os *links*, representando um certo grau de autonomia e de participação ativa dos indivíduos;
- Possibilitar diferentes análises das redes sociais:
 - Quanto ao tempo de formação;
 - Quanto às comunidades/*clusters* formados;
 - Quanto ao grau de proeminência de um ator na rede.

- Proposta de uma metodologia que utiliza técnicas de Mineração de Redes Sociais que permitem a identificação de comunidades, por meio de um algoritmo computacional;
- Permitir o confronto/comparação de *clusters* definidos por um atributo e *clusters* detectados pelo algoritmo computacional, a fim de verificar a predição entre as diferentes classificações;
- Aplicação da metodologia em intervenções reais de formação, caso os *links* representem o aspecto-chave a ser analisado;
- Permitir a aplicação de todos os níveis da metodologia ou apenas os níveis necessários em um determinado contexto;
- Aplicação em um programa de inclusão digital no Brasil, país de dimensões continentais e com carência de ferramentas/metodologias relacionadas à análise de programas e projetos públicos/privados em larga escala, especificamente no curso de formação em TICs, promovido pelo programa Telecentros.BR (BRASIL, 2016b);
- Produção de resultados que podem ser manuseados por técnicas de Inteligência Computacional, a fim de melhorar os processos de descoberta de conhecimento nas bases de dados analisadas;
- Produção de resultados que podem servir como referencial de análise para ser utilizado e/ou comparado com outras intervenções de natureza similar.

1.6 Organização do Trabalho

Este documento está estruturado em seis capítulos, sendo esse o primeiro capítulo e os demais organizados como segue abaixo:

No capítulo 2 (Técnicas para Análise e Mineração de Redes Sociais), apresentam-se definições e conceitos relacionados às redes e redes sociais, técnicas relacionados à SNA e mineração de redes sociais — nesse caso, especificamente relacionados à identificação de comunidades.

No capítulo 3 (Trabalhos Relacionados), o foco são os trabalhos que fazem o relato de instrumentos/procedimentos usados para realizar a avaliação dos indivíduos/curso de

formação em TICs; e os que estão relacionados ao desenvolvimento de modelos, *frameworks* e ferramentas para medir o fosso digital.

No capítulo 4 (Metodologia para Acompanhamento e Monitoramento de Intervenções de Formação em TICs), contextualiza-se este presente trabalho de tese em uma abordagem proposta pelo grupo de pesquisa em inclusão digital, do Laboratório de Tecnologias Sociais (LTS) da Universidade Federal do Pará (UFPA) e descreve-se a metodologia desenvolvida nesta tese com base na análise e mineração de redes sociais.

No capítulo 5 (Aplicação da Metodologia), apresenta-se o cenário de estudo usado para a aplicação da Metodologia proposta nesta tese — a formação do Programa Telecentros.BR, as fontes de dados usadas e os resultados obtidos a partir da aplicação da metodologia desenvolvida.

No capítulo 6 (Conclusões), apresentam-se as contribuições desse trabalho, dificuldades encontradas, trabalhos futuros decorrentes dos estudos aqui realizados e os resultados em termos de publicação, alcançados com os estudos e pesquisas provenientes do processo de doutorado.

Por fim, são apresentadas as referências bibliográficas que fundamentaram o desenvolvimento desta tese.

2. TÉCNICAS PARA ANÁLISE E MINERAÇÃO DE REDES SOCIAIS

Neste capítulo, apresentam-se definições e conceitos relacionados às Redes e Redes Sociais. Em seguida, são apresentados aspectos relacionados à Análise de Redes Sociais: Análise Visual ou Visualização; métricas/medidas mais difundidas e também usadas nesta tese; e exemplos de softwares usados para realizar manipulações e operações com redes sociais. Por fim, apresentam-se algumas tarefas que realizam mineração de redes sociais, mais especificamente a identificação de comunidades de duas diferentes formas: através do uso de um atributo e através do uso de um algoritmo para detecção automática de comunidades. Para realizar a associação entre essas diferentes formas são apresentados dois diferentes índices estatísticos.

2.1 Prolegômenos

A Análise e Mineração em Redes Sociais (do inglês: *Social Network Analysis and Mining – SNAM*) é um campo de pesquisa multidisciplinar em rápido crescimento, que possui técnicas próprias, que são utilizadas quando as técnicas usuais de SNA demonstram ser ineficientes para serem aplicadas em grandes conjuntos de dados. Desta forma, nessa tese, foram usadas as técnicas convencionais de SNA para a realização da análise dos *links* estabelecidos pelos participantes em uma intervenção de formação, bem como técnicas de mineração de redes sociais, adequadas para a identificação de comunidades/clusters nas redes sociais analisadas.

2.2 Redes

O estudo das redes tem tido avanços importantes nas mais diversas áreas de pesquisa, como na física, matemática, biologia, psicologia, medicina, computação, sociologia, linguística, engenharia, telecomunicações e astronomia, dentre outras áreas. As redes podem ser representadas por objetos tangíveis no espaço euclidiano (ou não). No caso de objetos tangíveis no espaço euclidiano, alguns exemplos de rede são (Boccaletti et al., 2006): as redes de energia elétrica, Internet, estradas, sistemas de metrô e redes neurais e por outro lado, caso sejam entidades definidas em um espaço abstrato, pode-se citar como exemplo as colaborações entre indivíduos.

O estudo de redes tem sido principalmente o domínio de um ramo da matemática discreta, conhecido como teoria dos grafos. A teoria dos grafos originou-se em 1736, com o problema clássico das sete pontes de Königsberg (Prússia no século XVIII e atual Kaliningrado, na Rússia). Naquela cidade, sete pontes cruzavam o rio Pregel, estabelecendo ligações entre duas ilhas e entre as ilhas e as margens opostas do rio. Os moradores buscavam uma solução a respeito da possibilidade de fazer um passeio pela cidade, começando e terminando no mesmo lugar, cruzando cada ponte exatamente uma vez. Isso já havia se tornado uma lenda, quando o matemático Leonhard Euler provou que não existia tal caminho. Para isso, o matemático apresentou a modelagem do problema através de um grafo, representando cada porção de terra por um ponto (vértices ou nós) e as pontes por linhas (arestas) e com isso solucionou o dilema dos moradores da cidade. A solução usada por Euler possibilitou que muitos outros problemas importantes da matemática aplicada fossem possíveis de serem modelados. Desde então, a teoria dos grafos tem fornecido soluções a um conjunto de questões práticas como: qual a rota mais curta para atingir um determinado ponto da cidade? Como colorir as regiões de um mapa utilizando o menor número de cores, de modo que as regiões vizinhas recebam diferentes cores?

Na teoria dos grafos, uma rede não direcionada é representada por um grafo não direcionado G , onde $G = (N, E)$, sendo que o primeiro conjunto é composto por um conjunto de vértices/nós, representado por $N = \{n_1, n_2, \dots, n_g\}$ e o segundo conjunto é composto de um conjunto de linhas ou arestas, representada por E , definidos em função dos elementos de N . O conjunto $E = \{e_1, e_2, e_3, \dots, e_l\}$, onde cada aresta e pertencente ao conjunto E é denotado pelo par de vértices (n_i, n_j) que a forma. O número total de vértices existente em um grafo é representado por g e o número total de arestas por l . Dois vértices quaisquer do grafo são ditos adjacentes ou vizinhos se existe uma aresta e unindo-os. No caso de uma rede direcionada, representada por um grafo direcionado, o conjunto L contém pares ordenados de vértices, que são chamados de arcos/arestas direcionadas/setas, embora muitas vezes o termo aresta seja também utilizado, sem distinção. Um grafo pode ser representado esquematicamente ou visualmente como apresentado nas figuras 2.1 e 2.2, com nós rotulados em círculos e linhas (direcionadas ou não) representado as arestas/arcos entre eles.

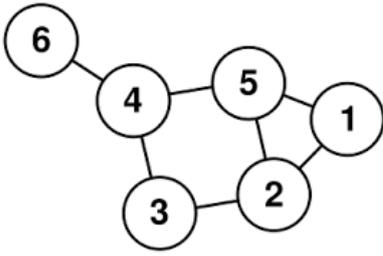


Figura 2.1 Exemplo de grafo não orientado.

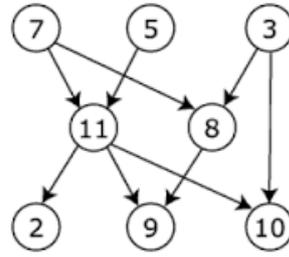


Figura 2.2 Exemplo de grafo orientado.

Historicamente, as redes também passaram a ser extensivamente estudadas nas ciências sociais. Por volta de 1930, os sociólogos notaram a importância dos padrões de conexão entre as pessoas para a compreensão do funcionamento da sociedade humana. Dessa forma, quando se trata de analisar interações entre os indivíduos, estes passam a ser os nós/vértices da rede e as interações são representadas pelas arestas entre os nós. Estudos de redes sociais típicos envolvem questões relacionadas à centralidade (quais indivíduos são mais conectados a outros ou têm mais influência) e conectividade (se/como os indivíduos estão ligados um ao outro através da rede). A linguagem formal para a representação de redes, seja qual for o tipo da rede, encontra-se na teoria dos grafos, mas segundo Newman M. (2003) novas abordagens estatísticas e novas métricas têm sido adotadas para lidar com grandes redes (com milhões ou bilhões de vértices), a fim de mensurá-las. Dessa forma, em seu clássico estudo de revisão sobre redes, Newman M. (2003) identifica quatro categorias: as redes biológicas, as redes de informação, as redes sociais e as redes tecnológicas. Nesta tese, o foco são as redes sociais.

2.3 Redes Sociais

Na definição clássica de Wasserman e Faust (1994), uma *rede social*⁸ consiste de um conjunto finito de *atores* e as *relações* definidas entre eles. Neste caso, um *ator* é uma unidade discreta ou uma unidade coletiva que pode corresponder a diferentes entidades (e.g. aluno, professor, escola infantil, associação de voluntários para ajudar aos animais de rua, tribo dos índios Caetés, países da África). Assim, o conceito de *ator* permite diferentes níveis de agregação, sendo adequado a diferentes problemas de pesquisa. Um *ator* possui associado

⁸ Cada conceito/definição inerente às Redes Sociais está grafado nesta seção em itálico. A maior parte do texto da seção foi formulado após a leitura de um clássico da área de SNA — *Social Network Analysis: Methods and Applications* (WASSERMAN e FAUST, 1994).

a ele características, denominadas de *atributos* (e.g. sexo, escolaridade, língua falada). *Atores* estão associados entre si por *laços relacionais* (do inglês: *relational tie*), também denominados simplesmente de *laço ou ligação* (do inglês: *link/linkage*) (e.g. a transferência de recursos materiais: uma transação de compra e venda, a transferência de recursos não materiais: a troca de mensagens eletrônicas ou não, laços biológicos: relações de parentesco e laços formais: laços de autoridade entre chefe e subordinado em uma empresa). Um *laço* pode possuir *propriedades* (e.g. para os laços de amizade em uma classe, o valor um pode ser usado para representar a escolha de quem é a primeira pessoa que o estudante convidaria para jantar e o valor dois para a segunda pessoa) e pode ocorrer em torno do mesmo *ator*, denominado de *loop*. Uma *relação* em uma rede consiste em um conjunto de *laços* de uma classe específica, dado um conjunto de *atores* (e.g. o conjunto de laços de amizade mantido por pares de estudantes em uma sala). Mas, é possível ter diferentes *relações* para qualquer conjunto de *atores*, (e.g. adicionalmente aos laços de amizade entre os estudantes de uma classe, é possível verificar os laços de parentesco entre os mesmos estudantes). Quanto ao *direcionamento*, as *relações* podem ser *direcionadas*, caso exista a necessidade de identificar os atores como transmissores e receptores (e.g. envio/recebimento de mensagens de e-mail) ou *não direcionadas*, quando não existe necessidade em identificar quem é o transmissor e quem é o receptor (e.g. laço de casamento entre duas pessoas). Se as linhas forem *direcionadas* são chamadas de *arcos*, caso contrário são *arestas*. As relações na rede também podem possuir *peso*, ou seja, um valor é associado a cada um dos *laços* estabelecidos na rede. Na pesquisa associada às redes sociais, os *laços relacionais* são o alvo principal da análise, embora os *atributos* também possam ser analisados

2.4 Análise de Redes Sociais

A SNA proporciona uma metodologia para analisar as relações sociais existentes entre as entidades em uma rede social, sendo amplamente usada nas ciências sociais e comportamentais, bem como nas ciências políticas, economia, ciência organizacional e engenharia industrial, dentre outras áreas.

Um sumário do progresso da SNA, organizado cronologicamente é realizado por Freeman (2004). Em seu estudo, o autor examina as contribuições ou trabalhos realizados na área em quatro momentos distintos: i) no período que antecede 1920; ii) na década de 1930; iii) de 1940 até 1969; e iv) de 1970 em diante. Mas, conforme Freeman (2004), a maioria dos

autores que examina as questões relacionadas às origens da SNA, cita o ano de 1934 como marco inicial, com o trabalho de Jacob Moreno. J. Moreno introduziu a sociometria através de sua publicação “*Who Shall Survive?*” (MORENO, 1934). A sociometria estuda as relações interpessoais. Para os pesquisadores envolvidos nessa ciência, a sociedade não é um agregado de indivíduos e suas características, como a estatística assume, mas uma estrutura de relacionamentos interpessoais, onde o indivíduo não é a unidade social básica. O átomo social consiste de um indivíduo e seus relacionamentos econômicos, sociais e culturais. Os átomos sociais são “*linkados*” em grupos e a sociedade consiste de grupos inter-relacionados.

Outro marco importante aconteceu em 1967, quando uma importante propriedade presente nas redes sociais foi apresentada por Stanley Milgram (MILGRAM, 1967) — um pesquisador da área de sociologia — descobriu que a distância média entre duas pessoas quaisquer nos Estados Unidos era próxima de seis. Para descobrir isso, foi realizado o envio de centenas de cartas a pessoas residentes em Wichita (Kansas) e Omaha (Nebraska). Os indivíduos foram escolhidos aleatoriamente e a pergunta da carta era se elas conheciam a esposa de um aluno de graduação que morava em Sharon (Massachusetts) ou se conheciam um corretor de fundos públicos em Boston. No caso de conhecerem, as cartas deveriam ser enviadas aos respectivos destinatários, senão, a pessoa deveria preencher sua identificação e enviá-la a outras pessoas que supostamente os conheciam. Essas cartas passavam por diversas pessoas até chegarem ao seu destino, mas Milgram poderia saber a rota pela qual elas teriam passado. Das 160 cartas enviadas, 42 chegaram ao destino, e com isso o pesquisador determinou o caminho médio que separava duas pessoas quaisquer nos Estados Unidos. A distância foi determinada como sendo 5,5 e arredondada para seis. A partir desta descoberta, surgiu o famoso termo “seis graus de separação”. Conforme esse princípio, a distância média entre um morador da Inglaterra e um morador de Belém, no estado do Pará, no Brasil, deve ser em torno de seis, embora, atualmente acredite-se que a distância média que separa duas pessoas em qualquer local do planeta seja menor do que a distância calculada por Milgram (BARABÁSI, 2003). Esse fenômeno é denominado como efeito de mundo pequeno (do inglês: *small world*) e acontece devido ao fato da distância de separação crescer mais lentamente do que o tamanho da rede. Como exemplo, Backstrom et al. (2012), revela que a distância entre duas pessoas no *Facebook* — uma das mais populares redes sociais nos últimos anos é de apenas 4,74 passos.

Segundo Lewis et al. (2008), o aumento do interesse na área de SNA como um campo acadêmico coincidiu com o crescimento do interesse dos usuários de Internet em sites de redes sociais⁹ (do inglês: *Social Network Sites - SNSs*), que permitem principalmente que os usuários possam: (i) construir um perfil completamente ou parcialmente público dentro de um sistema fechado; (ii) articular lista de usuários com quem compartilham conexões; e (iii) visualizar as suas listas de conexões e de outros indivíduos. SNSs possuem uma quantidade e uma singularidade de detalhes de informações pessoais que os usuários fornecem, representando uma valiosa fonte de dados para a pesquisa. Adicionalmente, periódicos especializados (e.g. *Social Network Analysis*, da Elsevier) e conferências (e.g. *International Conference on Advances in Social Networks Analysis and Mining*, promovida pela ACM/IEEE e *International Sunbelt Social Network Conference*) têm contribuído para o rápido desenvolvimento da área. Dessa forma, os pesquisadores já reconheceram o potencial dos SNSs (e.g. *Facebook*, *Linkedin*, *Twitter*, *MySpace*, *Badoo*) como um atrativo para melhorar a pesquisa e desenvolvimento na área de SNA.

Nas seções seguintes (2.4.1 a 2.4.3) são apresentados aspectos relacionados à SNA, como a visualização, medidas e softwares que podem ser usados para realizar a análise.

2.4.1 Análise Visual ou Visualização das Redes Sociais

A base da visualização de redes sociais são os sociogramas, provenientes da sociometria e otimizados matematicamente pela utilização da Teoria dos Grafos. Um sociograma é uma imagem que permite a visualização da estrutura de conexão entre os atores de uma rede. No sociograma apresentado na figura 2.3 (NOOY et al., 2005), cada garota em um dormitório de estudantes é representada visualmente por um círculo (vértice). As garotas escolhem duas possíveis parceiras para o jantar, sendo o arco rotulado com o número “1” para a primeira opção de escolha e “2” para a segunda opção. Na figura 2.3, os atributos dos vértices (ou seja, o nome da garota) e das relações (ou seja, a ordem da opção da escolha) estão representados graficamente próximos destes.

⁹ Sites de redes sociais são também denominados de redes sociais on-line.

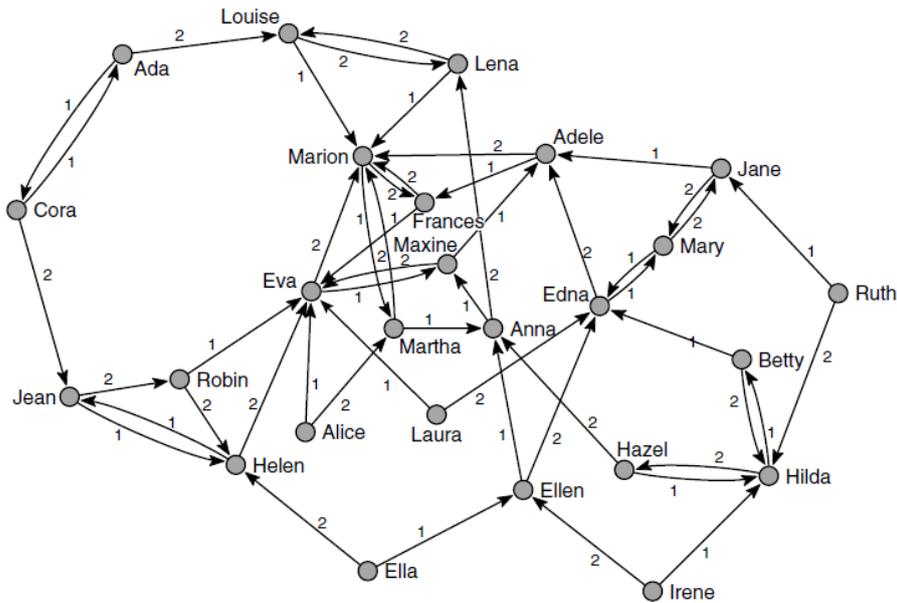


Figura 2.3 Exemplo de Sociograma. Fonte: Nooy et al., (2005).

O exemplo da figura demonstra não somente quais garotas são populares, indicado pelo número de escolhas que elas recebem, mas também se as escolhas vêm de garotas populares ou não. Exemplo: Hilda é escolhida por quatro garotas e ela retribui às escolhas de Betty e Hazel. Ruth e Irene que também escolheram Hilda, não são escolhidas por nenhuma das garotas. Mas Hilda está na margem do sociograma, enquanto Frances é mais central porque é escolhida por garotas “populares” como Adele e Marion, embora tenha sido escolhida apenas por estas duas garotas. Isso mostra que uma simples contagem das escolhas feitas para uma garota não revela a centralidade dos atores, enquanto um sociograma pode revelar.

A representação visual de um sociograma é similar à representação da estrutura de um grafo e por isso herda características inerentes à visualização de grafos. Algoritmos como os de Spritzer e Freitas (2008), Sugiyama et al. (1981), Walker (1990), foram propostos para lidar com a representação visual de grafos. Esses algoritmos procuram encontrar a melhor posição para os vértices e arestas de forma que o resultado seja de fácil entendimento, permitindo ao usuário extrair informações e compreender melhor a estrutura do grafo. Segundo Batitini et al. (1985), várias investigações realizadas com usuários mostram que certos aspectos podem facilitar a leitura de um grafo (e.g. evitar o cruzamento entre as arestas, limitar o número de vértices por aresta, evitar a sobreposição de vértices).

2.4.2 Métricas para análise de Redes Sociais

Segundo Wasserman e Faust (1994) o formalismo inicial de análise das redes esteve influenciado conceitualmente por uma visão estruturalista da sociedade, que focava a sociedade não apenas a partir de seus elementos, mas a partir da estrutura formada pelas relações entre esses elementos. Assim, em análise de redes, essa visão busca investigar qual a estrutura formada por uma determinada rede e as posições ocupadas por seus diversos atores. Essas posições revelam a influência de um ator nessa rede, ou seja, se um ator é central, intermediário ou periférico, por exemplo. Essa análise está apoiada em modelos matemáticos e computacionais, essencialmente baseados no uso da Teoria dos Grafos e contempla um conjunto de métricas/medidas. Segundo Wasserman e Faust (1994), um nó da rede é considerado importante ou proeminente se os seus relacionamentos o tornam particularmente visível aos outros nós da rede. Basicamente, existem duas classes de proeminência: centralidade e prestígio do ator (nó). Um nó de prestígio é objeto (receptor) de muitos relacionamentos e um nó central é aquele que está intensivamente envolvido em relacionamentos com outros nós, seja como transmissor ou como receptor (WASSERMAN e FAUST, 1994). Outra importante medida em redes sociais, que indica a probabilidade de dois amigos quaisquer A e B terem um amigo C em comum¹⁰ é a medida de agrupamento ou transitividade, representada pelo *coeficiente de aglomeração* (do inglês: *clustering coefficient* – cc), que basicamente mede a densidade dos vizinhos de um nó da rede. Nas seções 2.4.2.1 à 2.4.2.4 são apresentadas as medidas relevantes em SNA e utilizadas nesta tese.

2.4.2.1 Centralidade

Existem três medidas de centralidade amplamente usadas em SNA: grau de centralidade (do inglês: *degree centrality*), centralidade por proximidade (do inglês: *closeness centrality*) e centralidade por intermediação (do inglês: *betweenness centrality*).

O **grau de centralidade** — também denominado de grau de conectividade, diz respeito à visibilidade do nó na rede, ou seja, o nó mais central é aquele que possui o maior grau. Para Wasserman e Faust (1994), essa medida permite identificar os nós que estão em

¹⁰ Amigos em redes sociais tem conotação distinta da realidade, pois nem sempre duas pessoas que estabelecem uma conexão na rede social online, precisam se conhecer. Assim, um amigo em comum é quando um ator na rede social possui um amigo, que é amigo de outra pessoa, que não necessariamente é amiga do primeiro, ou seja, eles possuem um amigo em comum, mas não são amigos.

contato direto com muitos outros nós, portanto, ocupam um lugar central na rede. Por outro lado, um nó menos central indica que o ator ocupa uma posição periférica na rede. No exemplo apresentado na figura 2.4, de uma rede não direcionada, que representa a parceria na produção de artigos científicos de um grupo de pesquisadores, o nó mais central é o nó j , pois possui o maior número de conexões com outros nós (no caso, são seis), enquanto o nó i é o menos central, com apenas uma conexão na rede.

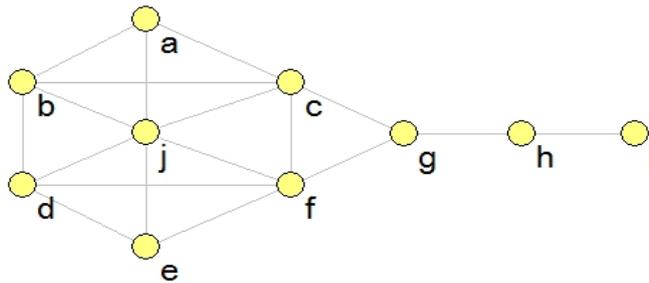


Figura 2.4 Exemplo de rede de pesquisadores com parceria em artigos científicos.

A métrica grau de centralidade pode ser definida em termos do número de arestas de um nó, ou seja, em um grafo não direcionado, o grau denotado por $d(n_i)$ é o número de linhas incidentes em um nó, ou ainda, de forma equivalente, o número de nós adjacentes a ele (Equação 2.1). O grau de um nó pode variar de 0, caso no qual o nó é isolado, até $g-1$, caso no qual o nó está em contato com todos os demais nós do grafo.

$$C_D(n_i) = d(n_i) \quad (\text{Equação 2.1})$$

Para permitir a comparação entre atores de redes diferentes, a fórmula pode ser normalizada, dividindo-se o grau do nó pelo grau máximo que um nó pode ter, ou seja, o número de nós no grafo menos 1 (o próprio nó), chegando-se à equação 2.2 (WASSERMAN e FAUST, 1994).

$$C'_D(n_i) = d(n_i)/(g - 1) \quad (\text{Equação 2.2})$$

No caso de um grafo direcionado, usualmente consideram-se duas medidas separadas de grau de centralidade: o grau de entrada (do inglês: *indegree*) e o grau de saída (do inglês: *outdegree*). O grau de entrada é o número de ligações direcionadas ao nó (Equação 2.3), enquanto o grau de saída (Equação 2.4) é a quantidade de ligações que o nó direciona a outros.

$$C_{DI}(n_i) = d_I(n_i) \quad (\text{Equação 2.3})$$

$$C_{DO}(n_i) = d_o(n_i) \quad (\text{Equação 2.4})$$

O grau de centralidade do nó, apesar de ser uma medida simples, pode ser usada na identificação de vértices denominado *hubs* — que são vértices altamente conectados. Os *hubs* têm papel fundamental na formação da estrutura das redes complexas, uma vez que a sua saída pode causar a fragmentação da rede. No caso do exemplo da figura 2.4, o único *hub* é o nó j . A partir das medidas dos graus de centralidade dos nós, é possível achar a medida da rede, ou medida global, chamada de grau médio da rede — calculada a partir da medida dos graus de centralidade dos nós da rede.

A métrica **centralidade por proximidade**, baseada em distância, mede a proximidade de um nó em relação aos demais nós da rede (WASSERMAN e FAUST, 1994). A centralidade por proximidade está inversamente relacionada à distância, ou seja, quanto maior a distância de um vértice para o restante da rede, menor é o valor da sua centralidade por proximidade. Na figura 2.4, o nó c e o nó f possuem a maior centralidade por proximidade, pois estão mais próximos dos outros nós da rede, enquanto o nó i possui o menor valor. Embora os nós c e f possuam menos ligações que o nó j , suas relações diretas e indiretas permitem acessar todos os nós da rede mais rapidamente do que qualquer outro nó, pois possuem os caminhos mais curtos para todos os outros vértices. Vértices nesta posição podem melhor monitorar o tráfego de informação na rede.

Para calcular a centralidade por proximidade é utilizada a fórmula que soma a distância geodésica (comprimento do menor caminho entre dois nós) do nó em relação a todos os demais nós do grafo. Em seguida, inverte-se o resultado deste somatório, uma vez que, quanto maior a distância, menor a proximidade (WASSERMAN e FAUST, 1994), chegando-se à equação 2.5:

$$C_C(n_i) = \left[\sum_{j=1}^g d(n_i, n_j) \right]^{-1} \quad (\text{Equação 2.5})$$

Esta medida também pode ser calculada para as relações direcionadas (FAUST e WASSERMAN, 1992) e a fórmula é exatamente a mesma que para relações não direcionadas. Especificamente, o índice de centralidade por proximidade do ator é obtido através da uniformização dos índices de modo que o valor máximo seja igual a unidade. Para fazer isso, $C_C(n_i)$ deve ser multiplicado por $g-1$ (FAUST e WASSERMAN, 1992), conforme mostrado na equação 2.6:

$$C'_C(n_i) = (g - 1) / \left[\sum_{j=1}^g d(n_i, n_j) \right]$$

$$= (g - 1) \cdot C_C(n_i) \quad (\text{Equação 2.6})$$

A **centralidade por intermediação** representa a habilidade de um indivíduo estar entre outros, ou seja, para um nó ter uma alta centralidade de intermediação ele deve estar no caminho entre diversos outros nós. Essa medida pode ser útil para apontar o quanto um ator é importante para fazer informações circularem na rede. No exemplo da figura 2.4, o nó g possui a maior centralidade por intermediação, pois está entre duas importantes circunscrições na rede. Neste caso, o nó g ocupa o papel de ponte (do inglês: *broker*) na rede, pois sem ele os nós h e i estariam desconectados do restante da rede. Outrossim, os nós a , e e i apresentam valores nulos de centralidade por intermediação.

A centralidade por intermediação do ator n_i é a soma das probabilidades estimadas sobre todos os pares de atores não incluindo o i -ésimo ator. Esta medida pode ser calculada pela equação 2.7 (WASSERMAN e FAUST, 1994), onde g_{jk} é o número de caminhos geodésicos (mais curtos e de mesmo tamanho) que ligam os nós j e k , e $g_{jk}(n_i)$ são os caminhos, no total de g_{jk} , que passam pelo nó n_i . A centralidade por intermediação para um nó n_i mede a soma de probabilidades de o mesmo estar no caminho geodésico entre todos os demais nós do grafo.

$$C_B(n_i) = \sum_{j < k} g_{jk}(n_i) / g_{jk} \quad (\text{Equação 2.7})$$

Para Batagelj e Mrvar (2008), a normalização de tal índice para redes direcionadas e não direcionadas deve ser calculada de forma diferente. Para redes não direcionadas, deve-se dividi-lo pelo seu máximo possível, que é o número de pares de nós no grafo que não incluem n_i , ou seja, $(g - 1)(g - 2)/2$ chegando-se à centralidade por intermediação normalizada (C'_B) para n_i (Equação 2.8).

$$C'_B(n_i) = C_B(n_i) / [(g - 1)(g - 2)/2] \quad (\text{Equação 2.8})$$

A normalização da centralidade por intermediação em redes direcionadas é definida pela equação 2.9 (Batagelj e Mrvar, 2008):

$$C'_B(n_i) = C_B(n_i) / [(g - 1)(g - 2)] \quad (\text{Equação 2.9})$$

A centralidade por proximidade e intermediação também são medidas de rede, ou medidas globais — calculadas a partir das centralidades individuais de cada nó da rede.

2.4.2.2 Prestígio

Medidas de prestígio podem ser calculadas apenas para redes direcionadas. A medida de prestígio mais simples no nível do ator é o grau de entrada, denotado por $d_i(n_i)$ (cf. Equação 2.3). Nesse caso, os atores com prestígio tendem a receber muitas escolhas. No exemplo apresentado na figura 2.5, de uma rede direcionada que pode representar a troca de mensagens entre um orientador (nó o) e os estudantes sob sua orientação (nós $a-d$), o nó com mais prestígio é o nó o , pois recebeu mensagens de outros três nós, só não recebendo mensagens do nó b .

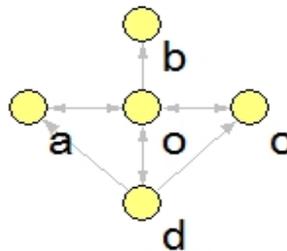


Figura 2.5 Exemplo de rede de troca de mensagens entre orientador e estudantes.

A normalização desse índice é definida pela equação 2.10 (FAUST E WASSERMAN, 1992):

$$P'_D(n_i) = d_i(n_i)/(g - 1) \quad (\text{Equação 2.10})$$

A equação 2.10 dá a proporção de atores que escolhem o ator i . O prestígio máximo ocorre quando $P'_D(n_i)=1$, ou seja, quando o ator i é escolhido por todos os outros atores (WASSERMAN e FAUST, 1994).

2.4.2.3 Densidade da rede

A densidade de uma rede é calculada através da razão entre o número de conexões de um grafo em relação ao número máximo de conexões que o mesmo grafo poderia ter. Outrossim, o cálculo da densidade para redes direcionadas e sem *loops* pode ser realizado conforme equação 2.11 (NOOY et al., 2005) ou conforme equação 2.12, caso o grafo seja

não-direcionado, onde m é o número de linhas existentes no grafo e g o número de vértices de um grafo.

$$Densidade = m/g (g - 1) \quad (\text{Equação 2.11})$$

$$Densidade = m/g (g - 1)/2 \quad (\text{Equação 2.12})$$

Matematicamente, um grafo é considerado denso quando o número de arestas está próximo do número máximo de arestas possíveis, e o contrário — um grafo com apenas algumas arestas é denominado de grafo esparso. O valor da densidade pode ser usado para comparar diferentes redes.

2.4.2.4 Coeficiente de agrupamento

O coeficiente de agrupamento mede o grau da densidade de ligações da vizinhança de um determinado nó, isto é, corresponde ao grau com que os vizinhos de um nó se interligam. O cálculo é realizado pela razão entre o número de arestas entre os vizinhos de um dado nó i , (denotado por e_i) e o número máximo possível de arestas entre esses vizinhos. Outrossim, considerando o número de arestas/arcs entre a vizinhança do nó i e o grau do nó ($d(n_i)$), o coeficiente de agrupamento pode ser calculado como a seguir, conforme Watts e Strogatz (1998), para redes direcionadas (Equação 2.13) e não direcionadas (Equação 2.14):

$$cc_i = e_i / d_{(ni)} (d_{(ni)} - 1) \quad (\text{Equação 2.13})$$

$$cc_i = 2 e_i / d_{(ni)} (d_{(ni)} - 1) \quad (\text{Equação 2.14})$$

Na figura 2.6 apresentam-se duas configurações possíveis para a vizinhança do nó a . Na primeira, todos os vizinhos do nó a estão conectados entre si, portanto, o $cc_a = 1$, enquanto na segunda existe apenas uma aresta entre os vizinhos do nó a (dentre as três possíveis), portanto, o $cc_a = 1/3$. A configuração apresenta um exemplo de *clique*, onde todos os nós estão conectados entre si. O cálculo também é possível de ser realizado com a equação 2.14 — considerando $e_i = 3$ e $k_i = 3$, na primeira configuração e $e_i = 1$ e $k_i = 3$, na segunda.

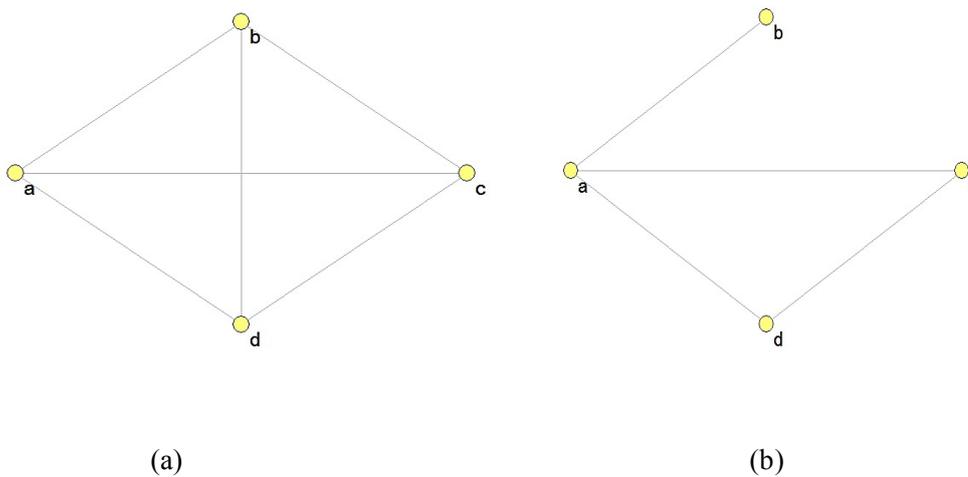


Figura 2.6 Exemplo de duas possíveis configurações para a vizinhança do nó *a*.

2.4.3 Softwares para Análise de Redes Sociais

Softwares para SNA são utilizados para representar, analisar, manipular ou visualizar redes sociais, com seus vértices e possíveis relações e permitem que os pesquisadores investiguem redes de diferentes tamanhos.

A utilização de softwares para SNA provê um conjunto de operações úteis para manipulação de redes sociais (e.g. é possível extrair partes significativas e relevantes para serem analisadas; escolher o que será analisado primeiramente em redes com mais de um tipo de relação entre os nós; representar um grupo de atores por um único nó), ou seja, várias operações de manipulação de grafos são facilitadas pelo uso de softwares específicos para lidar com SNA, que de outro modo levariam muito tempo para serem executadas.

Outro aspecto relacionado aos *softwares* diz respeito a possibilidade de diferentes visualizações das redes sociais formadas, uma vez que estas são importantes para a compreensão dos dados e transmissão dos resultados da análise. Em muitos casos, a visualização é usada como um método de análise de dados único ou pode ser utilizada junto com outras técnicas. Com relação à visualização, essas ferramentas de SNA são usadas para alterar a disposição, as cores, o tamanho e outras propriedades úteis de representação de redes.

Exemplos de software dessa natureza são: Ucinet (2016), Pajek (2016), ORA (2016) e Gephi (2016). Nesta tese, a manipulação das redes foi realizada utilizando o software Pajek¹¹.

2.5 Mineração de Redes Sociais

A mineração de redes sociais inclui várias tarefas, como: a descoberta e análise de comunidades (LI e SONG, 2013); recomendação de amigos (AGARWAL e BHARADWAJ, 2013); análise de comportamento do usuário em fóruns (BARBIER e LIU, 2011; LANG e WU, 2012); modelagem de rede social (CHUA e LIM, 2011; KAZIENKO et al., 2011; YAGER, 2008); identificação e descoberta de redes complexas, crescimento e padrões de evolução das redes com o uso de aprendizado de máquina, de mineração de dados ou de sistemas baseados em agentes (AIELLO, 2012; MUSIAL, et al., 2013; READ, 2010; SHANG e YUAN, 2012; SRIVASTAVA, 2008).

Periódicos, como *Social Network Analysis and Mining* (SNAM), da editora *Springer* e conferências, como *IEEE/ACM International conference on Advances in Social Network Analysis and Mining* (ASONAM) e *Brazilian Workshop on Social Network Analysis and Mining* (BraSNAM) surgiram a partir do avanço e reconhecimento da área de Análise e Mineração em Redes Sociais — um campo de pesquisa multidisciplinar em rápido crescimento. A recomendação feita pelos cientistas sociais é que as técnicas de mineração de redes sociais sejam usadas para grandes conjuntos de dados, quando a possibilidade de investigar com técnicas convencionais de SNA torna-se difícil ou impossível de ser aplicada. A principal preocupação desses cientistas é que o uso das técnicas de mineração de redes sociais proporcione uma regressão da pesquisa, gerando trabalhos puramente descritivos, que não realizem considerações a respeito do seu verdadeiro significado. Scott (2011) relata com ênfase essa questão, quando conclui que, seria uma catástrofe se os usos das técnicas de mineração de redes sociais fizessem que retornássemos para situações anteriores, onde os pesquisadores estavam mais interessados em encontrar padrões do que na verdadeira interpretação deles.

Nas seções seguintes (2.5.1 a 2.5.3) são realizadas descrições do processo de identificar/descobrir comunidades ou *clusters* empregados nesta tese, bem como os índices

¹¹ Este software foi usado devido a experiência adquirida pela autora desta tese junto aos grupos com os quais interagiu no processo de doutorado.

estatísticos usados para encontrar a associação entre os dois tipos diferentes de *clusters* encontrados.

2.5.1 Identificação de Comunidades ou *Clusters*

Segundo Fortunato (2010), comunidades podem ser chamadas de *clusters* ou módulos e são compostas por um conjunto de vértices que compartilham propriedades comuns e/ou possuem papéis similares dentro de um grafo. A estrutura de comunidades ou agrupamento é uma característica relevante a ser observada em grafos, nos mais diversos contextos (e.g. na biologia, ciência da computação, engenharia, economia, política). Por exemplo, em uma rede de interação entre proteínas, as comunidades podem representar grupos de proteínas com as mesmas funções específicas dentro das células, já em um grafo na *World Wide Web*, comunidades podem corresponder a grupos de páginas com tópicos idênticos ou similares. De fato, comunidades sociais, tem sido uma área de estudo há bastante tempo (FREEMAN, 2004; KOTTAK, 2004), entretanto com a difusão da Internet, uma nova definição para o termo comunidades ganhou evidência — os grupos virtuais ou comunidades virtuais ou *on-line*, que interagem principalmente pela Internet (e.g. SNSs, salas de bate-papo, fóruns ou listas de discussão, jogos, blogs, mundos virtuais). Segundo Porter (2004), uma comunidade virtual pode ser definida como um conjunto de membros ou parceiros profissionais, que interagem em torno de um interesse comum, cujas interações são suportadas e/ou mediadas por tecnologia e guiada por alguns protocolos/regras.

Detectar comunidades, com a identificação dos módulos e seus vértices, permite uma classificação dos vértices quanto à sua posição estrutural. Por exemplo, se um vértice possui uma posição central em um *cluster*, possuindo um grande número de conexões com outros *clusters*, este pode ter uma importante função de controle e estabilidade dentro desse grupo de vértices, mas se um vértice está localizado estruturalmente no limite do módulo, este pode exercer um importante papel de mediação com diferentes comunidades. Na figura 2.7, ilustra-se um exemplo de um grafo com três comunidades, identificadas pelos círculos tracejados, onde essas estruturas de comunidades identificadas revelam uma forte conexão entre seus vértices, com uma quantidade maior de arestas em um mesmo *cluster*, ou seja, conexões internas densas, e poucas arestas entre os *clusters*, ou seja, conexões externas mais esparsas.

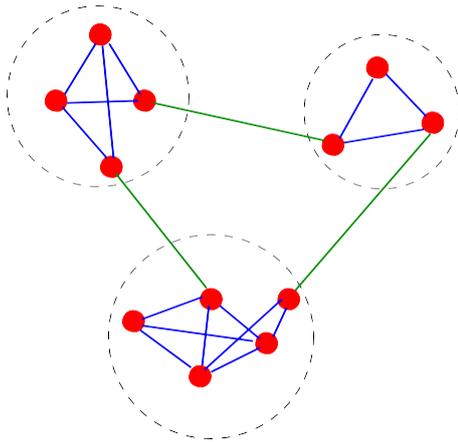


Figura 2.7 Grafo simples com três comunidades, identificadas pelos círculos tracejados. Fonte: Fortunato e Castellano (2009).

Em SNA, é possível que as linhas/arestas/arcs e os vértices possam ter valores associados. No caso dos vértices, além da denominação que recebem, também podem representar um atributo/característica específica do vértice para a análise (e.g. vértices que representam médicos em um programa do governo e que estão rotulados com a região do país na qual o médico pertence).

Com os atributos, é possível classificar ou agrupar os vértices de tal forma que cada vértice possa ser assinalado a exatamente uma classe ou *cluster* (NOOY et al., 2005). Dessa forma, as classes passam a ser o atributo do vértice — representando uma característica deles (e.g. classe “1” para a Região Norte, “2” – Nordeste, “3” – Centro-Oeste, “4” – Sul e “5” – Sudeste). Desta forma, classificar significa criar diferentes *clusters* definidos por um atributo.

Esta abordagem é útil, pois permite: (i) a análise de apenas um *cluster*, representado por uma sub-rede da rede social de todos os participantes, que deve revelar as medidas estruturais apenas do *cluster* analisado; (ii) comparação dos diferentes *clusters* em questão, utilizando medidas de SNA; e (iii) visualização de componentes da rede mais fortemente conectados em um determinado *cluster*.

2.5.2 Detecção automática de comunidades

O problema de detecção de comunidades requer a partição de uma rede em comunidades de nós com ligações internas mais densas, e com poucas ligações para os nós que pertencem a outras comunidades. Segundo Fortunato (2010), este problema ainda não foi resolvido de forma satisfatória, apesar do esforço de um grande comunidade interdisciplinar

de cientistas que trabalham nele ao longo dos últimos anos, embora seja possível encontrar vários algoritmos/métodos propostos para encontrar boas partições de um modo razoavelmente rápido. De fato, em redes de pequena escala, a identificação das estruturas de comunidades ou *clusters* pode ser feita por uma simples inspeção visual, mas no caso de redes mais fortemente conectadas e densas (com um grande número de nós), é imprescindível o uso de algoritmos computacionais para a descoberta de *clusters*.

Na literatura, podem ser encontrados vários tipos de algoritmos para detecção de comunidades, como aqueles que realizam divisões sucessivas na rede para detectar arestas que conectam vértices de diferentes comunidades, a fim de removê-los da rede, de forma que as comunidades fiquem desligadas umas das outras. Os algoritmos mais populares desta categoria são os propostos por Girvan e Newman (GIRVAN e NEWMAN, 2002; NEWMAN e GIRVAN, 2004), que usam a medida de centralidade por intermediação para realizar a remoção sucessivas das arestas com maiores valores. Existem os algoritmos aglomerativos, que mesclam nós/comunidades similares de forma recursiva, a partir de uma medida de semelhança entre os vértices. Para isso diferentes métodos de similaridade podem ser usados, dependendo da forma com que os algoritmos escolhem com que as comunidades se juntem. O algoritmo proposto por Pons e Latapy (2006) usa os caminhos aleatórios entre os vértices como medida de similaridade. Também existem os algoritmos de otimização, baseados na maximização de uma função objetivo, como os propostos por Clauset et al. (2004), Newman M. (2006) e Wu e Huberman (2004).

Nas duas seções seguintes, são tratados especificamente da função modularidade (2.5.2.1), uma medida que permite a avaliação da divisão de uma rede em comunidades e o método/algoritmo Louvain (2.5.2.2), bastante conhecido por implementar a função modularidade.

2.5.2.1 Função Modularidade

Em detecção de comunidades, é importante usar medidas que avaliem a divisão da rede em comunidades. Para isso, diferentes critérios podem ser usados, (e.g. tamanho de corte, condutância, fração de grau de saída, volume, modularidade) (LESKOVEC et al., 2009; LESKOVEC et al., 2010). A modularidade é uma função de qualidade bastante conhecida e utilizada (BLONDEL et al., 2008; FORTUNATO, 2010), proposta por Newman e Girman

(2004). A modularidade de uma partição é um valor escalar (entre -1 e 1) que mede a densidade das conexões dentro das comunidades. O mecanismo dessa função é utilizar um modelo aleatório da rede tomada como alvo do estudo, considerando-se o grau dos vértices que a compõem. Uma partição é considerada de melhor qualidade, conforme suas conexões se diferenciam do modelo aleatório da rede. A divisão da rede em comunidades será considerada ótima no momento em que o valor da função de modularidade for máximo. Existem diversos trabalhos que visam a otimização desta função, dentre eles os de Clauset et al. (2004), Leon-Suematsu e Yuta (2010) e Newman M. (2006). Nesta tese, o algoritmo usado é o Louvain, pois é um método embutido em alguns softwares (e.g. Pajek e Gephi), que permitem uma divisão de rede em comunidades com alta modularidade.

2.5.2.2 Método Louvain

Blondel et al. (2008) propôs um algoritmo fácil de implementar e extremamente rápido, que proporciona partições de alta modularidade em grandes redes. Este algoritmo proporciona diferentes resoluções de estruturas de comunidades. O funcionamento do algoritmo é baseado em duas fases que são repetidas iterativamente. Inicialmente, o número de comunidades corresponde ao número de nós na rede, e para cada nó i é calculado o aumento/ganho de modularidade obtido ao mover este nó i para a comunidade de seu nó vizinho j . Assim, o nó i é movido para a comunidade em que o aumento de modularidade for máximo, mas somente se este aumento for positivo, e caso não seja, o nó permanece na comunidade em que estava. Este processo é aplicado repetidamente e sequencialmente para todos os nós, até que nenhum aumento mais possa ser alcançado na modularidade. Quando isso acontecer, o algoritmo considera a primeira fase como executada. O ganho da modularidade ΔQ obtido ao mover um nó isolado i para uma comunidade C é calculado pelo algoritmo conforme a equação 2.15:

$$\Delta Q = \left[\frac{\sum_{in} + k_{i,in}}{2m} - \left(\frac{\sum_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \quad (\text{Equação 2.15})$$

Onde \sum_{in} é a soma dos pesos das ligações internas à comunidade C ; \sum_{tot} é a soma dos pesos dos *links* incidentes aos nós da comunidade C ; k_i é a soma dos pesos das ligações incidentes ao nó i ; $k_{i,in}$ é a soma dos pesos das ligações do nó i para os nós em C ; e m é a soma dos pesos de todos os *links* da rede.

A segunda fase do algoritmo consiste em construir uma nova rede, cujos nós são as comunidades detectadas na primeira fase, e os pesos das ligações entre os novos nós é dado pela soma do peso das ligações entre os nós nas duas comunidades correspondentes. Após completar essa fase, o algoritmo executa a primeira fase para a rede resultante e esse processo é repetido até que não haja mais mudanças. Nesse momento, espera-se que o máximo de modularidade seja alcançado. Esse algoritmo foi aplicado e testado por Blondel et al. (2008) para identificar comunidades em uma rede com 118 milhões de nós, levando cerca de 152 minutos para apresentar os resultados. Segundo os autores, a maior dificuldade enfrentada para o algoritmo é quanto à capacidade limitada de armazenamento, e não quanto ao tempo de execução.

2.5.3 Medida de associação entre diferentes classificações

Com a possibilidade de criar diferentes *clusters*, a partir de diferentes classificações, é possível analisar a associação entre elas, usando diferentes índices estatísticos como os índices V de Cramer e Rajski¹² (NOOY et al., 2005). Esses índices variam de 0 a 1, mas se os valores variam entre 0 e 0,05 — não existe associação; de 0,05 a 0,25 — fraca associação; de 0,25 a 0,60 — associação moderada e acima de 0,60 — forte associação. O índice V de Cramer mede a dependência estatística entre as duas classificações. Esse índice não é confiável caso a tabulação cruzada possua muitas células com valores iguais ou próximos de 0 ou contenha diversos vértices presentes em uma classificação, mas não na outra.

O índice Rajski mede o grau com que as informações em uma classificação são preservadas na outra classificação. Para isso, esse índice deve ser medido em três variações: (i) a medida de simetria entre as duas classificações; (ii) a medida em que a primeira classificação pode ser prevista pela segunda; e (iii) a medida em que a segunda classificação pode ser prevista pela primeira.

Na tabela 2.1, é apresentada uma classificação dos países no comércio de fabricantes de metal, conforme os anos de 1980 e 1994. Na tabulação cruzada, as linhas contêm as quatro classes em 1980: 1 – países centrais; 2 – semiperiféricos fortes; 3 – semiperiféricos fracos; e 4 – periféricos, enquanto as colunas contêm as três classes em 1994: 1 – países centrais, 2 – semiperiféricos; 3 – periféricos. Os resultados mostram que os países que eram

¹² Esses índices são calculados pelo software Pajek, após operações de manipulação de clusters.

semiperiféricos fortes ou fracos (em 1980) constituem a maior parte dos países semiperiféricos na classificação de 1994. Os índices de Cramer V e Rajski foram computados para o exemplo da tabela 2.1 e o índice de Cramer não demonstrou ser confiável porque muitas células na tabulação cruzada possuem valor 0. Entretanto, quanto ao índice de Rajski, a classificação dos países em 1994 pode ser prevista partir das posições dos países em 1980, pois o valor computado para esse índice é 0,68; o valor para a medida de simetria entre as duas classificações é 0,34; e o valor para que a classificação relacionada a 1980 possa ser prevista pela classificação de 1994 é 0,40. A conclusão é que as duas classificações estão fortemente associadas.

Tabela 2.1 Distribuição dos países nas categorias de 1980 e 1994 – tabela adaptada de Nooy et al. (2005).

Categorias em 1980	Categorias em 1994			Total
	1	2	3	
1	10	1	0	11
2	1	16	0	17
3	0	15	0	15
4	0	4	5	9
Total	11	36	5	52

2.6 Comentários Finais

Neste capítulo, foram apresentados conceitos da teoria dos grafos, importantes para entender as definições e conceitos básicos da área de Redes Sociais, bem como definições importantes dessa área. Na sequência, foram destacadas técnicas/ferramenta de SNA importantes para quem deseja investigar as redes formadas, como a visualização; principais medidas possíveis de serem extraídas em uma rede social, como centralidade, prestígio e coeficiente de agrupamento; e a importância do uso de softwares específicos, que permitem a execução de operações e visualizações das redes sociais. Essas técnicas/ferramentas foram utilizadas nos processos de SNA embutidos na metodologia desenvolvida nesta tese.

No caso de lidar com tarefas que necessitam do uso de técnicas sofisticadas, abordou-se a área de Mineração de Redes Sociais, tratando especificamente da identificação de comunidades. Assim, por ser fundamental para o desenvolvimento da metodologia desenvolvida nesta tese, foram descritas duas formas básicas de encontrar

cluster/comunidades nas redes: a primeira, mais simples, em que os vértices são classificados/agrupados usando-se um atributo específico e a segunda, que utiliza algoritmos computacionais para a descoberta de *clusters*. Algumas medidas para avaliar a divisão da rede em comunidades, usando diferentes critérios foram citadas, entretanto, foi descrita a função modularidade, por ser uma função de qualidade bastante conhecida e utilizada no algoritmo Louvain, que foi usado nas simulações desenvolvidas nos experimentos realizados nesta tese. Para finalizar, descrevem-se dois índices estatísticos (V de Cramer e Rajska) para realizar a associação entre esses diferentes tipos de *clusters*.

3. TRABALHOS RELACIONADOS

Neste capítulo, apresenta-se uma descrição de trabalhos que envolvem a avaliação/acompanhamento dos participantes de intervenções de formação em TICs e aqueles que propõem métodos, frameworks ou instrumentos para analisar/medir a exclusão digital. Por fim, realiza-se uma síntese desses trabalhos.

3.1 Prolegômenos

Na literatura especializada, encontram-se vários relatos de intervenções de formação em TICs que também realizam a avaliação/acompanhamento dos aprendizes, bem como os estudos que se dedicam exclusivamente à avaliação dos participantes de uma intervenção que foi conduzida por terceiros. Esses dois tipos de trabalhos foram analisados, para investigar-se aspectos e métodos que poderiam ser usados no desenvolvimento da metodologia proposta nesta tese. Nessa pesquisa, também foram considerados *frameworks*, modelos e instrumentos, desenvolvidos para realizar a medida de *digital/divide* em um determinado contexto, por serem considerados importantes instrumentos para a construção do modelo conceitual adotado nesta tese. Em suma, descrevem-se dois tipos de estudos¹³, que são: (i) aqueles cujo o foco são os relatos dos instrumentos/procedimentos usados para realizar a avaliação dos indivíduos aprendizes (seção 3.2); e (ii) aqueles relacionados ao desenvolvimento de modelos, *frameworks* e ferramentas para medir a exclusão/inclusão digital/*digital divide* (seção 3.3).

3.2 Intervenções para avaliação ou acompanhamento de indivíduos

No estudo apresentado por Ashraf et al. (2009), são examinados os impactos e as consequências de uma intervenção de TICs em uma comunidade localizada na aldeia Chandanbari — área rural de Bangladesh. A intervenção considera um centro de acesso comunitário multipropósito, denominado *Gonokendra*, que oferta formação em informática básica, dentre outras atividades. Segundos os pesquisadores, o alvo do estudo foi selecionado

¹³ Na descrição dos trabalhos, procura-se manter os termos e as descrições usadas pelos pesquisadores em seus estudos. Os comentários e conclusões sobre esses trabalhos são feitos nas seções 3.4 e 3.5, pela autora desta tese.

em virtude da continuidade das ações do projeto no local (mínimo de cinco anos) e pelas facilidades de infraestrutura para recebê-los. O estudo é classificado como exploratório, qualitativo e interpretativo baseado em aspectos socioeconômicos.

Os pesquisadores conduziram entrevistas pessoais no local, discussões em grupos específicos e observação das atividades, no contexto das ações de formação em TICs. Para demonstrar seus resultados, são utilizadas narrações de histórias dos seguintes grupos: (i) indivíduos envolvidos com o projeto *Gonokendra* desde a sua concepção (e.g. gerente, funcionário); (ii) indivíduos que concluíram com sucesso o programa de formação em TICs ou que terminaram pelo menos um curso de informática básica; (iii) bibliotecários ou formadores; (iv) população local que tinha conhecimento da atuação do projeto; e (v) responsáveis pelos menores que frequentam o *Gonokendra*. As análises dos resultados são baseadas no modelo de Heeks (2005), que envolve quatro estágios (acesso, avaliação, aplicação e ação) e quatro tipos de recursos (de dados, econômicos, sociais e de ação). O estudo demonstra o impacto de uma intervenção em um contexto social específico, relata perspectivas dos membros da comunidade (nível micro) e identifica aspectos de interesse dos cidadãos na formação e na utilização das TICs; dificuldades de aceitação a longo prazo; questões de discriminação de gênero; e barreiras socioculturais, dentre outros.

Em outro estudo, realizado por Berger e Croll (2012) é feita uma avaliação de cursos de educação não-formal, financiados pelo Ministério de Economia, do Governo Federal Alemão. Os cursos foram oferecidos para prover habilidades de uso de TICs aos seguintes grupos específicos: (i) idosos com deficiência visual; (ii) jovens imigrantes da Rússia; e (iii) mulheres com baixa renda salarial.

Na formação para idosos com deficiência visual, o objetivo do curso foi desenvolver habilidades básicas de uso de tecnologias e tecnologias assistidas, no período de duas semanas. A metodologia compreende teoria e prática; exploração livre e assistida dos conteúdos; e trabalho em grupo. No curso para jovens imigrantes da Rússia, o objetivo foi desenvolver habilidades básicas em tecnologia de rede, no período de onze semanas. O curso é teórico e prático, baseado em materiais *on-line*, fornecidos pela *Cisco Systems*¹⁴ — Empresa Multinacional, sediada na Califórnia. A metodologia do curso envolve trabalhos individuais e

¹⁴ O site da empresa Cisco Systems é <http://www.cisco.com/>

utiliza exemplos de práticas de negócios. No curso para mulheres com baixa renda salarial, o objetivo foi o desenvolvimento de habilidades básicas de uso de tecnologias, no período de uma semana. Embora seja orientado pelo interesse das participantes, o curso também é estruturado, teórico e prático.

Nesse estudo, cada grupo respondeu dois tipos de questionários, o primeiro preenchido durante o curso, com a presença dos avaliadores, e outro por telefone, após dois meses do final do curso. Os questionários possuem questões relacionadas à frequência do acesso ao computador e Internet, bem como uma autoavaliação sobre as habilidades de uso de TICs antes/durante e após o curso. O primeiro questionário possui questões adicionais sobre os objetivos e motivações dos participantes. Os formadores/instrutores também preenchem um questionário e recebem um *feedback* sobre o curso ministrado. Segundo os autores, esse modelo de avaliação privilegia a identificação de aspectos motivacionais no início do curso e o potencial de permanência no curso, além da satisfação com o treinamento, os resultados da aprendizagem e a relação dos participantes com os instrutores dos cursos.

Ferrari (2012) realizou uma análise de quinze casos com os seguintes objetivos: (i) identificar os principais componentes de competência digital em termos de conhecimentos, habilidades e atitudes necessárias para ser digitalmente competente; (ii) desenvolver descritores de competência digital para alimentar um *framework* conceitual que deve ser analisado no contexto do continente europeu. Essa análise, foi realizada para o serviço de consultoria em ciências — *Joint Research Centre*, da Comissão Europeia. Dessa pesquisa, foram selecionados quatro estudos, descritos nos parágrafos seguintes.

O *framework* iDCA (do inglês: *Instant Digital Competence Assessment*) (CALVANI et al., 2012) é parte do projeto "Internet e Escolas: Problemas de Acessibilidade, Políticas de Igualdade e Gestão da Informação". Esse *framework* compreende uma série de testes, para alunos entre 14 e 16 anos e prevê a coexistência de dimensões caracterizadas em níveis tecnológicos, cognitivos e éticos, integrados entre si.

A avaliação de competências digitais no *framework* iDCA é o principal objetivo do projeto. Os testes são realizados a partir da plataforma Moodle (2016), disponíveis mediante inscrição, compostos por questionários e tarefas, representando problemas do cotidiano. Esses questionários e tarefas possuem atividades como (FERRARI, 2012): exploração tecnológica, com a finalidade de usar e dominar uma interface tecnológica desconhecida; simulações, a

fim de formular hipóteses; pesquisas, para levantar informações relativas a um tema pré-definido; colaboração com outros participantes, seguindo critérios de atividade colaborativa; e participação em redes sociais, para demonstrar a consciência dos riscos e deveres relacionados a essa atividade.

Outro estudo explorado por Ferrari (2012) é o ACTIC (*Acreditación de competencias en tecnologías de la información y la comunicación*) (ACTIC, 2016) — com o objetivo de certificar a competência digital de jovens a partir de 16 anos. A certificação é baseada em um modelo de três níveis (básico, médio e avançado). Segundo Ferrari (2012), as competências avaliadas são: cultura digital e participação cívica na Web; tecnologia digital; uso do computador e do sistema operacional; navegação e comunicação; destreza com dados; informações escritas, numéricas e não escritas (áudio, vídeo e gráfico); e apresentação de conteúdo. Essas competências são avaliadas por meio de testes automatizados com limite de tempo.

A avaliação Iskills (2016), da *Educational Testing Service*, propõe-se a analisar as habilidades de resolução de problemas em um ambiente digital e o pensamento crítico dos indivíduos. As competências são medidas por meio de um teste *on-line* supervisionado, contendo tarefas que simulam cenários do mundo real em contextos acadêmico, empresarial e pessoal. Os conteúdos são equilibrados entre as ciências humanas, sociais e naturais; e assuntos práticos e de cultura popular. Esses testes podem ser contratados por instituições/organizações. Essa avaliação mede a competência digital através de sete tipos de tarefas, que representam diferentes maneiras de lidar com as TICs (FERRARI, 2012): (i) definir (compreender e articular o escopo de um problema a fim de facilitar a busca eletrônica de informações); (ii) acessar (coletar e/ou recuperar informações em ambientes digitais); (iii) avaliar (julgar se as informações coletadas satisfazem o problema); (iv) gerenciar (organizar as informações a fim de recuperá-las posteriormente); (v) integrar (interpretar e representar a informação, utilizando ferramentas digitais para sintetizar, resumir, comparar e confrontar informações de múltiplas fontes); (vi) criar (adaptar, aplicar, projetar ou construir informação em formato digital de acordo com as especificações); e (vii) comunicar (divulgar informações em formato digital para públicos específicos).

O modelo de avaliação de Newman T. (2008), com foco em professores e alunos (até 16 anos) de escolas primárias e secundárias para Alfabetização Digital, baseia-se no

entendimento de que a *literacia digital*¹⁵ é composta de competências digitais e habilidades de pensamento crítico. Tal modelo propõe uma série de perguntas que podem ser usadas para autoavaliação e dispõe de um pacote de recursos. Esses recursos podem ser adquiridos a partir da página da empresa Timmus Limited (2016).

No curso de formação à distância disponibilizado na plataforma Moodle, desenvolvido para os monitores¹⁶ de telecentros (BRASIL, 2016b) do Programa Telecentros.BR, foram usados questionários *on-line*, preenchidos mensalmente por tutores¹⁷ para realizar a avaliação dos monitores. Os questionários possuem um conjunto de questões objetivas que os tutores preenchem a respeito dos monitores sob sua responsabilidade, como: (i) uso da plataforma de aprendizagem, pelo monitor; (ii) uso de software livre; (iii) uso de tecnologias disponíveis; (iv) acesso à Internet; (v) qual fase de desenvolvimento do projeto comunitário que o monitor está; (vi) área de articulação do projeto comunitário desenvolvido pelo monitor; e (vii) ausência/não do monitor durante o mês na formação. O tutor também descreve suas observações com respeito aos avanços e dificuldades do monitor; tema do projeto; dificuldades relacionadas à execução do projeto; e no caso de ausência do monitor, quais os motivos. Como o curso de formação de monitores representa o cenário de estudo da aplicação da metodologia desenvolvida nesta tese, mais detalhes sobre esse curso estão no capítulo 5.

Em Buzato (2008), a intervenção escolhida aconteceu na Casa Brasil¹⁸ Água e Vida, localizada em Guarulhos, na cidade de São Paulo — Brasil. O estudo utiliza uma fundamentação teórico-metodológica proposta por Certeau (1994), considerando três pares de conceitos relevantes, que são: (i) produtores e consumidores; (ii) estratégias e táticas; (iii) lugar e espaço. Assim, Buzato (2008), considera que os sujeitos que utilizam o espaço analisado podem ser vistos como produtores e consumidores. Os produtores desse lugar (seus idealizadores, financiadores e administradores) criam uma estratégia de inclusão que pode ser traduzida na forma de um sistema disciplinar. Os consumidores (pessoas atendidas pelo projeto) defletem e subvertem essa estratégia por meio de usos e táticas, as quais por sua vez

¹⁵ A expressão “literacia digital” diz respeito ao uso eficaz de tecnologias digitais.

¹⁶ Monitores são indivíduos responsáveis pelo atendimento da comunidade, no telecentro, auxiliando e propondo processos que permitam à comunidade fazer uso das TICs.

¹⁷ Tutores são responsáveis por mediar o curso de formação de monitores.

¹⁸ O Casa Brasil é um Programa de Inclusão Digital da Secretaria de Ciência e Tecnologia para Inclusão Social do Ministério da Ciência e Tecnologia do Governo Brasileiro (BRASIL, 2016c).

podem constituir formas de inclusão digital não previstas e eventualmente não desejadas pelos produtores, mas que têm potencial para a transformação das relações de hegemonia-subalternidade.

A avaliação foi realizada por meio de: (i) observações de participantes em uma sessão de “uso livre”¹⁹ de aproximadamente uma hora e de uma aula do curso de introdução à informática, com duração de duas horas; (ii) entrevista casual, não estruturada, com uma voluntária da Casa Brasil Água e Vida; (iii) coleta de depoimentos de diversos usuários do serviço de “uso livre”; e (iv) duas entrevistas em profundidade com o coordenador do projeto, uma não estruturada, no início da pesquisa, e uma quase-estruturada, após a preparação do esboço dos resultados. A partir dos resultados coletados, o autor conclui que certos letramentos digitais analisados apontam para uma concepção de inclusão digital fundamentada na transformação, na agentividade e na subversão de mecanismos de ampliação e reforço de certas hierarquias de poder.

3.3 Intervenções para a medição da Exclusão/Inclusão digital

Chang et al. (2012) propõem um modelo para identificar e medir a exclusão digital/*digital divide* no contexto de governos locais em Taiwan. Os autores classificam 301 conceitos em 49 domínios que são organizados nas cinco dimensões do modelo proposto por Mutula e Van Brakel (2006), que são: infraestrutura de TICs; recursos humanos; informação; e ambientes externo e interno da organização. A construção do modelo foi realizada em duas etapas: (i) aplicação de questionários para especialistas (da indústria, do governo e da comunidade acadêmica) e o ajuste dos domínios para as dimensões, segundo os especialistas; e (ii) atribuição de pesos para as dimensões após a aplicação dos questionários para os especialistas, por meio do uso do método *Analytic Hierarchy Process* (AHP).

Após a construção do modelo, foram aplicados 32 questionários aos funcionários públicos e usuários de serviços eletrônicos (do inglês: *e-service*), do Governo da Cidade de *Chiayi*. Esses indivíduos responderam questionários sobre o grau de *digital divide* na sua própria organização. Para os pesquisadores, os resultados da aplicação do modelo permitiram a determinação do grau de *digital divide* nos aspectos pesquisados.

¹⁹ Para o autor, uma sessão de uso livre é onde o usuário/participante utiliza o telecentro com objetivos próprios de navegação na Internet ou uso de um software (e.g. editores, planilhas).

No estudo de Teles e Joia (2011), uma abordagem baseada na Teoria Ator-Rede (do inglês: *Actor Network Theory – ANT*) foi aplicada para avaliar o programa Pirai Digital, no município de Pirai, no Rio de Janeiro — Brasil. Os autores utilizam conceitos da ANT (tradução, inscrição e irreversibilidade) para desenvolver um modelo que estabelece condições essenciais para alcançar a inclusão digital, que são: (i) sustentabilidade econômica, política, ética e legal do processo; (ii) infraestrutura e acesso a computadores e Internet; (iii) produção de conteúdo local e/ou adaptável para uso local; e (iv) educação com autonomia e qualificação dos cidadãos.

O modelo de Teles e Joia (2011) foi concretizado através da aplicação de entrevistas semiestruturadas, baseada nos elementos estabelecidos pelos pesquisadores. Para isso, foram coletadas as informações de doze atores, selecionados pela sua importância na cidade e/ou no programa. Segundo os pesquisadores, os resultados revelam que a inclusão digital em Pirai tem sido uma realidade na administração pública local, na educação e na saúde, entretanto, foi observado que o setor produtivo e a sociedade civil local, pouco ou nada tem se beneficiado do programa em questão.

A ITU propõe um índice de Desenvolvimento de TICs (do inglês: *ICT Development Index – IDI*), que mede diferenças digitais/*digital divide* entre países. Para isso, o processo de desenvolvimento das TICs pelo qual um país precisa passar para fazer parte da sociedade da informação é descrito utilizando o seguinte modelo de três estágios (ITU, 2014) (Figura 3.1): (i) disponibilidade às TICs — refletindo o nível de infraestrutura de rede e acesso às TICs; (ii) uso de TICs — refletindo o nível de utilização das TICs na sociedade; (iii) impacto das TICs — refletindo os resultados/saídas da utilização eficiente e eficaz das TICs. Mas para chegar à fase final, e maximizar o impacto das TICs, é necessário analisar o terceiro componente do IDI, composto pelas habilidades de uso de TICs.

Dessa forma, o IDI é um índice composto baseado em três índices — disponibilidade de TICs, uso de TICs e habilidades de uso de TICs, que podem ser calculados da seguinte forma (ITU, 2014): (i) a infraestrutura de rede e de acesso pode ser medida pela largura de banda de Internet por usuário, assinaturas de telefonia móvel e fixa, número de computadores e computadores com acesso à Internet nas famílias; (ii) habilidades, podem ser medidas pelas taxas de alfabetização de adultos, taxas de escolarização secundárias e superior; e (iii) a utilização das TICs pode ser medida pela percentagem de indivíduos que utilizam a Internet, assinaturas de banda larga fixa e assinaturas de banda larga móvel (para cada 100 habitantes).

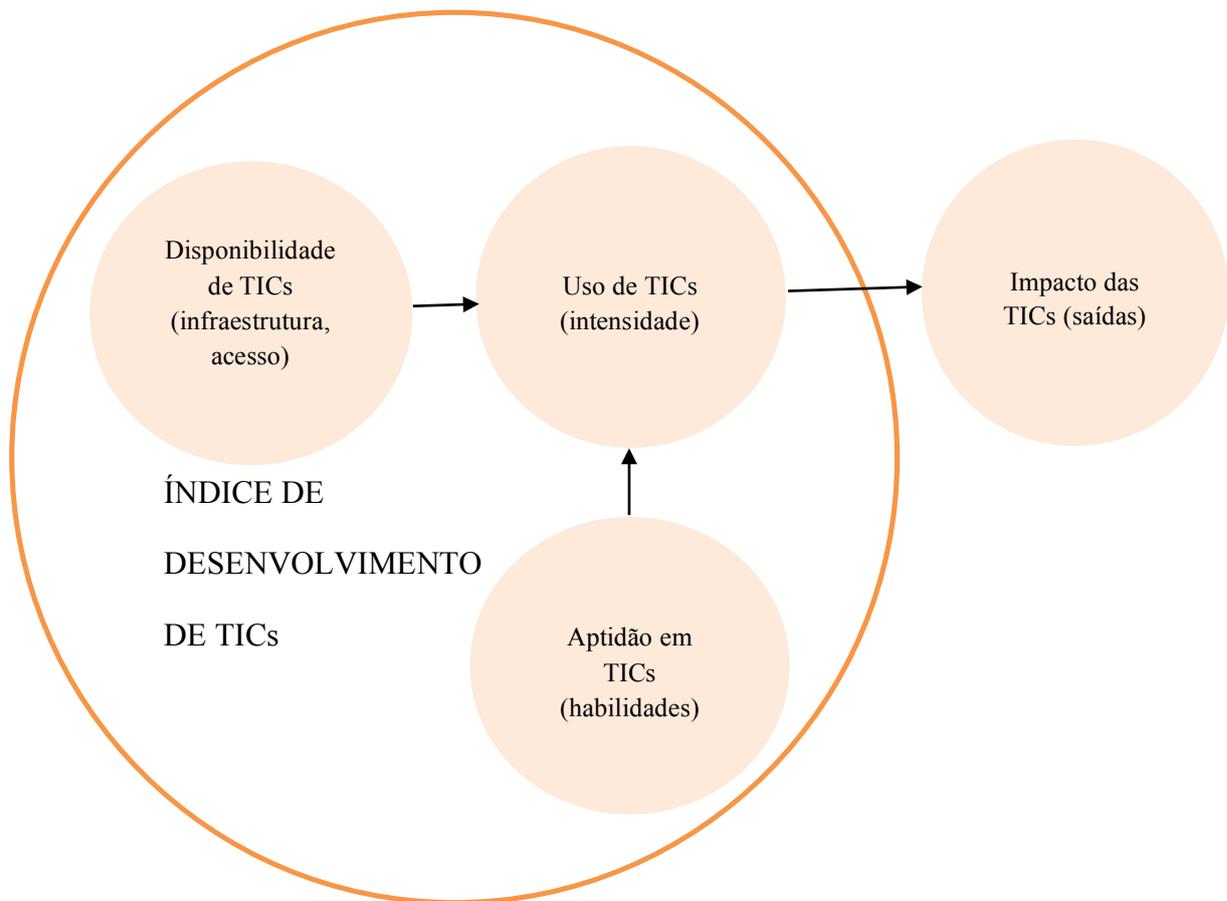


Figura 3.1 Modelo de três estágios proposto pelo ITU. Fonte: Adaptado de ITU (2014).

3.4 Síntese dos Trabalhos Relacionados

Na seção anterior, foram descritos trabalhos que avaliam habilidades de uso de TICs em diferentes níveis (e.g. de indivíduo, de grupos, cidade, países); público-alvo diferenciado (e.g. adultos, crianças, jovens, idosos, servidores públicos); e apresentam diferentes objetivos (e.g. aprendizagem do indivíduo, identificação de barreiras, uso de serviços de governo, utilização das TICs por indivíduos de uma cidade/país). Esses trabalhos utilizam diferentes métodos/instrumentos para realizar a avaliação dos aprendizes. Assim, para sintetizar, na tabela 3.1, apresentam-se os trabalhos quanto aos aspectos avaliados e métodos usados.

Tabela 3.1 Trabalhos relacionados, aspectos avaliados e métodos utilizados.

Trabalhos	Aspectos avaliados	Métodos
Buzatto (2008)	Apropriação das tecnologias pelos usuários, subversão e pró-atividade.	(i) observações de participantes; (ii) entrevista casual, não estruturada, com uma voluntária; (iii) coleta de depoimentos de diversos usuários do serviço de “uso livre”; e (iv) duas entrevistas em profundidade com o coordenador do projeto.
Newman T. (2008)	Competências digitais e habilidades de pensamento crítico.	Perguntas que podem ser usadas para autoavaliação
Ashraf et al., (2009)	Aspectos de interesse dos cidadãos na formação e na utilização das TICs, dificuldades de aceitação a longo prazo, questões de discriminação de gênero e barreiras socioculturais.	Entrevistas pessoais no local, discussões em grupos específicos e observação das atividades, no contexto das ações de formação em TICs.
Teles e Joia (2011)	Grau de inclusão do programa Piraí digital; Sustentabilidade; infraestrutura e acesso; educação e conteúdo local; e aumento da conscientização e demanda.	Realização de entrevistas semiestruturadas.
Berger e Croll, (2012)	Objetivos e motivações dos participantes Frequência de acesso ao computador e Internet, Habilidades antes/durante e após o curso.	Questionário aplicado durante o curso; Questionários aplicados por telefone, depois de dois meses do final do curso; Formadores/instrutores também são avaliados <i>Feedback</i> para os formadores/instrutores.
iDCA (CALVANI et al., 2012)	Dimensões caracterizadas em níveis tecnológico, cognitivo e ético, integrados entre si.	Testes na plataforma Moodle (2016), representando problemas do cotidiano, como: exploração tecnológica, simulação, colaboração e participação em redes sociais.
ACTIC (2016)	Cultura digital e participação cívica na Web; tecnologia digital; uso do computador e do sistema operacional; navegação e comunicação; destreza com dados, informações escritas, numéricas e não escritas (áudio, vídeo e gráfico); e apresentação de conteúdo.	Testes automatizados com limite de tempo.
Brasil (2016b)	Comunicação comunitária, redes, cultura digital, comunidade, telecentros.	Questionários online para cursistas, tutores e supervisores de tutoria. Os questionários contemplam aspectos do desenvolvimento dos projetos comunitários.
Chang et al. (2012)	Grau de exclusão digital da organização considerando aspectos de infraestrutura de TICs; recursos humanos; informação; e ambiente externo e interno da organização.	Aplicação de questionários.
ITU (2014)	Disponibilidade de TICs, uso de TICs e habilidades de uso de TICs nos países	Aplicação do Índice IDI
Skills (2016), da Educational Testing Service	Habilidades de Resolução de problemas em um ambiente digital e pensamento crítico.	Testes <i>on-line</i> supervisionado, contendo tarefas que simulam cenários do mundo real em contextos acadêmico, empresarial e pessoal. Os conteúdos são equilibrados entre as ciências humanas, sociais e naturais; e assuntos práticos e de cultura popular.

Na pesquisa dos trabalhos relacionados é possível enumerar um conjunto de observações e reflexões, que são:

- Como participantes possuem perfil diferenciado, pode ser necessário mapear os perfis para ajustar conteúdos e estratégias de aprendizagem. Além disso, conhecer o perfil do cursista antes da intervenção pode auxiliar em avaliações mais consistentes. Berger e Croll (2012), por exemplo, aplicam questionários iniciais para conhecer o perfil dos participantes;
- Barreiras relacionadas às questões de discriminação de gênero, sociais, econômicas e culturais podem impactar nos resultados de participação e de aprendizagem. No trabalho de Ashraf et al. (2009), por exemplo, são realizadas discussões em grupos específicos para identificação de barreiras sob o ponto de vista de membros da comunidade. Com isso, os pesquisadores observaram discriminação contra mulheres, pois para participar do programa, era necessária aprovação dos pais ou do marido;
- Dependendo da abordagem utilizada na formação, as avaliações podem ser diferenciadas: baseada em projetos e conteúdo (BRASIL, 2016b), orientada a problemas (CALVANI et al., 2012; ISKILLS, 2016) ou envolvendo conteúdos multidisciplinares (ISKILLS, 2016);
- A avaliação pode considerar os diferentes níveis de aprendizagem abordados na formação. Por exemplo, ACTIC (2016) contempla diferentes competências na avaliação: (i) cultura digital e participação cívica; (ii) uso do computador e sistema operacional; (iii) navegação e comunicação; (iv) informação escrita, não-escrita (áudio, vídeo e gráfico), numéricas; (vii) destreza com dados; e (h) apresentação de conteúdo;
- Um aspecto importante é a perspectiva da avaliação: autoavaliação (BERGER e CROLL, 2012; NEWMAN, T., 2008), avaliação pelo tutor (BRASIL, 2016b), testes automatizados (ACTIC, 2016) ou supervisionado (ISKILLS, 2016) e *feedback* para o tutor (BERGER e CROLL, 2012).

3.5 Comentários Finais

Na literatura, é possível encontrar vários estudos que avaliam o impacto das intervenções de TICs em um nível micro, ou seja com foco principal no beneficiário de uma

ação de intervenção. Estas investigações são geralmente realizadas por meio de questionários e/ou entrevistas, aplicados a um pequeno número de participantes, favorecendo, assim, uma visão detalhada do impacto das intervenções de TICs dentro de um contexto social particular. Essa abordagem, que é dispendiosa em termos de recursos físicos e humanos, e que necessitaria de um longo tempo para ser concluída, quando a escala dos participantes cresce, é inadequada para investigar intervenções em um extenso território geográfico. Outrossim, as avaliações em larga escala são muitas vezes padronizadas e não consideram as diferentes condições de acesso ou os níveis socioeconômicos e culturais dos participantes.

Em suma, existe uma lacuna nas estratégias, modelos, ou frameworks necessários para avaliar o impacto e monitoramento de intervenções de TICs em grande escala. A pesquisa do grupo de inclusão digital do LTS/UFPA busca preencher um *gap* existente no acompanhamento dos programas de formação em TICs em larga escala, atuando na pesquisa, propostas e desenvolvimento de modelos, frameworks e metodologias, adequadas para esses casos.

Desse modo, no próximo capítulo, é apresentado o contexto atual do desenvolvimento alcançado por este grupo nos últimos anos (2011–2016) e especificamente um dos frutos desta pesquisa — representado por uma metodologia. Essa metodologia foi aplicada em um curso de formação em TICs e os seus resultados foram divulgados em formato de artigos científicos em conferências, periódicos e capítulo de livro.

4. METODOLOGIA PARA ACOMPANHAMENTO E MONITORAMENTO DE INTERVENÇÕES DE FORMAÇÃO EM TICS

Neste capítulo, situa-se a proposta desta tese em uma abordagem proposta pelo grupo de pesquisa em inclusão digital, do Laboratório de Tecnologias Sociais da UFPA, apresenta-se o modelo conceitual que norteia o desenvolvimento dessas pesquisas e o conjunto de componentes que fazem parte da abordagem proposta pelo grupo. Finalmente, descreve-se a metodologia proposta nesta tese com base na análise e mineração de redes sociais.

4.1 Prolegômenos

Um elevado nível de acesso às TICs é visto pelas Nações Unidas e outros pesquisadores como um direito humano básico (UNITED NATIONS, 2012) que deve ser assegurado em condições de igualdade para todos os cidadãos numa sociedade globalizada. Entretanto, apesar dos avanços na qualidade e disponibilidade de TICs, o nível de acesso ou habilidades varia amplamente entre os indivíduos. Segundo (ITU, 2014), os custos elevados dos serviços e a falta de habilidades necessárias foram identificados como os principais obstáculos para uma maior utilização das TICs.

Frequentemente, em países em desenvolvimento, organizações internacionais (e.g. ONU, Banco Mundial) ou organizações nacionais (governamentais ou não-governamentais) investem esforços humanos e financeiros em programas ou intervenções que permitam o acesso às TICs ou o desenvolvimento de habilidades em TICs para os indivíduos. Essas intervenções são geralmente executadas por outras instituições e implementadas em pelo menos um grupo, comunidade ou organização (ASHRAF et al. 2008). Nesta tese foi apontado (cf. capítulos 1 e 3) que existe uma lacuna quanto aos instrumentos para realizar o acompanhamento e monitoramento de intervenções de TICs, principalmente quando elas acontecem em grande escala, para contemplar um grande número de participantes. Assim, em direção às abordagens, *frameworks*, arquiteturas, sistemas informatizados, propostas metodológicas e indicadores para acompanhar e monitorar programas de formação, atuam um conjunto de pesquisadores associados ao Laboratório de Tecnologias Sociais da UFPA. A

abordagem apresentada neste capítulo é fruto do trabalho desse grupo e tem como base um modelo conceitual de referência (seção 4.2) a partir do qual são estabelecidos os conceitos-chave das pesquisas desenvolvidas no LTS/UFPA, situadas no contexto da inclusão digital. Esta tese, está situada nesta abordagem (cf. seção 4.3) e apresenta-se como uma metodologia para realizar o acompanhamento, monitoramento de uma intervenção de formação em TICs.

4.2 Modelo Conceitual

A definição de um modelo conceitual (com suas entidades e relacionamentos) é um passo crítico para a proposta de uma abordagem para avaliação e acompanhamento de intervenções em TICs. De fato, a habilidade para especificar, modelar, projetar algoritmos e técnicas, a fim de medir resultados das intervenções depende essencialmente da especificação de um modelo conceitual de referência. Esse modelo deve enfatizar conceitos-chave que identificam fenômenos relevantes do processo de utilização das TICs pelos indivíduos nas comunidades.

Com base em intervenções de formação em TICs, que propõem instrumentos, *frameworks*, modelos ou outros artefatos que contemplam o acompanhamento ou avaliação dos indivíduos e comunidades e seus impactos (ASHRAF et al., 2008; ASHRAF et al., 2009; BERGER e CROLL, 2012; BRITO et al., 2013; CALVANI et al., 2012; CHANG et al., 2012; ITU, 2014; TELES e JOIA, 2011; VENKATESH e SYKES, 2013), foi desenvolvido o modelo conceitual proposto por pesquisadores do LTS/UFPA (BRITO et al., 2014b). O modelo (Figura 4.1) está sendo adotado nesta tese para guiar a metodologia proposta pela autora.

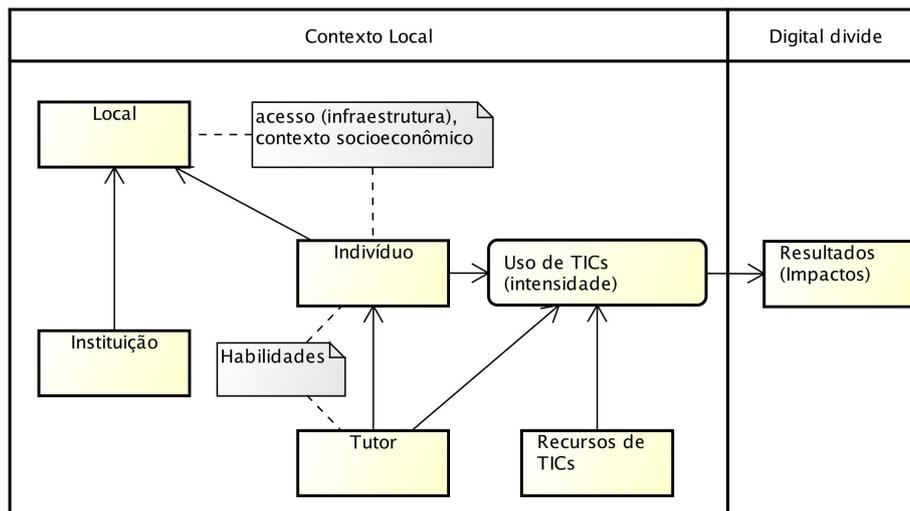


Figura 4.1 Modelo Conceitual.

Nesse modelo, a utilização das TICs está associada com a intensidade com que estas são usadas por indivíduos, em um local específico. Para melhorar a utilização das TICs, o foco principal consiste em fornecer melhorias nas condições de acesso (e.g. infraestrutura com custo reduzido) e habilitar/capacitar esses indivíduos para a utilização. A atividade de capacitar é um processo que pode ser conduzido por tutores/mediadores/instrutores/professores com habilidades suficientes para orientar os indivíduos. Outrossim, as instituições são as responsáveis pela gerência e consecução da intervenção em um local. O impacto de uma intervenção, ou seja, seus resultados, refere-se ao grau em que a utilização das TICs são intensificadas. Com essas medidas, deve ser possível identificar se houveram reduções das diferenças nas habilidades digitais entre os indivíduos/grupos analisados.

Em suma, o modelo conceitual desenvolvido é composto das seguintes entidades:

- *Indivíduo*: é o principal beneficiário das ações de intervenção. Esses indivíduos podem pertencer à diferentes localidades em uma ação de intervenção. Cada indivíduo tem características determinadas por diferentes aspectos (e.g. étnico, cultural, socioeconômico);
- *Instituição*: consiste de uma entidade pública, privada ou do terceiro setor, com ou sem história prévia de ação na comunidade e que tem a responsabilidade de promover localmente a intervenção;

- *Recursos de TICs*: são os recursos que podem ser utilizados pelos membros da comunidade para finalidades variadas (e.g. uso de serviços de governo eletrônico, conteúdos digitais, redes sociais *on-line*, cursos);
- *Tutores*: são agentes que atuam nas intervenções para difundir, disseminar e auxiliar os indivíduos das comunidades no uso dos recursos de TICs;
- *Local*: identifica onde o indivíduo vive e/ou onde as intervenções estão beneficiando esses indivíduos (e.g. aldeia, bairro, ilha, município). *Local* tem indicadores demográficos e socioeconômicos variados (e.g. educação, ocupação, renda e níveis de infraestrutura);
- *Resultados*: são identificados por medidas do uso de recursos de TICs (e.g. medidas do uso das redes sociais *on-line*, dos serviços de governo eletrônico, taxa de *download* de conteúdo digital, número de acessos nos cursos).

4.3 Abordagem para monitoramento e acompanhamento de programas de formação em larga escala

O acompanhamento apropriado dos participantes (e.g. indivíduos de uma comunidade, tutores) de um programa de formação em larga escala requer metodologias, *frameworks* e o desenvolvimento de ferramentas automatizadas, que permitam que gestores e financiadores possam tomar decisões no decorrer do processo de formação ou mesmo que permitam que novos programas de formação possam ser remodelados baseados em tais análises.

A abordagem desenvolvida objetiva o acompanhamento dos participantes de uma intervenção apoiada por uma plataforma de aprendizagem com diferentes recursos de interação (e.g. salas de bate-papo, fóruns ou listas de discussão, mensagens instantâneas, *blogs*, *wikis*). As interações com essas ferramentas podem configurar conexões/*links* entre os participantes (e.g. envio/recebimento de mensagens através do chat) ou *links* entre os participantes e os recursos (e.g. *links* entre um participante e conteúdos acessados), indicando que os participantes estão engajados ativamente no uso das mídias sociais ou recursos úteis de TICs, e não são apenas receptores passivos de informação. Por isso, a abordagem proposta é apropriada nos casos em que a base da análise sejam os *links* entre os participantes.

Esta abordagem considera os registros de interação entre os participantes durante a formação. Dessa forma, cada interação entre os participantes — atores dessa rede, constituem

os *links* das redes a serem analisadas. Nessas redes, consideram-se que os atores, representados por nós ou vértices, possuem atributos (e.g. *papel* ou *função* na formação, *escolaridade*, *localidade* a que pertencem) que permitem diferentes tipos de refinamentos na análise das redes sociais dos participantes. Adicionalmente, na mineração de dados são investigadas hipóteses que a instituição executora e/ou contexto local (e.g. aspectos de infraestrutura da localidade do participante, indicadores socioeconômicos e/ou de educação, ocupação, renda) influenciam o alto/baixo grau de conexão dos indivíduos na formação. Na figura 4.2, apresenta-se a visão dos componentes da abordagem proposta por Brito et al. (2014b), baseada em técnicas de SNA, mineração de redes sociais (especificamente a identificação de comunidades) e mineração de dados.

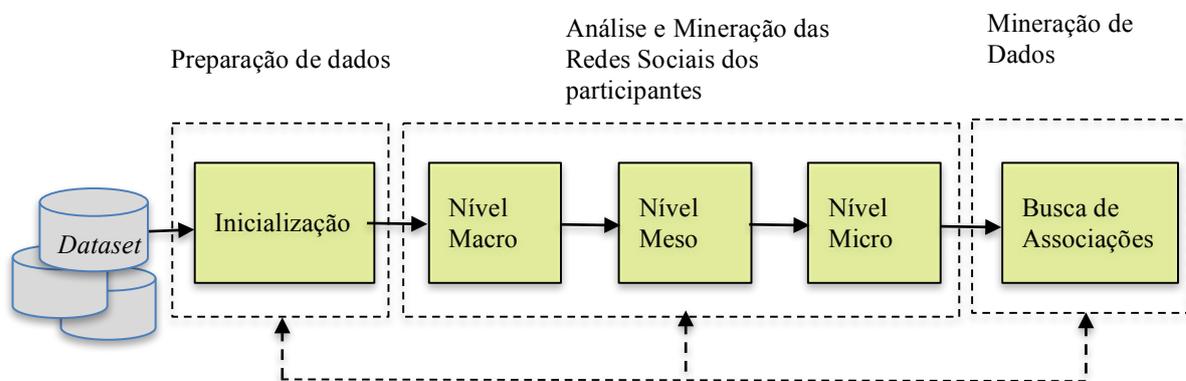


Figura 4.2 Abordagem para monitorar/acompanhar intervenções de formação em TICs em larga escala.

A abordagem é estruturada em três componentes principais: preparação de dados, análise e mineração das redes sociais dos participantes e mineração de dados. Na preparação de dados, a base é preparada e integrada, incluindo a rede social dos participantes a ser analisada nas etapas seguintes.

Para a análise e mineração das redes sociais dos participantes, que representa um componente da abordagem proposta por Brito et al. (2014b), esta tese propõe uma metodologia com a perspectiva de análise multinível, em função de permitir diferentes níveis de agregação entre os indivíduos. O componente está estruturado nos seguintes níveis, com propósitos específicos:

- Nível Macro: a fim de analisar se existe intensificação das relações de acordo com o tempo decorrido na formação;

- Nível Meso: com a finalidade de analisar a influência do papel ou da localidade do participante nas redes sociais formadas;
- Nível Micro: para identificar os atores com mais alto grau de conexão nas redes sociais da formação ou a identificação de variáveis associadas com a baixa conexão.

A perspectiva de mineração de dados, em busca de associações das redes sociais com outros parâmetros, está fora do escopo desta tese, entretanto, tem sido alvo de pesquisas (BRITO et al., 2013; BRITO et al., 2014b; SILVA et al., 2014; SILVA et al., 2016) realizadas no âmbito do grupo de inclusão digital do LTS/UFPA.

4.4 Metodologia para Análise das Redes Sociais dos participantes

Para Duch e Arenas (2005), na literatura existem redes descritas de diferentes formas e diferentes níveis, como os níveis microscópico, macroscópico e mesoscópico. Em um nível microscópico, cada nó é analisado individualmente (e.g. através do coeficiente de agrupamento, da centralidade, prestígio). Enquanto em um nível macroscópico, a rede é descrita em termos de propriedades gerais (e.g. através do seu número de nós, densidade, grau de centralidade da rede). Entre esses níveis existe a descrição mesoscópica da rede, que define a sua estrutura de comunidades.

Assim, esta metodologia está estruturada em três diferentes níveis (macro, meso e micro), em que são empregadas técnicas de SNAM para permitir diferentes análises: (i) quanto ao tempo de formação; (ii) quanto ao papel e local dos participantes; e (iii) quanto à intensidade individual de uso dos participantes nas redes sociais analisadas. As atividades propostas em cada um dos níveis estão ilustradas na figura 4.3 e detalhadas nas seções 4.4.1 a 4.4.4.

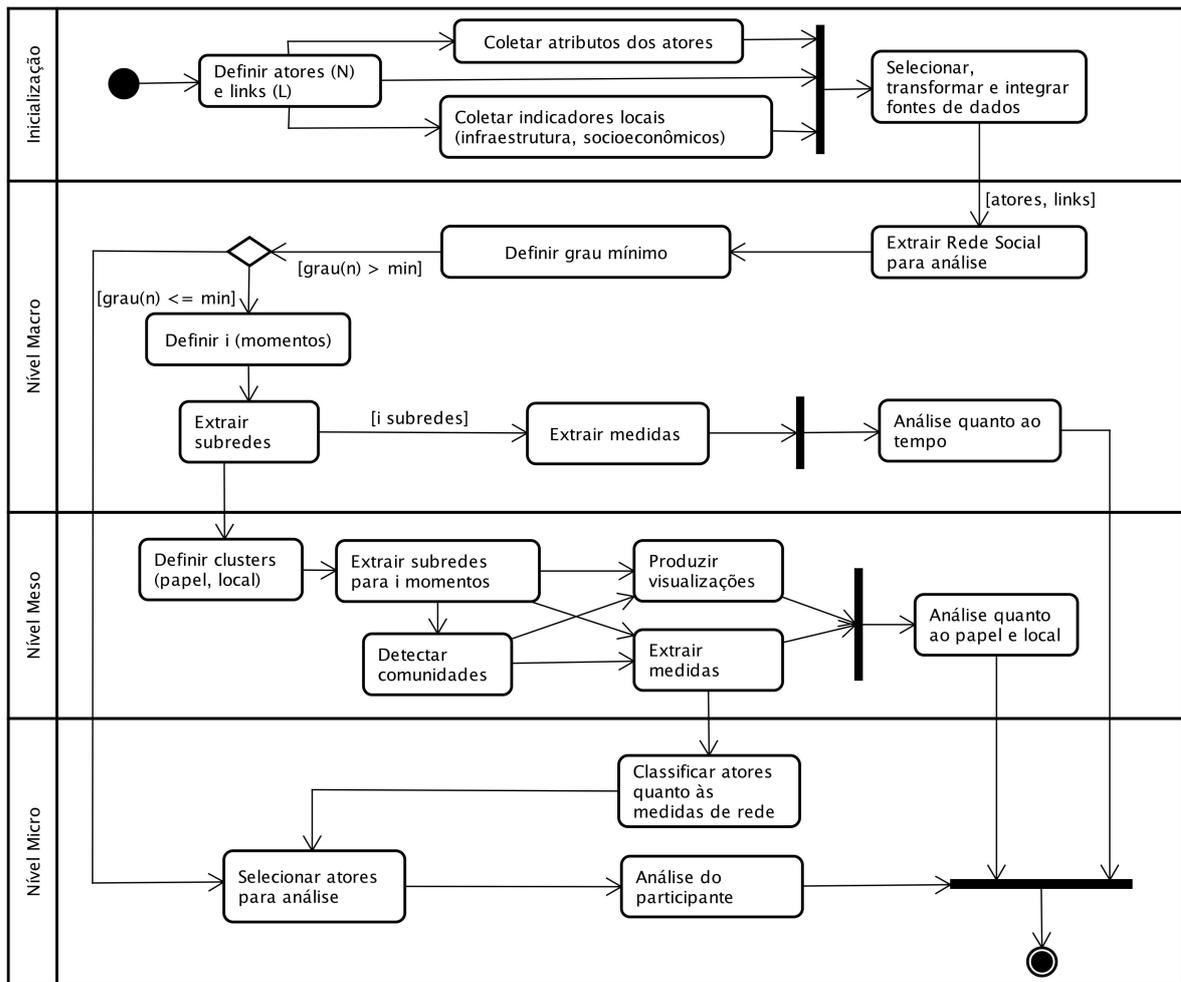


Figura 4.3 Atividades da metodologia para analisar as redes sociais dos participantes.

4.4.1 Inicialização

Na inicialização, é necessário definir a estrutura social a ser analisada, com seus atores e *links*, bem como todas as fontes de dados que serão usadas, a fim de coletar os dados necessários para que possam ser preparados e integrados. O conjunto de dado preparado e integrado serve como fonte de dados para a análise das redes sociais dos participantes e para a mineração de dados.

4.4.2 Nível Macro

Nesta fase, a rede social dos participantes é extraída a partir da fonte de dados e preparada para servir como entrada para os procedimentos que manipulam, operam, transformam e permitem a visualização das relações de uma rede social. A partir da Rede

Social preparada, deve-se identificar os critérios para a realização da análise das redes sociais dos participantes. Para isso, é possível utilizar o grau de centralidade do nó — ou simplesmente grau —, como uma medida de atividade e assim definir qual o critério para o nó/vértice prosseguir na análise do nível seguinte (meso). O resultado disso é que, os atores que não tenham satisfeito o critério mínimo de atividade, possam ser analisados no nível micro. Uma vez que o propósito central deste nível é verificar se existe uma intensificação das interações, de acordo com o tempo decorrido da formação, recomenda-se a definição dos períodos precisos de tempo para cada momento de observação (e.g. momento um: meses 1–2 e momento dois: meses 3–4). Em sequência, recomenda-se: (i) preparar as redes para cada período de tempo analisado, incluindo os nós e *links* estabelecidos naquele período; (ii) extrair as medidas estruturais para cada rede (e.g. densidade e grau médio); e (iii) realizar a comparação entre as medidas das diferentes redes, a fim de verificar se houve intensificação ou redução das interações nos diferentes períodos de tempo analisados. Para verificar se as redes ficaram mais densas com o progresso do tempo na formação, as medidas de densidade e grau médio da rede são as mais apropriadas.

4.4.3 Nível Meso

Neste nível, a fim de entender melhor a influência dos atributos *papel* e *local* do participante na formação, os seguintes passos podem ser adotados: (i) definir *clusters* por atributos — por *papel* e por *local*; (ii) detectar automaticamente comunidades nas redes desejadas (e.g. nas redes por momento ou na rede completa); (iii) confrontar os dois tipos de *clusters* — os definidos pelos atributos *local* e *papel* com os automaticamente encontrados; e (iv) analisar a estrutura de rede das comunidades encontradas. Esses passos são detalhados a seguir:

- Definir *clusters* por atributos: Quando se define *clusters* pelo atributo *papel*, cada nó é atribuído exatamente a um *cluster*. Com esses *clusters*, é possível analisar as redes separadamente, apenas com relações entre pares. Isso torna possível a comparação de medidas das redes dos diferentes *clusters*, a fim de realizar observações relacionadas ao papel que o participante possui na formação. As medidas possíveis de serem comparadas estão na tabela 4.1. Quando se define o *cluster* pelo atributo *local*, é possível visualizar e quantificar as mensagens trocadas entre locais diferentes e mesmo local, como também é

possível identificar e visualizar os componentes da rede mais fortemente conectados. Para criar *cluster* por *local*, é necessário adotar um atributo para isso, como por exemplo a região geográfica do país: Norte, Sul, Centro-Oeste, Nordeste e Sudeste.

- Encontrar automaticamente comunidades: é possível utilizar um ou mais algoritmos de detecção para encontrar comunidades nas redes.
- Confrontar *clusters*: Com o objetivo de identificar a formação de comunidades dentro dos polos ou de forma independente (e.g. uma comunidade com a maioria de participantes de localidades diferentes), realiza-se a tabulação cruzada a partir dos *clusters* definidos pelo atributo *local* e as comunidades automaticamente detectadas. Para isso, recomenda-se o uso de índices estatísticos (e.g. V de Cramer ou Rajska — cf. seção 2.5.3) que possam medir a associação entre os diferentes tipos de *clusters*. Na sequência, é realizada a tabulação cruzada entre as comunidades detectadas e os *clusters* definidos pelo *papel*, com o objetivo de identificar diferentes atuações/funções nas comunidades formadas (e.g. uma comunidade mista com metade dos participantes de um papel e metade de outro).
- Visualizar as estruturas de rede das comunidades automaticamente detectadas (e.g. visualizar as relações entre diferentes papéis através da redução de um *cluster* para um único vértice).

4.4.4 Nível Micro

Para este nível, recomenda-se que alguns atores sejam selecionados para análise. Nesta metodologia, duas situações são recomendadas para que alguns atores sejam selecionados para análise, que são: (i) atores com grande proeminência na rede social. Para isso, uma ou mais medidas estruturais podem ser usadas para que se realize uma classificação de atores em faixa de valores (e.g. uso do grau e/ou do prestígio). Esta estratégia permite a identificação de atores em evidência, que podem possuir habilidades para atuar como instrutores/mediadores em novas intervenções de formação, e também uma análise dos atributos desses atores (e.g. qual a localidade desses atores); e (ii) atores que foram considerados sem atividade na rede social no nível macro, bem como aqueles que pararam de interagir de um momento para outro da intervenção. Esses atores devem ser analisados por meio de instrumentos que possam dar

subsídios sobre o participante (e.g. usando autoavaliação ou avaliações a respeito do participante).

4.4.5 Índices sugeridos para análise

Na tabela 4.1, sintetizam-se os principais índices a serem calculados e extraídos dos atores ou das diferentes redes, em todos os níveis desta metodologia. Os cálculos de como encontrar tais índices estão definidos na seção 2.4.2.

Tabela 4.1 Índices extraídos das redes.

Índices	Significado	Ator/Rede
Nós	Número de nós ou vértices de uma rede.	Rede
<i>Links</i>	Número de <i>links</i> de uma rede específica. Caso a rede corresponda a um determinado período de tempo significa o número corresponde apenas aos <i>links</i> estabelecidos no período.	Rede
Densidade	Medida para avaliar a estrutura da rede quanto à densidade (muitas conexões, poucas conexões). A densidade é calculada pela razão entre o número de arestas e o número máximo de arestas possível em uma rede.	Rede
Grau de centralidade	Medida do número de conexões com outros nós da rede (em redes não direcionadas). Em redes direcionadas, o grau é calculado a partir de duas diferentes medidas: o grau de entrada e o grau de saída do nó.	Ator
Grau médio	O grau médio é a média de conexões (grau de centralidade) de todos os nós de uma rede.	Rede
Maior Grau	O maior grau de uma rede é o maior grau de centralidade encontrado em um nó dessa rede.	Rede
Centralidade por proximidade	Medida da proximidade de um nó em relação aos demais nós da rede. Quanto maior a distância de um vértice para o restante da rede, menor é o valor de sua centralidade por proximidade.	Ator
Centralidade por proximidade (média)	A centralidade por proximidade média de uma rede é a média dos graus de centralidade por proximidade dos nós da rede.	Rede
Centralidade por intermediação	Medida de quanto um nó está no caminho entre diversos outros nós.	Ator
Centralidade por intermediação (média)	A centralidade por intermediação média de uma rede é a média dos graus de centralidade por intermediação dos nós da rede.	Rede
Coefficiente de agrupamento (CC)	Medida do grau de densidade de ligações da vizinhança de um determinado nó.	Ator
Coefficiente de agrupamento (médio)	O coeficiente de agrupamento médio de uma rede é a média do coeficiente de agrupamento dos nós da rede.	Rede

4.5 Comentários Finais

Neste capítulo, foi descrito o contexto desta tese — a abordagem proposta pelo grupo de trabalho de inclusão digital nas comunidades, de pesquisadores do LTS/UFPA. Inicialmente, apresenta-se o modelo conceitual de referência tomado como base nas pesquisas desenvolvidas até o momento por este grupo. Em seguida, apresenta-se a abordagem que usa os *links* entre os participantes como base de análise e possui três componentes principais: preparação de dados, análise e mineração das redes sociais dos participantes e mineração de dados. Para a análise e mineração das redes sociais dos participantes, nesta tese foi desenvolvida uma metodologia com a perspectiva de análise multinível.

Os níveis propostos na metodologia buscam investigar aspectos relacionados ao tempo de formação, a importância do local dos participantes nas redes formadas, a análise da distribuição dos papéis nas comunidades formadas e a intensidade/não de uso dos participantes nas redes sociais analisadas. Para isso, um conjunto de atividades são sugeridas para que essa análise possa ser realizada. Essa metodologia pode ser seguida em todos os seus níveis, a fim de investigar os mesmos aspectos supracitados ou apenas parcialmente, caso deseje-se investigar alguns desses aspectos.

Uma vez que as principais motivações para elaboração desta abordagem e desta metodologia tenham sido provenientes das necessidades de avaliação de cursos de formação executado particularmente por grupos de pesquisa e/ou extensão do LTS/UFPA, no próximo capítulo apresenta-se o cenário de estudo, que representa o alvo da aplicação da metodologia desenvolvida, caracterizado por um curso de formação em TICs, no qual a Universidade Federal do Pará foi uma das instituições executoras nos anos de 2010–2013.

5. APLICAÇÃO DA METODOLOGIA

Neste capítulo, apresenta-se o cenário de estudo usado para a aplicação da Metodologia desenvolvida nesta tese — o curso de formação do Programa Telecentros.BR, bem como as Fontes de Dados usadas e os resultados obtidos a partir da aplicação da Metodologia.

5.1 Prolegômenos

No Brasil, consideradas as peculiaridades de caráter geográfico e as dimensões territoriais extraordinárias, além das dificuldades de acesso entre as diversas regiões do país, são inúmeros os desafios para que uma política pública de inclusão digital contemple efetivamente todas as regiões. Nessa direção, os telecentros foram projetados para proporcionar acesso público e gratuito às TICs, com a proposta de computadores conectados à Internet, disponíveis para múltiplos usos. Assim, surgiu uma ação proposta pelo Governo Federal Brasileiro — o Programa Telecentros.BR, instituído mediante Decreto específico (BRASIL, 2009) no âmbito da política de inclusão digital nas comunidades, para apoiar a implantação de novos telecentros e fortalecer os já existentes no país, bem como atuar na formação/capacitação de indivíduos para usar e disseminar o uso das TICs nesses espaços públicos

A formação dos indivíduos do Programa Telecentros.BR é a intervenção analisada nas seções a seguir, configurando o cenário de estudo analisado nesta tese, em razão de sua grande escala em relação à abrangência geográfica, número de participantes e diferentes instituições envolvidas.

5.2 Cenário de Estudo: Curso de Formação do Programa Telecentros.BR

Na concepção do Programa Telecentros.BR, os seguintes aspectos foram previstos (BRASIL, 2009): (i) disponibilização de equipamentos de informática e mobiliário necessários ao funcionamento dos telecentros; (ii) disponibilização e manutenção do serviço de conexão em banda larga à Internet; (iii) concessão de bolsas para auxílio financeiro de indivíduos para atuar como monitores nos telecentros; e (iv) constituição de rede de formação para monitores de telecentros apoiados pelo programa. No programa Telecentros.BR,

monitores são os indivíduos responsáveis pelo atendimento da comunidade no espaço do telecentro, auxiliando e propondo processos que permitam à comunidade fazer uso das TICs.

Para a consecução da rede de formação para os monitores de telecentros, o Ministério do Planejamento, Orçamento e Gestão, por intermédio da Secretaria de Logística e Tecnologia da Informação, através de Avisos de Chamamento Público, publicados em 2010 (BRASIL, 2010a; BRASIL 2010b), definiu aspectos como: objetivos, metas, diretrizes, atividades, etapas, cronograma de execução e forma de gestão da rede de formação. Esses documentos públicos (BRASIL, 2010a; BRASIL 2010b) definiam que: (i) monitores deveriam participar de um curso de formação; (ii) o curso ficaria sob a responsabilidade de diferentes centros/polos regionais, para atender as diversas regiões/estados do país, coordenados por um polo nacional, com a função de coordenar, supervisionar e mediar as ações dos polos regionais; e (iii) os polos de formação deveriam ser representados por entidades da administração pública ou privada sem fins lucrativos.

Assim, inicialmente, a oferta do curso de formação do Programa Telecentros.BR contou com uma configuração composta por cinco polos regionais de formação (Norte, Nordeste, Centro-Oeste, Sul e Sudeste), dois polos estaduais (São Paulo e Ceará) e um polo nacional. O projeto do curso de formação envolveu: o *design* instrucional do curso, diretrizes pedagógicas, propostas de conteúdo, atividades, escolhas tecnológicas a serem utilizadas, e também as estratégias de monitoramento e avaliação usadas no decorrer da formação por gestores pedagógicos e coordenadores. É importante ressaltar que, embora a metodologia apresentada nesta tese não tenha sido adotada pelos gestores pedagógicos, coordenadores e/ou financiadores do programa Telecentros.BR, esta tem sido aplicada pelo grupo de inclusão digital do LTS/UFPA, cujos resultados têm sido apresentados em um conjunto de artigos científicos (BRITO et al., 2013; BRITO et al., 2014a; BRITO et al., 2014b; SILVA et al., 2013; SILVA et al., 2014; SILVA et al., 2015; SILVA et al., 2016).

No curso, foi utilizada uma única plataforma de aprendizagem — Moodle (MOODLE, 2016), onde conteúdos, atividades e estratégias foram definidos de forma colaborativa para atender a diversidade das características regionais do Brasil (população ribeirinha, indígenas, quilombolas, zonas rurais e urbana). O curso de formação de monitores foi estruturado em 480 horas, para ser realizado à distância, com três fases principais: fase um, cujo objetivo é a familiarização com a plataforma de aprendizagem e uma visão geral sobre os conteúdos abordados na formação; fase dois, cujo objetivo é o aprofundamento dos assuntos relevantes

do curso (e.g. informática básica, inclusão digital nas comunidades, comunicação e compartilhamento em redes, cultura digital, *software* livre); e fase três, com o objetivo de formular, desenvolver e dar visibilidades aos projetos envolvendo as comunidades no entorno dos telecentros. Cada polo de formação possuía uma equipe composta por coordenadores, gestores pedagógicos, especialista em conteúdo, supervisores de tutoria e tutores — contratados para atuar no âmbito da região do polo, a fim de acompanhar a implementação do curso e os monitores na plataforma de aprendizagem selecionada. Embora monitores e tutores estivessem vinculados a um polo regional/estadual de formação, a integração de todos os participantes em um único ambiente de aprendizagem facilitou as interações entre esses indivíduos, oriundos de diferentes regiões do país. De fato, a plataforma de aprendizagem com suas tecnologias de interação (e.g. fóruns, mensagens instantâneas, *wiki*), visava facilitar a troca de experiências entre os monitores e compartilhamento de soluções que auxiliassem na solução de problemas do dia-a-dia dos telecentros.

Neste cenário de estudo, a aplicação da Metodologia desenvolvida nesta tese mostra-se adequada para:

- Analisar a intensidade das conexões, de acordo com o progresso do curso;
- Analisar as relações entre pares e entre todos os participantes, ou seja, apenas entre tutores, apenas entre monitores e entre tutores e monitores;
- Analisar as relações estabelecidas entre os participantes quanto à região em que se encontram;
- Detectar automaticamente comunidades na rede, a fim de compreender os atributos que foram determinantes nas estruturas de comunidades formadas;
- Identificar monitores com grande proeminência na rede social analisada;
- Identificar os monitores com baixa proeminência na rede social analisada.

5.3 Fontes de Dados

Na aplicação da Metodologia, foram coletados dados sobre os participantes de três fontes principais: (i) da plataforma *Moodle* utilizada durante o curso de formação; (ii) das planilhas fornecidas aos polos de formação pelo Ministério das Comunicações; e (iii) do Sistema de Avaliação sobre os monitores. Estes conjuntos de dados foram obtidos diretamente

da Universidade Federal do Pará, uma das instituições responsáveis pela execução do curso de formação de monitores, que concedeu a permissão necessária para que as análises fossem realizadas nas pesquisas realizadas pelos pesquisadores do LTS/UFPA.

O conjunto de dados utilizado nesta tese corresponde a um período de onze meses entre 2011–2012 e contém informações sobre as mensagens trocadas na plataforma de aprendizagem, atributos dos participantes e avaliação sobre os monitores, conforme descrito nas seções 5.3.1 a 5.3.3.

5.3.1 Mensagens instantâneas

A plataforma *Moodle* possui uma ferramenta de troca de mensagens instantâneas que permite que os usuários visualizem quem está *on-line* e enviem mensagens instantâneas ou que remetam mensagens para quem não está *on-line*, neste caso o envio é assíncrono e o usuário destinatário recebe a mensagem em um momento posterior. As mensagens trocadas nesta ferramenta representam a fonte utilizada para análise da rede social dos participantes do curso. Essa escolha deve-se ao fato de que essas mensagens eram trocadas espontaneamente no decorrer do curso — ou seja, sem nenhuma obrigatoriedade de uso em qualquer atividade do curso, o que corrobora com a possibilidade de investigar aspectos relacionados ao curso que signifiquem que esses indivíduos estão em um estágio de apropriação de recursos úteis de TICs, que podem possibilitar a eles mais autonomia e pró-atividade em suas ações.

Assim, foram extraídos registros de conexão de 4.382 participantes e 104.831 mensagens. Cada mensagem trocada entre os participantes na plataforma representa um *link* (ou arco) na rede social a ser analisada. Essa rede não considera os *loops*, ou seja, mensagens enviadas de um vértice para si mesmo; a troca de mensagens entre os vértices é tratada de forma direcional, ou seja, o conjunto de dados analisado contém informações sobre remetente e destinatário de cada mensagem no seguinte formato: $\langle id_envio, id_receptor \rangle$, onde *id_envio* é o código identificador do remetente da mensagem e *id_receptor* é o código identificador do destinatário da mensagem. O conteúdo das mensagens não foi analisado nas pesquisas realizadas pelo grupo de inclusão digital do LTS/UFPA.

5.3.2 Atributos dos participantes

Com o objetivo de analisar a influência do papel ou da localidade do participante nas redes sociais formadas ao longo da formação, foi necessário coletar os atributos *papel* e *local*

do participante. A fonte de dados foi a plataforma de aprendizagem *Moodle*, entretanto, uma vez que estas informações nem sempre estavam devidamente preenchidas na plataforma de aprendizagem, foram necessárias as planilhas fornecidas pelo Ministério das Comunicações aos polos de formação. Na tabela 5.1 são apresentadas a categoria, descrição e valores que o atributo pode assumir durante o curso de formação de monitores.

Tabela 5.1 Categoria, Descrição e Valores para o Atributo.

Categoria do Atributo	Descrição e Valores para o Atributo
Função do participante	Representado pelo atributo <i>papel</i> — identifica o papel do participante na plataforma de aprendizagem e pode assumir um dos seguintes valores: Tutor, Monitor ou não identificado.
Polo regional de formação	Representado pelo atributo <i>polo</i> — identifica o polo no qual os monitores e tutores participam da formação. O polo pode assumir um dos seguintes valores: Norte, Nordeste, Sul, Sudeste, Centro-Oeste, São Paulo e Ceará. Nesta tese, consideram-se que os membros da equipe do polo nacional pertencem ao polo <i>coordenação</i> , portanto, não corresponde a nenhuma região específica do país.

5.3.3 Avaliação sobre os Monitores

No sistema de avaliação, para cada monitor em acompanhamento, o tutor respondia um questionário mensal contendo um conjunto de questões objetivas sobre: (i) uso do Moodle; (ii) uso de software livre; (iii) uso de tecnologias disponíveis; (iv) acesso à Internet; (v) fase de desenvolvimento do projeto comunitário; (vi) área de articulação do projeto comunitário; e (vii) ausência/não do monitor durante o mês na formação. Adicionalmente, o tutor descrevia suas observações com respeito aos avanços e dificuldades do monitor; tema do projeto; dificuldades relacionadas à execução do projeto; e motivos de ausência do monitor, se fosse o caso.

Para o contexto desta tese, foram utilizadas as observações qualitativas do tutor quanto ao acesso à Internet e desempenho do monitor no curso.

5.4 Resultados da Aplicação da Metodologia

Nas seções 5.4.1 a 5.4.4 são apresentados os resultados da aplicação da metodologia no cenário de estudo adotado nesta tese, conforme atividades especificadas na metodologia.

5.4.1 Inicialização

Nesta fase, foram coletados os registros referentes às mensagens instantâneas, dados do usuário oriundos da plataforma de aprendizagem, bem como os registros referentes às

avaliações sobre os monitores no Sistema de Avaliação. Em sequência, realizou-se um pré-processamento desses dados, em busca de: (i) ausência de valores de registros; (ii) ruídos provenientes do preenchimento humano na plataforma de aprendizagem; (iii) descarte de dados desnecessários provenientes das tabelas importadas do banco de dados da plataforma. Assim, a integração dos dados foi realizada a partir das diferentes fontes, a fim de garantir que os campos necessários para a análise fossem todos preenchidos. Em vários registros foi necessário o preenchimento manual de identificação do papel e polo dos participantes.

5.4.2 Nível Macro

Nesta fase, os registros das mensagens instantâneas foram extraídos a partir da fonte de dados integrada e preparados no formato de arquivo de entrada para o *Pajek* (PAJEK, 2016) — software utilizado para processar e extrair indicadores dos atores e das redes sociais dos participantes. A rede social extraída das conexões dos participantes na plataforma de aprendizagem e analisada neste nível contém 4.382 vértices. Nesta etapa da análise, realizaram-se os seguintes passos:

- 1) Definição do valor um para o grau de centralidade mínimo, como critério para análise dos vértices neste nível e no nível seguinte. Os demais vértices devem ser analisados no nível micro. Com essa definição, o número de nós da rede social analisada neste nível foi reduzido para 2.303.
- 2) Em sequência, foram identificados três diferentes períodos de tempo para a análise: meses 1–2, 3–6 e 7–11. Esses diferentes períodos de tempo estão denominados nesta tese como momentos #1, #2 e #3, em que cada momento possui apenas os *links* estabelecidos no período de tempo considerado na análise. A escolha dos três momentos está associada às diferentes fases do curso: fase um, onde o objetivo maior era a familiaridade do participante com a plataforma de aprendizagem; fase dois, em que os participantes possuíam diferentes atividades relacionadas a diversos assuntos (e.g. informática básica, inclusão digital nas comunidades, comunicação e compartilhamento em redes, software livre, lixo eletrônico); e fase três, cujo objetivo era formular, executar e avaliar projetos envolvendo as comunidades do entorno dos telecentros.
- 3) Neste passo, foram extraídas as medidas de densidade e grau de centralidade para cada uma das redes associadas aos diferentes momentos analisados. A quantidade de *links*

existentes em cada uma dessas redes, bem como suas medidas de densidade e grau de centralidade estão ilustradas na tabela 5.2.

Tabela 5.2 Índices da Rede Telecentros.BR.

	#1	#2	#3
<i>Links</i>	19.025	37.681	48.125
Densidade	0,0036	0,0071	0,0091
Grau médio*	16,5	32,7	41,8

* Valor não normalizado

Em termos estruturais, a rede em #2 é 98% mais densa que a rede em #1 e a rede #3 é 153% mais densa que a rede no início da formação. Da mesma forma, o número de *links* e o grau médio dos atores também foi uma medida que evoluiu com o aumento do tempo de formação (Tabela 5.2). O aumento dessas medidas, de acordo com a evolução do tempo de formação era um aspecto desejável, dada a estrutura e objetivos de cada fase do curso.

5.4.3 Nível Meso

Nesta fase, definem-se *clusters* distintos por papel: monitores e tutores; e por polo: Norte, Sul, Sudeste, Centro-Oeste, Nordeste, São Paulo e Ceará. Na sequência, utiliza-se o algoritmo Louvain para encontrar comunidades na rede formada após os onze meses de formação. Nessas comunidades, são analisadas a distribuição dos participantes quanto ao papel e o polo — o qual representa a localidade do participante. Por fim, apresentam-se visualmente as estruturas de algumas comunidades formadas.

5.4.3.1 Definição de *clusters* por atributos: *papel e polo*

Com o objetivo de verificar se as conexões são influenciadas pelo papel do participante (tutor ou monitor) ou pela região geográfica a que pertencem, foram definidos *clusters* distintos, classificados por esses atributos. Definir *clusters* por *papel* permite uma análise apenas entre pares, uma vez que são excluídos os *links* de relações existentes entre participantes com diferentes posições hierárquicas. Para isso, foram geradas diferentes medidas para os diferentes momentos para cada *cluster* (tutor e monitor), bem como para a rede com todos os participantes. Essas medidas estão apresentadas na tabela 5.3.

Tabela 5.3 Índices das Redes de Tutores e/ou de Monitores.

	Tutores			Monitores			Tutores e Monitores		
	#1	#2	#3	#1	#2	#3	#1	#2	#3
Nós		210			2.085			2.303	
Links	2.616	1.077	2.093	3.917	6.020	8.504	19.025	37.681	48.125
Densidade	0,0596	0,0245	0,0477	0,0009	0,0014	0,002	0,0036	0,0071	0,0091
Grau médio *	24,9	10,3	19,9	3,8	5,8	8,2	16,5	32,7	41,8
Maior grau *	609	118	447	275	394	419	1275	1963	2735
Centralidade por proximidade (média)	0,1713	0,0661	0,0923	0,0109	0,0257	0,0496	0,0416	0,0787	0,1736
Centralidade por intermediação (média)	0,1183	0,0662	0,0782	0,0059	0,0160	0,0360	0,0167	0,0251	0,0425
Coefficiente de agrupamento	0,4724	0,3571	0,4724	0,0877	0,087	0,0599	0,3192	0,2618	0,2719

* Valores não normalizados

Em termos estruturais, a rede de tutores é mais densa do que a rede de monitores e do que a rede de todos os participantes, nos três diferentes momentos analisados. No entanto, a rede de monitores aumenta de densidade (122%) do momento #3 em relação ao #1, em contraste com a rede de tutores, que no mesmo período decresce 20%. Era desejável por gestores/coordenadores da formação que com o desenvolvimento de projetos por alguns monitores, bem como com o progresso do tempo, estes interagissem mais com seus pares. Esse crescimento/decrescimento nos valores da densidade das diferentes redes (tutores, monitores, tutores e monitores) pode ser melhor visualizado na figura 5.1.

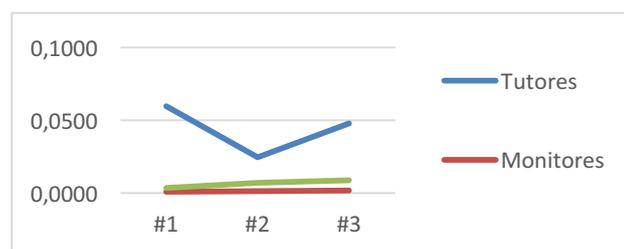


Figura 5.1 Valores de densidade das redes nos diferentes momentos analisados.

A centralidade por proximidade — que diz respeito à proximidade de um ator em relação a outros atores — é mais baixa na rede de monitores do que na rede de Tutores nos três diferentes momentos, entretanto esse índice aumenta 355% do momento #3 em relação ao #1, na rede de monitores. Isso é um resultado desejável, porque a formação também tinha como meta fazer com que os monitores ficassem mais próximos e interagissem para colaborar

com soluções para os problemas do dia-a-dia dos telecentros. Esse crescimento/decrescimento nos valores da centralidade por proximidade das diferentes redes (tutores, monitores, tutores e monitores) pode ser melhor visualizado na figura 5.2.

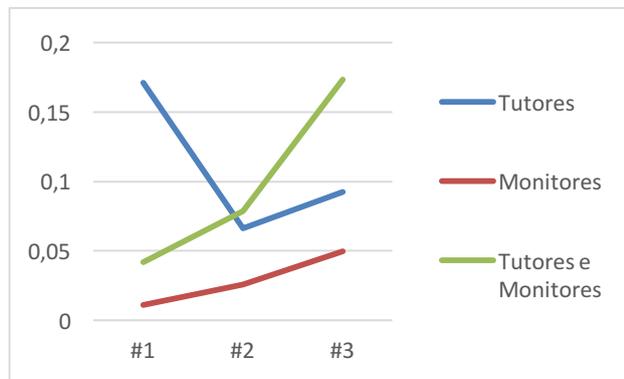


Figura 5.2 Valores de centralidade por proximidade das redes nos diferentes momentos analisados.

Na análise do coeficiente de agrupamento, a rede de tutores apresenta coeficiente mais alto que a rede de monitores nos três diferentes momentos observados (Tabela 5.3). Os valores mais altos na rede de tutores refletem uma maior densidade na vizinhança dos nós das redes dos tutores que nos nós das redes de monitores. A rede de tutores e a de monitores são visualizadas nas figuras 5.3 e 5.4 respectivamente. Nas figuras, as conexões correspondem aos onze meses da formação e os atores estão agrupados por polos — identificados por diferentes cores.

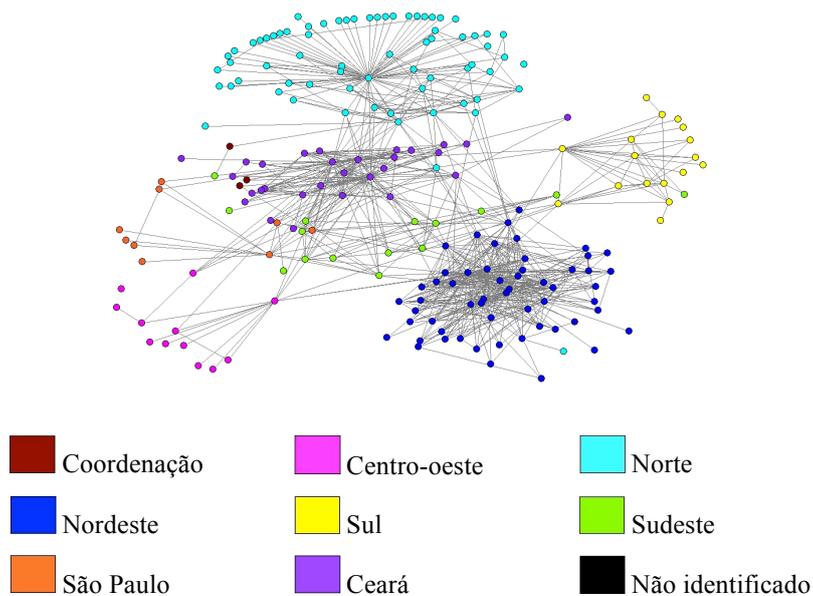


Figura 5.3 Rede de Tutores.

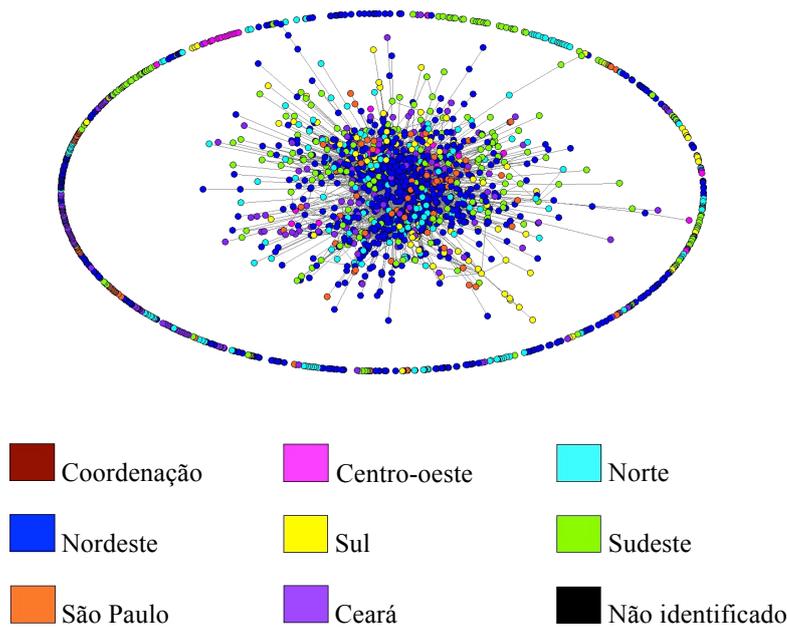


Figura 5.4 Rede de Monitores.

Na figura 5.3 é possível visualizar vértices do mesmo polo fortemente conectados, mas também existem conexões entre tutores de diferentes polos de formação. Hipóteses que podem ser levantadas para explicar a forte conexão entre os tutores são: a formação continuada pela qual passavam os tutores, o trabalho colaborativo para a construção adicional de conteúdos relacionados à formação, o compartilhamento das estratégias adotadas por seu polo e trocas de informações de cunho geral.

Os valores de centralidade por proximidade e centralidade por intermediação mais altos na rede de tutores nos três momentos analisados (Tabela 5.3) revelam a importância do papel dos tutores no curso de formação, para motivar as participações e as interações entre monitores, mas também pode indicar uma relação de hegemonia já esperada por parte dos tutores no curso, e nem tanto a autonomia esperada dos monitores. Quando visualizada a rede apenas de monitores, é possível observar que existem nós à margem do componente principal da rede, sem qualquer conexão (*links*) ou com pouca conexão com outros vértices (Figura 5.4), que indica que esses monitores possuíam apenas conexões com tutores, uma vez que era condição possuir o valor um como grau de centralidade mínimo.

A rede com todos os participantes pode ser visualizada na figura 5.5, onde é possível observar uma intensa troca de mensagens entre participantes de diferentes polos, o que sugere a possibilidade de encontrar a formação de *clusters* ou comunidades independentes de sua

região geográfica. Assim, segue-se para uma próxima etapa no nível meso, a fim de investigar as comunidades que podem ser automaticamente detectadas.

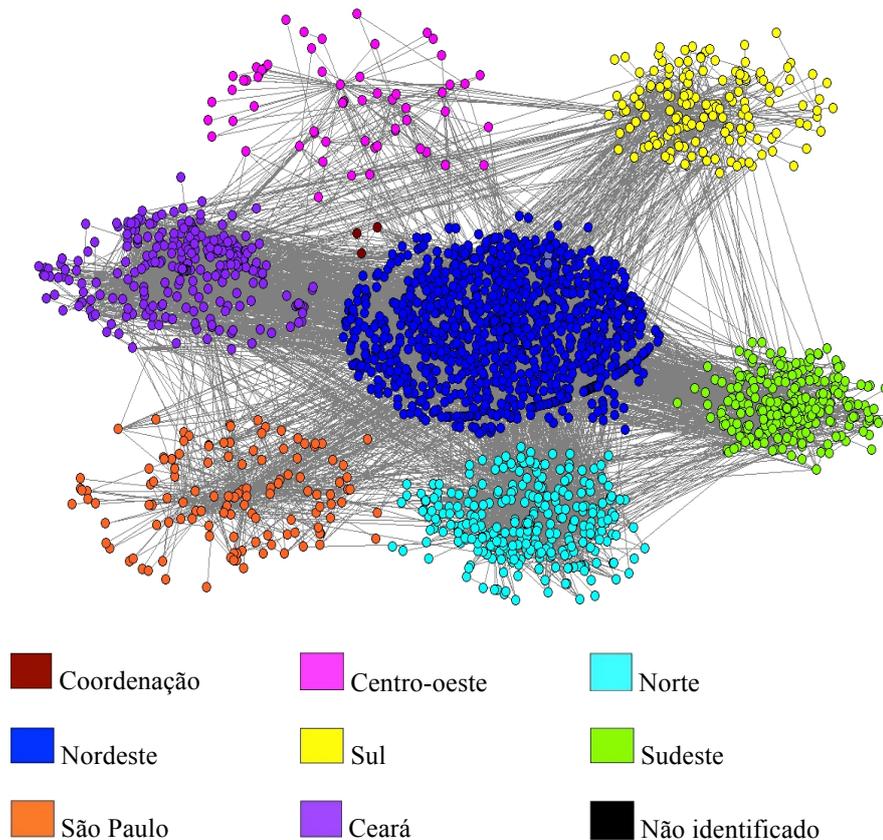


Figura 5.5 Rede de todos os participantes (tutores e monitores).

5.4.3.2 Detecção Automática de Comunidades e confronto entre *Clusters*

Nesta tese, utilizou-se o método/ algoritmo Louvain para detecção de comunidades (cf. descrito na seção 2.5.2.2), que utiliza uma função para maximização da função modularidade. O algoritmo utilizado nesta simulação tem como entrada a rede com todos os participantes (sem *loops* e sem múltiplas linhas) e um parâmetro de resolução que permite um controle sobre o tamanho e número de comunidades encontradas. Assim, quanto maior o parâmetro de resolução, maior é o número de comunidades/*clusters* encontrados; por outro lado, resoluções mais baixas produzem um menor número de comunidades. Com o parâmetro de resolução igual a um, o algoritmo executa o método de Louvain padrão (LOUVAIN, 2016). O algoritmo produz como saída: as comunidades encontradas e o valor alcançado para a função de modularidade.

Para seleccionar uma solução adequada para análise, o algoritmo foi executado com diferentes parâmetros de resolução — entre 0.19 até 40, encontrando entre 9 e 247 comunidades, com modularidade entre 0,868 e 0,049 respectivamente e índice de Rajska entre 0,434 e 0,858, quando realizado o confronto com os *clusters* definidos por polo. Cada solução foi comparada com os *clusters* definidos pelos polos dos participantes usando o índice de Rajska, em busca de reduzir o número de comunidades encontradas e maximizar o índice de Rajska. A associação entre os *clusters* definidos pelo atributo *polo* e os encontrados automaticamente pelo método Louvain é a mais alta (0,858) para a resolução igual a 40, entretanto, a modularidade é muito baixa (0,049) e o número de comunidades encontradas (247) é muito superior ao número de polos, dificultando a análise. Com o algoritmo Louvain padrão, ou seja, executado com parâmetro de resolução igual a um, foram encontrados 30 *clusters* (com modularidade igual a 0,8094) — valor que se mostrou adequado para a análise.

Quanto ao índice de Rajska — que mede a associação entre as diferentes classificações, a classificação de um *cluster* definido pelo atributo *polo* pode ser prevista pela classificação usando o método Louvain, pois o valor computado para esse índice foi de 0,7413. De fato, uma comunidade automaticamente detectada está frequentemente inserida em um *cluster* definido pelo atributo *polo*, conforme observado na tabela 5.4. Na tabela, apresentam-se os percentuais de nós dos *clusters* definidos pelo atributo *polo* que fazem parte dos *clusters* encontrado pelo método Louvain (ID-1 até ID-30). É possível observar as células com maior coincidência entre os *clusters* (em destaque), como por exemplo: no *cluster* ID-1 (Louvain) 96,8% pertencem ao polo Norte; no *cluster* ID-2 93,1% pertencem ao Ceará; no *cluster* ID-4 94,2% pertencem ao Sul; no *cluster* ID-5 98% pertencem ao Sudeste; no *cluster* ID-6 68,1% pertencem ao Centro-Oeste; no *cluster* ID-10 83,6% pertencem ao estado de São-Paulo. Foram também encontrados 21 *clusters* (*cluster* ID-3, ID-8, ID-9 e *cluster* ID-11 até ID-28) com predominância de membros do Nordeste, cujo polo possuía um maior número de participantes.

Tabela 5.4 Percentual de participantes (%) quanto ao polo nas comunidades detectadas pelo Louvain.

Identificação (ID) dos Clusters (Louvain)	ID dos clusters pré-definido por polo *								
	1	2	3	4	5	6	7	8	9
1	0,8	0,8	96,8	0,0	0,0	1,6	0,0	0,0	0,0
2	1,0	1,0	0,0	1,0	0,0	1,5	1,5	93,1	1,0
3	0,0	0,7	5,8	72,5	1,4	4,3	0,5	14,7	0,0
4	0,0	0,7	0,7	2,9	94,2	0,7	0,0	0,7	0,0

5	0,0	0,4	0,0	1,2	0,0	98,0	0,0	0,4	0,0
6	0,0	68,1	1,4	0,0	1,4	5,6	0,0	23,6	0,0
7	0,0	0,0	96,3	1,9	0,6	1,2	0,0	0,0	0,0
8	0,0	0,0	4,3	91,3	0,0	2,2	0,0	2,2	0,0
9	0,0	0,0	3,1	84,4	0,0	0,0	3,1	9,4	0,0
10	0,0	0,7	0,0	1,4	10,7	1,4	83,6	2,1	0,0
11	0,0	0,0	0,0	91,2	0,0	5,9	0,0	2,9	0,0
12	0,0	0,0	2,8	97,2	0,0	0,0	0,0	0,0	0,0
13	0,0	2,7	5,4	78,4	0,0	0,0	0,0	13,5	0,0
14	0,0	0,0	3,3	93,3	0,0	0,0	0,0	3,3	0,0
15	0,0	0,0	0,0	93,8	1,6	1,6	1,6	1,6	0,0
16	0,0	0,0	0,0	97,1	0,0	0,0	2,9	0,0	0,0
17	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0
18	0,0	0,0	0,0	94,1	0,0	5,9	0,0	0,0	0,0
19	0,0	0,0	0,0	96,9	0,0	3,1	0,0	0,0	0,0
20	0,0	3,2	0,0	96,8	0,0	0,0	0,0	0,0	0,0
21	0,0	0,0	2,4	95,2	0,0	2,4	0,0	0,0	0,0
22	0,0	0,0	1,7	86,4	5,1	6,8	0,0	0,0	0,0
23	0,0	0,0	0,0	93,5	3,2	3,2	0,0	0,0	0,0
24	0,0	1,8	0,0	98,2	0,0	0,0	0,0	0,0	0,0
25	0,0	0,0	0,0	96,6	0,0	3,4	0,0	0,0	0,0
26	0,0	0,0	0,0	88,6	0,0	5,7	0,0	5,7	0,0
27	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0
28	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0
29	0,0	0,0	0,0	50,0	0,0	50,0	0,0	0,0	0,0
30	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0

* (1) Coordenação; (2) Centro-oeste; (3) Norte; (4) Nordeste; (5) Sul; (6) Sudeste; (7) São Paulo; (8) Ceará; (9) não-identificado.

Nessas comunidades encontradas pelo método Louvain, também foi possível identificar que os Tutores estão presentes em 28 *clusters* (*cluster* ID-1 ao ID-28). Na tabela 5.5 é possível identificar a quantidade de tutores, monitores e participantes não-identificados em cada comunidade.

Tabela 5.5 Distribuição dos atores quanto aos papéis nas comunidades detectadas pelo Louvain.

ID do <i>cluster</i>	Tutores	Monitores	Não-identificados	Total
1	18	106	0	124
2	31	169	4	204
3	35	379	0	414
4	16	121	1	138
5	12	237	0	249
6	7	65	0	72
7	47	114	0	161

8	3	43	0	46
9	3	29	0	32
10	10	128	2	140
11	1	33	0	34
12	2	34	0	36
13	2	35	0	37
14	1	29	0	30
15	2	62	0	64
16	1	34	0	35
17	1	37	0	38
18	1	33	0	34
19	1	31	0	32
20	2	29	0	31
21	2	82	0	84
22	3	56	0	59
23	1	30	0	31
24	3	52	0	55
25	1	28	0	29
26	1	33	1	35
27	2	26	0	28
28	1	26	0	27
29	0	2	0	2
30	0	2	0	2
Total	210	2.085	8	2.303

5.4.3.3 Análise visual das estruturas de redes das comunidades

Para melhor visualizar a estrutura de rede formada nos *clusters* encontrados pelo método Louvain, foram usadas operações de redução de rede. Assim, os vértices com papel de monitor/tutor passam ser representados por um único vértice, favorecendo a análise da relação entre tutores e monitores. Como exemplo, na figura 5.6 mostra-se o *cluster* ID-3, o mais numeroso, com 414 nós (35 tutores e 379 monitores), com 72,5% dos seus membros pertencendo ao polo Nordeste (cf. tabela 5.4). Enquanto a figura 5.7 mostra o mesmo *cluster* ID-3, sendo que os vértices com papel de monitor passam a ser representado por um único vértice. Da mesma forma, a figura 5.8 mostra o *cluster* ID-1, com 124 nós (18 tutores e 106 monitores), com 96,8% dos seus membros pertencendo ao polo Norte e a figura 5.9 mostra os monitores representados por um único vértice e suas conexões com os demais tutores e as conexões entre os tutores. A cor do vértice determina o papel do ator nestas figuras. Nas

figuras 5.6 e 5.8 é possível observar grupos de monitores no entorno de alguns tutores, sem conexão com qualquer outro monitor, outrossim, as figuras 5.7 e 5.9 revelam alguns tutores cujas relações acontecem principalmente com seus pares, pois não possuem nenhum *link* com o vértice único, que representa os monitores.

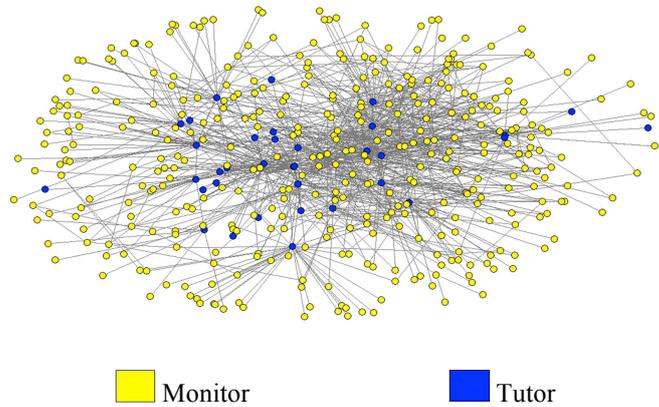


Figura 5.6 Rede do *Cluster ID-3* (Método Louvain).

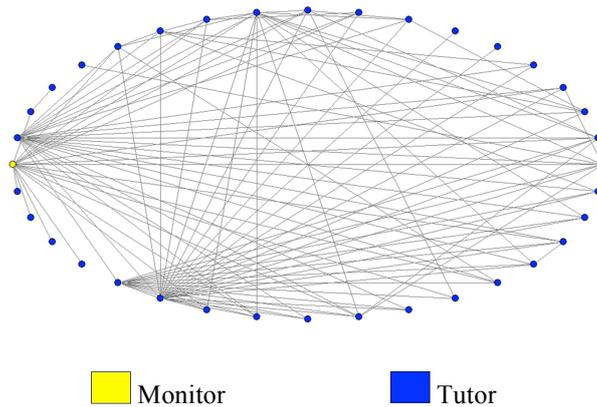


Figura 5.7 Rede do *Cluster ID-3* (Método Louvain) após operação de redução de vértices.

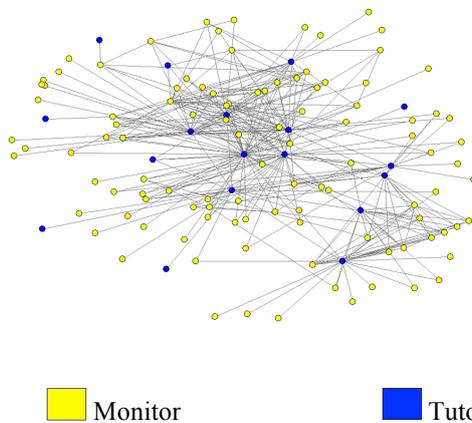


Figura 5.8 Rede do *Cluster ID-1* (Método Louvain).

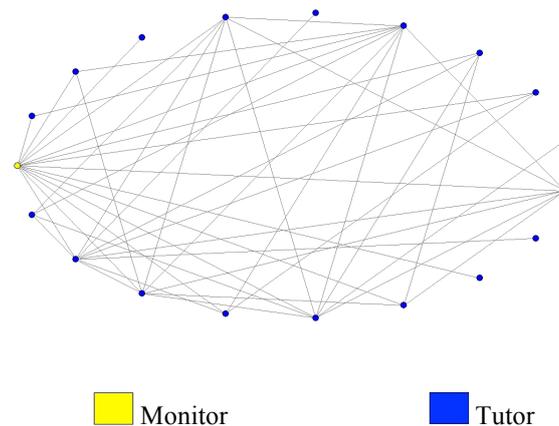


Figura 5.9 Rede do *Cluster* ID-1 (Método Louvain) após operação de redução de vértices.

5.4.4 Nível Micro

No nível micro, cada ator é analisado individualmente. Assim, com foco na identificação dos monitores com mais alto *grau de centralidade*, foi considerada, somente a faixa mais alta desse índice (25% dos registros), contendo 524 monitores. Desses, o maior número (387) e percentual de monitores na maior faixa (38,5) era proveniente do polo Nordeste — que também foi o polo que apresentou maior número de monitores na formação (1.004) durante os onze meses analisados, enquanto o polo sudeste foi o que apresentou o menor percentual de monitores na faixa mais alta (6,4%), conforme demonstrado na tabela 5.6.

Tabela 5.6 Distribuição dos Monitores com mais alto grau de centralidade por polos.

Polo	Monitores na formação	Monitores na formação (%)	Monitores na faixa de maior grau de centralidade	Monitores na faixa de maior grau de centralidade (%)
Centro-Oeste	49	1,3	7	14,3
Norte	242	6,7	35	14,5
Nordeste	1.004	73,9	387	38,5
Sul	141	3,1	16	11,3
Sudeste	281	3,4	18	6,4
São Paulo	115	4,0	21	18,3
Ceará	253	7,6	40	15,8

A fim de identificar os monitores na faixa mais alta de grau de centralidade e prestígio, aplicou-se o mesmo processo supracitado para o *grau de prestígio* — ou seja, identificou-se os atores na faixa mais elevada desse índice (25% dos registros) e realizou-se a combinação das listas de atores com maior grau de centralidade e de prestígio, removendo as entradas repetidas. A lista final resultante desse processo apresenta 572 monitores proeminentes, e

mostra-se na tabela 5.7 uma lista reduzida com os primeiros atores. Este procedimento pode ser útil para identificar monitores com potencial para atuar como tutores em novos programas de formação e que de fato, melhor se apropriaram de um recurso que não foi imposto a eles. Além disso, também serve como uma estratégia para explicitar, para gestores e investidores, a importância das interações/conexões em um programa de formação em larga-escala.

Tabela 5.7 Indicadores de Centralidade e Prestígio em ordem decrescente de grau de centralidade.

Id.	Polo	Grau de centralidade	Grau de prestígio	Id.	Polo	Grau de centralidade	Grau de prestígio
1	Nordeste	0.1681	0.2094	17	Sudeste	0.0854	0.0769
2	Nordeste	0.1620	0.1373	18	Nordeste	0.0773	0.1156
3	Nordeste	0.1596	0.0982	19	Nordeste	0.0771	0.0856
4	Nordeste	0.1410	0.1286	20	Sul	0.0760	0.0773
5	Nordeste	0.1371	0.1434	21	Nordeste	0.0749	0.1082
6	Nordeste	0.1188	0.1095	22	Nordeste	0.0749	0.0908
7	Nordeste	0.1156	0.1203	23	Nordeste	0.0730	0.0986
8	Nordeste	0.1066	0.0830	24	Nordeste	0.0708	0.0725
9	Nordeste	0.1049	0.0982	25	Nordeste	0.0702	0.0773
10	Norte	0.1027	0.1134	26	São Paulo	0.0699	0.0560
11	Nordeste	0.0990	0.1030	27	Nordeste	0.0693	0.0891
12	Nordeste	0.0971	0.1121	28	Nordeste	0.0686	0.0747
13	Nordeste	0.0954	0.0982	29	Nordeste	0.0684	0.0782
14	Nordeste	0.0945	0.1312	30	Nordeste	0.0643	0.1051
15	Nordeste	0.0919	0.0930	31	Nordeste	0.0643	0.0639
16	Nordeste	0.0897	0.0717	32	Sul	0.0632	0.0361

Outrossim, dentre os atores selecionados para análise neste nível, encontram-se 2.079 participantes sem qualquer conexão na ferramenta analisada ou que descontinuaram suas conexões entre os momentos observados — que são 328 monitores. Para elucidar as razões para isso, foram usadas as avaliações realizadas pelos tutores a respeito dos monitores, disponíveis no Sistema de Avaliação usado pelo programa de formação. Dessa forma, foram encontrados principalmente os seguintes cenários:

- Relatos sobre monitores, que apesar de registrados na plataforma de aprendizagem, não iniciaram a formação, porque o telecentro não recebeu os equipamentos necessários para o seu funcionamento ou porque apesar de ter recebido os

equipamentos, faltou mão de obra qualificada para realizar a instalação e deixá-lo operacional;

- Telecentros com conexões de Internet instáveis ou de baixa velocidade, como por exemplo, o relato sobre um monitor alocado no telecentro da Terra Indígena *Zoró*, no município de Rondolândia, divisa do estado de Mato Grosso e Rondônia. Neste caso, o tutor relata que para o monitor acessar a plataforma de aprendizagem *on-line*, este precisava se deslocar para um município vizinho e usar um *cibercafé*. Este cenário, com Internet de baixa qualidade, foi relatado principalmente por tutores do polo Norte (96), seguidos pelos do polo Nordeste (40), Ceará (34), Centro-Oeste (13), São Paulo (10), Sul (8) e Sudeste (5).
- Relatos sobre monitores com baixo desempenho (149), ocasionado pela falta de motivação para participar do curso de formação. Esses relatos foram efetuados principalmente por tutores do polo Norte (72). Como neste estudo não foi possível ter acesso a todos os monitores do polo Norte, não foram investigadas as razões para estas desmotivações.

Os distintos comportamentos dos monitores na rede social podem ser influenciados pelo contexto local, por diferentes estratégias usadas pelos polos de formação ou motivações pessoais. Conforme descrito na seção 4.3, esta fase faz parte da abordagem para monitoramento e acompanhamento de programas de formação em larga escala, sendo identificada pelo componente de mineração de dados.

5.5 Comentários Finais

Neste capítulo, foi realizada a aplicação da metodologia desenvolvida nesta tese no contexto de um curso realizado a distância para monitores localizados em diferentes regiões geográficas do Brasil. Nesse curso, foram identificados dois atores: os monitores de telecentros, principais beneficiários da intervenção e os tutores, que atuaram na mediação do curso. As mensagens trocadas na plataforma de aprendizagem do curso foram usadas para representar os *links* entre os atores das redes sociais formadas pelos participantes do curso.

Como a metodologia desenvolvida nesta tese foi criada com a perspectiva de análise multinível, em um primeiro nível, a rede de todos os participantes foi analisada em três momentos distintos, para a análise da evolução em termos de densidade, número de *links* e

grau médio dos participantes. Como conclusão, verificou-se que as medidas aumentaram, de acordo com o progresso do curso.

Em um segundo nível, a fim de investigar a influência dos atributos *papel* e *polo*, definiram-se *clusters* determinados por esses atributos. Investigou-se então as redes compostas apenas por monitores/tutores e a rede com todos os participantes nos três diferentes momentos escolhidos para a análise, utilizando os índices sugeridos no desenvolvimento da metodologia. Na visualização da rede com todos os participantes foi observada uma intensa troca de mensagens entre participantes de diferentes polos, o que sugeriu a possibilidade de encontrar a formação de *clusters*/comunidades independentes da região geográfica dos membros. Partiu-se então para a detecção automática de comunidades, utilizando o algoritmo Louvain, que demonstrou ser rápido e eficiente para as simulações de detecção de comunidades realizadas. A partir das comunidades detectadas, verificou-se forte associação entre elas e os *clusters* definidos pelo atributo *polo*, sendo que frequentemente a maior parte dos membros de uma comunidade detectada pelo algoritmo fazem parte quase em sua totalidade de um único polo, contrariando expectativas de que as conexões pudessem estar acontecendo independente do polo do monitor. Na análise da distribuição dos papéis nas comunidades detectadas, os resultados demonstraram que, em quase totalidade das comunidades sempre existia a presença de pelo menos um tutor, o que também contrariou expectativas de que as comunidades estivessem sendo formadas independentes da presença de um tutor, que no curso possuía uma diferente posição hierárquica em relação aos monitores.

Por fim, na análise do nível micro, foram identificados os atores mais proeminentes, bem como os seus polos; e para os monitores sem qualquer conexão ou que desistiram das conexões nos momentos analisados, recorreu-se ao Sistema de Avaliação para elucidar esses cenários. Os cenários mais frequentes foram: telecentros não-operacionais; conexões de Internet instáveis ou de baixa velocidade; e desmotivação do monitor.

Finalmente, por mais que a aplicação da metodologia tenha sido demonstrada especificamente sobre os dados de um curso de formação em inclusão digital, é importante compreender que existe uma carência por estratégias de acompanhamento de um grande número de participantes em regiões geográficas tão diversas como o Brasil. Assim, na metodologia desenvolvida buscou-se estabelecer mecanismos de visualizar aspectos dinâmicos do curso, que no caso são representadas pelas conexões em uma ferramenta da plataforma de aprendizagem, para tentar compreender as relações formadas.

6. CONCLUSÕES

Neste capítulo, apresentam-se algumas conclusões do trabalho de tese, as contribuições, dificuldades, trabalhos futuros e publicações geradas durante o processo de desenvolvimento.

6.1 Considerações

Para países em desenvolvimento como o Brasil, ainda persiste o interesse em intervenções de formação em larga escala para combater o fosso digital. O foco principal dessas intervenções tem sido contribuir para o desenvolvimento do país, reduzindo as desigualdades entre o Brasil e outros países, bem como para diminuir as desigualdades internas.

Acompanhar e avaliar intervenções de TICs e mesmo medir o fosso digital, representam atividades fundamentais, que podem ajudar gestores, pesquisadores e profissionais no processo de apoio à decisão. Entretanto, o acompanhamento dessas intervenções é uma questão crítica. De fato, há uma lacuna em abordagens que possam ser apropriadas nesses cenários de intervenção, principalmente quando a intervenção acontece um extenso território geográfico e atinge uma grande quantidade de participantes. Como foi apontado na revisão de trabalhos relacionados, diversas abordagens bem-sucedidas para avaliar o impacto da intervenção são baseadas em questionários e entrevistas. Entretanto, essas abordagens são apropriadas para um pequeno número de pessoas, mas torna-se inviável em cenários que envolvem um grande número de pessoas em regiões geograficamente distantes.

Assim, esta tese apresenta uma metodologia que permite acompanhar uma intervenção de formação — em larga escala, que tenha como objetivo prover habilidades/conhecimentos básicos de TICs para indivíduos. Os *links* são a base da abordagem utilizada nesta tese. Lidar com este conceito (*links*) mostrou ser mais apropriado do que as abordagens que somente consideram o acesso dos indivíduos à tecnologia, pois reflete de alguma forma a autonomia dos indivíduos. No desenvolvimento da metodologia, foram também usados a função/papel do participante na intervenção, bem como a localidade que esses indivíduos pertencem, como conceitos-chave investigados nas análises realizadas.

6.2 Contribuições

As contribuições deste trabalho podem ser organizadas em dois aspectos: (i) relacionados ao desenvolvimento da metodologia multinível; e (ii) referentes à construção de um trabalho científico desta natureza.

No que concerne ao primeiro aspecto, são consideradas as seguintes contribuições:

- Proposta de uma metodologia de acompanhamento/monitoramento de intervenções de formação em TICs;
- Proposta de uma metodologia multinível que permite diferentes níveis de agregação entre os participantes, possibilitando analisar: o nível da rede completa (nível macro), o nível de comunidades/*cluster* (nível meso), e o nível de um determinado participante (nível micro);
- Proposta de uma metodologia que considera como base os *links*, representando um certo grau de autonomia e de participação ativa dos indivíduos;
- Possibilitar diferentes análises das redes sociais:
 - Quanto ao tempo de formação;
 - Quanto às comunidades/*clusters* formados;
 - Quanto ao grau de proeminência de um ator na rede.
- Proposta de uma metodologia que utiliza técnicas de Mineração de Redes Sociais que permitem a identificação de comunidades, por meio de um algoritmo computacional;
- Permitir o confronto/comparação de *clusters* definidos por um atributo e *clusters* detectados pelo algoritmo computacional, a fim de verificar a predição entre as diferentes classificações;
- Aplicação da metodologia em intervenções reais de formação, caso os *links* representem o aspecto-chave a ser analisado;
- Permitir a aplicação de todos os níveis da metodologia ou apenas os níveis necessários em um determinado contexto;

- Aplicação em um programa de inclusão digital no Brasil, país de dimensões continentais e com carência de ferramentas/metodologias relacionadas à análise de programas e projetos públicos/privados em larga escala, especificamente no curso de formação em TICs, promovido pelo programa Telecentros.BR (BRASIL, 2016b);
- Produção de resultados que podem ser manuseados por técnicas de Inteligência Computacional, a fim de melhorar os processos de descoberta de conhecimento nas bases de dados analisadas;
- Produção de resultados que podem servir como referencial de análise para ser utilizado e/ou comparado com outras intervenções de natureza similar.

No que concerne à construção de um trabalho científico desta natureza, são consideradas as seguintes contribuições:

- Divulgação dos estudos realizados neste trabalho, junto às comunidades nacional e internacional, por meio da publicação de artigos em congressos e periódicos, o que confirma a importância e contribuições das investigações realizadas. Na seção 6.5, são apresentadas as publicações realizadas, por meio dos estudos e pesquisas provenientes do processo de doutorado.
- Elaboração deste documento de tese para disponibilização da metodologia, métodos, aplicação, bem como dos resultados obtidos.

6.3 Dificuldades Encontradas

Cabe ressaltar, que ao longo do desenvolvimento da tese, algumas dificuldades foram enfrentadas para a consecução deste trabalho. Dentre essas dificuldades, podem ser destacadas as que seguem:

- A contextualização dessa pesquisa foi um fator de entrave. No início do processo de doutorado (2011–2012), alguns artigos foram enviados na área de “computador e educação”, porém a maioria dos periódicos recusava, sem nem passar pelos revisores, por não considerarem o curso analisado na pesquisa como um curso formal;

- Existem poucos periódicos especializados na área de SNA e os existentes não fazem parte da lista de periódicos da área de Engenharia IV, conforme classificação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).
- Domínio das técnicas empregadas em SNA. O programa de pós-graduação em engenharia elétrica da UFPA não possui disciplinas específicas sobre a área, que poderiam ajudar a superar esse desafio;
- As bases de dados que compõem a pesquisa do grupo pertencem a fontes diversas;
- A base de dados proveniente da plataforma não apresentava dados sobre o papel e local de muitos participantes. Nesses casos, foi necessário o preenchimento manual, o que foi muito exaustivo;
- A análise dos registros de avaliação sobre os monitores também foi exaustivamente lida, a fim de identificar os principais problemas relatados pelos tutores sobre os monitores.

6.4 Trabalhos Futuros

Algumas novas aplicações/melhorias que podem ser realizadas, a partir da definição dessa metodologia, são:

- Estudar alternativas de empregar a metodologia em outras intervenções que não sejam de formação em TICs;
- Aplicar outros algoritmos computacionais para detectar comunidades, a fim de comparar com as comunidades detectadas pelo método Louvain;
- Investigar as atividades que fazem parte da metodologia, a fim de estabelecer uma arquitetura básica de componentes computacionais necessários para a execução da metodologia;
- Estabelecer novos cenários para a aplicação da metodologia (e.g. um curso básico de informática, com poucos participantes);

- Considerar outros aspectos a serem incorporados na metodologia (e.g. incluir *clusters* por gênero, aspectos socioeconômicos).

6.5 Publicações geradas

Os seguintes trabalhos diretamente relacionados à produção desta pesquisa ou com pesquisas afins foram submetidos e/ou publicados em periódicos, conferências ou como capítulo de livro, durante o processo de doutorado da autora desta tese:

- Periódicos publicados:
 - BRITO, S. R., SILVA, A. S., MARTINS, D. L., VIJAYKUMAR, N. L., ROCHA, C. A. J., COSTA, J. C. W. A., FRANCÊS, C. R. L. (2013) Employing online social networks to monitor and evaluate training of digital inclusion agents. **Social Network Analysis and Mining**, v. 3, n. 3, p. 497–519, September 2013.
 - SILVA, A. S.; BRITO, S. R.; MARTINS, D. L.; VIJAYKUMAR, N. L.; ROCHA, C. A. J.; COSTA, J. C. W. A.; FRANCÊS, C. R. L. (2014) Social Networks Analysis and Participation in Learning Environments to Digital Inclusion Based on Large-Scale Distance Education. **International Journal of Distance Education Technologies**, v. 12, p. 1–25, 2014. Engenharias IV (Classificação de Periódicos 2012 – *Qualis B3*).
 - SILVA, A. S.; BRITO, S. R.; VIJAYKUMAR, N. L.; ROCHA, C. A. J.; COSTA, J. C. W. A.; FRANCÊS, C. R. L. (2015) Social Network Analysis to Monitor Interactions in Virtual Learning Environment. **Revista IEEE América Latina**, v. 13, n. 10, p. 3482–3387, 2015. *Fator de Impacto* (JCR 2014): 0.326. Engenharias IV (Classificação Qualis 2014 – *Qualis B2*).
 - BRITO, S.R.; SILVA, A. S.; CRUZ, A. G.; BARROSO, R. F. F.; MONTEIRO, M. A.; COSTA, J. C. W. A.; FRANCÊS, C. R. L. (2016) Gravidez na adolescência e o acesso às Tecnologias de Informação e Comunicação na Amazônia. **Mundo Amazônico**, v. 6, n. 2., 2016. Saúde Coletiva (Classificação Qualis 2014 – *Qualis B4*).
 - BRITO, S. R.; SILVA, A. S.; CRUZ, A. G.; MONTEIRO, M. A.; VIJAYKUMAR, N. L.; SILVA, M. S.; COSTA, J. C. W. A.; FRANCÊS,

- C. R. L. (2016) Concentration of Access to Information and Communication Technologies in the Municipalities of the Brazilian Legal Amazon. **Plos One**, v. 11, p. e0152655, 2016. *Fator de Impacto* (JCR 2014): 3.234.
- SILVA, A. S.; BRITO, S. R.; VIJAYKUMAR, N. L. ROCHA, C. A. J; MONTEIRO, M. A.; COSTA, J. C. W. A; FRANCÊS, C. R. L. (2016) Social Network Analysis and Mining to Monitor and Identify Problems with Large-Scale Information and Communication Technology Interventions. **Plos One**, v. 11, n. 1, p. e0146220, 2016. *Fator de Impacto* (JCR 2014): 3.234.
 - Periódicos submetidos (em revisão):
 - BRITO, S. R.; SILVA, A. S.; VIJAYKUMAR, N. L; ROCHA, C. A. J; MONTEIRO, M. A.; COSTA, J. C. W. A.; FRANCÊS, C. R. L. (2014b). An approach to evaluate large-scale ICT training interventions. **Information Systems Frontiers**. (Submetido em 07 de novembro de 2014).
 - Conferências:
 - CORREA NETO, S. S.; BRITO, S. R.; SILVA, A. S.; ELIASQUEVICI, M. K.; TAVARES, O. L.; FAVERO, E. L. (2011) Aprendizagem de iniciantes em algoritmos e programação: foco nas competências de autoavaliação. In: Simpósio Brasileiro de Informática na Educação, 22º Simpósio Brasileiro de Informática na Educação e 17º WORKSHOP DE INFORMÁTICA NA ESCOLA, 2011, Aracaju, Sergipe, Brasil. p. 750-759.
 - BRITO, S. R.; SILVA, A. S.; TAVARES, O. L.; FAVERO, E. L.; FRANCES, C. R. L. (2011) Computer Supported Collaborative Learning for helping novice students acquire self-regulated problem-solving skills in computer programming. In: The 2011 International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS'11), 2011, Las Vegas.

- MARTINS, D. L.; BRITO, S. R.; SILVA, A. S. (2012) Painéis de controle para análise de redes sociais: uma forma de visualização integrada de múltiplas camadas de redes através de indicadores de centralidade e o uso de gráficos radares. In: Simpósio de Pesquisa Científico-Tecnológica (SPTC), 2012, São Paulo, Brasil.
- SILVA, A. S.; BRITO, S. R.; VIJAYKUMAR, N. L.; MARTINS, D. L.; ROCHA, C. A. J.; COSTA, J. C. W. A.; FRANCÊS, C. R. L. (2013) Análise de redes sociais para avaliação e monitoramento de programas de treinamento em larga escala baseados no uso de ambientes de aprendizagem e redes sociais online. In Proceedings of the 2nd Brazilian Workshop on Social Network Analysis and Mining (BRASNAM), 2013, Maceió, Alagoas, Brasil. Anais da Sociedade Brasileira de Computação. p. 1648–1659.
- BRITO, S. R.; SILVA, A. S.; COSTA, K. C. F.; MATA, E. C. (2013) Promoção, realização e gestão das atividades complementares e de extensão no curso de Licenciatura em Computação. In: Seminário TIC-PARFOR: Uso de Tecnologia de Informação e Comunicação na Formação de Professores no Estado do Pará (TIC-Parfor), 2013, Belém, Pará, Brasil. v. I.
- Capítulo de livro:
 - BRITO, S. R.; SILVA, A. S.; MARTINS, D. L.; ROCHA, C. A. J.; WEYL, J. C. W. A.; FRANCÊS, C. R. L. (2014a) Brazilian Government's Training Network for Digital Inclusion: Analysis of strategies for improving interactivity. **Handbook of Research on Enterprise 2.0: Technological, Social, and Organizational Dimension**. In: Cruz-Cunha, M. M.; Moreira, F.; Varajão, J. (Org.). Handbook of Research on Enterprise 2.0: Technological, Social, and Organizational Dimension. IGI Global, 2014, p. 618-643.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACTIC. Acreditación de competencias en tecnologías de la información y la comunicación. Disponível em: <http://www20.gencat.cat/portal/site/actic?newLang=es_ES>. Acesso em: 25 fev. 2016.
- ADAM, L.; WOOD, F. An investigation of the impact of information and communication technologies in sub-Saharan Africa. **Journal of Information Science**, v. 25, n. 4, p. 307–318, 1999.
- AGARWAL, V.; BHARADWAJ, K. K. A collaborative filtering framework for friends recommendation in social networks based on interaction intensity and adaptive user similarity. **Social Network Analysis and Mining**, v. 3, n. 3, p. 359–379, 2013.
- AIELLO, L. M. et al. Link creation and information spreading over social and communication ties in an interest-based online social network. **EPJ Data Science**, v. 1, n. 1, p. 1–31, 2012.
- ANDRADE, A. D.; URQUHART, C. Unveiling the modernity bias: a critical examination of the politics of ICT4D. **Information Technology for Development**, v. 18, n. 4, p. 281–292, 2012.
- ASHRAF, M.; SWATMAN, P.; HANISCH, J. An Extended Framework to Investigate ICT Impact on Development at the Micro (Community) Level. In Proceedings of the 16th European Conference on Information Systems (ECIS), 2008, Galway, Ireland.
- ASHRAF, M.; HANISCH, J.; SWATMAN, P. ICT intervention in the Chandanbari Village of Bangladesh: Results from a field study. **Information Systems Frontiers**, v. 11, n. 2, p. 155–166, 2009.
- AVGEROU, C. Information Systems in Developing Countries: A critical research review. **Journal of Information Technology**, v. 23, n. 3, p. 133–146, 2008.
- BABIC, J. et al. Prototype-driven software development process for augmentative and alternative communication application. In Proceedings of the 13th International Conference on Telecommunications (ConTEL), 2015, Graz. IEEE. p. 1–8.
- BACKSTROM, L. Four degrees of separation. In Proceedings of the 4th Annual ACM Web Science Conference (WebSci '12), 2012, New York. ACM. p. 33–42.
- BARABASI, A-L. **Linked: How Everything is Connected to Everything else and what it means for Business, Science, and Everyday Life**. Cambridge: Plume, 2003.
- BARBIER, G.; LIU, H. Data Mining in Social Media. In: AGGARWAL, C. C. (Ed.). **Social Network Data Analytics**. US: Springer, 2011. p. 327–352.
- BASU, S. E-government and developing countries: an overview. **International Review of Law, Computers and Technology**, v. 18, n. 1, p. 109–132, 2004.
- BATAGELJ, V.; MRVAR, A. Analysis of kinship relations with Pajek. **Social Science**

Computer Review. p. 224–246, 2008.

BATINI, C.; FURLANI, L.; NARDELLI, E. What is a Good Diagram? A Pragmatic Approach. In *Proceedings of the Fourth International Conference on Entity - Relationship Approach*, 1985, Washington, DC, USA. IEEE Computer Society. p. 312–319.

BEREND, A. *Technology Appropriation Revisited: Mediation Theory as a New Philosophy of Technology for Information Systems Research*. Thesis – University of Twente, 2013.

BERGER, A.; CROLL, J. Training in basic Internet skills for special target groups in non-formal educational settings – conclusions from three pilot projects. **Research in Learning Technology**, North America, 2012.

BLONDEL, V. D. et al. Fast unfolding of communities in large networks. **Journal of Statistical Mechanics: Theory and Experiment**, v. P10008, p. 1–12, 2008.

BOCCALETTI, S. et al. Complex Networks: Structure and Dynamics. **Physics Reports**, v. 424, n. 4–5, p. 175–308, 2006.

BRASIL. Decreto n. 6991, de 27 de outubro de 2009. Institui o Programa Nacional de Apoio à Inclusão Digital nas Comunidades Telecentros.BR. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2009, n. 206, p. 3, 28 out. 2009. Seção 1.

BRASIL. Aviso de Chamamento Público nº 1/2010, de 10 de março de 2010. Abertura do processo de seleção de entidades responsáveis pela condução do Polo Nacional da Rede Nacional de Formação para Inclusão Digital nas Comunidades - Telecentros.BR. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2010a, n. 46, p. 121, 10 mar. 2010. Seção 3.

BRASIL. Aviso de Chamamento Público nº 2/2010, de 10 de março de 2010. Abertura do processo de seleção de entidades responsáveis pela condução dos Polos Regionais da Rede Nacional de Formação para Inclusão Digital nas Comunidades - Telecentros.BR. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2010b, n. 46, p. 125, 10 mar. 2010. Seção 3.

BRASIL. Ministérios das Comunicações, Secretaria de Inclusão Digital. 2016a. Disponível em: <<http://www.mc.gov.br/estrutura-institucional/142-institucional/secretaria-de-inclusao-digital/32325-secretaria-de-inclusao-digital-sid>>. Acesso em: 25 fev. 2016.

BRASIL. Ministérios das Comunicações, Telecentros. 2016b. Disponível em: <<http://www.mc.gov.br/institucional/405-temas/telecentros/>>. Acesso em: 25 fev. 2016.

BRASIL. Ministério da Ciência e Tecnologia do Governo Brasileiro, Secretaria de Ciência e Tecnologia para Inclusão Social. 2016c. Disponível em: <<http://www.governoeletronico.gov.br/acoes-e-projetos/inclusao-digital/casa-brasil>>. Acesso em: 01 mar. 2016.

BRITO, S. R. et al. Employing online social networks to monitor and evaluate training of digital inclusion agents. **Social Network Analysis and Mining**, v. 3, n. 3, p. 497–519, 2013.

BRITO, S. R. et al. Brazilian Government's Training Network for Digital Inclusion: Analysis of strategies for improving interactivity. In: Cruz-Cunha, M. M.; Moreira, F.; Varajão, J. (Eds.). **Handbook of Research on Enterprise 2.0: Technological, Social, and Organizational Dimension**. IGI Global, 2014a. p. 618–643.

BRITO, S. R. et al. (submetido). An approach to evaluate large-scale ICT training interventions. **Information Systems Frontiers**, 2014b.

BUZATO, M. E. K. Inclusão digital como invenção do cotidiano: um estudo de caso. **Revista Brasileira de Educação**, v. 13, p. 325–342, 2008.

CALVANI, A. et al. Are young generations in secondary school digitally competent? A study on Italian teenagers. **Computers & Education**, v. 58, n. 2, p. 797–807, 2012.

CERTEAU, M. **A invenção do cotidiano: artes de fazer**. Tradução: Ephraim Ferreira Alves. 8. ed. Petrópolis: Vozes, 1994.

CHANG, S.-I. et al. Study of the digital divide evaluation model for government agencies—a Taiwanese local government's perspective. **Journal Information Systems Frontiers**, v. 14, n. 3, p. 693–709, 2012.

CHINA INTERNET NETWORK INFORMATION CENTER (CNNIC). The 35th Statistical report on internet development in China. Disponível em: <<http://www1.cnnic.cn/IDR/ReportDownloads/201507/P020150720486421654597.pdf>>. Acesso em: 06 mar. 2016.

CHUA, F. C. T.; LIM, E.-P. Modeling Bipartite Graphs Using Hierarchical Structures. In Proceedings of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2011, Kaohsiung City, Taiwan. IEEE. p. 94–101.

CLAUSET, A.; NEWMAN, M. E. J.; MOORE, C. Finding community structure in very large networks. **Physical Review E**, v. 70, n. 6, 2004.

DINIZ, E.; BIROCHI, R.; POZZEBON, M. Triggers and barriers to financial inclusion: The use of ICT-based branchless banking in an Amazon county. **Electronic Commerce Research and Applications**, v. 11, n. 5, p. 484-494, 2012.

DUCH, J.; ARENAS, A. Community detection in complex networks using extremal optimization. **Physical Review E**, v. 74, n. 3, 2005.

FAUST, K.; WASSERMAN, S. Centrality and prestige: A review and synthesis. **Journal of Quantitative Anthropology**, v. 4, n. 1, p. 23-78, 1992.

FEIJÓO, C.; GÓMEZ-BARROSO, J. L.; RAMOS, S. Techno-economic implications of the mass-market uptake of mobile data services: Requirements for next generation mobile networks. **Telematics and Informatics**, v. 33, n. 2, p. 600-612, 2016.

FERRARI, A. **Digital Competence in Practice: An Analysis of Frameworks**. European Commission. Luxembourg: Publications Office of the European Union, 2012. Technical Report.

FORTUNATO, S. Community detection in graphs. **Physics Reports**, v. 486, n. 3–5, p. 75–174, 2010.

FORTUNATO, S.; CASTELLANO, C. Community structure in graphs. In: Meyers, R. A. (Ed.). **Encyclopedia of Complexity and Systems Science**. New York: Springer, 2009. p. 1141–1163.

FREEMAN, L. C. **The Development of Social Network Analysis: A Study in the Sociology of Science**. Vancouver: Empirical Press, 2004. 208 p.

GEPHI. Gephi – The Open Graph Viz Platform. Disponível em: <<https://gephi.org/>>. Acesso em: 13 jan. 2016.

GIRVAN, M.; NEWMAN, M. E. J. Community structure in social and biological networks. **Proceedings of the National Academy of Sciences of the USA (PNAS)**. v. 99, n. 12, p. 7821–7826, 2002.

HAGE, E. et al. Implementation factors and their effect on e-Health service adoption in rural communities: a systematic literature review. **BMC Health Services Research**, v. 13, n. 1, p. 1–16, 2013.

HEEKS, R. Foundation of ICTs in development: The information chain. **eDevelopment Briefing**, v. 3, n. 1, p. 1–2, 2005.

HEEKS, R. ICT4D 2.0: The next phase of applying ICT for international development. **Computer**, v. 41, n. 6, p. 26–33, 2008.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). PNAD, Pesquisa Nacional por Amostra de Domicílios. 2016. Disponível em: <<http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?z=pnad&o=3&i=P&c=2387>>. Acesso em: 01 Fev. 2016.

INTERNATIONAL TELECOMMUNICATION UNION (ITU). **Measuring the information society report 2014**. Geneva, Switzerland: International Telecommunication Union, 2014.

ISKILLS. ETS – **iSkills**. Disponível em: <<http://www.ets.org/iskills/>>. Acesso em: 25 de fevereiro de 2016.

JORGENSON, D. W.; VU, K. M. The ICT revolution, world economic growth, and policy issues. **Telecommunications Policy**, Published online 5 feb 2016.

KAHEN, G. Assessment of Information Technology for developing countries: appropriateness, local, constraints, IT characteristics and Impacts. *International Journal of Computer and Applications Technology*. v. 8, n. 5/6, 1995.

KAZIENKO, P. et al. Multidimensional Social Network: Model and Analysis. . In: Jedrzejowicz, P.; Nguyen, Ngoc-Thanh; Hoang, K. (Eds.). **Computational Collective Intelligence: Technologies and Applications**. Springer, 2011. p. 378–387.

KENNEDY, C. Communicating disparity: how social design can create public engagement

with issues of inequality. **Journal of Development and Communication Studies**, v. 3, n. 1, p. 24–34, 2014.

KHAN, G. F. et al. A socio-technical perspective on e-government issues in developing countries: a scientometrics approach. **Scientometrics**, v. 87, n. 2, p. 267–86, 2011.

KOTTAK, C. P. **Cultural Anthropology**. New York: McGraw-Hill, 2004.

LANG, J.; WU, S. F. Social network user lifetime. **Social Network Analysis and Mining**, v. 3, n. 3, p. 285–297, 2012.

LEON-SUEMATSU, Y. I.; YUTA, K. Framework for Fast Identification of Community Structures in Large-Scale Social Networks. In: MEMON, N.; XU, J. J.; HICKS, D. L.; CHEN, H (Eds.). **Data Mining for Social Network Data**. US: Springer, 2010. p. 149–175.

LESKOVEC, J. et al. Community Structure in Large Networks: Natural Cluster Sizes and the Absence of Large Well-Defined Clusters. **Internet Mathematics**, v.6, n.1, p. 29-123, 2009.

LESKOVEC, J.; LANG, K. J.; MAHONEY, M. Empirical comparison of algorithms for network community detection. 19th international conference on World wide web (WWW '10), 2010. ACM. p. 631–640.

LEWIS, K. Tastes, ties, and time: A new social network dataset using Facebook.com. **Social Networks**, v. 30, n. 4, p. 330–342, 2008.

LI, J.; SONG, Y. Community detection in complex networks using extended compact genetic algorithm. **Soft Comput.**, v. 17, n. 6, p. 925–937, 2013.

LOUVAIN. Detecting Communities with Louvain Method and VOS Clustering. Disponível em: < <http://mrvar.fdv.uni-lj.si/pajek/community/CommunityDrawExample.htm>>. Acesso em: 06 jan. 2016.

LUCAS, H. Information and communications technology for future health systems in developing countries. **Sociel science & medicine**, v. 66, n. 10, p. 2122–32, 2008.

LUNA-REYES L. F.; GIL-GARCIA J. R.; ROMERO G. Towards a multidimensional model for evaluating electronic government: proposing a more comprehensive and integrative perspective. **Government Information Quarterly**, v. 29, n. 3, p. 324–34, 2012.

MILGRAN, S. The small world problem. **Psychology Today**, v. 1, n. 1, p. 60–67, 1967.

MOODLE. Moodle Community. Disponível em: <<https://www.moodle.org/>>. Acesso em: 01 mar. 2016.

MORENO, J. L. Who Shall Survive: A New Approach to the Problem of Human Interrelations. Washington: Nervous and Mental Disease Publishing Company, 1934.

MORI, C. K. **Políticas públicas para inclusão digital no Brasil: aspectos institucionais e efetividade em iniciativas federais de disseminação de telecentros no período 2000–2010**. Tese – Universidade de Brasília, DF, Brasil, 2011.

MUSIAL, K.; BUDKA, M.; JUSZCZYSZYN, K. Creation and growth of online social network. **Journal World Wide Web**, v. 16, n. 4, p. 421–447, 2013.

MUTULA, S. M.; VAN BRAKEL, P. An evaluation of e-readiness assessment tools with respect to information access: towards an integrated information rich tool. **International Journal of Information Management**, v. 26, p. 212–223, 2006.

NEWMAN, M. E. J. The structure and function of complex networks. **SIAM review**, v. 45, n. 2, p. 167–256, 2003.

NEWMAN, M. E. J. Finding community structure in networks using the eigenvectors of matrices. **Physical Review E**, v. 74, n. 3, 2006.

NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E**, v. 69, n. 2, 2004.

NEWMAN, T. **A review of digital literacy in 3 – 16 year olds: evidence, developmental models, and recommendations**. London: Becta, 2008.

NOOY, W.; MRVAR, A.; BATAGELJ, V. **Exploratory Social Network Analysis with Pajek**. 1 ed. Cambridge University Press, 2005. 362 p.

ORA. ORA Project. Disponível em: <<http://www.casos.cs.cmu.edu/>>. Acesso em: 13 jan. 2016.

PAJEK. Pajek: analysis and visualization of large networks. Disponível em: <<http://pajek.imfm.si/doku.php>>. Acesso em: 13 jan. 2016.

PONS, P.; LATAPY, M. Computing Communities in Large Networks Using Random Walks. **Journal of Graph Algorithms and Applications**, v.10, n.2, p. 191–218, 2006.

PORTER, C. E. A Typology of Virtual Communities: A Multi-Disciplinary Foundation for Future Research. **Journal of Computer-Mediated Communication**, v. 10, n. 1, 2004.

PSCHEIDT, M. **Appropriate information system development: A methodology for sustainable cross-cultural information system production and use**. PhD thesis – Institute for Computing and Information Sciences (iCIS), Radboud Universiteit Nijmegen, Netherlands, 2012.

READ, D. Agent-based and multi-agent simulations: coming of age or in search of an identity?. **Journal Computational and Mathematical Organization Theory**, v. 16, n. 4, p. 329–347, 2010.

SALINAS, A.; SÁNCHEZ, J. Digital inclusion in Chile: Internet in rural schools. **International Journal of Educational Development**, v. 29, n. 6, p. 573–582, 2009.

SASSI, S.; GOAIED, M. Financial development, ICT diffusion and economic growth: lessons from MENA region. **Telecommunications Policy**, v. 37, n. 4–5, p. 252–261, 2013.

SCOTT, J. Social network analysis: developments, advances, and prospects. **Social Network**

Analysis and Mining, v. 1, n.1, p. 21–26, 2011.

SHANG, X.; YUAN, Y. Social Network Analysis in Multiple Social Networks Data for Criminal Group Discovery. In Proceedings of the International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CYBERC), 2012, Sanya, China.

SILVA, A. S. et al. Análise de redes sociais para avaliação e monitoramento de programas de treinamento em larga escala baseados no uso de ambientes de aprendizagem e redes sociais online. In Proceedings of the 2nd Brazilian Workshop on Social Network Analysis and Mining (BRASNAM), 2013, Maceió, Alagoas, Brasil. Anais da Sociedade Brasileira de Computação. p. 1648–1659.

SILVA, A. S. et al. Social network analysis to extract indicators from interaction networks and their correlation with indicators of participation in learning environments. **International Journal of Distance Education Technologies**, v. 12, p. 1–25, 2014.

SILVA, A. S. et al. Social Network Analysis to Monitor Interactions in Virtual Learning Environment. **Revista IEEE América Latina**, v. 13, n. 10, p. 3482–3387, 2015.

SILVA, A. S. et al. Social Network Analysis and Mining to Monitor and Identify Problems with Large-Scale Information and Communication Technology Interventions. **Plos One**, v. 11, n.1, p. e0146220, 2016.

SPRITZER, A. S.; FREITAS, C. M. D. S. A physics-based approach for interactive manipulation of graph visualizations. Proceedings of the working conference on Advanced visual interfaces (AVI '08), 2008, New York. ACM. p. 271–278.

SRINIVASAN, R. Re-thinking the cultural codes of new media: The question concerning ontology. **New Media & Society**, v. 15, n. 2, p. 203–223, 2013.

SRIVASTAVA, J. Data mining for social network analysis. In Proceedings of the IEEE International Conference on Intelligence and Security Informatics (ISI 2008), 2008, Taipei, Taiwan. IEEE.

SUGIYAMA, K.; TAGAWA, S.; TODA, M. Methods for Visual Understanding of Hierarchical System Structures. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 11, n. 2, p. 109–125, 1981.

TELES, A.; JOIA, L. A. Assessment of digital inclusion via the actor-network theory: The case of the Brazilian municipality of Piraiá. **Telematics and Informatics**, v. 28, n. 3, p. 191–203, 2011.

TIMMUSLIMITED. Timmus – Social and Educational Research. Disponível em: <<http://www.timmuslimited.co.uk>>. Acesso em: 08 mar. 2016.

UCINET. UCINET Software. Disponível em: <<http://www.analytictech.com/ucinet/ucinet.htm>>. Acesso em: 13 jan. 2016.

UNITED NATIONS. **Human development report 2001: making new technologies work for human development**. New York: United Nations Development Programme, 2001.

UNITED NATIONS. **The promotion, protection and enjoyment of human rights on the Internet, DocumentA/HRC/20/L.13**. New York: United Nations General Assembly, 2012.

UNITED NATIONS. **The Millenium Development Goals Report**. New York, 2015.

VENKATESH, V.; SYKES, T. A. Digital Divide Initiative Success in Developing Countries: A Longitudinal Field Study in a Village in India. **Information Systems Research**, v. 24, n. 2, p. 239–260, 2013.

WALKER, J. Q. A node-positioning algorithm for general trees. **Journal of Software: Practice and Experience**, v. 20, n. 7, p. 685–705, 1990.

WASSERMAN, S.; FAUST, K. **Social Network Analysis: Methods and Applications**. In M. Granovetter (Ed.). USA: Cambridge University Press, 1994. 825 p.

WATTS, D.; STROGATZ, S. Collective dynamics of ‘small-world’ networks. **Nature**. n. 393, p.440–442, 1998.

WILLIAMS, K. H. **Social networks, social capital, and the use of information and communications technology in socially excluded communities: a study of community groups in Manchester, England**. PhD thesis – Universidade de Michigan, 2005.

WOUTERS, J. et al. Improving delivery in development: the role of voice, social contract, and accountability. In: *The World Bank Legal Review*. 6th ed. Washington DC: World Bank, 2015.

WU, F.; HUBERMAN, B. A. Finding communities in linear time: a physics approach. **Eur. Phys. J. B** **38**, 331– 338, 2004.

YAGER, R. R. Intelligent Social Network Modeling and Analysis for Security Informatics. In *Proceedings of the 1st European Conference on Intelligence and Security Informatics (EUROISI '08)*, 2008, Berlin, Heidelberg. Springer-Verlag.

ZHOU, Y.; SINGH, N.; KAUSHIK, P. D. The digital divide in rural South Asia: Survey evidence from Bangladesh, Nepal and Sri Lanka. **IIMB Management Review**, v. 23, n. 1, p. 15-29, 2011.