Representação, extração e avaliação de interações
entre usuários de redes sociais online

Alan Keller Gomes

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito: 18/01/2013

Assinatura: _____

Representação, extração e avaliação de interações entre usuários de redes sociais online

Alan Keller Gomes

***Orientadora:* Profa. Dra. Maria da Graça Campos Pimentel**

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências - Ciências de Computação e Matemática Computacional. *EXEMPLAR DE DEFESA.*

**USP – São Carlos
Janeiro de 2013**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

G 633r Gomes, Alan Keller
 Representação, extração e avaliação de interações
entre usuários de redes sociais online / Alan Keller
Gomes; orientadora Maria da Graça Campos Pimentel. --
São Carlos, 2013.
150 p.

 Tese (Doutorado - Programa de Pós-Graduação em
Ciências de Computação e Matemática Computacional) --
Instituto de Ciências Matemáticas e de Computação,
Universidade de São Paulo, 2013.

 1. Sistemas Web e Multimídia. 2. Redes Sociais
Online. 3. Interação Social. 4. Extração Automática de
Interações. 5. Regras if-then. I. Pimentel, Maria da
Graça Campos , orient. II. Título.

Agradecimentos

Agradeço a Deus por cada benção, cada oportunidade, cada porta que se abriu e tantas outras que serão abertas com a conquista do doutorado em Ciência de Computação. A realização do sonho de criança de me tornar um cientista começa a se concretizar aqui.

Agradeço a minha orientadora, Profa. Maria da Graça Campos Pimentel, que me guiou pelo ‘caminho das pedras’. Meu mais profundo e sincero agradecimento. A dimensão da transformação que a oportunidade de trabalharmos juntos causou em minha vida é incomensurável. Guardarei para sempre esse aprendizado.

Agradeço ao meu pai que se fez presente. A minha mãe que se manteve forte. Ao meu irmão Victor Magnus e a minha tia Lourdes por terem cuidado de minha mãe. Ao meu irmão Rubens e meus sobrinhos Yago e Yan.

Meu agradecimento ao programa de Pós-Graduação do ICMC-USP, à agência de fomento CAPES, ao programa RH-TVD, ao IFG e ao programa PIQS. Agradeço ao Prof. Paulo César - Reitor do IFG, e também ao Prof. Cleiton José - Diretor do Campus de Inhumas. Agradeço o apoio, o incentivo e a oportunidade.

Agradeço aos professores do ICMC, Profa. Renata Pontin, Prof. Rudinei, Profa. Sandra Aluísio, Profa. Agma e Prof. Masiero. Agradeço também a Profa. Maria Carolina Monard, que me orientou no mestrado, e ao Prof. Paulo André Sant’Anna Peres, que me orientou na iniciação científica.

Agradeço aos colegas do laboratório Intermídia, Flávia, Renan, Fagá, Tim, Lilian, Vivian, Didier, Diogo Pedrosa, Diogo Martins, Claudia, Raoni, Raiza, Olíbio, Bruna.

Meu agradecimento ao Sandro Camargo, companheiro e companhia em momentos importantes e únicos, principalmente nas horas mau humor.

Agradeço à minha grande amiga Rachel Aires e às colegas Kamila Rios da UFSCar, Fernanda SAD e Camila Araujo. Aos amigos do GTI, firmes desde a graduação: Giovanna, Felix, Marcus, Leandro, Antônio e Guilherme, Pedro, Isaac,

e também a suas esposas e filhos.

Agradeço aos eternos grandes amigos do ICMC, Marcos Geromini e Jorge Pelizoni. Agradeço também aos meus amigos, os irmãos que eu escolhi, Giovanne, Fábio, Glauco e Bruno, Luiz Lurst e Reinaldo (Nico), Ewerton, Irdho, Ricardo Soriah, Leonardo França, Rafael Gervásio e Iure, Thiago Carbonel, Greg, Welinton e Alessandro, Vinicius Novais, Caio César, Israel e Paulo Andrade, Adriano Martins, Dona Solange e Roger. Meu muito obrigado por terem me entendido, me acompanhado e me incentivado nesse momento de crescimento.

Resumo

Com a popularidade e o crescimento das Redes Sociais Online, o interesse pelo entendimento de como seus usuários interagem entre si também tem crescido, configurando assim um cenário rico no qual são criadas oportunidades para melhorar o design de interfaces, compreender a organização de movimentos sociais, o marketing viral e a distribuição de conteúdos multimídia, dentre outros. Na investigação dessas oportunidades de pesquisa, um modelo de representação da atividade dos usuários amplamente empregado é baseado na construção de um grafo. Embora as atividades dos usuários em uma Rede Social Online sejam variadas, assim como as possibilidades de interação entre usuários, a construção de um grafo normalmente considera uma interação específica, que é então analisada a partir da interpretação de medidas baseadas em grafos ou de medidas estatísticas. Técnicas de mineração de dados podem ser empregadas de forma alternativa e complementar ao modelo baseado em grafos, possibilitando a extração e a avaliação de padrões das atividades de usuários. Entretanto, tanto na análise baseada em grafos quanto na utilização de técnicas de mineração de dados, a literatura reporta trabalhos em que são implicitamente representados elementos associados com a atividade dos usuários como ações executadas, mídias compartilhadas, aplicações e tipos de dispositivos utilizados. Nesse cenário, existe uma demanda por um modelo descritivo que permita a representação explícita dos elementos associados com as atividades dos usuários – representação essa que possa ser utilizada na extração e na avaliação das interações entre usuários. Para atender essa demanda, nesta tese é apresentada uma técnica e um método para representar, extrair e avaliar interações entre usuários de Redes Sociais Online. São também reportados resultados de experimentos da aplicação da técnica e do método a partir de dados obtidos de uma Rede Social Online.

Palavras-chave: Redes Sociais Online, Contingências Comportamentais, Interações Sociais, Regras *if-then*.

Abstract

The interest in understanding user interactions is increasing with the popularity of Online Social Networks. This scenario leads to rich opportunities to improve the design of interfaces, to understand the organization of social movements, to model the dynamics of viral marketing and to characterize the distribution of multimedia content, among others. In the investigation of these research opportunities, a widely used representation of users' activities is based on graphs as underlying data models. Graph data models usually consider a specific interaction, which is analyzed based on the interpretation of statistical measures and graph-based measures. On the other hand, user activities and possibilities of interaction in social networks are multifaceted, so that single-interaction graph-based approaches might prove to be unfeasible to properly model the problem. In order to deal with these limitations, data mining techniques may be employed as an alternative to graph-based modeling approaches, by enabling extraction and evaluation of users' activities as patterns. However, both in the use of the model graph-based as in the use of data mining techniques, the literature reports an implicit representation of elements associated with the users' activities such as executed actions, shared media, used applications and devices. In this scenario, there is a demand for a descriptive model which allows explicit representation of the elements associated with users activities - this representation can be used in the extraction and evaluation of interactions among users. To meet this demand, this thesis presents a technique and a method to represent, to extract and to evaluate interactions among users of Online Social Networks. Results of experimentation of applying the technique and method based on data obtained from an Online Social Network are also reported.

Keywords: Online Social Networks, Behavioral Contingencies, Social Interactions, *If-then* Rules.

Sumário

Resumo	vii
Abstract	ix
Sumário	xiv
Publicações Resultantes	xv
Glossário de Termos	xix
Lista de Abreviaturas	xxi
Lista de Figuras	xxiv
Lista de Tabelas	xxvi
Lista de Listagens	xxvii
1 Introdução	1
1.1 Contextualização	2
1.2 Motivação	3
1.3 Objetivo	6
1.4 Metodologia e Resultados	7
1.5 Contribuições	10
1.6 Estrutura da Tese	11
2 Representação de Redes Sociais Online: Ligações e Interações entre Usuários	13
2.1 Conceitos Básicos de Análise de Redes Sociais	14
2.1.1 Propriedades de Redes Complexas Observadas em Redes Sociais Online	14
2.1.2 Softwares para a Análise de Redes Sociais	15
2.2 Caracterização do Grafo Social de Redes Sociais Online	15
2.3 Caracterização do Grafo de Interação de Redes Sociais Online	18
2.4 Tipos de Interações entre Usuários de Redes Sociais Online	20
2.4.1 Interações Sociais baseadas em Mídias	20
2.4.2 Interações Sociais baseadas em Aplicativos de Terceiros	27

2.4.3	Outros Tipos de Interações Sociais	28
2.5	Trabalhos Relacionados	28
2.5.1	Propriedades Estruturais das Redes Sociais Online	28
2.5.2	Interações Sociais Baseadas em Mídias	30
2.5.3	Interações Sociais Baseadas no Uso de Aplicativos de Terceiros	32
2.6	Considerações Finais	33
3	Uma Técnica para Representação, Extração e Avaliação de Interações entre Usuários	35
3.1	A Linguagem de Mechner para Representação de Interações Sociais	36
3.2	Representação de Interações Sociais	37
3.3	Extração de Interações Sociais	40
3.4	Avaliação de Interações Sociais	44
3.5	Experimentações da Técnica sem Extração Automática	47
3.5.1	Interações Sociais Genéricas no Facebook	47
3.5.2	Interações Sociais Baseadas em Mídia	50
3.5.3	Interações Sociais Providas via Aplicações de Smartphones .	53
3.5.4	Síntese das Observações	55
3.6	Experimentação da Técnica com Extração Automática	56
3.6.1	Interações Sociais baseadas em Mídias vinculadas a Movimentação Social	57
3.6.2	Síntese das Observações	61
3.7	Trabalhos Relacionados	61
3.7.1	Técnicas de Mineração de Dados na Análise Relações Sociais	61
3.7.2	Ligações Sociais Ativadas pelo Uso de Aplicações Disponíveis em <i>Smartphones</i>	62
3.7.3	Engajamento de Usuários de Redes Sociais Online em Mobilizações Cívicas e Políticas	63
3.8	Considerações Finais	64
4	Um Método Guia para a Aplicação da Técnica	65
4.1	Um Procedimento para Análise de Interações Sociais baseadas em Mídias	66
4.2	Experimentação do Procedimento de Análise	68
4.2.1	Coleta e Organização de Dados de Interação	69
4.2.2	Representação das Interações Sociais	69
4.2.3	Mensuração de Interações Sociais	71
4.2.4	Interpretação dos Resultados e Melhorias	72
4.2.5	Síntese das Observações	73

4.3	Um Método para Representação, Extração e Avaliação de Interações Sociais	73
4.4	Experimentação do Método com Extração Manual de Contingências Comportamentais	77
4.4.1	Observação e Captura	77
4.4.2	Representação e Extração Manual	78
4.4.3	Avaliação, Validação e Interpretação	79
4.4.4	Especialização	80
4.4.5	Representação e Extração Manual de Novas Interações Sociais	80
4.4.6	Avaliação e Interpretação de Novas Interações Sociais	81
4.4.7	Síntese das Observações	82
4.5	Experimentação do Método na Análise de Comportamento Social .	83
4.5.1	Observação da Rede Social e Captura dos Dados	84
4.5.2	Preparação de Dados	85
4.5.3	Representação e Extração de Interações Sociais	85
4.5.4	Avaliação de Interações Sociais	86
4.5.5	Validação e Interpretação	88
4.5.6	Síntese das Observações	88
4.6	Trabalhos Relacionados	88
4.6.1	Análise do Comportamento Social dos Usuários da rede Twitter	89
4.6.2	Análise do Comportamento Social dos Usuários da rede Facebook	90
4.7	Considerações Finais	91
5	Avaliações da Aplicação da Técnica e do Método	93
5.1	Recrutamento dos Potenciais Usuários para as Avaliações	94
5.2	Avaliação da Técnica e do Método com Dados Coletados de Diferentes Unidades de Observação do Facebook	94
5.2.1	Observação e Captura	95
5.2.2	Preparação dos Dados	95
5.2.3	Representação e Extração de Interações Sociais Genéricas .	96
5.2.4	Avaliação, Validação e Interpretação de Interações Sociais Genéricas	98
5.2.5	Especialização	99
5.2.6	Extração de Interações Sociais Baseadas em Mídia	99
5.2.7	Avaliação das Interações Sociais baseadas em Mídia	101
5.2.8	Validação e Interpretação das Interações Sociais baseadas em Mídia	101
5.2.9	Síntese das Observações	102

5.3	Avaliação Qualitativa com Potenciais Usuários	103
5.3.1	Compilação dos Resultados das Sessões de Entrevista	103
5.3.2	Síntese dos Principais Resultados da Entrevista	107
5.4	Trabalhos Relacionados	109
5.4.1	Avaliação da Sociabilidade de Aplicações	109
5.4.2	Participação dos Usuários na Avaliação de Redes Sociais	109
5.5	Considerações Finais	110
6	Conclusão	113
6.1	Síntese das Contribuições	114
6.2	Síntese dos Resultados	115
6.3	Limitações	116
6.4	Trabalhos Futuros	116
6.5	Considerações Finais	118
	Referências	120
A	Experimentação da Técnica com Dados de Interação da CWaCTool	139
A.1	A Ferramenta CWaCTool em Uso	139
A.2	Representando e Avaliando Interações Sociais da CWaCTool	141
A.3	Trabalhos Relacionados	145
B	Planejamento da Avaliação com Engenheiros de Software	147
C	Roteiro da Entrevista com Engenheiros de Software	149

Publicações Resultantes

Artigos Resumidos em Conferências Internacionais

1. Gomes, A. K., Pedrosa, D. d. C., and Pimentel, M. G. C. (2011). Evaluating asynchronous sharing of links and annotation sessions as social interactions on internet videos. In Proceedings of the IEEE/IPSJ International Symposium on Applications and the Internet (SAINT '11), pages 184 – 189 [Gomes et al., 2011].
2. Gomes, A. K. and Pimentel, M. G. C. (2011c). Social interactions representation as users behavioral contingencies and evaluation in social networks. In Proceedings of the 2011 IEEE International Conference on Semantic Computing (ICSC '11), pages 275 – 278 [Gomes and Pimentel, 2011e].

Artigos Completos em Conferências Internacionais

1. Gomes, A. K. and Pimentel, M. G. C. (2011a). Measuring media-based social interactions provided by smartphone applications in social networks. In Proceedings of the 2011 ACM Workshop on Social and Behavioral Networked Media Access (SBNMA '11), pages 59 – 64 [Gomes and Pimentel, 2011c].
2. Gomes, A. K. and Pimentel, M. G. C. (2011b). Measuring synchronous and asynchronous sharing of collaborative annotations sessions on ubi-videos as social interactions. In Proceedings of the 2011 International Conference on Ubi-Media Computing (U-MEDIA '11), pages 122 – 129. ACM DL [Gomes and Pimentel, 2011d].
3. Gomes, A. K. and Pimentel, M. G. C. (2012e). A media-based social interactions analysis procedure. In Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC '12), pages 1018 – 1024 [Gomes and Pimentel, 2012e].

4. Gomes, A. K. and Pimentel, M. G. C. (2012a). Media-based Social Interaction Patterns: a Case Study in an Online Civic Mobilization. In Proceedings of the 2012 International Workshop on Socially-Aware Multimedia (IWSAM'12) in conjunction with ACM Multimedia, pages 1018 – 1024 [Gomes and Pimentel, 2012d].

Artigo Completo Premiado em Conferência Nacional

1. Gomes, A. K. and Pimentel, M. G. C. (2011f). A technique for human-readable representation and evaluation of media-based social interactions in social networks. In Proceedings of the 17th Brazilian Symposium on Multimedia and the Web (WebMedia '11), pages 119 – 126. SBC (Menção Honrosa) [Gomes and Pimentel, 2011f].

Artigos Completos em Conferências Nacionais

1. Gomes, A. K. and Pimentel, M. G. C. (2011e). Um Método para Análise de Interações Sociais na Web Social como Regras se-então. In Proceedings of the Workshop sobre Aspectos da Interação Humano-Computador para a Web Social (WAIHCWS'11) no X Simpósio Brasileiro de Fatores Humanos em Sistemas de Computação (IHC'11), pages 119 – 126. SBC [Gomes and Pimentel, 2011b].
2. Gomes, A. K. and Pimentel, M. G. C. (2012c). Measuring Media-based Social Interactions in Online Civic Mobilization Against Corruption in Brazil. In Proceedings of the 18th Brazilian Symposium on Multimedia and the Web (WebMedia '12). pages 173 – 180. ACM DL [Gomes and Pimentel, 2012b].

Artigo Resumido em Conferência Nacional

1. Gomes, A. K. and Pimentel, M. G. C. (2011d). Análise de Interações Sociais em Ambientes Sociais Multimídia. In Proceedings of the 17th Brazilian Symposium on Multimedia and the Web (WebMedia '11), pages 119 – 126. SBC [Gomes and Pimentel, 2011a].

Capítulo de Livro em Colaboração com Grupo de Pesquisa Internacional

1. Astudillo, H. and Codocedo, V. and Canals, G. and Torres, D. and Diaz, A. and Napoli, A. and **Gomes, A. K.** and Pimentel, M. G. C. (2009). Combining knowledge discovery, ontologies, annotations, and semantic wikis. Webmedia Minicourses, In Proceedings of the 15th Brazilian Symposium on Multimedia and the Web (WebMedia 2009). SBC [Astudillo et al., 2009].

Capítulo de Livro Internacional (em revisão)

1. Gomes, A. K. and Pimentel, M. G. C. (2012b). Evaluation of media-based social interactions: Linking collective actions to media types, applications and devices in social networks – Chapter (under review). In Nitin Agarwal, M. L. and Wigand, R., editors, Online Collective Action: Dynamics of the Crowd in Social Media – Lecture Notes in Social Networks. Springer [Gomes and Pimentel, 2012a].

Artigos Completos em Periódicos Internacionais (em avaliação)

1. Gomes, A. K. and Pimentel, M.G.C. (2012f). Representing, Mining and Measuring Social Interactions in Social Media Systems: A Rule-based Approach (submitted). In Journal of the Brazilian Computer Society. Springer [Gomes and Pimentel, 2012f].
2. Gomes, A. K. and Pimentel, M.G.C. (2012g). Media-based Application-induced Collective User Behavior Change (submitted). In IEEE Pervasive Computing [Gomes and Pimentel, 2012c].

Glossário de Termos

Análise de Redes Sociais: É campo de pesquisa que trata do estudo de conceitos, métodos e aplicações para analisar Redes Sociais [Scott, 2000] (pág. 14).

Contingências Comportamentais: São definidas como interações cotidianas entre os indivíduos (quaisquer tipo de interações sociais) que correspondem a observações do que as pessoas fazem ou não fazem em uma variedade de situações [Skinner, 1953] (pág. 7).

Elementos Associados com as Atividades dos Usuários: Para o propósito desta tese, são mídias compartilhadas e provedores de conteúdo correspondentes, dispositivos móveis, aplicativos de terceiros e ações executadas pelos usuários de Redes Sociais Online (pág. 3).

Estímulo Social: É a ação que dá início a uma interação social [Skinner, 1953] (pág. 37).

Grafo de Interação: É um grafo que representa as atividades específicas que os usuários executam na Rede Social Online. Também denominado de *grafo da atividade dos usuários* [Wilson et al., 2009], [Viswanath et al., 2009] e [Nazir et al., 2012] (pág. 5).

Grafo Social: É um grafo que representa uma Rede Social Online. Também pode ser denominado de *grafo de amizades* [Wasserman and Faust, 1994], [Scott, 2000] e [Freeman, 2004] (pág. 5).

Interações Sociais: São ações ou práticas de duas ou mais pessoas mutuamente orientadas, ou seja, são comportamentos de pessoas que afetam as experiências subjetivas ou as intenções de outras pessoas [Rummel, 1976] (pág. 3). No escopo desta tese, interações sociais são contingências comportamentais representadas como implicações lógicas que podem ser descritas como regras *if-then*.

Interações Sociais baseadas em Aplicativos de Terceiros: São interações entre usuários de Redes Sociais Online que ocorrem a partir da utilização aplicativos de terceiros (pág. 27).

Interações Sociais baseadas em Mídias: São interações entre usuários de Redes Sociais Online que ocorrem a partir do compartilhamento de diferentes tipos de mídia (pág. 21).

Interações Sociais Genéricas: São situações que ocorrem com frequência na Rede Social Online envolvendo os usuários em interações sociais. Na representação dessas interações são detalhadas apenas as ações que os usuários executam nas interações sociais (pág. 48).

Redes small word: São redes que atendem a hipótese de que a distância para conectar duas pessoas quaisquer é, em geral, pequena [Kleinberg, 2000] (pág. 14).

Redes Sociais Online: São sistemas nos quais (a) os usuários são as entidades principais e possuem um perfil público ou semipúblico, (b) os usuários podem criar ligações explícitas (conexões) com usuários ou com itens de conteúdo, e (c) os usuários podem navegar visitando as ligações e os perfis de outros usuários [Mislove et al., 2007] (pág. 1).

Redes power law: São redes que têm uma abundância de nós com poucas ligações e uma pequena mas significativa minoria de nós que têm um grande número de ligações [Magnien et al., 2011] (pág. 15).

Regra if-then: É uma declaração no formato *if-then* associada a uma implicação lógica que corresponde a uma interação social (pág. 8).

Unidade de Observação: É uma sub-rede da Rede Social Online a partir da qual podem ser coletados dados [Wasserman and Faust, 1994] (pág. 47).

Lista de Abreviaturas

AlaR Alavancagem da Regra

AM Aprendizado de Máquina

API *Application Programming Interface*

CWaC *Colaborative Watch-and-Comment*

ConR Confidência da Regra

CosR Correlação Cosseno da Regra

DM *Data Mining*

DSI *Dataset of Social Interactions*

GanR Ganho da Regra

GSP *Generalized Sequential Patterns*

KDD *Knowledge Discovery in Databases*

IA Inteligência Artificial

ICMC Instituto de Ciências Matemáticas e de Computação

LapR Laplace da Regra

MIS Método de Inspeção Semiótica

OSN *Online Social Network*

SBC Sociedade Brasileira de Computação

SenR Sensitividade da Regra

SIREMr *Social Interaction Representation Extraction Measuring rule-based*

SNA *Social Network Analysis*

SPADE *Sequential PAtterns Discovery using Equivalence classes*

SupR Suporte da Regra

TVD TV Digital

UtiR Utilidade da Regra

USP Universidade de São Paulo

URL *Uniform Resource Locator*

XML *Extensible Markup Language*

Lista de Figuras

2.1	Um Grafo Social de uma Rede Social Online	16
2.2	Interação Social baseada em Texto nas redes A) Facebook e B) Twitter	21
2.3	Interação Social baseada em Foto nas redes A) Facebook e B) Instagram	22
2.4	Interação Social baseada em Vídeo nas redes A) Facebook e B) YouTube	23
2.5	Interação Social baseada em Música nas redes A) Facebook e B) SoundCloud	24
2.6	Interação Social baseada em Link nas redes A) Facebook e B) Google+	25
2.7	Interação Social baseada em Checkin nas redes A) Facebook e B) Foursquare	25
2.8	Aplicativos de Terceiros instalados nas plataformas A) <i>Apple</i> e B) <i>Android</i>	27
3.1	Frequência de Mídias Compartilhadas na rede Facebook em Março de 2011 (fonte: [Gomes and Pimentel, 2011f])	51
3.2	Frequência de Uso das Aplicações de <i>Smartphone</i> no Compartilhamento de Mídias na rede Facebook em Março de 2011 (fonte: [Gomes and Pimentel, 2011c])	53
3.3	Frequências de Mídias Compartilhadas na Movimentação Social (fonte: [Gomes and Pimentel, 2012d])	57
3.4	Suporte das Interações Baseadas em Mídia vinculadas ao Movimento [Gomes and Pimentel, 2012d]	60
3.5	Correlação Cosseno das Interações Baseadas em Mídia vinculadas ao Movimento [Gomes and Pimentel, 2012d]	60
4.1	Visão Geral do Procedimento de Análise de Interações baseadas em Mídias (fonte: Gomes and Pimentel [2012e])	66

4.2	Compartilhamento de Vídeo da rede Youtube para a rede Facebook	68
4.3	O Método SIREMr (fonte: Gomes and Pimentel [2012b])	74
4.4	Frequência de Compartilhamento das Mídias (topo), Frequência dos Aplicativos Utilizados no Compartilhamento das Mídias (inferior) (fonte: Gomes and Pimentel [2012c])	83
4.5	Suporte das Interações baseadas em Mídia: Análise do Comportamento Social (fonte: Gomes and Pimentel [2012c])	87
4.6	Suporte das Interações baseadas em Mídia Providas por Aplicativos de Terceiros: Análise do Comportamento Social (fonte: Gomes and Pimentel [2012c])	87
5.1	Suporte das Interações Sociais Genéricas das Unidades de Observação	98
5.2	Frequências das Mídias Compartilhadas obtidas do Conjunto de Dados UOb	100
5.3	Suporte das Interações Baseadas em Mídias obtidas das Unidades de Observação	101
5.4	Correlação Cosseno das Interações Baseadas em Mídias obtidas das Unidades de Observação	102
A.1	Iniciando a <i>CWaCTool</i>	140
A.2	<i>CWaCTool</i> em uso [Gomes and Pimentel, 2011d].	141

Lista de Tabelas

1.1	Redes Sociais Online Populares e Ligações entre Usuários	2
2.1	Medidas do Grafo Social do Orkut, Flickr, Facebook, Twitter e Google+	17
2.2	Redes Sociais Online Populares e Mídias	20
3.1	Ilustração do Resultado da Execução do Algoritmo GSP	40
3.2	Informações sobre as Unidades de Observação: Técnica sem Extração Automática	47
3.3	Ações, Consequências e Usuários da rede Facebook na Experimentação da Técnica	48
3.4	Medidas de Contingências Comportamentais dos Usuários da rede Facebook: Interações Sociais Genéricas [Gomes and Pimentel, 2011e].	50
3.5	Medidas para Contingências baseadas em Mídia [Gomes and Pimentel, 2011f]	52
3.6	Medidas das Interações Sociais baseadas em Mídias e Aplicações de <i>Smartphone</i> [Gomes and Pimentel, 2011c]	55
3.7	Informações sobre a Unidade de Observação: Técnica com Extração Automática	57
3.8	Especialização da ação A_1 apresentada anteriormente na Tabela 3.3 (fonte: [Gomes and Pimentel, 2012d])	58
3.9	Comparação entre Algoritmos de Extração de Regras	59
4.1	Ações Executadas no Compartilhamento de Mídias entre Provedores de Conteúdo (fonte: Gomes and Pimentel [2012e])	69
4.2	Consequências e Usuários no Compartilhamento de Mídias entre Provedores de Conteúdo(fonte: Gomes and Pimentel [2012e])	70
4.3	Informações sobre a Unidade de Observação: Procedimento de Análise	70

4.4	Medidas para as Interações Sociais R_5 a R_{10} (fonte: [Gomes and Pimentel, 2012e])	72
4.5	Informações sobre as Unidades de Observação: Método de Análise	77
4.6	Ações, Consequências e Usuários da rede Facebook na Experimentação do Método	78
4.7	Especialização da ação A_1 apresentada anteriormente na Tabela 3.3 (fonte: Gomes and Pimentel [2012a])	78
4.8	Contingências $R_{4.1}$ a $R_{4.7}$ e as Medidas Computadas	79
4.9	Medidas para as Contingências $R_{4.10}$ to $R_{4.18}$	81
4.10	Ações, Consequências e Usuários: Análise do Comportamento Social	85
5.1	Informações sobre Perfil dos Potenciais Usuários	94
5.2	Informações sobre as Unidades de Observação: Avaliação da Técnica e do Método	95
5.3	Ações, Consequências e Usuários (replicação da Tabela 4.10)	96
5.4	Pergunta 3 e Respostas de Cada Potencial Usuário	104
5.5	Pergunta 4 e Respostas de Cada Potencial Usuário	105
5.6	Pergunta 5 e Respostas de Cada Potencial Usuário	105
5.7	Pergunta 6 e Respostas de Cada Potencial Usuário	105
5.8	Pergunta 7 e Respostas de Cada Potencial Usuário	106
5.9	Pergunta 8 e Respostas de Cada Potencial Usuário	106
A.1	Medidas para grupo G2 e grupo G3 [Gomes et al., 2011]	143
A.2	Medidas para grupo G4 e todos Grupos [Gomes et al., 2011]	143

Lista de Listagens

3.1	Contingências Comportamentais de Usuários da rede Facebook: Interações Sociais Genéricas [Gomes and Pimentel, 2011e].	49
3.2	Contingências Comportamentais de Usuários do Facebook: Interações Sociais baseadas em Mídia [Gomes and Pimentel, 2011f] . . .	52
3.3	Mídias e Aplicações de <i>Smartphones</i> Representadas na Contingência R3 [Gomes and Pimentel, 2011c]	54
3.4	Interações Sociais baseadas em Mídia vinculadas à Movimentação Social [Gomes and Pimentel, 2012d]	59
4.1	Contingências Comportamentais Associadas ao Compartilhamento de Mídias entre Provedores de Conteúdo [Gomes and Pimentel, 2012e]	70
4.2	Interações Sociais baseadas em Mídia (fonte: Gomes and Pimentel [2012a])	79
4.3	Aplicações Web e de <i>Smartphones</i> nas Interações Sociais baseadas em Mídia (fonte: [Gomes and Pimentel, 2012a])	80
4.4	Interações Sociais baseadas em Mídia na Análise do Comportamento Social [Gomes and Pimentel, 2012c]	86
4.5	Interações Sociais baseadas em Mídia providas via Aplicativos na Análise do Comportamento Social [Gomes and Pimentel, 2012c] . .	86
5.1	Interações Sociais Genéricas Recorrentes nas Unidades de Observação	96
5.2	Interações Sociais Baseada em Mídias para Avaliação com Usuários	100
B.1	Planejamento da Avaliação com Engenheiros de Software	147
C.1	Questionário aplicado na Avaliação com Engenheiros de Software .	149

Introdução

A proliferação de sistemas de computação social tais como as Redes Sociais Online (em inglês *Online Social Networks* – OSNs) motivam pesquisas na área de Web e Multimídia em aspectos que incluem a documentação, a indexação, a recuperação e a mineração de conteúdo baseado em múltiplas mídias [Cesar et al., 2011], [Tilson et al., 2010].

A popularidade e o crescimento das Redes Sociais Online estão relacionados com as facilidades no estabelecimento de ligações entre usuários e também com as facilidades de compartilhamento de conteúdo. Por exemplo, Gilbert and Karahalios [2009] argumentam que a força das ligações entre os usuários varia amplamente em uma Rede Social Online, de acordo com o tipo de conteúdo compartilhado pelos usuários.

Segundo Mislove et al. [2007], *Redes Sociais Online* são sistemas nos quais (a) os usuários são as entidades principais e possuem um perfil público ou semi-público, (b) os usuários podem criar ligações explícitas (conexões) com usuários ou com itens de conteúdo, e (c) os usuários podem navegar visitando as ligações e os perfis de outros usuários.

As Redes Sociais Online dependem essencialmente das interações entre seus usuários, a partir das quais o conteúdo compartilhado é construído por meio do uso de tecnologias disponíveis na Web 2.0 [King et al., 2009], em plataformas móveis e ubíquas [Wang and Chin, 2010] e também em plataformas de TV digital [Cesar and Geerts, 2011]. Os usuários de redes sociais podem manter e fortalecer laços sociais que existem na vida real ou ainda fazer novos contatos sociais. Dessa forma, os usuários articulam e tornam visíveis suas redes soci-

Tabela 1.1: *Redes Sociais Online Populares e Ligações entre Usuários*

Nome	URL	Ligação entre Usuários
Facebook	http://www.facebook.com	amizade – bidirecional assinante/seguidor – direcional
Twitter	http://www.twitter.com	seguidor/seguido – direcional
Google+	http://plus.google.com	amizade – direcional
YouTube	http://www.youtube.com	assinante – direcional
Digg	http://www.digg.com	assinante – direcional
Flickr	http://www.flickr.com	amizade – bidirecional
FourSquare	http://www.foursquare.com	amizade – bidirecional
Instagram	http://www.instagram.com	seguidor/seguido – direcional
Soundcloud	http://soundcloud.com	seguidor/seguido – direcional
Myspace	http://www.myspace.com	amizade – bidirecional
Orkut	http://www.orkut.com	amizade – bidirecional
Last.fm	http://www.last.fm	amizade – bidirecional

ais, e ainda, mantêm a comunicação com pessoas que fazem parte da sua Rede Social da vida real e da vida virtual.

1.1 Contextualização

As conexões entre usuários de Redes Sociais Online podem atender a diferentes interesses, como por exemplo, para estabelecimento de contatos profissionais (e.g., LinkedIn) ou pessoais – como no caso de contatos com amigos ou conhecidos (e.g., Facebook, Google+). O tipo de conexão entre usuários varia de acordo com a Rede Social Online: uma ligação pode ser recíproca ou bidirecional (por exemplo, relação de amizade), ou ainda, unilateral ou direcional (por exemplo, relação de seguidor ou assinante) [Wasserman and Faust, 1994].

Na Tabela 1.1 são apresentadas URLs de Redes Sociais Online populares¹ e também o tipo de ligação estabelecida entre os usuários dessas redes. Atualmente, uma das redes mais populares no mundo é Facebook, tanto em termos de usuários quanto em termos de quantidade de tempo que os usuários permanecem autenticados no site, de acordo com o site Alexa.com². Os usuários do Facebook são também os mais conectados entre si: 57% dos usuários têm mais de 100 amigos.³ De acordo com informações disponíveis na própria rede, o Facebook contava com mais de 1 bilhão de usuários em Novembro de 2012.⁴

Um conceito importante área de Ciências Sociais é o de *interação social*: Rumel [1976] define interações sociais como ações ou práticas de duas ou mais

¹Estatísticas importantes e dados sobre o crescimento de 20 redes sociais populares estão disponíveis no link <http://ijustdid.org/2012/11/top-social-networks/> acessado em 10/01/2013.

²<http://www.alexacom.com/siteinfo/facebook.com> acessado em 10/01/2013.

³<http://www.bloggingprweb.com/infographic-facebook-twitter-pr-social-media> acessado em 10/01/2013.

⁴newsroom.fb.com/News/One-Billion-People-on-Facebook acessado em 10/01/2013.

pessoas mutuamente orientadas, ou seja, são comportamentos de pessoas que afetam as experiências subjetivas ou as intenções de outras pessoas. No contexto de Redes Sociais Online, a possibilidade de utilização de recursos variados estimula atividades por parte de seus usuários, ou seja, a utilização dos recursos disponibilizados na rede tem por objetivo provocar o engajamento de usuários em interações sociais [Wilson et al., 2009].

Nas interações sociais entre usuários de Redes Sociais Online são trocados conteúdos baseados em mídias como texto, fotos, vídeos e músicas. Conteúdos baseados em mídias podem ser criados ou compartilhados via dispositivos móveis, provedores de conteúdo ou aplicativos de terceiros. Os usuários podem se associar aos conteúdos baseados em mídias por meio da execução de ações pré-definidas na rede como curtir, comentar e compartilhar.

Para o propósito desta tese, mídias compartilhadas e provedores de conteúdo correspondentes, dispositivos móveis, aplicativos de terceiros e ações executadas pelos usuários⁵ são *elementos associados com as atividades dos usuários* de Redes Sociais Online.

Na Rede Social Online Facebook, por exemplo, no ano de 2012 foram compartilhados de mais de 250 milhões de fotos, foram utilizados de mais de 29 mil aplicativos de terceiros, e ainda, foram executadas de mais de 2.7 bilhões de ações *Curtir* a partir do compartilhamento de diferentes conteúdos baseados em mídias.⁶

1.2 Motivação

O interesse pelo estudo das atividades dos usuários nas Redes Sociais Online é crescente. O estado da arte reporta pesquisas que tratam de temas como:

- Propriedades estruturais das Redes Sociais Online. Nos trabalhos de Mislove et al. [2007], Ahn et al. [2007], Wang and Chin [2010], Amaral et al. [2000], e Moradi et al. [2012] são examinadas propriedades estruturais em diversas redes sociais. A representação das redes Twitter e Facebook na forma de um grafo é investigada nos trabalhos de Wilson et al. [2009], Viswanath et al. [2009], Kivran-Swaine and Naaman [2011], Silva et al. [2012] e [Wilson et al., 2012].
- Interações sociais baseadas em mídias. Interações sociais baseadas em texto são estudadas nos trabalhos de De Choudhury et al. [2009], Burke

⁵Alguns termos usados nesta tese referem-se a ações dos usuários de Redes Sociais Online, por exemplo, os termos *Post*, *tweet* e *retweet*, que foram mantidos em inglês. O termo *Post* equivale a ação de publicar ou compartilhar um conteúdo nas redes Facebook e Google+. O termo *tweet* equivale a ação de publicar ou compartilhar uma mensagem de texto com no máximo 140 caracteres na rede Twitter. O termo *retweet* equivale a ação de encaminhar um *tweet* já publicados.

⁶<http://infographiclabs.com/infographic/facebook-2012/> em 10/01/2013.

et al. [2010] e Comarela et al. [2012]. Interações sociais baseadas em fotos são investigadas nos trabalhos de Malinen [2010], Rabbath et al. [2010], Rabbath et al. [2012] e Pesce et al. [2012]. Interações sociais baseadas em vídeo são tratadas nos trabalhos de Benevenuto et al. [2009a], Susarla et al. [2012] e Doerr et al. [2012]. Já estudos de interações sociais baseadas em outros tipos de mídia como, por exemplo, música, links para notícias e *checkin* são reportados, respectivamente, nos trabalhos de Mechant and Evens [2011], Park and Chung [2012] e Vasconcelos et al. [2012].

- Interações sociais baseadas em aplicativos de terceiros. O uso de aplicativos de terceiros que permitem a interação entre usuários do Facebook é analisado nos trabalhos de Gjoka et al. [2008] (caracterização do uso de aplicações de terceiros), Nazir et al. [2008] (demanda pelo uso da rede), Egele et al. [2012] (proteção de usuários contra aplicações maliciosas), Anthonysamy et al. [2012] (controle de privacidade) e Nazir et al. [2012] (caracterização do grafo de interação com base no uso de aplicações de envio de presentes).
- Processamento e análise do comportamento social dos usuários. A modelagem, a análise e a visualização do comportamento social são exploradas nos trabalhos de Vinciarelli et al. [2009], Tsai et al. [2012], Wilkinson and Thelwall [2011] e Wang et al. [2011]. A análise do comportamento dos usuários do Twitter é realizada nos trabalhos de Silva et al. [2011], De Choudhury et al. [2011], Bigonha et al. [2010], Diakopoulos et al. [2012], Liao et al. [2012] e Barbosa et al. [2012]. Nos trabalhos de Gilbert and Karahalios [2009], Rabbath et al. [2010], Hew [2011], Kramer [2012] e Moore and McElroy [2012] são reportadas análises do comportamento dos usuários do Facebook.
- Ligações sociais ativadas a partir de dispositivos móveis. As ligações sociais ativadas a partir de *smartphones* são foco de estudo nos trabalhos de Barkhuus and Tashiro [2010], Quercia et al. [2010] e Quercia et al. [2011]. O uso de aplicativos que são habilitados em dispositivos móveis, e ativam ligações sociais entre usuários, é analisado nos trabalhos de Lane et al. [2010], Anthonysamy et al. [2012], Egele et al. [2012] e Wang [2012].
- Técnicas de mineração de dados voltadas para a análise relações sociais. Técnicas de agrupamento são aplicadas nos trabalhos de Maia et al. [2008], Stroele et al. [2009], Negoescu et al. [2009] e Abrol and Khan [2010]. Técnicas de classificação são empregadas nos trabalhos de Fortuna et al. [2007]

e Kajdanowicz et al. [2010]. Técnicas de regressão são usadas nos trabalhos de Jin et al. [2010] e Golder and Yardi [2010]. Técnicas baseadas em regras de associação são utilizadas nos trabalhos de Leung et al. [2010], Esmaeili et al. [2011] e Braghin et al. [2011].

- Engajamento de usuários em mobilizações cívicas online. As redes Twitter e Facebook são utilizadas pelos usuários no engajamento para mobilizações cívicas e políticas, conforme apresentam em seus trabalhos Williams [2007], Robertson et al. [2009], Zhou et al. [2010], Mascaro and Goggins [2011], Starbird and Palen [2011] e Starbird and Palen [2012].
- Análise de interações sociais com a avaliação dos usuários. A participação dos usuários de diversas redes na avaliação de interações sociais é tratada nos trabalhos de De Choudhury et al. [2009], Benevenuto et al. [2009a], Benevenuto et al. [2012], Lin et al. [2012] e Xu et al. [2012].

Em pesquisas da área de *Análise de Redes Sociais* (em inglês *Social Network Analysis* – SNA), Redes Sociais são estudadas em caráter multidisciplinar, envolvendo disciplinas de Computação e de Ciências Sociais [Wasserman and Faust, 1994]. Nessas pesquisas, Redes Sociais são geralmente representadas como um grafo (ou sociograma), no qual os nós do grafo correspondem aos indivíduos (entidades sociais) e as arestas correspondem às ligações entre os indivíduos [Scott, 2000], [Freeman, 2004].

No caso de Rede Social Online, o grafo que representa a rede é denominado *grafo social* ou *grafo de amizades*. A caracterização de um grafo social é realizada a partir de métricas comumente empregadas na Análise de Redes Complexas [Newman, 2003], [Costa et al., 2005].

Conforme destaca Nazir et al. [2012], o crescente interesse acadêmico no estudo de Redes Sociais Online tem produzido resultados cujo foco está no estudo do grafo social, no entanto, pouco se sabe sobre a atividade dos usuários de OSNs que emerge a partir da formação das ligações sociais. A representação baseada em grafos de Redes Sociais Online tem uma limitação apontada por Scott [2000], que advoga que essa representação não priorizada a análise da ação coletiva ou individual dos usuários.

No modelo baseado em grafos, as atividade dos usuários em Redes Sociais Online são representadas em um grafo denominado de *grafo de interação* [Wilson et al., 2009], [Viswanath et al., 2009], ou ainda, de *grafo da atividade dos usuários* [Nazir et al., 2012].

Embora a atividade dos usuários em uma Rede Social Online seja variada, assim como as possibilidades de interação social, o grafo de interação é construído a partir da análise de uma interação específica entre usuários, por exemplo, a

adição de comentários a uma foto [Viswanath et al., 2009], [Wilson et al., 2012], a manifestação de opinião a favor ou contra determinados assuntos [Bigonha et al., 2010], [Sun and Ng, 2012], e ainda, o convite para utilizar um aplicativo social [Nazir et al., 2008], [Nazir et al., 2009].

Alternativamente e também de forma complementar ao modelo baseado em grafos, técnicas de mineração de dados são usadas na extração e na avaliação de padrões da atividade dos usuários de Redes Sociais Online, por exemplo, utilizando-se técnicas de agrupamento (em inglês *clustering*) [Maia et al., 2008] [Stroele et al., 2009], [Negoescu et al., 2009], [Abrol and Khan, 2010], técnicas de classificação [Fortuna et al., 2007], [Kajdanowicz et al., 2010], técnicas de regressão [Jin et al., 2010] [Golder and Yardi, 2010] e também técnicas baseadas em regras de associação [Esmaeili et al., 2011], [Anastopoulos et al., 2011].

Os padrões de atividades dos usuários extraídos pelas técnicas de mineração de dados não têm foco na representação explícita das ligações entre elementos associados com a atividade dos usuários como ações executadas, mídias compartilhadas, aplicações e tipos de dispositivos utilizados. A ausência dessa representação explícita pode ser um problema para a interpretação dos padrões de atividades identificados.

Os trabalhos que exploram um modelo baseado em grafos ou técnicas de mineração de dados, referenciados acima, fazem uso de uma representação implícita para os elementos associados com as atividades dos usuários como ações executadas, mídias compartilhadas, aplicações e tipos de dispositivos utilizados. Nesse cenário, existe uma demanda por um modelo descritivo que permita a representação explícita dos elementos associados com as atividades dos usuários – representação essa que possa ser utilizada na extração e na avaliação de interações sociais. Para atender essa demanda, nesta tese é apresentada uma técnica e um método para representar, extrair e avaliar interações entre usuários de Redes Sociais Online.

1.3 Objetivo

Os objetivos desta tese são:

- propor uma técnica que permita a representação, a extração e a avaliação de interações entre usuários de Redes Sociais Online,
- propor um método que guie a aplicação da técnica, e permita a análise de interações sociais especializadas.

A técnica e o método propostos devem possibilitar a representação explícita dos elementos associados com as atividades de usuários em Redes Sociais Online,

e incluem as ações executadas pelos usuários, as mídias compartilhadas, os servidores, os tipos de dispositivos móveis e os tipos de aplicativos de terceiros utilizados para compartilhar essas mídias.

1.4 Metodologia e Resultados

No campo da Ciência Comportamental, *contingências comportamentais* são definidas como interações cotidianas entre os indivíduos (quaisquer tipo de interações sociais) que correspondem a observações do que as pessoas fazem ou não fazem em uma variedade de situações [Skinner, 1953]. Por exemplo, as regras de um jogo de xadrez que determinam a forma como os jogadores podem participar do jogo, são contingências comportamentais.

Resultados de Psicologia Social Experimental argumentam que interações sociais podem ser especificadas como contingências comportamentais, que podem ser descritas como relações condicionais no formato de declarações *if-then* [Mechner, 1959]. Como exemplo de declaração *if-then* pode ser citada a situação:

if jogador a faz jogada 1 *then* jogador b faz jogada 2.

A linguagem formal simbólica denominada *linguagem de Mechner* é usada para codificar contingências comportamentais como implicações lógicas a partir de declarações *if-then* [Mechner, 2008]. Por exemplo, utilizando a linguagem de Mechner para representar a declaração apresentada no parágrafo anterior, é possível identificar os usuários a e b , a ação $A = \text{faz jogada 1}$ e a consequência $C = \text{faz jogada 2}$. Identificados usuários, ação e consequência, a declaração pode ser representada como a implicação lógica $aA \rightarrow bC$.

No escopo desta tese, contingências comportamentais associadas a interações sociais entre usuários de Redes Sociais Online são representadas na linguagem de Mechner, ou seja, uma interação social é representada como uma implicação lógica que corresponde a uma regra (declaração) no formato *if-then*. Ainda no escopo desta tese, interações sociais representadas como contingências comportamentais são extraídas de forma manual e automática. Com o suporte de um especialista em Psicologia Social, após a identificação de elementos associados a linguagem de Mechner presentes na Rede Social Online, a extração manual de contingências comportamentais é realizada a partir da observação das interações entre usuários. O suporte desse especialista também é solicitado na validação das contingências comportamentais extraídas de forma automática.

A partir da coleta de dados associados com os elementos da linguagem de Mechner, a extração automática de contingências comportamentais é realizada a partir de uma implementação da modificação de um algoritmo de aprendizado de padrões sequenciais. Em seguida, as contingências comportamentais ex-

traídas como padrões sequenciais são convertidas para implicações lógicas que correspondem a regras *if-then*.

A avaliação de contingências comportamentais que correspondem a regras *if-then* é realizada utilizando-se procedimentos típicos do campo de pesquisa da mineração de dados. Cada interação social no formato de uma regra *if-then* é representada como uma implicação lógica na forma geral $B \rightarrow H$. Para cada regra na forma geral, são computados os valores $bh + b\bar{h} + \bar{b}h + \bar{b}\bar{h}$. Partindo-se desses valores, são calculadas medidas de avaliação de regras como Suporte, Confiança e Correlação Cosseno [Lavrac et al., 1999], [Fürnkranz et al., 2011].

No desenvolvimento da técnica e do método propostos nesta tese, os passos seguidos foram:

- Primeiramente foi realizado um levantamento bibliográfico com o propósito de caracterizar o estado da arte das pesquisas relacionadas com a representação, a extração e a avaliação da atividade dos usuários em Redes Sociais Online. A partir desse levantamento, foi identificada a oportunidade de pesquisa tratada no escopo desta tese.
- A seguir, foi realizada a computação de contingências comportamentais associadas com o compartilhamento síncrono [Gomes and Pimentel, 2011d] e assíncrono [Gomes et al., 2011] de anotações em vídeo a partir do uso da ferramenta *CWaCTool* [Fagá et al., 2010]. As experimentações correspondentes permitiram fazer observações iniciais sobre a utilização da linguagem de Mechner na representação de contingências comportamentais que ocorrem a partir do uso da ferramenta. Também foi possível realizar observações iniciais sobre a utilização de medidas de regras na avaliação dessas contingências comportamentais.
- Uma técnica para a representação e a avaliação de interações sociais foi desenvolvida como consequência dos trabalhos anteriores [Gomes and Pimentel, 2011f]. A técnica combina a representação das interações sociais na linguagem de Mechner com a avaliação dessas interações por meio de um procedimento clássico de mineração de dados. A técnica para a representação e avaliação de interações sociais foi aplicada na identificação de situações frequentes que envolvem usuários do Facebook em interações sociais [Gomes and Pimentel, 2011e], e também na análise de interações sociais baseadas em mídias providas por aplicações disponíveis em *smartphones* [Gomes and Pimentel, 2011c].

Nas experimentações correspondentes, foi possível observar que interações sociais frequentes podem ser especializadas em interações baseadas em mí-

dia. Por sua vez, interações sociais baseadas em mídia podem ser especializadas em interações sociais baseadas em mídia providas por aplicações de *smartphones*. Esse resultado sugere que a técnica pode ser aplicada de forma iterativa e ainda, mediante sucessivas especializações, a técnica permite que mídias e aplicações de *smartphones* que provêm essas mídias possam ser detalhadas nas interações sociais.

- Em seguida, um procedimento de análise a partir da aplicação da técnica foi desenvolvido. O procedimento de análise permite que interações sociais genéricas sejam especializadas em interações sociais baseadas em mídia. O procedimento permite, por exemplo, a identificação das interações sociais baseadas em mídia mais frequentes, e a identificação dos provedores de conteúdo utilizados no compartilhamento dos diferentes tipos de mídia [Gomes and Pimentel, 2012e].

Nas experimentações correspondentes, foi possível observar algumas preferências dos usuários no uso das opções de compartilhamento de mídia disponíveis dentro de uma Rede Social Online, e também, preferências no uso das opções de compartilhamento de mídias a partir dos provedores de conteúdo para dentro da rede.

- Um método foi desenvolvido a partir da expansão do procedimento de análise. O método guia a aplicação da técnica de representação e a avaliação de interações sociais e foi especificado em termos da ligação entre mídias e outros elementos associados com as ações coletivas executadas pelos usuários de OSNs [Gomes and Pimentel, 2011b], [Gomes and Pimentel, 2012a].

Nas experimentações correspondentes foi possível observar o impacto da inclusão de um novo tipo de mídia na rede como opção para interação entre os usuários. Foi possível constatar que os resultados da aplicação da técnica e do método foram consistentes e robustos, mostrando que interações sociais genéricas podem ser especializadas em interações sociais baseadas em mídia, que por sua vez podem ser especializadas em interações sociais baseadas em mídia providas por aplicações Web e dispositivos móveis.

- No passo seguinte, uma etapa de extração automática de contingências comportamentais na forma de interações sociais foi incluída na técnica de representação e avaliação. Um algoritmo clássico para extração de padrões sequenciais foi modificado para a extração automática de interações sociais. Os padrões sequenciais extraídos são representados no formato de regras *if-then*. A técnica de representação e avaliação expandida com a etapa de extração foi aplicada na análise das manifestações, iniciadas por

usuários da rede Facebook, em dois momentos diferentes de mobilização cívica contra corrupção no Brasil [Gomes and Pimentel, 2012d].

Nas experimentações correspondentes foi possível observar que a técnica expandida pode ser empregada na análise do comportamento coletivo de usuários de Redes Sociais Online. A técnica permite que o comportamento coletivo seja representado como interações sociais baseadas em mídias, e permite também que essas interações sejam extraídas de forma automática e avaliadas com diferentes medidas.

- Em seguida, o método foi reestruturado para guiar a aplicação da técnica de representação, de extração e de avaliação de interações sociais. O método foi aplicado na análise das manifestações iniciadas por usuários da rede Facebook, em dois momentos diferentes de mobilização cívica contra corrupção no Brasil [Gomes and Pimentel, 2012b], e na análise do comportamento dos usuários no uso de aplicativos de terceiros em interações sociais [Gomes and Pimentel, 2012c].

Nas experimentações correspondentes foi possível observar um padrão de comportamento dos usuários do Facebook, em termos de interações sociais genéricas, durante a mobilização social. Também foi possível observar uma mudança de comportamento dos usuários da rede em termos do compartilhamento de mídias e do uso de aplicativos para o compartilhamento dessas mídias.

- Na etapa final, o método foi aplicado na análise de dados colhidos de diferentes unidades de observação do Facebook. Também foi realizada uma avaliação qualitativa do método com engenheiros de software [Gomes and Pimentel, 2012f].

Nas experimentações correspondentes foi possível observar que a aplicação do método manteve-se estável e consistente com resultados anteriores. Questões relevantes foram avaliadas com relação a aplicação da técnica e do método, e também, foram levantados três cenários para a construção de aplicações sociais.

1.5 Contribuições

As principais contribuições desta tese são a técnica e o método propostos para a representação, a extração e a avaliação de interações sociais nas quais os elementos associados com as atividades dos usuários em Redes Sociais Online são explicitamente descritos. Essas contribuições foram alcançadas mediante sucessivas experimentações com dados coletados da rede Facebook. Essa Rede

Social foi escolhida devido ao grande número de usuários por meio dela conectados, a riqueza de possibilidades de interação entre usuários e a variedade de mídias disponíveis para interação. Como a aplicação da técnica e do método é independente da Rede Social utilizada, a princípio as contribuições desta tese podem ser generalizadas e aplicadas a outras Redes Sociais Online.

Resultados anteriores que contribuíram para a proposição da técnica de representação e avaliação foram obtidos a partir da análise de interações sociais associadas com o compartilhamento assíncrono [Gomes et al., 2011] e síncrono [Gomes and Pimentel, 2011d] de anotações. Essas interações sociais são resultantes da atividade dos usuários da ferramenta de anotação em vídeo CWaCTool [Fagá et al., 2010], uma aplicação que explora a abordagem social do paradigma *Watch-and-Comment* [Cattelan et al., 2008]. Uma síntese desses resultados é apresentada no Apêndice A.

1.6 Estrutura da Tese

Os demais capítulos desta tese estão organizados da seguinte forma: no Capítulo 2 são tratados os principais conceitos relacionados com a representação das ligações e das interações entre usuários, e também são destacados os principais tipos de interações sociais que existem atualmente nas Redes Sociais Online; no Capítulo 3 é apresentada a técnica de representação, extração e avaliação de interações sociais, e são reportados resultados da aplicação da técnica; no Capítulo 4 é apresentado o procedimento de análise de interações sociais baseadas em mídia, bem como o método que guia a aplicação da técnica; no Capítulo 5 são reportados o resultado da avaliação da técnica e do método com engenheiros de software, e resultados da aplicação da técnica e do método na análise de interações sociais coletadas a partir de diferentes unidades de observação da Facebook; o Capítulo 6, além de sintetizar as contribuições da tese, discute as limitações do trabalho atual e aponta direções para trabalhos futuros.

Representação de Redes Sociais Online: Ligações e Interações entre Usuários

Um modelo tipicamente utilizado na representação e na análise de Redes Sociais Online é o modelo baseado em grafos. O grafo que representa as ligações entre usuários é analisado utilizando-se medidas típicas da Análise de Redes Sociais. Essa representação viabiliza a análise de propriedades estruturais da rede, no entanto, não permite que as ligações que estão ativas nas interações sociais sejam representadas e analisadas. Para superar essa limitação, é utilizado um grafo que representa as ligações ativas a partir de eventos de interação entre usuários. Por sua vez, diferentes tipos de interações sociais são tratados implicitamente a partir da identificação dos eventos de interação. A identificação das formas de interações sociais é importante para a indicação precisa de quais são as ligações ativas na rede.

Neste capítulo são tratadas questões fundamentais para a compreensão da utilização do modelo baseado em grafos na análise de ligações e interações sociais em Redes Sociais Online. Também são caracterizados os diferentes tipos de interações sociais que não são contemplados explicitamente no modelo baseado em grafos.

Este capítulo está organizado da seguinte forma: conceitos básicos da Análise de Redes Sociais são apresentados na Seção 2.1; a caracterização do grafo social de Redes Sociais Online e a apresentação de algumas medidas utilizadas para analisar essa representação são tratadas na Seção 2.2; os principais aspectos relacionados com a caracterização do grafo de interação de Redes Sociais

Online são reportados na Seção 2.3; na Seção 2.4 são apresentados alguns tipos de interações sociais que emergem de características das OSNs e não são explicitamente tratadas na representação baseada em grafos; na Seção 2.5 são apresentados os trabalhos relacionados com a análise da estrutura e das interações sociais em Redes Sociais Online; finalmente, na Seção 2.6 são apresentadas as considerações finais.

2.1 Conceitos Básicos de Análise de Redes Sociais

Como mencionado anteriormente, o campo de pesquisa que trata do estudo de conceitos, métodos e aplicações para analisar Redes Sociais é denominado *Análise de Redes Sociais* (SNA). A abordagem empregada nessa área, é útil para explicar muitos fenômenos do mundo real tais como as relações que são estabelecidas nas Redes Sociais Online. Nesses estudos, os atributos dos indivíduos são menos importantes do que as ligações entre os atores. Essa abordagem deixa menos espaço para a análise da ação individual ou coletiva da atividade dos atores na rede, além de ser dependente da estrutura interna da rede [Scott, 2000].

Redes Sociais Online são representadas como um grafo (ou sociograma) na Análise de Redes Sociais [Wasserman and Faust, 1994], [Scott, 2000], [Freeman, 2004]. Na análise de Redes Sociais Online são empregadas métricas da Análise de Redes Complexas. Uma breve análise dos principais modelos e medidas de Redes Complexas pode ser encontrado no trabalho de Costa et al. [2005]. No trabalho de Newman [2003] pode ser encontrada uma revisão mais detalhada de Redes Complexas e suas medidas.

2.1.1 Propriedades de Redes Complexas Observadas em Redes Sociais Online

Algumas propriedades específicas comumente associadas a Redes Complexas tais como propriedades *small word* e *power law* podem ser observadas em Redes Sociais Online.

Redes *small word* recebem esse nome por analogia ao fenômeno de igual nome também conhecido como o fenômeno dos “seis graus de separação”. Esse fenômeno corresponde à hipótese de que, em uma Rede Social, a distância para conectar duas pessoas quaisquer em qualquer lugar do mundo é, em geral, pequena. Maiores detalhes sobre propriedades de redes *small word* podem ser obtidas no trabalho de Kleinberg [2000].

Redes Sociais Online com propriedade *small word* têm alto coeficiente de agrupamento dos nós e pequeno comprimento médio de caminhos entre nós [Watts, 1999]. Por exemplo, Mislove et al. [2007] verificaram a existência da proprie-

dade *small word* em quatro redes sociais online: LiveJournal, Flickr, Orkut, e YouTube.

Redes *power law* têm distribuições dos graus dos nós que seguem a lei de potência, ou seja, são redes que têm uma abundância de nós com poucas ligações, e uma pequena mas significativa minoria de nós que têm um grande número de ligações. No trabalho de Magnien et al. [2011] é apresentado um *survey* sobre redes *power law*.

2.1.2 Softwares para a Análise de Redes Sociais

Softwares voltados para a Análise de Redes Sociais são utilizados para identificar, representar, analisar, visualizar ou simular nós e arestas (relações) a partir de conjuntos de dados de entrada que podem, por exemplo, estar na forma de tabelas ou arquivos XML. Esses softwares permitem a construção de representações visuais das Redes Sociais que são importantes para o entendimento da organização dos dados da rede, facilitando a interpretação qualitativa desses dados. Alguns softwares que podem ser utilizados na Análise de Redes Sociais são listados na Wikipédia¹ conforme observado por Scott [2000].

O material disponível na Wikipédia também discute vários aspectos das ferramentas listadas, por exemplo, em relação a tipo de interface gráfica, aos formatos de dados manipulados e ao tipo de licença de cada ferramenta. Em relação a interface gráfica, Algumas das ferramentas *opensource* são amplamente utilizadas são NetMiner, UCINET, Pajek, GUESS, ORA, e Cytoscape. Algumas ferramentas comerciais disponíveis são Orgnet, Keyhubs, e KXEN. Entre as ferramentas que permitem a criação de scripts estão: NetMiner que utiliza scripts em Python; IGRAPH, que utiliza scripts nas linguagens R e Python; a biblioteca NetworkX, que utiliza scripts em Python e em C++. Essas ferramentas possuem tutoriais e documentação extensos.

Com relação à visualização, as ferramentas de Análise de Redes Sociais permitem alterar a disposição dos nós, as cores, o tamanho e outras propriedades da representação de rede. As ferramentas NetMiner, IGRAPH, Cytoscape, NetworkX possuem funcionalidades que permitem de produção de gráficos com alta qualidade e grande variedade de perspectivas de visualização da organização dos dados da rede.

2.2 Caracterização do Grafo Social de Redes Sociais Online

Uma Rede Social Online pode ser representada como um *grafo social*, ou seja, um grafo $G = (V, E)$ onde os vértices V são os usuários, e as arestas E são as ligações entre os usuários [Wasserman and Faust, 1994]. Na Figura 2.1 é apre-

¹http://en.wikipedia.org/wiki/Social_network_analysis_software em 20/11/2012.

Figura 2.1: *Um Grafo Social de uma Rede Social Online*

sentado um grafo social que representa uma Rede Social Online². Geralmente, os usuários com mais ligações são representados no centro do grafo e os usuários com menos ligações são representados nas bordas.

Um grafo social pode ser avaliado a partir da utilização de medidas típicas da Análise de Redes Sociais. Essa avaliação tem o propósito de caracterizar algumas propriedades estruturais do grafo social, e consecutivamente, da Rede Social Online. Na avaliação das propriedades estruturais de um grafo social são utilizadas medidas como *Raio*, *Diâmetro*, *Caminho Médio*, *Coefficiente de Agrupamento* e *Assortatividade*. Essas medidas são explicadas a seguir de acordo com o exposto no trabalho de Wilson et al. [2012].

As medidas de *Raio* e *Diâmetro* são calculadas a partir da excentricidade de cada nó do grafo social. A excentricidade é definida como a distância máxima entre um nó e outro nó qualquer do grafo. *Raio* é definido como o valor mínimo de todas as excentricidades, enquanto que o *Diâmetro* é o valor máximo. O *Caminho Médio* é a média de todos os pares de caminhos do grafo social. Conforme apresentam Mislove et al. [2007], Wilson et al. [2009] e Magno et al. [2012], os

²Essa figura foi obtida do endereço <http://www.orgnet.com/email.html> em 20/11/2012. As cores dos nós representam comunidades que os usuários participam. Os usuários de uma mesma cor fazem parte de uma mesma comunidade.

Tabela 2.1: *Medidas do Grafo Social do Orkut, Flickr, Facebook, Twitter e Google+*

Rede Social Online	Raio	Diâmetro	Caminho Médio	Coeficiente de Agrupamento	Assortatividade
Orkut [Mislove et al., 2007]	6	9	4.25	0.171	0.072
Flickr [Mislove et al., 2007]	13	27	5.67	0.313	0.202
Facebook [Wilson et al., 2009]	9.8	13.4	4.8	0.164	0.17
Twitter [Magno et al., 2012]	-	18	4.1	$<\{\text{Fb}, \text{G}+\}^a$	-
Google+ [Magno et al., 2012]	-	19	5.9	$>\{\text{Fb}\}^b$	-

^a Magno et al. [2012] mencionam que o Coeficiente de Agrupamento do Twitter é menor que os coeficientes do Facebook e Google+ e não apresentam o valor exato.

^b Magno et al. [2012] mencionam que o Coeficiente de Agrupamento do Google+ é maior que o coeficiente do Facebook e não apresentam o valor exato.

valores do *Raio*, *Diâmetro* e *Caminho Médio* são estimativas baseadas na determinação da excentricidade de 1000 nós aleatórios em cada grafo.

O *Coeficiente de Agrupamento* é uma medida utilizada para determinar se um grafo social está em conformidade com o princípio *small word*. Essa medida é definida como a razão entre o número de ligações da vizinhança imediata de um nó e o número máximo de ligações que podem existir no grafo. Ela é calculada sob um grafo não direcionado.

Intuitivamente, um coeficiente de agrupamento alto significa que os nós tendem a formar agrupamentos fortemente conectados com seus vizinhos imediatos. O coeficiente de agrupamento para um grafo social é a média de todos os coeficientes de agrupamento dos nós individuais, portanto, coeficiente de agrupamento do grafo social varia de acordo com o grau social de um nó (número de amigos de um usuário).

Na Tabela 2.1 são apresentados os valores dessas medidas para o grafo social das Redes Sociais Online Orkut, Flickr, Facebook, Twitter e Google+, conforme alguns trabalhos reportados na literatura.

Redes Sociais Online que apresentem *Coeficiente de Agrupamento* entre 0.133 e 0.211 apresentam níveis elevados de agrupamento local, portanto, os nós das bordas do grafo social dessas redes também estão fortemente agrupados. Este fato, combinado com *Diâmetro* e *Caminho Médio* relativamente baixos indicam que um grafo social tem propriedades *small word*, conforme apresentam Mislove et al. [2007].

A *Assortatividade* de um grafo mede a capacidade de que nós de um grafo se liguem a outros nós de grau social similar. É calculada como o coeficiente de correlação de Pearson dos graus sociais de pares de nós para todas as arestas em um grafo, e retorna os resultados no intervalo de -1 e 1 . *Assortatividade* maior que zero indica que os nós tendem a se conectar com outros nós de grau semelhante,

enquanto *Assortatidade* menor que zero indica que os nós se conectam a outros nós de graus diferentes.

Conforme argumentam Wilson et al. [2009], os valores da *Assortatidade* de grandes Redes Sociais Online se assemelham, como pode ser observado na Tabela 2.1 para redes como Facebook e Flickr. Esses resultados indicam a presença de nós fortemente agrupados tanto no núcleo do grafo social quanto nas bordas do grafo.

2.3 Caracterização do Grafo de Interação de Redes Sociais Online

As ligações entre usuários que são representadas em um grafo social não são indicativas de que os usuários conectados na rede interajam entre si, e nem de como esses usuários mantêm suas atividades e evoluem nas suas interações. Além disso, a distribuição das interações dos usuários está fortemente orientada para uma fração dos amigos de cada usuário, ou seja, os usuários interagem efetivamente apenas com uma fração dos seus amigos na rede. Isso implica que as ligações sociais e os grafos sociais que se formam não são indicadores precisos da interação entre os usuários.

Com o propósito de representar com mais precisão as relações sociais entre os usuários, levando em conta as interações reais dos usuários de uma Rede Social Online, é modelado um grafo denominado de *grafo de interação* [Wilson et al., 2009] e [Wilson et al., 2012] ou também *grafo da atividade dos usuários* [Nazir et al., 2012].

As interações entre usuários representadas em um grafo de interação referem-se a atividades específicas que os usuários executam na Rede Social Online. Essas atividades são eventos registrados na rede (com uma fonte, um destino, e um *timestamp*) gerados por usuários, sem qualquer ambiguidade. Por exemplo, na modelagem de um grafo de interação para a rede Facebook, Wilson et al. [2009] e Wilson et al. [2012] consideram *publicação de mensagens* no Mural e *comentários de fotos* como atividades entre usuários. No trabalho de Nazir et al. [2012], o evento de interação considerado é a troca de presentes entre usuários da rede Facebook.

Um grafo de interação pode ser direcionado ou não direcionado. Wilson et al. [2009] e Wilson et al. [2012] argumentam que um grafo de interação não direcionado permite compará-lo diretamente com o grafo social de uma Rede Social Online usando o mesmo conjunto de métricas. Em seus trabalhos, eles apresentam uma definição formal para o grafo de interação não direcionado. Um grafo de interação não direcionado é definido como um $G'(n, t) = (V, I)$. Um grafo social $G = (V, E)$ e um grafo de interação G' compartilham o mesmo conjunto de

vértices V . No entanto, G usa conjunto de arestas E (as ligações sociais entre os usuários), enquanto G' usa conjunto de arestas I (as interações entre usuários) onde $I \subseteq E$. Embora I seja um conjunto múltiplo, G' não é um multigrafo: arestas duplicadas são simplesmente filtradas [Wilson et al., 2009].

Um grafo de interação não direcionado é parametrizado por duas constantes n e t . Estas constantes filtram as arestas do conjunto de interações I , n define um número mínimo de eventos de interação (atividades dos usuários) para a admissão de uma aresta $i_{u,v}$, tal que $|i_{u,v}| \geq n$. Por exemplo, considerando $n = 2$ para os eventos de interação *publicações de mensagens* e *comentários de fotos*, então uma aresta entre os usuários u e v só existe no grafo de interação se houver duas ou mais *publicações de mensagens* e *comentários de fotos* entre u e v . Um grau de interação entre usuários é o número de nós adjacentes a u em G' . A métrica equivalente em G é o grau de um nó ou $\deg(u)$. A constante t estipula uma janela de tempo durante a qual as interações devem ocorrer. Tomados em conjuntamente, n e t delineiam um limite de taxa de interação. Intuitivamente, um grafo de interação é o subconjunto do gráfico social onde a taxa de interação é maior do que ou igual à taxa estipulada por n e t .

Conforme apresentam Wilson et al. [2012], para modelar os eventos de interação como um grafo de interação não direcionado, é preciso primeiro demonstrar que conjuntos de pares de amigos sociais executam interações recíprocas um com o outro. Intuitivamente, isto significa que se a escreve uma mensagem para b , b irá responder a mensagem de a , satisfazendo assim as condições para uma ligação não direcionada (reciprocidade de interação). Em algumas Redes Sociais Online, a avaliação da reciprocidade das interações entre usuários é um desafio. Por exemplo, a rede Facebook registra apenas eventos de entrada para um usuário específico, ou seja, o evento a escreve mensagem para b só é registrado no Mural de b , não no Mural de a .

Uma suposição implícita subjacente à formulação de grafos de interação não direcionado é que a maioria dos eventos de interação entre usuários ocorrem por meio de ligações sociais, isto é, $I \subseteq E$. Por exemplo, a rede Facebook só permite que amigos possam comentar fotos publicadas no Mural, no entanto, outras redes sociais não possuem essa restrição. Nesse caso, Wilson et al. [2009] e Wilson et al. [2012] argumentam que pode ser benéfico não definir grafos de interação como um subconjunto do grafo social mas, em vez disso, definir um grafo totalmente novo baseado unicamente em dados de interação.

Nazir et al. [2012] argumentam que as interações entre os usuários podem não ser recíprocas e, com base nesse argumento, fazem uso de um grafo de interação direcionado. No entanto, uma definição formal para a construção de um grafo de interação direcionado não é detalhada em seu trabalho. O grafo de

Tabela 2.2: *Redes Sociais Online Populares e Mídias*

	texto	foto	vídeo	música	links	checkin	swf	question
Facebook	x	x	x	x	x	x	x	x
Twitter	x	x	x		x			
Google+	x	x	x		x			
YouTube			x					
Digg			x					
Flickr		x						
FourSquare						x		
Instagram		x						
Soundcloud								
Myspace		x	x	x	x			
Orkut	x	x	x		x			
Last.fm				x				

interação direcionado proposto por Nazir et al. [2012] considera um evento de interação que não está restrito ao tipo de ligação que os usuários possuem na rede. O envio de um presente virtual pode ser realizado entre usuários que não são amigos no Facebook. De acordo com os argumentos apresentados em seu trabalho, Nazir et al. [2012] afirmam que o grafo de interação direcionado tem variação temporal nas suas propriedades estruturais diferentes do grafo social da rede.

2.4 Tipos de Interações entre Usuários de Redes Sociais Online

As atividades dos usuários de Redes Sociais Online possuem características comuns que permitem identificar formas comuns de interações entre usuários (interações sociais). Na modelagem baseada em grafos, as diferentes formas de interações sociais são tratadas implicitamente a partir da identificação dos eventos de interação. A identificação das formas de interações sociais é importante para a indicação precisa de quais são as ligações ativas. A seguir, são apresentados exemplos de tipos de interações sociais que ocorrem nas Redes Sociais Online.

2.4.1 Interações Sociais baseadas em Mídias

As Redes Sociais Online estão voltadas para o compartilhamento de conteúdos baseados em diferentes tipos de mídias. Algumas redes permitem o compartilhamento de um tipo específico de mídia, por exemplo, YouTube permite o compartilhamento de vídeos, Soundcloud permite o compartilhamento de músicas, Foursquare permite o compartilhamento da marcação de presença em locais específicos (*checkin*). Outras OSNs permitem o compartilhamento de mais de um

Figura 2.2: Interação Social baseada em Texto nas redes A) Facebook e B) Twitter

tipo de mídia. Por exemplo, Facebook permite que os usuários compartilhem vídeos, fotos, texto, músicas, *checkin*, links; Google+ permite que os usuários compartilhem fotos, vídeos e links.

Na Tabela 2.2 são apresentadas algumas redes sociais populares e o tipo de mídia que pode ser compartilhado para se iniciar uma interação social baseada em mídia. Dentro do escopo desta tese, *interações sociais baseadas em mídias* são interações entre usuários de Redes Sociais Online que ocorrem a partir do compartilhamento de diferentes tipos de mídia.

Os tipos comuns de interações sociais baseadas em mídias são:

- Interações Sociais baseadas em Texto

São interações sociais que ocorrem por meio do compartilhamento de texto. Elas podem ocorrer na forma de comunicação *um-para-um*, como mensagens privadas em um bate-papo, ou na forma *um-para-muitos*, como mensagens enviadas para grupos específicos ou todos os usuários ligados a um determinado usuário.

Outra forma de interação textual ocorre por meio do compartilhamento de mensagens curtas que contêm um indicativo do status do individuo ou que contêm informações pessoais. Mensagens curtas tornaram-se populares com a rede Twitter [Savage, 2011] e também podem ser usadas no nas redes Facebook e Google+.

Figura 2.3: *Interação Social baseada em Foto nas redes A) Facebook e B) Instagram*

Na Figura 2.2 são apresentados exemplos de interação social baseada em texto originada por meio do compartilhamento de mensagem curta. Na Figura 2.2 A é apresentado um exemplo de interação social baseada em texto na rede Facebook e na Figura 2.2 B é apresentado um exemplo de interação social baseada em texto na rede Twitter. Os nomes e as imagens pessoais dos usuários envolvidos nas interações são omitidos. É possível observar que a partir do compartilhamento de uma mensagem curta, algumas ações podem ser executadas pelos usuários da rede, por exemplo, na rede Facebook podem ser executadas as ações curtir, comentar e compartilhar.

- **Interações Sociais baseadas em Fotos**

São interações sociais que ocorrem a partir do compartilhamento de fotos. Por exemplo, uma interação baseada em foto ocorre quando um usuário é marcado (ou anotado) na foto de um amigo. Sistemas que permitem que os usuários sejam marcados, como nas redes Facebook [Rabbath et al., 2012], Flickr [Valafar et al., 2009], dedicam uma seção especial no álbum de cada usuário para as fotos nas quais ele foi marcado. Além disso, o usuário marcado pode desfazer a ação, se assim lhe for conveniente.

Na Figura 2.3 A é apresentado um exemplo de interação social baseada em foto da rede Facebook e na Figura 2.3 B é apresentado um exemplo

Figura 2.4: *Interação Social baseada em Vídeo nas redes A) Facebook e B) Youtube*

de interação social baseada em foto do aplicativo Instagram. Os nomes e as imagens pessoais dos usuários envolvidos nas interações são omitidos. Pode ser observado que a partir do compartilhamento de uma foto na rede, algumas ações podem ser executadas pelos usuários. Por exemplo, na rede Facebook, os usuários podem executar as ações curtir, comentar e compartilhar após o compartilhamento de uma foto.

- Interações Sociais baseadas em Vídeo

São interações sociais que ocorrem a partir do compartilhamento de vídeos. As redes sociais que possibilitam esse tipo de interação entre usuários, além de permitir o compartilhamento de vídeos, também oferecem funcionalidades baseadas em vídeo como alternativa ao texto. Como exemplo, o recurso de resposta de vídeo da rede YouTube permite que os usuários interajam e conversem por meio de vídeo, por meio da criação de uma sequência que começa com a abertura de um vídeo, e é seguida de respostas em vídeo de outros usuários [Benevenuto et al., 2009a].

Na Figura 2.4 A é apresentado um exemplo de interação social baseada em vídeo na rede Facebook e na Figura 2.4 B é apresentado um exemplo de interação social baseada em vídeo na rede Youtube. Os nomes e as imagens pessoais dos usuários envolvidos nas interações são omitidos. Pode ser observado que em uma interação baseada em vídeo na rede Facebook, por exemplo, os usuários podem executar ações curtir, comentar e compartilhar.

Figura 2.5: *Interação Social baseada em Música nas redes A) Facebook e B) SoundCloud*

- Interações Sociais baseadas em Música

São interações sociais que ocorrem a partir do compartilhamento de músicas. Em uma interação baseada em música, os usuários podem adicionar comentários, marcar como favorita ou ainda baixar o respectivo arquivo de áudio.

Por exemplo, a rede Soundcloud permite que seus usuários adicionem comentários clicando em qualquer ponto da interface de execução da música. Esse comentário se torna uma marcação naquele momento da música e outros usuários da rede podem visualizar o comentário [Bogdanov et al., 2011]. A rede Soundcloud ainda permite que a música seja compartilhada em outras redes como Facebook, Google+ e Twitter.

Na Figura 2.5 A é apresentado um exemplo de interação social baseada em musica da rede Facebook e na Figura 2.5 B é apresentado um exemplo de interação social baseada em musica da rede SoundCloud. Os nomes e as imagens pessoais dos usuários envolvidos nas interações são omitidos. Pode ser observado que em uma interação social baseada em música na rede Facebook, por exemplo, os usuários podem executar as ações curtir e comentar após o compartilhamento de uma música.

- Interações Sociais baseadas em Links

São interações sociais que ocorrem a partir do compartilhamento de links Web. Alguns sites disponibilizam seus conteúdos para as Redes Sociais Online por meio de links que podem ser acessados por um ícone específico. Quando um site não tem esse recurso, o usuário pode copiar o endereço disponível no browser e colar esse endereço no campo de compartilhamento na rede.

Figura 2.6: *Interação Social baseada em Link nas redes A) Facebook e B) Google+*

Figura 2.7: *Interação Social baseada em Checkin nas redes A) Facebook e B) Foursquare*

Por exemplo, um usuário pode compartilhar um link na rede Twitter no corpo de um tweet. Esse link tem seu tamanho reduzido e ele é incorporado a mensagem de 140 caracteres [Park and Chung, 2012]. Redes como Facebook e Google+ também permitem interações baseadas em links.

Na Figura 2.6 A é apresentado um exemplo de interação social baseada em link da rede Facebook e na Figura 2.6 B é apresentado um exemplo de interação social baseada em musica da rede Google+. Os nomes e as imagens pessoais dos usuários envolvidos nas interações são omitidos. Pode ser observado que em uma interação social baseada em link na rede Facebook, por exemplo, os usuários podem executar as ações curtir e comentar após o compartilhamento de um link.

- Interações Sociais baseadas em *Checkin*

São interações sociais que ocorrem a partir do compartilhamento de uma

marcação de presença em uma localidade ou em um lugar (*checkin*). Quando os usuários visitam lugares e registram presença usando aplicativos instalados em dispositivos móveis, essa marcação indica que eles estiveram nesse lugar.

Por exemplo, na rede Foursquare os usuários voluntariamente podem fazer *checkin* em lugares que visitam usando um dispositivo móvel [Colombo et al., 2012]. Na rede Facebook também disponibiliza essa funcionalidade na sua interface para dispositivos móveis. Tanto na rede Foursquare quanto na rede Facebook, após o compartilhamento de um *checkin*, os usuários podem executar ações como *Curtir* e *Comentar*.

Na Figura 2.7 A é apresentado um exemplo de interação social baseada em musica da rede Facebook e na Figura 2.7 B é apresentado um exemplo de interação social baseada em musica da rede SoundCloud. Os nomes e as imagens pessoais dos usuários envolvidos nas interações são omitidos. Pode ser observado que após o compartilhamento de um *checkin* na rede Facebook, por exemplo, os usuários podem executar as ações curtir, comentar e compartilhar.

Além das interações sociais baseadas nas mídias mencionadas, a rede Facebook ainda oferece interações baseadas em mídias do tipo *swf* e *question*. As interações baseadas em mídias *swf* são interações que ocorrem a partir do compartilhamento de animações em linguagem *Flash*. As interações baseadas em mídias *question* são interações que ocorrem a partir do compartilhamento de uma pergunta ou de uma enquete com respostas pré-definidas.

Dependendo de algumas restrições, os usuários de redes sociais podem interagir uns com os outros por meio da execução de ações definidas na rede. Por exemplo, na rede Facebook os assinantes de um usuário podem não ter permissão para *Curtir* ou *Comentar* uma publicação. A interface da plataforma utilizada para acesso aos recursos da Rede Social Online também pode impor restrições.

Por exemplo, na interface Web da rede Twitter, os usuários podem ver uma conversa ou conteúdo multimídia contido na mensagem, e têm permissão para responder, repassar um tweet ou ainda marcar uma mensagem curta como favorito. Já na interface para dispositivos móveis, a rede Twitter permite que o usuário possa ver um conteúdo multimídia contido em um link Web, responder, repassar um tweet, citar, marcar como favorito, ou ainda compartilhar uma mensagem curta para outras redes.

Figura 2.8: Aplicativos de Terceiros instalados nas plataformas A) Apple e B) Android

2.4.2 Interações Sociais baseadas em Aplicativos de Terceiros

Aplicativos de terceiros (em inglês, *third-party applications*) podem oferecer variadas formas de interação entre usuários, que podem variar de acordo com o aplicativo. Esses aplicativos geralmente estão disponíveis em plataformas Web e móveis. No escopo desta tese, as interações entre usuários de Redes Sociais Online que ocorrem a partir da utilização aplicativos de terceiros são denominadas *interações sociais baseadas em aplicativos de terceiros*.

Na Figura 2.8 são apresentados ícones de aplicativos de terceiros instalados em plataformas móveis. Alguns desses aplicativos permitem a interação dos usuários nas Redes Sociais Online. Por exemplo, na Figura 2.8 A são apresentados ícones dos aplicativos das redes SoundCloud, Instagram, Youtube e Facebook, que estão instalados na plataforma *Apple* (em um *iPhone*). Na Figura 2.8 B são apresentados ícones dos aplicativos das redes Twitter, SoundCloud, Instagram, Youtube, Vimeo e Facebook que estão instalados na plataforma *Android*.

Redes Sociais Online como Facebook, Google+, Youtube e Twitter permitem que seus usuários possam interagir usando aplicativos específicos para cada rede [Nazir et al., 2009]. Por exemplo, usuários do aplicativo Youtube podem compartilhar vídeos para as redes Facebook e Twitter. Usuários do aplicativo SoundCloud podem compartilhar músicas para as redes Facebook e Twitter.

Algumas redes sociais notificam os usuários sobre os aplicativos que foram instalados e usados por seus amigos. Alguns aplicativos solicitam aos seus usuários que convidem seus amigos para instalá-los, provocando comunicação direta entre usuários [Wang, 2012]. Por exemplo, quando um usuário da rede Facebook instala o aplicativo Instagram, os amigos desse usuário são notificados da instalação. O usuário que instalou o aplicativo Instagram pode convidar amigos da rede Facebook que não possuem a ferramenta para instalá-la.

2.4.3 Outros Tipos de Interações Sociais

Outros dois tipos de interações presentes em Redes Sociais Online são citadas por Benevenuto [2010]: interações sociais oportunistas e interações sociais silenciosas ou latentes. Benevenuto et al. [Benevenuto et al., 2009a] definem interação oportunista como qualquer tipo de comunicação não solicitada ou não desejada. Exemplos de interações oportunistas são os anúncios ou propagandas em vídeos disponíveis no Youtube. Diferentemente das respostas textuais ou comentários, é preciso que o usuário veja o anúncio ou execute a ação de pular o anúncio para que o vídeo que o usuário deseja possa ser visto.

Conforme apresenta Jiang et al. [2010], interações silenciosas ou latentes são ações em que apenas uma das partes está ativamente envolvida, por exemplo, a navegação em um perfil ou a navegação em um álbum de fotos. Geralmente, essas ações não são registradas na Rede Social³.

Dentro do escopo desta tese, interações sociais oportunistas e interações sociais silenciosas ou latentes não são tratadas e podem ser objeto de análise em trabalhos futuros.

2.5 Trabalhos Relacionados

Nesta seção são apresentados outros trabalhos relacionados com a análise de propriedades estruturais de Redes Sociais Online, interações entre usuários baseadas em mídias e baseadas no uso de aplicativos de terceiros.

2.5.1 Propriedades Estruturais das Redes Sociais Online

Seguindo a modelagem de Redes Sociais como grafos, Mislove et al. [2007] estudaram as propriedades das redes Orkut, Flickr, YouTube e LiveJournal, e confirmaram a existência de propriedades *small world* e *power law* nessas redes sociais. Complementarmente, Ahn et al. [2007] estudaram as propriedades de rede da rede Cyworld, uma OSN popular na Coreia do Sul. Eles mostraram que Cyworld tem propriedades *small world* e *power law*, e ainda, grande coeficiente de agrupamento.

³A rede Orkut registra visitas ao perfil do usuário, sem detalhamento de quais campos do perfil foram lidos [Benevenuto et al., 2009b].

Wang and Chin [2010] compararam redes sociais formadas por meio do uso de smartphones e conexões sem fio com Redes Sociais Online. Eles verificaram que as redes sociais móveis tem propriedades *small world* e *power law*, e ainda, um coeficiente de agrupamento inferior ao observado nas Redes Sociais Online.

Moradi et al. [2012] investigaram as propriedades de redes sociais em uma rede de grande escala gerada pelo tráfego de e-mails, apontando características que distinguem o tráfego de spams do tráfego regular esperado. Eles confirmaram que o tráfego de e-mail regular gera uma rede com propriedades *small world* e *power law*, e seus resultados revelam que as anomalias nas propriedades estruturais das redes de e-mail se devem ao comportamento antissocial de spams.

Magno et al. [2012] estudaram as propriedades estruturais da rede Google+ e apresentam uma descrição abrangente da plataforma, destacando as principais diferenças dessa rede com outras redes sociais populares.

Com base na representação de redes sociais como um grafo Wilson et al. [2009] estudaram a interação entre os usuários da rede Facebook mostrando que, em uma Rede Social Online, o grafo de interação dos usuários é um subconjunto do grafo de relações sociais.

Fazendo uma análise estrutural da rede Facebook por meio do uso de uma série de medidas topológicas, Viswanath et al. [2009] estudaram a evolução das atividades com os usuários do Facebook para capturar, ao longo tempo, laços sociais que podem se desenvolver como mais fortes ou mais fracos.

Kivran-Swaine and Naaman [2011] analisaram a associação de propriedades do Twitter, como tamanho e densidade, com expressão de emoção dos usuários. Seus resultados sugerem que a expressão da emoção pode explicar parte da variação da estrutura da rede, e que o uso da emoção nas interações entre os usuários é um forte fator para explicação dessa variação.

Silva et al. [2012] estudaram a correlação entre conjuntos de atributos e a ocorrência de subgrafos densos em grandes grafos que representam redes sociais. Eles apresentaram um algoritmo eficiente que combina as estratégias de busca e poda na identificação dos padrões mais relevantes de correlação estrutural desses grandes grafos.

Os estudos identificados nesta seção examinam propriedades de redes *small world* e *power law* em Redes Sociais Online, e ainda, verificam a predominância de componentes fracamente ligados. Outros trabalhos que tratam da força das ligações em Redes Sociais Online são apresentados por [Amaral et al., 2000] e por Gilbert and Karahalios [2009].

2.5.2 Interações Sociais Baseadas em Mídias

Os esforços pesquisa que estudam os tipos de interações sociais entre usuários de Redes Sociais Online são numerosos, sendo as interações baseadas em texto, foto e vídeo as mais investigadas.

Interações Sociais Baseadas em Texto

Na análise de interações baseadas em texto, De Choudhury et al. [2009] investigaram comentários de vídeos na rede Youtube para determinar o que torna uma conversa na Rede Social interessante. O método que eles desenvolveram maximiza a informação mútua, e é significativamente melhor do que outros métodos usados com esse propósito.

Burke et al. [2010] investigaram o papel da interação entre amigos da rede Facebook e o consumo de conteúdos produzidos pelos usuários, incluindo atualizações de status e conversas entre os usuários. Eles identificaram ligações entre a percepção do capital social e o sentimento de solidão dos usuários têm implicações para o design de interfaces darem suporte ao bem-estar dos usuários.

Comarella et al. [2012] estudaram a importância das mensagens na rede Twitter, apresentando evidências de que o excesso de informação está presente: os usuários muitas vezes leem centenas de mensagens antes de identificar aquelas que eles têm interesse em responder ou encaminhar. Eles exploraram métodos de aprendizado de máquina para reordenar tweets e seus resultados mostraram que é essa reordenação aumenta a fração de mensagens respondidas ou encaminhadas.

Interações Sociais Baseadas em Fotos

Interações sociais baseadas em fotos são estudadas predominantemente a partir de dados colhidos das redes Flickr e Facebook. Malinen [2010] explora o papel da interação social na rede Flickr, e os resultados mostram que essa rede é mais do que um álbum de fotos online, pois tem muitas funções sociais importantes. Essa rede é usada para registrar a vida de seus usuários, como uma forma de divulgar habilidades fotográficas, bem como o reconhecimento de usuários que disponibilizam conteúdo de boa qualidade. Além disso, vários comportamentos sociais para a construção e a manutenção de uma comunidade podem ser identificados.

Rabbath et al. [2010] analisaram tags associadas a fotos, a interação entre usuários com base em fotos, e informações semânticas exploradas para a seleção de fotos para um álbum. Em um trabalho seguinte, Rabbath et al. [2012] exploraram recursos da rede Facebook que ligam fotos relacionadas a um mesmo

evento. Por meio da análise de relações de amizades entre usuários e também de características visuais, marcações de áreas e estrutura de fotos, eles definiram uma abordagem probabilística semissupervisionada que dá suporte a detecção de eventos baseados em fotos.

Recentemente, Pesce et al. [2012] demonstraram quantitativamente como o simples ato de marcar fotos na rede Facebook pode revelar atributos particulares do usuário. Seus resultados sugerem que as marcações de fotos podem ser utilizadas para ajudar a previsão de alguns dos atributos por eles investigados.

Interações Sociais Baseadas em Vídeo

As interações sociais baseadas em vídeo são estudadas principalmente como resposta a publicação de vídeos em redes como Youtube e Digg. Benevenuto et al. [2009a] caracterizaram interações sociais que permitem que usuários da rede YouTube possam discutir temas e fornecer opiniões sobre produtos ou lugares. Essa caracterização tem duas perspectivas: a resposta da visualização do vídeo e a visualização da interação na rede.

Susarla et al. [2012] encontraram que as interações sociais são influentes não só para determinar quais vídeos da rede Youtube podem se tornar populares, mas também sobre a magnitude desse impacto. Eles também encontraram evidências de uma série de mecanismos pelos quais a influência social é transmitida.

Doerr et al. [2012] apresentaram resultados de um estudo de vários anos da rede Digg.com, observando a importância dos amigos e da rede de amigos na propagação de informações. Seus resultados indicam a presença de fatores anteriormente desconsiderados, como o alinhamento temporal entre as atividades do usuário e a existência de outras relações lógicas além da topologia do grafo social, que são capazes de conduzir e orientar a dinâmica da rede.

Interações Sociais Baseadas em Músicas, Links e *Checkin*

Alguns trabalhos têm analisado as interações baseadas no compartilhamento de músicas, links e *checkin*. Mechant and Evens [2011] realizaram entrevistas face-a-face com usuários da rede Last.fm e observaram seus perfis nessa rede. Os resultados por eles apresentados sugerem que a maioria dos usuários interagem mais entre em si do que com as músicas, entretanto, as músicas compartilhadas são o principal elemento para a identificação de afinidades entre os usuários e o estabelecimento de comunidades de usuários com gostos musicais semelhantes.

Park and Chung [2012] coletaram dados das redes Twitter e Groupon para compreender as características dos tweets gerados pelo usuário que contêm links URL para promoções diárias. Eles descobriram as demandas dos usuários refletidas através de tweets e caracterizaram padrões de comportamento

dos clientes no compartilhamento de informação na rede Twitter, e ainda forneceram evidências de que o compartilhamento de ofertas diárias na rede Twitter contribui para a melhoria do desempenho de vendas na rede Groupon.

Vasconcelos et al. [2012] analisaram como os usuários da rede Foursquare exploram características da rede, revelando perfis de comportamento diferentes. Esse estudo revela a existência de usuários muito ativos e influentes, alguns dos quais são empresas e marcas famosas, que parecem empenhados em publicar dicas em uma grande variedade de locais, recebendo uma grande quantidade de comentários dos usuários. Eles também fornecem evidências da ocorrência de spam na rede, mostrando a existência de usuários que dão dicas que não estão relacionadas com a natureza ou com o domínio tratados nas mensagens trocadas.

2.5.3 Interações Sociais Baseadas no Uso de Aplicativos de Terceiros

O estudo das interações entre usuários que utilizam aplicativos de terceiros é foco de alguns trabalhos. Gjoka et al. [2008] investigaram o uso e a popularidade desse tipo de aplicativo na rede Facebook. Eles mostraram que, apesar do número de usuários desses aplicativos aumentarem com o tempo, o nível de atividade dos usuários que utilizam esses aplicativos diminui.

Nazir et al. [2008] analisaram interações baseadas no uso de aplicações para o Facebook. Ao analisar o grafo de interação, eles encontraram agrupamento elevado dos nós, e também a indicação de formação de comunidades online a partir das interações analisadas. Eles descobriram que uma pequena fração dos usuários é responsável pela maioria das atividades originadas pelo uso dos aplicativos de Facebook, e ainda que um pequeno número de aplicações é usada pela maioria dos usuários.

Em um trabalho posterior, Nazir et al. [2012] apresentam um modelo de crescimento probabilístico para o grafo que representa a atividade dos usuários em Redes Sociais Online a partir do uso de aplicativos sociais.

Egele et al. [2012] apresentaram uma extensão para a rede Facebook que permite com que usuários possam ter controle refinado sobre o acesso que aplicativos podem ter aos dados do seu perfil. Anthonysamy et al. [2012] também investigaram as requisições para dados privados de aplicativos da rede Facebook. Eles elaboraram uma abordagem de gestão privacidade, útil para determinar que tipo da aplicação possa ser utilizado, chamando a atenção para questões de privacidade de dados resultantes da utilização desses aplicativos.

2.6 Considerações Finais

Neste capítulo foram apresentados conceitos fundamentais da Análise de Redes Sociais, que fazem uso do modelo baseado em grafos e é utilizado na análise de interações sociais em Redes Sociais Online. Também foram apresentados os tipos de interações sociais que emergem em Redes Sociais Online, as quais não são explicitamente tratadas no modelo baseado em grafos. Trabalhos relacionados com a análise da estrutura da rede e com os tipos de interações sociais em Redes Sociais Online também são apresentados.

A partir da análise dos trabalhos reportados neste capítulo, pode ser observado que a literatura não reporta trabalhos que tratam da representação explícita dos elementos relacionados com a atividade dos usuários de Redes Sociais Online. A pesquisa apresentada nesta tese trata de uma técnica e um método para a representação, extração e avaliação de interações sociais em Redes Sociais Online — representação essa que trata explicitamente os elementos associados com as atividades dos usuários.

No próximo capítulo é apresentada a técnica proposta, bem como resultados obtidos a partir da sua experimentação. Os primeiros resultados foram obtidos da aplicação da técnica com regras (associadas a interações sociais) identificadas manualmente, portanto, sem a utilização da etapa de extração automática de regras. Em um passo seguinte, são realizadas experimentações com a técnica incluindo a extração automática de interações sociais.

Uma Técnica para Representação, Extração e Avaliação de Interações entre Usuários

A análise das interações entre usuários de Redes Sociais Online está geralmente baseada na representação em um grafo das atividades dos usuários. Esse modelo não prioriza a representação explícita dos elementos associados com as atividades dos usuários na rede, e tem o foco na análise de um evento de interação específico.

Resultados de Psicologia Social mostram que as interações sociais podem ser especificadas como *contingências comportamentais*: implicações lógicas associadas a declarações *if-then* que correspondem a observações do que as pessoas fazem ou não em diversas situações. Em mineração de dados, regras *if-then* são implicações lógicas que podem ser extraídas de um conjunto de dados e avaliadas utilizando-se medidas do aprendizado de regras.

Uma técnica para representação, extração e avaliação de interações sociais em Redes Sociais Online é apresentada neste capítulo. A linguagem de Mechner é utilizada para representar contingências comportamentais como implicações lógicas. Explorando a natureza sequencial das atividades dos usuários nas Redes Sociais, um algoritmo clássico para a descoberta de padrões sequenciais foi modificado para a extração de contingências comportamentais. A avaliação dessas contingências comportamentais é realizada a partir do uso de um procedimento típico de mineração de dados para a avaliação de regras.

Este capítulo está organizado da seguinte forma: na Seção 3.1 é apresentada a linguagem de Mechner utilizada para representar contingências comportamen-

tais; o uso da linguagem de Mechner na representação de interações sociais é apresentado na Seção 3.2; a extração automática de interações sociais como contingências comportamentais no formato de regras *if-then* é descrita na Seção 3.3; a avaliação de interações sociais no formato de regras é apresentada na Seção 3.4; resultados da experimentação da técnica são apresentados nas Seções 3.5 e 3.6; trabalhos relacionados são apresentados na Seção 3.7 e, finalmente, na Seção 3.8 são apresentadas as considerações finais.

3.1 A Linguagem de Mechner para Representação de Interações Sociais

No campo de pesquisa da Psicologia Social, contingências comportamentais são definidas como interações cotidianas entre os indivíduos, que correspondem a observações do que as pessoas fazem ou não fazem em uma variedade de situações Skinner [1953]. Por exemplo:

- Uma lei pode ser escrita no formato de uma regra tal como “Se uma pessoa pratica determinados atos, então essa pessoa pode sofrer certas consequências”. Em essência, as leis são contingências comportamentais destinadas para regular, modificar ou influenciar o comportamento de uma sociedade;
- Em sistemas de ensino, contingências comportamentais governam as interações entre alunos, professores, pais, administradores e membros da comunidade;
- As regras de um jogo, por exemplo, *tic-tac-toe* (ou “jogo da velha”), são contingências comportamentais que determinam a forma como os jogadores participam do jogo.

Mechner [1959] apresentou um dos primeiros sistemas de notação para a codificação de contingências comportamentais no qual são combinadas álgebra booleana com um conjunto de diagramas. Weingarten and Mechner [1966] estenderam a obra original de Mechner, representando interações sociais como variáveis independentes no formato de regras *if-then*. Em um trabalho recente, Mechner [2008] apresentou uma linguagem formal simbólica para a codificação de variados tipos de contingências comportamentais que envolvem vários indivíduos. A importância metodológica da linguagem de Mechner para ciências comportamentais é discutida em Mechner [2011].

Na *Linguagem de Mechner*, contingências comportamentais são declarações *if-then* no formato de implicações lógicas, que podem ser avaliadas como variáveis independentes. Alguns elementos importantes dessa linguagem são:

1. *Ação (ou ações)*: corresponde ao antecedente de uma contingência, por exemplo, $A \rightarrow$. Se houver mais de uma ação executada uma após a outra, elas são representadas como $A_1 \cap A_2 \cdots \rightarrow$.
2. *Agente(s) da(s) ação(ões)*: representados por letras minúsculas colocadas na frente de uma ação A . Por exemplo, um agente a de uma ação A realizada é representado por aA . Uma letra minúscula pode representar um único agente ou um grupo de agentes que executam uma ação.
3. *Consequência(s)*: corresponde ao consequente da contingência, por exemplo, $\rightarrow C$. Se houver mais de uma consequência percebida consecutivamente, elas são representadas como $\rightarrow C_1 \cap C_2$.

Contingências comportamentais codificadas na Linguagem de Mechner podem ser descritas como declarações *if-then* tais como:

- $aA_1 \cap bA_2 \rightarrow abC_1$. *If* a executa a ação A_1 e b executa a ação A_2 *then* a consequência C_1 é percebida por a e b .
- $\bar{a}A_1 \cap bA_2 \rightarrow \bar{a}bC_2$. *If* a não executa a ação A_1 e b executa a ação A_2 *then* a consequência C_2 não é percebida por a mas é percebida por b .
- $aA_1 \cap bA_2 \rightarrow aC_1 \cap bC_2$. *If* a executa a ação A_1 e b executa a ação A_2 *then* a consequência C_1 é percebida por a e a consequência C_2 é percebida por b .

A ação que dá início a uma interação social é chamada de *estímulo social* [Skinner, 1953]. Nos exemplos acima o estímulo sociais corresponde a ação A_1 .

Embora tenham sido propostos outros sistemas de notação para codificação de comportamentos em processos de análise experimental (por exemplo, Mat-
taini [1995] e Sidman [2000]), a linguagem de Mechner Mechner [2008] permite a representação de contingências comportamentais como expressões booleanas na forma normal disjuntiva, ou seja, expressões booleanas que possuem conectivos *not*, *and* e *or* ($\neg$, $\cap$ e $\cup$). Essa propriedade lógica é exigida pelo processo de mineração de dados que é adotado na técnica, como detalhado nas próximas seções.

3.2 Representação de Interações Sociais

Na técnica apresentada neste capítulo, a linguagem de Mechner é usada para representar situações que envolvem usuários de Redes Sociais Online em interações sociais. Essa representação é realizada a partir da identificação dos

elementos da linguagem de Mechner, em outras palavras, a partir da identificação de *Ações*, *Usuários* (atores das ações) (por exemplo, o usuário a , ou grupo de usuários k e l), e *Consequências*.

Cada Rede Social tem conjuntos específicos de *Ações*, *Usuários* e *Consequências* que podem ser mapeados para a representação de contingências comportamentais.

Por exemplo, as seguintes ações podem ser mapeadas nas redes Facebook, Google+ e Twitter:

- na rede Facebook:
 - A_1 = criar um *Post* no Mural,
 - A_2 = *Curtir* um *Post*,
 - A_3 = *Comentar* um *Post*,
 - A_4 = *Compartilhar* um *Post*.
- na rede Google+:
 - A_1 = criar um *Post*,
 - A_2 = marcar com *+1* um *Post*,
 - A_3 = *Compartilhar* um *Post*,
 - A_4 = *Comentar* um *Post*.
- na rede Twitter:
 - A_1 = criar um *Tweet*,
 - A_2 = *Responder* um *Tweet*,
 - A_3 = *Retweet* um *Tweet*,
 - A_4 = marcar como *Favorito* um *Tweet*.

Na linguagem de Mechner, usuários de Redes Sociais são *Usuários* que podem executar uma ou mais ações, individualmente (por exemplo, o usuário a ou b) ou em grupos (por exemplo, o grupo k ou l), de acordo com permissões fornecidas pelo sistema. Usuários da rede Facebook¹ podem ser assim representados:

- a = usuário que cria um *Post*,
- l = grupo de usuários que executam *Curtir* naquele *Post*,
- k = grupo de usuários que executam *Comentar* naquele *Post*,

¹Uma representação semelhante também pode ser realizada para usuários das redes Twitter e Google+.

- s = grupo de usuários que executam *Compartilhar* naquele *Post*.

Como resultado da execução de uma ou mais ações, os usuários podem ser notificados de uma ou mais *Consequências*, que são respostas automaticamente geradas pela Rede Social Online (respostas do sistema). Ainda considerando a rede Facebook², podem ser representadas as seguintes consequências:

- C_1 = notificação do *Post* no Mural,
- C_2 = notificação de *Curtir* naquele *Post*,
- C_3 = notificação de *Comentar* naquele *Post*,
- C_4 = notificação de *Compartilhar* naquele *Post*.

Além disso, dependendo da permissão que os usuários têm, eles podem também executar ações a partir do resultado da execução de uma ou mais ações. Por exemplo, o usuário b pode executar a ação de *Curtir* um *Post* (C_2), ou pode executar a ação de *Comentar* um *Post* (C_3), após ser notificado que o usuário a executou a ação de criar um *Post* no Mural (C_1).

Após a identificação de *Ações*, *Usuários* (usuários), e *Consequências* para representação de interações sociais, declarações *if-then* podem ser realizadas. Por exemplo, considerando usuários a e b sendo amigos na rede Facebook:

- *if* o usuário a executa a ação A_1 = criar um *Post* no Mural,
 - *then* os usuários a e b recebem C_1 = notificação do *Post* no Mural,
 - *depois if* o usuário b executa a ação A_2 = *Curtir* aquele *Post* (depois de ser notificado ação do usuário a),
 - * *then* usuário a e b recebem C_2 = notificação de *Curtir* naquele *Post*

usando a linguagem de Mechner, essa interação social pode ser representada pela sequência

$$aA_1 \rightarrow abC_1 \rightarrow bA_2 \rightarrow abC_2,$$

ou ainda, aplicando operações de lógicas para separar as *Ações* das *Consequências*,

$$aA_1 \cap bA_2 \rightarrow abC_1 \cap abC_2.$$

No âmbito da técnica, contingências comportamentais codificadas na linguagem de Mechner são representadas como declarações *if-then* sob o formato *Corpo* $\rightarrow$ *Cabeça*, em inglês, *Body* $\rightarrow$ *Head* (de forma curta, $B \rightarrow H$).

²Uma representação semelhante também pode ser realizada para consequências nas redes Twitter e Google+.

Tabela 3.1: *Ilustração do Resultado da Execução do Algoritmo GSP*

Passo	Sequências	Frequência das Sequências
1	$\langle A1 \rangle, \langle A2 \rangle, \langle A3 \rangle$	$f(\langle A1 \rangle) = 35\%$, $f(\langle A2 \rangle) = 30\%$, $f(\langle A3 \rangle) = 30\%$
2	$\langle A1, A2 \rangle, \langle A1, A3 \rangle,$ $\langle A2, A3 \rangle$	$f(\langle A1, A2 \rangle) = 30\%$, $f(\langle A1, A3 \rangle) = 25\%$, $f(\langle A2, A3 \rangle) = 25\%$
3	$\langle A1, A2, A3 \rangle$	$f(\langle A1, A2, A3 \rangle) = 30\%$

Por exemplo, considerando a contingência comportamental $aA_1 \cap bA_2 \rightarrow abC_1 \cap abC_2$, ou seja, considerando a declaração *if* $aA_1 \cap bA_2$ *then* $abC_1 \cap abC_2$, tem-se $B = aA_1 \cap bA_2$ e $H = abC_1 \cap abC_2$.

O formato $B \rightarrow H$ é usado tanto para extração quanto para avaliação de contingências comportamentais, como será apresentado a seguir.

Na representação de interações sociais usando a linguagem de Mechner, a identificação de um usuário ou um grupo de usuários, a granularidade das Ações e suas respectivas Consequências, são definições realizadas por quem observa uma Rede Social, que pode contar com o apoio de especialistas para identificar interações sociais.

Na forma de declarações *if-then*, contingências comportamentais podem ser usadas na codificação de relações de causalidade (ações levam a consequências). Entretanto, dentro do escopo de aplicação da técnica, pretende-se observar correlações entre as atividades dos usuários da rede, ou seja, atividades realizadas pelos usuários correlacionadas com as respostas do sistema percebidas pelos usuários.

3.3 Extração de Interações Sociais

As situações que envolvem os usuários de Redes Sociais Online em interações sociais são sequencias temporalmente ordenadas de ações e consequências. Por exemplo, *conversações* são sequências de comentários temporalmente ordenados em torno de um objeto de mídia compartilhado, conforme apresentam De Choudhury and Sundaram [2011] e Shamma et al. [2011]. Na técnica proposta, a ordem temporal das ações e consequências é mantida na organização de conjuntos de dados associados com atividades dos usuários de Redes Sociais. A partir dessa organização, padrões sequenciais podem ser extraídos.

Uma abordagem comumente utilizada na extração de padrões sequenciais é conhecida como *força bruta* ou abordagem *Apriori-like* [Tan et al., 2005], como é o caso do algoritmo *GSP* — *Generalized Sequential Pattern* [Agrawal and Srikanth, 1995], que efetua várias passagens (*scans*) por um conjunto de dados para realizar a extração das sequências.

A Tabela 3.1 apresenta um resultado da execução do algoritmos GSP na extração de sequências a partir dos itens $A1$, $A2$ e $A3$. O limiar utilizado para manter uma sequência para o próximo passo é de 27%.

No primeiro passo da execução do algoritmo GSP é realizada uma passagem pelo conjunto de dados, todos os itens individuais (sequências de tamanho 1) são contados e aqueles que possuem frequência maior que um limiar (suporte mínimo) são mantidos para uma extração de sequências de tamanho 2. Com um conjunto de sequências candidatas de tamanho 2 formadas, outra passagem pelo conjunto de dados é realizada para identificar frequência das sequências de tamanho 2. As sequências que têm frequência acima do limiar estipulado são utilizadas para gerar as candidatas de tamanho 3. Esse processo é repetido até todas as sequências que satisfaçam o suporte mínimo sejam encontradas.

Em outras palavras, são extraídas sequências que possuem tamanho mínimo 1 e o tamanho máximo pode corresponder ao número máximo de itens contidos no conjunto de dados. Além disso, o GSP não é capaz de iniciar a geração sequências a partir de outras sequências de tamanho maior que 1. Vale observar que o algoritmo GSP não é capaz de manipular atributos com nulos ou valores ausentes.

Outros algoritmos de extração de sequências apresentam melhores desempenhos do que GSP, sendo que os mais conhecidos são *SPADE* [Zaki, 2001] e *PrefixSpan* [Pei et al., 2004]. Outras abordagens recentes e mais eficientes do GSP são propostas na literatura, concentrando-se principalmente na extração de padrões sequenciais sem *scans* redundantes no conjunto de dados [Cho et al., 2009], [Dong et al., 2011].

Mesmo nas abordagens mais recentes do GSP, os algoritmos executam uma busca global no conjunto de dados, usando uma medida de suporte mínimo como limiar que é empregado na redução do espaço de busca. Além disso, um grande número de sequências é gerado e muitas dessas sequências podem ser inúteis, ou ainda pior, sequências importantes podem ser eliminadas ou simplesmente nunca serem extraídas.

A fim de ultrapassar essas limitações, algumas características particulares podem ser exploradas na tarefa de extração automática de contingências comportamentais como sequências. Por exemplo,

- contingências comportamentais que têm frequência maior que um suporte mínimo são tão interessantes quanto aquelas que têm frequência inferior;
- a identificação de *Ações*, *Usuários* (atores das ações) e *Consequências* pode ser útil para a redução do espaço de busca;
- um número finito de *Ações* e suas respectivas *Consequências* é conhecido

quando contingências comportamentais são codificadas na linguagem de Mechner, ou seja, os tamanhos máximo e mínimo das contingências comportamentais extraídas como sequências são conhecidos *apriori*;

- *Ações* não realizadas pelos usuários não têm *Consequências*. Isso significa dizer que as *Ações* não realizadas pelos usuários têm *Consequências* com valor nulo.

Considerando tais características particulares, uma modificação do algoritmo GSP é proposta nesta tese para extrair contingências comportamentais representadas na linguagem de Mechner. Um passo importante antes do início da extração das contingências comportamentais como sequências é a organização do conjunto de dados. Primeiramente *Ações* e *agente de ações* são organizados como atributos e, em segundo lugar, são organizadas as *Consequências*.

As *Ações* são organizadas na ordem em que elas estão disponíveis na rede. Por exemplo, *Ações* e *Usuários*³ da rede Facebook podem ser organizados como atributos em um conjunto de dados projetado na ordem:

$$\{Post, Curtir, Comentar, Compartilhar\} = \{aA_1, lA_2, kA_3, sA_4\}.$$

onde a é um usuário e l, k e s são conjuntos de um ou mais usuários. As *Consequências* seguem a ordem de organização da ações correspondentes:

$$\{Post, Curtir, Comentar, Compartilhar\} = \{alksC_1, alksC_2, alksC_3, alksC_4\}.$$

Essa preparação de dados tem impacto sobre a redução do espaço de busca, pois o conjunto de dados original é transformado em um conjunto de dados projetado com *Ações*, *agente de ações* e *Consequências*.

O algoritmo 1 foi implementado em *Python*⁴ e utilizado na extração de sequências. Nesta tese esse algoritmo é denominado *GSP-like*. No primeiro passo, cada linha do conjunto de dados com *Ações*, *agente de ações* e *Consequências* (com valores nulos ou faltantes) é computada como uma sequência candidata. São contados o número de contingências comportamentais observadas n (cada linha no conjunto de dados) e a frequência de ocorrência de cada sequência candidata. Se uma sequência extraída não fizer parte da lista de sequências candidatas (linha 3), ela é inserida na lista de sequências candidatas (linha 5). Se a sequência fizer parte, é feita uma busca pela posição da sequência na lista de sequências

³Os usuários representados por a, l, k, s se relacionam como amigos, portanto, eles conseguem perceber as notificações geradas pela Rede Social.

⁴www.python.org

Algoritmo 1 Extração de Contingências Comportamentais como Sequências

Entradas:

DsOb: conjunto de dados com observações (conjunto de dados projetado);

nA: número de *Ações*;nC: número de *Consequências*;**Saídas:***Seq*: lista com sequências candidatas;*Bseq*: lista com Corpos de cada sequência candidata;*Hseq*: lista com Cabeças de cada sequência candidata;*Bcount*: lista de frequências de Corpo das sequencias;*Hcount*: lista de frequências de Cabeça das sequencias;*Count*: lista com frequências das sequências;

n: número de observações;

Descrição:

ob = observação (linha);

 p, t = índices;**Pseudo código Passo 1**

```

1:  $t = -1; n = 0; Seq = \phi; Count = \phi;$ 
2: para cada  $ob \in DsOb$  { // scan do conjunto de dados DsOb
3:   se  $ob \notin Seq$ {
4:      $t++$ ;
5:      $Seq_t = ob$ ;
6:   }
7:   senão{
8:     busque posição  $p$  de  $ob$  em  $Seq$ ;
9:      $Count_p++$ ;
10:  }
11:  $n++$ 
12: }
13: retorne  $Seq, Count, n$ ;
14: Procedure Passo 1(DsOb, nA, nC)

```

Descrição: s = sequencia candidata;**Pseudo código Passo 2**

```

1:  $s = \phi; Seq = \phi; Bcount = \phi; Hcount = \phi;$ 
2: para cada  $s \in Seq$  { // scan a lista de sequencias
3:    $Bseq = s$  posição 0 até  $nA - 1$ ;
4:   remova  $null \in Bseq$ ;
5:    $Hseq = s$  posição  $nA - 1$  até  $nA + nC - 1$ ;
6:   remova  $null \in Hseq$ ;
7:    $Bcount = Hcount = Count$ ;
8: }
9: }
10: retorne  $Bseq, Bcount, Hseq, Hcount$ ;
11: Procedure Passo 2(Seq, Count, nA, nC)

```

candidatas (linha 8) e, logo após, a respectiva frequência é incrementada (linha 9). No final da execução deste passo, todas as sequências com comprimento máximo são geradas, por exemplo:

$$< aA_1, lA_2, kA_3, sA_4, alksC_1, alksC_2, alksC_3, alksC_4 > e$$

$$< aA_1, \bar{l}A_2, \bar{k}A_3, \bar{s}A_4, alksC_1, null, null, null >.$$

Na segunda etapa, cada sequência candidata é percorrida, dividida em Corpo e Cabeça e reescrita como uma implicação. Enquanto cada sequência é dividida, os valores nulos são removidos do Corpo e da Cabeça, por exemplo:

$< aA_1, lA_2, kA_3, sA_4, alksC_1, alksC_2, alksC_3, alksC_4 >$ é percorrida e dividida em

$$aA_1 \cap lA_2 \cap kA_3 \cap sA_4 \rightarrow alksC_1 \cap alksC_2 \cap alksC_3 \cap alksC_4,$$

$< aA_1, \bar{l}A_2, \bar{k}A_3, \bar{s}A_4, alksC_1, null, null, null >$ é percorrida e dividida em

$$aA_1 \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow alksC_1.$$

O passo final é um ajuste de frequências de b e h , ou seja, as frequências de ocorrência de cada Corpo e cada Cabeça são ajustadas. Se um Corpo aparece em mais de uma sequência, a frequência de ocorrência é aumentada; da mesma forma, se uma Cabeça aparece em mais de uma sequência.

Diferentemente do algoritmo GSP tradicional e demais algoritmos com abordagens mais eficientes, o algoritmo apresentado nesta Seção gera sequências com tamanho máximo e, a partir da remoção de valores nulos ou ausentes, gera sequências com outros tamanhos. No final da execução do algoritmo, cada sequência está representada na forma $B \rightarrow H$ e possui os valores bh , b ($bh + b\bar{h}$) e h ($bh + \bar{b}h$) computados. Usando esses valores, as medidas de regra são computadas como explicado a seguir.

3.4 Avaliação de Interações Sociais

Utilizando um procedimento de mineração de dados, uma regra na forma $B \rightarrow H$ pode ser medida por comparação com um conjunto de observações [Lavraç et al., 1999]. A partir da comparação de uma regra com n observações, valores como $bh, b\bar{h}, \bar{b}h, \bar{b}\bar{h}$ podem ser computados de modo que

$$n = bh + b\bar{h} + \bar{b}h + \bar{b}\bar{h}$$

onde

bh número de observações para as quais corpo b e cabeça h são verdadeiras,

$b\bar{h}$ número de observações para as quais corpo b é verdadeiro e cabeça h é falsa,

$\bar{b}h$ número de observações para as quais corpo b é falso e cabeça h é verdadeira,

$\bar{b}\bar{h}$ número de observações para as quais corpo b e cabeça h são falsos.

Usando esses valores, medidas de avaliação de regras podem ser computadas. Essas medidas são usadas, por exemplo, na análise objetiva e subjetiva de regras *if-then* extraídas de classificadores simbólicos [Gomes, 2002] [Gomes and Monard, 2003], ou ainda, na avaliação de regras no formato $B \rightarrow H$ com baixa frequência de ocorrência [Gomes, 2007].

Na técnica proposta, medidas de avaliação de regras são usadas para a avaliação de interações sociais representadas como contingências comportamentais no formato $B \rightarrow H$. Uma variedade de medidas para regras no formato $B \rightarrow H$ são propostas na literatura [Lavrac et al., 1999] [Martínez-Ballesteros and Riquelme, 2011]. Algumas dessas medidas foram usadas na experimentação da técnica:

- *Suporte*

$$SupR = P(BH) \equiv \frac{bh}{n} \quad (3.1)$$

determina o quanto um regra é aplicável a um determinado conjunto de observações, determinando a frequência com que H juntamente com B vão aparecer no conjunto de observações.

- *Confidência*

$$ConR = P(H|B) \equiv \frac{P(BH)}{P(B)} \equiv \frac{bh}{b} \quad (3.2)$$

determina a confiabilidade da inferência que pode ser realizada a partir de uma regra, determinando a frequência com que H aparece em observações que contêm B .

- *Correlação Cosseno*

$$CosR = \frac{P(BH)}{\sqrt{P(B) * P(H)}} \equiv \frac{bh}{n * \sqrt{\frac{b*h}{n^2}}} \quad (3.3)$$

determina a força (ou falta de força) de associação entre B e H . O valor da medida igual a 100% significa que existe uma correlação positiva entre B e H , ou seja, se uma das variáveis aumentar, a outra também aumenta. O valor igual a 0% significa que não há uma correlação entre B e H .

- *Sensibilidade*

$$SenR = P(B|H) \equiv \frac{P(BH)}{P(H)} \equiv \frac{bh}{bh + \bar{b}h} \quad (3.4)$$

determina a frequência com que B aparece em observações que contêm H .

- *Laplace*

$$LapR = \frac{bh + 1}{b + 2} \quad (3.5)$$

determina o equilíbrio entre as medidas de Suporte e Confidência para uma regra com base na estimativa de probabilidade de Laplace [Cestnik, 1990].

- *Utilidade baseada no Suporte*

$$UtiR = wSupR = w * SupR \quad (3.6)$$

determina a utilidade de uma implicação $B \rightarrow H$, por meio da associação de um peso w ao Suporte da regra. Esse peso w pode ser arbitrado tanto por usuários quanto por experimentadores.

- *Alavancagem*

$$AlaR = P(BH) - (P(B) * P(H)) \equiv \frac{bh}{n} - \left(\frac{h}{n} * \frac{b}{n}\right) \quad (3.7)$$

determina a diferença entre o quanto B e H aparecem conjuntamente, em detrimento da frequência de ocorrência de B e de H em separado.

Medidas de avaliação de regras pode ser *simétricas* ou *assimétricas* [Tan et al., 2005]. As medidas *Suporte*, *Correlação Cosseno*, *Utilidade* e *Alavancagem* são medidas *simétricas*, pois seus valores são idênticos tanto para inferências $B \rightarrow H$ quanto $H \rightarrow B$. Em contraste, as medidas *Confidência*, *Sensibilidade* e *Laplace* são *assimétricas*, pois seu valores podem não ser o mesmo para inferências $B \rightarrow H$ e $H \rightarrow B$. Medidas *simétricas* são geralmente utilizadas para avaliar B e H como valores independentes, enquanto que as medidas *assimétricas* permitem a análise da inferência $B \rightarrow H$. Por convenção, essas medidas têm seus valores expressos de 0 até 100%, ao invés de valores de 0 até 1 [Han, 2005].

Nesta tese, as medidas *Confidência*, *Suporte* e *Correlação Cosseno* são utilizadas nas experimentações apresentadas neste capítulo e, além dessas medidas, *Sensibilidade*, *Laplace* e *Alavancagem* são utilizadas nas experimentações apresentadas no Capítulo 4. Nas experimentações do Capítulo 5 são utilizadas as medidas *Confidência*, *Suporte*, *Correlação Cosseno*. Nas experimentações apresentadas no Apêndice A são utilizadas as medidas *Confidência*, *Suporte* e *Utilidade baseada no Suporte*. Experimentações da técnica com extração manual de interações sociais é apresentada na seção que segue.

Tabela 3.2: *Informações sobre as Unidades de Observação: Técnica sem Extração Automática*

Seção	Período de Coleta	# de Usuários	# de Contingências
Seção 3.5.1	Março de 2011	1.183	30.175
Seção 3.5.2	Março de 2011	1.264	38.517
Seção 3.5.3	Março de 2011	1.391	21.935

3.5 Experimentações da Técnica sem Extração Automática

A técnica proposta nesta tese foi elaborada primeiramente com as etapas de representação e avaliação (sem a extração automática de contingências comportamentais) e, em seguida, a etapa de extração foi incorporada. Os resultados da experimentação da técnica sem a etapa de extração automática são apresentados nesta seção.

Com o propósito de coletar dados de interação, foi criado um perfil na rede Facebook. A cada solicitação de amizade atendida, os usuários foram informados do propósito da coleta. Aqueles que não concordavam com o procedimento eram adicionados em outro perfil. Ao longo do tempo de desenvolvimento desta tese, o número de usuários que autorizaram a coleta aumentou de 1.000 para mais de 3.000.

Utilizando essa estratégia, foi construída uma sub-rede do Facebook que pode ser acessada por meio do perfil. Uma sub-rede corresponde a uma unidade de observação da Rede Social e é acessada a partir do perfil de um usuário [Wasserman and Faust, 1994]. Foi implementado um *crawler* usando a linguagem *Python* para a coleta dos dados de interação.

Na Tabela 3.2 são sumarizadas informações sobre as unidades de observação da rede Facebook utilizadas nas experimentações reportadas nesta seção. A partir dos dados coletados, foram representadas e avaliadas contingências comportamentais associadas com interações sociais. A validação das interações sociais extraídas foi realizada por um especialista em Psicologia Social.

3.5.1 Interações Sociais Genéricas no Facebook

Em um primeiro estudo [Gomes and Pimentel, 2011e] foram representadas e computadas situações frequentes que envolvem os usuários da rede Facebook em interações sociais. Foram coletadas informações associadas às interações sociais de mais de 1.000 usuários que são de diferentes localidades. Cerca da metade desses usuários é do Brasil, os demais são de países como Estados Unidos, Canadá, Argentina, Inglaterra, Portugal, Espanha, França, Rússia, Israel, Tailândia, dentre outros. Tabela 3.2 são sumarizadas informações sobre a unidade de observação da rede Facebook utilizada na experimentação reportada nesta seção.

Tabela 3.3: *Ações, Consequências e Usuários da rede Facebook na Experimentação da Técnica*

Ações	
A_1	ação de criar um Post
A_2	ação de Curtir um Post
A_3	ação de Comentar um Post
A_4	ação de Marcar o nome de um amigo no comentário daquele <i>Post</i>
Consequências	
C_1	ser notificado de uma ação de criar um Post
C_2	ser notificado de uma ação de Curtir um Post
C_3	ser notificado de uma ação de Comentar um Post
C_4	ser notificado de uma marcação do nome de um amigo no comentário daquele <i>Post</i>
Usuários	
a	usuário que executam a ação de criar um Post
l	usuários que executam a ação de Curtir
k	usuários que executam a ação de Comentar
$k1$	usuários que executam a ação de Comentar com marcação do nome de um amigo
$\bar{l}$	usuários que não executam a ação de Curtir
$\bar{k}$	usuários que não executam a ação de Comentar
$\bar{k}1$	usuários que não executam a ação de Comentar com marcação do nome de um amigo

Na representação das interações sociais desses usuários são detalhadas apenas as ações dos usuários nas interações sociais. No escopo desta tese, contingências comportamentais associadas a essas interações sociais são denominadas de *interações sociais genéricas*.

Nas interações sociais genéricas reportadas nesse primeiro estudo, os usuários do Facebook foram representados como *Usuários* da seguinte forma:

- O usuário que executa a ação de criar um *Post* é representado como usuário a . Por exemplo, na Figura 2.4 A que apresenta um exemplo de interação social baseada em video na rede Facebook, o usuário *news-of-madonna* é representado como usuário a . Assim que um usuário executa cria um *Post*, esse usuário e seu grupo de amigos recebem uma notificação da postagem.
- O grupo de usuários que, em seguida, executam um *Curtir* nessa postagem são representados por l , e o grupo de usuários que executam um *Comentar* nessa postagem são representados por k . Por exemplo, na Figura 2.4 A, os 25 usuários que executaram um *Curtir* são representados por l e o usuário que executou um *Comentar* é representado por k .

- Ao comentar uma postagem, os usuários podem marcar o nome de um amigo. O grupo de usuários marcados por amigos em um comentário é representado por $k1$. Por exemplo, na Figura 2.4 A o nome de um amigo do usuário que executou o comentário foi omitido. O usuário que executou a marcação do nome do amigo é representado por $k1$.

<p>R1 $aA_1 \cap alA_2 \cap akA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$</p> <p>R2 $aA_1 \cap alA_2 \cap \overline{ak}A_3 \rightarrow alkC_1 \cap alkC_2$</p> <p>R3 $aA_1 \cap \overline{al}A_2 \cap akA_3 \rightarrow alkC_1 \cap alkC_3$</p> <p>R4 $aA_1 \cap akA_3 \cap ak1A_4 \rightarrow alkC_1 \cap alkC_3 \cap alkC_4$</p>
--

Listagem 3.1: *Contingências Comportamentais de Usuários da rede Facebook: Interações Sociais Genéricas [Gomes and Pimentel, 2011e].*

Considerando ações, consequências e usuários (atores das ações) da Tabela 3.3, foi realizada a representação de interações sociais no Facebook, conforme apresentado na Listagem 3.1.

As contingências comportamentais da Listagem 3.1 foram representadas manualmente a partir das Ações, Usuários e Consequências.

- R1 *if* usuário a realiza ação A_1 e os usuários al executam a ação A_2 e os usuários ak executam a ação A_3 (usuário a fornece o estímulo social que recebe *Curtir* e recebe *Comentar*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem as notificações C_1 e C_2 e C_3 .
- R2 *if* usuário a realiza ação A_1 e os usuários al executam a ação A_2 e os usuários ak não executam a ação A_3 (usuário a fornece o estímulo social que recebe apenas *Curtir*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem as notificações C_1 e C_2 .
- R3 *if* usuário a realiza ação A_1 e os usuários al não executam a ação A_2 e os usuários ak executam a ação A_3 (usuário a fornece o estímulo social que recebe apenas *Comentar*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem as notificações C_1 e C_3 .
- R4 *if* usuário a realiza ação A_1 e os usuários ak executam a ação A_3 e os usuários $k1$ executam a ação A_4 (usuário a fornece o estímulo social que recebe *Comentar* e nomes de amigo(s) são marcados no comentário)
then usuário a e grupo de usuários l e k (usuário a e seus amigos) recebem as notificações C_1 e C_3 e C_4 .

Tabela 3.4: *Medidas de Contingências Comportamentais dos Usuários da rede Facebook: Interações Sociais Genéricas [Gomes and Pimentel, 2011e].*

	Confidência	Suporte	Correlação Cosseno
R1	100%	64.28%	80.17%
R2	100%	88.89%	94.28%
R3	100%	75.39%	86.83%
R4	100%	50.97%	71.39%

Um conjunto de dados contendo as ações, consequências e usuários (atores das ações) apresentados na Tabela 3.3 foi construído para a avaliação das contingências apresentadas na Listagem 3.1. Esses dados foram coletados no mês de Março de 2011 a partir das interações sociais de mais de 1.000 usuários. O conjunto de dados obtido foi utilizado na avaliação das interações sociais $R1$ até $R4$ e, para essa avaliação foram computadas as medidas de Confidência (Equação 3.2), Suporte (Equação 3.1) e Correlação Cosseno (Equação 3.3).

O resultado da avaliação das interações sociais genéricas é apresentado na Tabela 3.4. Deve ser observado que o valor da medida de Confidência = 100% significa que a totalidade de observações que tem associação a uma determinada regra foi avaliada. Os resultados do computo das medidas de Suporte e de Correlação Cosseno foram utilizados na construção de uma escala das regras, as quais foram ordenadas do maior para o menor valor de cada medida. A escala obtida é: $R3 > R2 > R1 > R4$.

Por meio dessa escala foi possível observar que, considerando os usuários que tiveram seus dados de interação coletados, é possível observar que esses usuários estão frequentemente envolvidos em interações sociais nas quais um *Post* recebe apenas *Curtir* (interação $R3$). Esse resultado é esperado, tendo em vista que a execução da ação *Curtir* é uma forma de interação simples e não demanda esforço dos usuários como, por exemplo, para a digitação de um texto.

Em outras palavras, considerando os mais de 1.000 usuários que permitiram a coleta de dados, os resultados obtidos mostram que esses usuários então engajados em interações sociais utilizando preferencialmente recursos mais simples de interação.

Esse resultado também permite observar que a técnica pode ser empregada para detectar padrões que refletem comportamentos coletivos esperados.

3.5.2 Interações Sociais Baseadas em Mídia

No estudo reportado no trabalho [Gomes and Pimentel, 2011f] foram representadas e avaliadas interações sociais baseadas em diferentes tipos de mídia, em complementação ao estudo apresentado na seção anterior.

Uma nova coleta de dados de interação da rede Facebook foi realizada em

Figura 3.1: *Frequência de Mídias Compartilhadas na rede Facebook em Março de 2011 (fonte: [Gomes and Pimentel, 2011f])*

Março de 2011. Novamente foram coletadas informações associadas às interações sociais de mais de 1.000 usuários que são de diferentes localidades. Diferentes tipos de mídia foram identificados nos dados coletados. Na Figura 3.1 são apresentadas as frequências das mídias que foram identificadas. Tabela 3.2 são sumarizadas informações sobre a unidade de observação da rede Facebook utilizada na experimentação reportada nesta seção.

As mídias identificadas são: tipo *status* que são mensagens de texto (Figura 2.2, página 21); tipo *photo* que são arquivos contendo imagens ou fotos (Figura 2.3, página 22); tipo *youtube* que são links da rede YouTube (Figura 2.4, página 23); *link* que são links da Web oriundos de outros repositórios que não da rede YouTube (Figura 2.6, página 25); tipo *video* que são arquivos de vídeo; tipo *music* que são links para arquivos de áudio (Figura 2.5, página 24); tipo *swf* que são arquivos de contendo animações criadas em *Flash*.

Considerando a ação A_1 apresentada na Tabela 3.3, esses tipos de mídia foram representadas como uma especialização de A_1 da seguinte forma:

- $A_1.status$ = ação de criar um *Post* do tipo *status*,
- $A_1.photo$ = ação de criar um *Post* do tipo *photo*,
- $A_1.YouT$ = ação de criar um *Post* do tipo *youtube*,
- $A_1.link$ = ação de criar um *Post* do tipo *link*,
- $A_1.video$ = ação de criar um *Post* do tipo *video*,
- $A_1.music$ = ação de criar um *Post* do tipo *music*,
- $A_1.swf$ = ação de criar um *Post* do tipo *swf*.

Tabela 3.5: *Medidas para Contingências baseadas em Mídia [Gomes and Pimentel, 2011f]*

	Confidência	Suporte	Correlação Cosseno
R3.1	100%	44.61%	92.89%
R3.2	100%	21.33%	97.06%
R3.3	100%	12.72%	95.69%
R3.4	100%	7.53%	92.14%
R3.5	100%	2.50%	96.19%
R3.6	100%	0.04%	93.88%
R3.7	100%	0.15%	100%

A partir da especialização de A_1 foram representadas interações sociais baseadas em mídias dentro de cada uma das contingências comportamentais da Listagem 3.1. Na experimentação reportada nesta seção foi realizada a especialização da interação social genérica mais frequente, ou seja, interação $R3$ apresentada na Listagem 3.1. As interações sociais baseadas nos diferentes tipos de mídia representadas a partir de $R3$ são apresentadas na Listagem 3.2.

- R3.1. $aA_1.status \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$
R3.2. $aA_1.photo \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$
R3.3. $aA_1.YouT \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$
R3.4. $aA_1.link \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$
R3.5. $aA_1.video \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$
R3.6. $aA_1.music \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$
R3.7. $aA_1.swf \cap \overline{ak}A_2 \cap alA_3 \rightarrow akIC_1 \cap akIC_3$

Listagem 3.2: *Contingências Comportamentais de Usuários do Facebook: Interações Sociais baseadas em Mídia [Gomes and Pimentel, 2011f]*

A avaliação das interações sociais $R3.1$ até $R3.7$ foi computada usando as medidas de Confidência (Equação 3.2), Suporte (Equação 3.1) e Correlação Cosseno (Equação 3.3). O resultado dessa avaliação é apresentado na Tabela 3.5. Deve ser observado que o valor da medida de Confidência = 100%.

Um ranqueamento do maior para o menor valor das medidas de Suporte e de Correlação Cosseno foi construído. Considerando a medida de Suporte, pode ser obtida a ordem $R3.1 > R3.2 > R3.3 > R3.4 > R3.5 > R3.6 > R3.7$, ou seja, considerando os dados de interação que foram analisados por meio da aplicação da técnica nessa experimentação, os usuários estão envolvidos em interações sociais baseadas em mídias do tipo *status*, *photo*, *youtube* e *link*.

Considerando a medida de Correlação Cosseno, pode ser obtida a ordem $R3.2 > R3.3 > R3.5 > R3.7$, ou seja, considerando os dados de interação que foram analisados por meio da aplicação da técnica nessa experimentação, os usuários estão envolvidos em interações sociais baseadas em mídias do tipo *photo*, *youtube*, *video* e *swf*.

Figura 3.2: *Frequência de Uso das Aplicações de Smartphone no Compartilhamento de Mídias na rede Facebook em Março de 2011 (fonte: [Gomes and Pimentel, 2011c])*

Algumas observações importantes podem ser feitas a partir desses resultados. Era esperado que os usuários para os quais os dados de interação foram coletados estivessem envolvidos com mais frequência em interações sociais baseadas em mídias do tipo *status*, *photo* e *youtube*, pois esse resultado está alinhado com a frequência de compartilhamento das mídias apresentada Figura 3.1. O ranqueamento baseado na medida de Suporte confirma essa expectativa. Os resultados obtidos com a medida de Correlação Cosseno refletem outra tendência que pode ser usada, por exemplo, em um desempate entre interações sociais com mesmo valor para a medida de Suporte.

Em resumo, considerando os usuários que tiveram seus dados de interação coletados para a realização da experimentação reportada nesta seção, os resultados a aplicação da técnica na análise do engajamento desses usuários em interações sociais baseadas em mídias estão alinhados com os resultados obtidos a partir observação das frequências de compartilhamento de mídias na rede.

3.5.3 Interações Sociais Providas via Aplicações de Smartphones

Em outro estudo [Gomes and Pimentel, 2011c] foi elaborada a representação do uso de aplicações disponíveis em dispositivos do tipo *smartphone* pelos usuários da rede Facebook para o compartilhamento de mídias em interações sociais. Novamente foram coletadas informações associadas às interações sociais de mais de 1.000 usuários de diferentes localidades, ainda no mês de Março de 2011. Tabela 3.2 são sumarizadas informações sobre a unidade de observação da rede Facebook utilizada na experimentação reportada nesta seção

Na Figura 3.2 são apresentadas as frequências do uso de aplicações dis-

poníveis em dispositivos do tipo *smartphone* no compartilhamento de mídias identificadas nos dados coletados. A ação A_1 apresentada na Tabela 3.3 foi especializada com mídias e aplicações disponíveis em *smartphones* da seguinte forma:

- $A_1.status_{facebook.for.iphone} = A_1.s_{fb.ip} =$ criar um *Post* do tipo *status* provido pela aplicação *facebook for iphone*,
- $A_1.status_{mobile} = A_1.s_{mob} =$ criar um *Post* do tipo *status* provido pela aplicação *mobile*,
- $A_1.photo_{facebook.for.iphone} = A_1.p_{fb.ip} =$ criar um *Post* do tipo *photo* provido pela aplicação *facebook for iphone*,
- $A_1.status_{facebook.for.blackberry} = A_1.s_{fb.bb} =$ criar um *Post* do tipo *status* provido pela aplicação *facebook for blackberry*,
- $A_1.status_{facebook.for.android} = A_1.s_{fb.an} =$ criar um *Post* do tipo *status* provido pela aplicação *facebook for android*,
- $A_1.photo_{facebook.for.android} = A_1.p_{fb.an} =$ criar um *Post* do tipo *photo* provido pela aplicação *facebook for android*,

A regra $R3$ da Listagem 3.1 foi reescrita tornando explícitas as mídias e as aplicações disponíveis em *smartphones*, como apresentado na Listagem 3.3.

As interações sociais $R3.11$ até $R3.15$ foram avaliadas utilizando as medidas de Confidência (Equação 3.2), Suporte (Equação 3.1) e Correlação Cosseno (Equação 3.3). O resultado dessa avaliação é apresentado na Tabela 3.6. Deve ser observado que o valor da medida de Confidência = 100%.

R3.11	$aA_1.s_{fb.ip} \cap \bar{k}A_2 \cap lA_3 \rightarrow aklC_1 \cap aklC_3$
R3.12	$aA_1.s_{mob} \cap \bar{k}A_2 \cap lA_3 \rightarrow aklC_1 \cap aklC_3$
R3.13	$aA_1.p_{fb.ip} \cap \bar{k}A_2 \cap lA_3 \rightarrow aklC_1 \cap aklC_3$
R3.14	$aA_1.s_{fb.bb} \cap \bar{k}A_2 \cap lA_3 \rightarrow aklC_1 \cap aklC_3$
R3.15	$aA_1.s_{fb.an} \cap \bar{k}A_2 \cap lA_3 \rightarrow aklC_1 \cap aklC_3$
R3.16	$aA_1.p_{fb.an} \cap \bar{k}A_2 \cap lA_3 \rightarrow aklC_1 \cap aklC_3$

Listagem 3.3: *Mídias e Aplicações de Smartphones Representadas na Contingência $R3$ [Gomes and Pimentel, 2011c]*

Um ranqueamento do maior para o menor valor das medidas de Suporte e de Correlação Cosseno foi construído. Considerando a medida de Suporte, pode ser obtida a ordem $R3.11 > R3.13 > R3.12 > R3.14 > R3.15 > R3.16$, ou seja, considerando os dados de interação que foram analisados por meio da aplicação da técnica nessa experimentação, os usuários estão envolvidos preferencialmente

Tabela 3.6: Medidas das Interações Sociais baseadas em Mídias e Aplicações de Smartphone [Gomes and Pimentel, 2011c]

	Confidência	Suporte	Correlação Cosseno
R3.11	100%	0.67%	38.31%
R3.12	100%	0.24%	28.91%
R3.13	100%	0.37%	43.30%
R3.14	100%	0.18%	44.49%
R3.15	100%	0.14%	39.95%
R3.16	100%	0.07%	36.36%

em interações sociais baseadas em mídias do tipo *status* providas via aplicações do tipo *facebook.for.iphone*.

Considerando a medida de Correlação Cosseno, pode ser obtida a ordem $R3.14 > R3.13 > R3.15 > R3.11 > R3.16 > R3.12$. Considerando os dados de interação que foram analisados pela aplicação da técnica nessa experimentação, é possível observar que os usuários estão envolvidos preferencialmente em interações sociais baseadas em mídias do tipo *status* providas via aplicações *facebook.for.blackberry*.

Algumas observações importantes podem ser feitas a partir desses resultados. Era esperado que os mais de 1.000 usuários para os quais os dados de interação foram coletados estivessem envolvidos com mais frequência em interações sociais baseadas em mídias do tipo *status* providas por aplicações do tipo *facebook.for.iphone*, dado que esse resultado está alinhado com os resultados observados para as frequências do uso de aplicações disponíveis em dispositivos do tipo *smartphone* no compartilhamento das mídias. O ranqueamento baseado na medida de Suporte confirma essa expectativa. Os resultados obtidos com a medida de Correlação Cosseno refletem outra tendência que pode ser usada, por exemplo, em um desempate entre interações sociais com mesmo valor para a medida de Suporte.

3.5.4 Síntese das Observações

As experimentações reportadas nesta seção apresentam resultados que foram obtidos a partir de conjuntos de dados coletados da rede Facebook no mês de Março de 2011. A aplicação da técnica de representação e avaliação permite que observações importantes sejam realizadas. Considerando os usuários cujos dados de interação foram coletados, foi possível observar que existem interações sociais genéricas que se repetem com regularidade dentro da rede Facebook.

Os resultados obtidos a partir da aplicação da técnica mostram que os usuários então engajados em interações sociais utilizando preferencialmente recursos mais simples de interação. A partir dessa análise, é possível observar que a apli-

cação da técnica é consistente e reporta resultados esperados obtidos a partir da análise de comportamentos coletivos.

Foi possível observar também que interações sociais genéricas podem ser especializadas em interações baseadas em mídia. Por sua vez, interações sociais baseadas em mídia podem ser especializadas em interações sociais baseadas em mídia providas por aplicações disponíveis em dispositivos do tipo *smartphone*. A análise sugere que a técnica pode ser aplicada de forma iterativa e, ainda, mediante sucessivas especializações, a técnica permite que o compartilhamento de mídias e o uso de dispositivos específicos possam ser estudados por meio do detalhamento das interações sociais.

Nesta seção foram apresentadas experimentações da técnica de representação e avaliação sem a extração automática de interações sociais. Foram representadas e avaliadas interações sociais genéricas, interações sociais baseadas em mídia, e também, interações sociais baseadas em mídia providas via aplicações disponíveis em *smartphones*. Na seção que segue é apresentada a experimentação da técnica de representação e avaliação incluindo a etapa de extração automática de interações sociais.

3.6 Experimentação da Técnica com Extração Automática

Na experimentação da técnica com extração automática de contingências comportamentais [Gomes and Pimentel, 2012d], foram extraídas e avaliadas interações sociais baseadas em mídia vinculadas em mobilizações sociais contra corrupção no Brasil organizadas na rede Facebook entre Setembro e Outubro de 2011. O objetivo foi utilizar a técnica extrair de forma automática e medir interações sociais baseadas em mídia, explicitando as mídias compartilhadas e as ações executadas pelos usuários nas duas oportunidades. Dessa forma, o comportamento coletivo dos usuários nessas mobilizações foi analisado. Uma extensão dos resultados apresentados nesta Seção é reportada no próximo capítulo.

As mobilizações nas ruas aconteceram em dois momentos diferentes: em 7 de Setembro de 2011⁵ com a participação de mais de 30.000 manifestantes, e em 12 outubro de 2011⁶ com a presença de pelo menos 50.000 manifestantes em 25 diferentes grandes cidades brasileiras.

Quatro coletas de dados foram realizadas com quinze dias de diferença, entre 1 setembro e 14 Outubro, a partir de 1.400 perfis onde a maioria dos usuários (954) são do Brasil e os demais são de países como Estados Unidos, Canadá, Ar-

⁵ <http://veja.abril.com.br/blog/augusto-nunes/o-pais-quer-saber/movimento-contr-a-corrupcao-volta-as-ruas-em-25-cidades-neste-12-de-outubro-veja-no-mapa-o-local-e-a-hora-das-manifestacoes/> acessado em 10/01/2013.

⁶ <http://negacaologica.com/marcha-contr-a-corrupcao-12-de-outubro/> acessado em 10/01/2013.

Tabela 3.7: Informações sobre a Unidade de Observação: Técnica com Extração Automática

Seção	Período de Coleta	# de Usuários	# de Contingências
Seção 3.6	entre Setembro e Outubro de 2011	1.423	57.156

Figura 3.3: Frequências de Mídias Compartilhadas na Movimentação Social (fonte: [Gomes and Pimentel, 2012d])

gentina, Inglaterra, Portugal, Espanha, França, Rússia, Israel, Tailândia, dentre outros. Na Tabela 3.7 são sumarizadas informações sobre a unidade de observação da rede Facebook utilizada na experimentação reportada nesta seção.

Foi construído um conjunto de palavras-chave (*bag-of-words*)⁷ a partir do qual foram selecionadas as interações sociais relacionadas ao movimento.

O primeiro conjunto de dados de interações sociais (DSI 1) dispõe de dados sobre 50.887 interações sociais publicadas entre 30 de Agosto e 7 de Setembro de 2011. O segundo conjunto de dados de interações sociais (DSI 2) dispõe de dados sobre 57.156 interações sociais publicadas entre 8 de Setembro e 12 de Outubro de 2011. Dois conjunto de dados DSI 1 e DSI 2 foram construídos para a extração automática de interações sociais baseadas e mídia, usando uma implementação do Algoritmo 1.

Nas experimentações reportadas nesta seção, a validação das interações sociais extraídas de forma automática foi realizada por um especialista em Psicologia Social.

3.6.1 Interações Sociais baseadas em Mídias vinculadas a Movimentação Social

A Figura 3.3 apresenta a frequência de quatro tipos de mídia mais compartilhados nos períodos de movimentação social online em que foi executada a coleta de dados. A Figura 3.3 A reporta a frequência das mídias compartilhadas relacionadas ou não à movimentação social contra a corrupção. Pode ser observado

⁷Foram usadas palavras como *corrupção*, *corrupto*, *marcha*, *manifestação*, *político*, *movimento*, *protesto*. Uma abordagem mais sofisticada para identificação de palavras relacionadas ao movimento, chamada *bag-of-related-words* [Rossi and Rezende, 2011], pode ser usada em trabalhos futuros.

Tabela 3.8: Especialização da ação A_1 apresentada anteriormente na Tabela 3.3 (fonte: [Gomes and Pimentel, 2012d])

Ação	Descrição da Ação
$A_1.status$	ação de criar um Post no Mural do tipo <i>status</i>
$A_1.photo$	ação de criar um Post no Mural do tipo <i>photo</i>
$A_1.video$	ação de criar um Post no Mural do tipo <i>video</i>
$A_1.link$	ação de criar um Post no Mural do tipo <i>link</i>

nessa figura que há um pequeno aumento nas frequências das mídias *status* e *link*, e um pequeno decréscimo na frequência das mídias *photo* e *video* do primeiro período para o segundo.

A Figura 3.3 B reporta a frequência das mídias compartilhadas relacionadas com a movimentação social online. Também é reportado na Figura 3.3B um pequeno decréscimo na frequência das mídias *status* e *link*, e também um pequeno acréscimo na frequência das mídias *photo* e *video* do primeiro período para o segundo.

Com exceção da ação A_4 , usuários $k1$ e consequência C_4 , as demais Ações, Usuários e Consequências apresentadas na Tabela 3.3 foram representadas nos conjuntos de dados DSI 1 e DSI 2. Em outras palavras, não foram consideradas a marcação de nomes nos comentários dos *Posts*.

Os tipos de mídia apresentados na Figura 3.3 foram utilizados na especialização da ação aA_1 , como apresentado na Tabela 3.8. Essa especialização foi inserida nos conjuntos de dados DS 1 e DS 2, dos quais as interações sociais baseadas em mídia foram extraídas automaticamente.

Na Listagem 3.4 são apresentadas as interações sociais $R4$ e $R1$ baseadas nos quatro tipos de mídia mais frequentes vinculadas à movimentação social que está em análise.

As regras $R4.1$ e $R1.1$ são descritas como:

R4.1 *if* usuário a executa a ação de criar um Post do tipo *status* e os usuários l executam a ação A_2 e os usuários k executam a ação A_3 (usuário a fornece o estímulo social que recebe *Curtir* e recebe *Comentar*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem as notificações C_1 e C_2 e C_3 .

R1.1 *if* usuário a executa a ação de criar um Post do tipo *status* e os usuários l não executam a ação A_2 e os usuários k não executam a ação A_3 (usuário a fornece o estímulo social que não recebe *Curtir* e não recebe *Comentar*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem a notificação C_1 .

Tabela 3.9: Comparação entre Algoritmos de Extração de Regras

Algoritmo	Tempo (s)	# Regras Extraídas	# Regras Úteis
JRIP	1.58	2	0
PART	0.3	2	0
RIDOR	0.26	3	0
Apriori	2.0	10	0
GSP	10.0	57	10
GSP-like	13.0	27	27

- R4.1. $\{status, L, K\} : aA_1.status \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
R4.2. $\{photo, L, K\} : aA_1.photo \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
R4.3. $\{video, L, K\} : aA_1.video \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
R4.4. $\{link, L, K\} : aA_1.link \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
R1.1. $\{status\} : aA_1.status \cap \bar{l}A_2 \cap \bar{k}A_3 \rightarrow alkC_1$
R1.2. $\{photo\} : aA_1.photo \cap \bar{l}A_2 \cap \bar{k}A_3 \rightarrow alkC_1$
R1.3. $\{video\} : aA_1.video \cap \bar{l}A_2 \cap \bar{k}A_3 \rightarrow alkC_1$
R1.4. $\{link\} : aA_1.link \cap \bar{l}A_2 \cap \bar{k}A_3 \rightarrow alkC_1$

Listagem 3.4: Interações Sociais baseadas em Mídia vinculadas à Movimentação Social [Gomes and Pimentel, 2012d]

Na Tabela 3.9 é apresentada uma comparação entre os algoritmos tipicamente utilizados na extração de regras, disponíveis na ferramenta WEKA [Hall et al., 2009], com o algoritmo apresentado nesta Seção denominado de *GSP – like*. É possível observar que, apesar de levar o maior tempo de execução, o algoritmo *GSP – like* extraiu o maior número de regras úteis, e ainda, um maior número de regras que os demais algoritmos a menos do GSP original. Algumas estratégias podem ser implementadas para que o desempenho do algoritmo, em termos de tempo de execução, possa ser melhorado. Por exemplo, o armazenamento do conjunto de dados em memória primária com o levantamento do número de linhas diferentes durante esse carregamento⁸.

As interações sociais baseadas em mídia apresentadas na Listagem 3.4 foram avaliadas usando as medidas de Suporte (Equação 3.1), Confidência (Equação 3.2) e Correlação Cosseno (Equação 3.3). Do mesmo modo que nas experimentações apresentadas na seção anterior, os valores da medida de Confidência foram 100% para as interações sociais baseadas em mídia analisadas nesta seção. Esse resultado significa que a totalidade de observações que tem associação a uma determinada regra foi avaliada.

Na Figura 3.4 são apresentados os valores da medida de Suporte para as

⁸De fato, o WEKA carrega os dados em memória primária e, durante esse procedimento, as frequências dos atributos e a contagem do número de linhas total do conjunto de dados são computadas.

Figura 3.4: Suporte das Interações Baseadas em Mídia vinculadas ao Movimento [Gomes and Pimentel, 2012d]

Figura 3.5: Correlação Cosseno das Interações Baseadas em Mídia vinculadas ao Movimento [Gomes and Pimentel, 2012d]

contingências apresentadas na Listagem 3.4. As interações sociais baseadas em mídias $\{status, L, K\}$ e $\{photo\}$ são as mais frequentes. Na Figura 3.5 são apresentados os valores da medida de Correlação Cosseno para as contingências apresentadas na Listagem 3.4. As interações sociais baseadas em mídias $\{status, L, K\}$ e $\{photo\}$ são as mais fortes.

Pela comparação dos resultados apresentados nas Figuras 3.4 e 3.5, considerando os usuários que tiveram seus dados de interação coletados, pode ser observado que o comportamento coletivo dos usuários na movimentação social realizada na rede Facebook é muito semelhante e consistente em dois momentos diferentes da mobilização social online.

A partir dessa análise é possível concluir que os usuários que tiveram seus dados de interação coletados participaram de interações sociais executando ações *Curtir* e *Comentar*, principalmente a partir da publicação de mensagens de texto relacionadas ao movimento. Pode ser observado que, mesmo com opções mais

ricas de mídias para compartilhamento, os usuários estão propensos a se envolver na mobilização social online a partir de publicações do tipo *texto*. A partir da publicação desse tipo de mídia, os usuários executam de ambas as ações *Curtir* e *Comentar*. Esse resultado é consistente com outros que reportam o compartilhamento de mídia tipo *texto* em mobilizações organizadas na rede Twitter [Zhou et al., 2010] [Starbird and Palen, 2012] [Starbird and Palen, 2011].

A publicação de fotos relacionadas ao movimento foi expressiva ($\text{Sup(photo)}=18\%$ das publicações, Figura 3.4), entretanto, essas fotos não motivaram a participação dos usuários em interações sociais por meio da execução de ações, isto é, os usuários não executaram a ação *Curtir* e nem a ação *Comentar*.

3.6.2 Síntese das Observações

A experimentação reportada nesta seção apresenta resultados que foram obtidos a partir de conjuntos de dados coletados do Facebook, em um período de tempo específico, com a finalidade de verificar os resultados obtidos com a aplicação da técnica na análise de interações sociais baseadas em mídia associadas a uma movimentação social. Foi possível observar que a técnica de representação, extração e avaliação pode ser empregada na análise do comportamento coletivo de usuários de Redes Sociais Online. A técnica permite que o comportamento coletivo seja representado como interações sociais baseadas em mídias, e permite também que essas interações sejam extraídas de forma automática e avaliadas com diferentes medidas.

3.7 Trabalhos Relacionados

Os esforços de pesquisa relacionados com a investigação apresentada neste capítulo contemplam a aplicação de técnicas de mineração de dados na análise de relações entre usuários de Redes Sociais Online, bem como o estudo das relações sociais viabilizadas por tecnologia móvel, e ainda, a investigação das atividades de usuários de Redes Sociais em movimentações políticas e sociais.

3.7.1 Técnicas de Mineração de Dados na Análise Relações Sociais

Técnicas de mineração de dados são usadas na análise de relações em Redes Sociais Online, por exemplo, em tarefas de agrupamento para a identificação de grupos de usuários na rede YouTube [Maia et al., 2008], na plataforma Lattes⁹ [Stroele et al., 2009], na rede Flickr [Negoescu et al., 2009], e na rede Twitter [Abrol and Khan, 2010].

Os resultados da aplicação da técnica apresentada neste capítulo estão alinhados aos resultados apresentados nesses trabalhos, pois permite a análise das interações sociais que são estabelecidas a partir das relações dos usuários

⁹www.lattes.cnpq.br

de Redes Sociais Online. A técnica permite a representação explícita de elementos associados com as atividades dos usuários. Essa representação explícita não é explorada na utilização de técnicas de agrupamento estudadas nos trabalhos mencionados.

Técnicas de classificação são empregadas na análise das atividades dos usuários de Redes Sociais Online, por exemplo, na classificação de mensagens de grupos de notícias, identificando concordância e desacordo com a mensagem original, perguntas e padrões de resposta [Fortuna et al., 2007], e na extração de características da Redes Sociais a partir de informações da estrutura da rede e de rótulos designados para nós da rede [Kajdanowicz et al., 2010]. Técnicas de regressão são usadas em tarefas de previsão a partir da análise do grafo social de redes como Flickr [Jin et al., 2010] e Twitter [Golder and Yardi, 2010]. Técnicas baseadas em regras de associação são empregadas também na extração de padrões de formação de ligações entre usuários de Redes Sociais [Leung et al., 2010] e políticas que regulam a divulgação de informações na Rede Social [Braghin et al., 2011]. Os padrões das atividades dos usuários reportados nesses trabalhos são extraídos por meio da utilização de técnicas de classificação e regressão não exploram a descrição dos padrões de atividades dos usuários. A aplicação da técnica reportada neste capítulo permite a extração de um padrão descritivo das atividades dos usuários de Redes Sociais Online.

Os resultados das experimentações apresentados neste capítulo estão alinhados com os resultados desses trabalhos, pois as interações sociais analisadas nas experimentações da técnica são extraídas como regras de associação. A técnica reportada permite explicitar os elementos associados com as atividades dos usuários da rede, resultado que não se observa nos trabalhos relacionados mencionados.

3.7.2 Ligações Sociais Ativadas pelo Uso de Aplicações Disponíveis em Smartphones

O estudo das ligações sociais ativadas a partir de aplicações disponíveis em dispositivos do tipo *smartphone* tem permitido o desenvolvimento de novas tecnologias que dão suporte a interações sociais.

Modelando os padrões de uso da rede Facebook, Barkhuus and Tashiro [2010] apresentaram um estudo do ponto de vista de plataformas móveis e fixas relacionadas com as práticas sociais dos usuários.

Quercia et al. [2010] desenvolveram uma nova tecnologia que silenciosamente produz informações de atualizações dos usuários para ajudar os adolescentes a cultivar suas relações sociais. Em um trabalho posterior, Quercia et al. [2011] propuseram um sistema que combina a localização dos telefones celulares com

anúncios de eventos em Redes Sociais para criar estimativas do número de pessoas em um determinado local.

Nesses trabalhos, os padrões de uso de aplicações disponíveis em dispositivos do tipo *smartphone* não detalham o tipo de interação social que envolve os usuários, e também, não detalham quais os tipos de mídia são compartilhados a partir do uso dessas aplicações. Em contrapartida, a aplicação da técnica reportada na Seção 3.5.3 permite explicitar tanto o tipo de interação social que envolve os usuários quanto o tipo de mídia compartilhada nessas interações.

3.7.3 Engajamento de Usuários de Redes Sociais Online em Mobilizações Cívicas e Políticas

Redes Sociais Online como as redes Twitter e Facebook têm sido exploradas pelos usuários no engajamento em mobilizações cívicas e políticas. Exemplos recentes incluem os protestos eleitorais no Irã [Zhou et al., 2010], os levantes políticos no Egito [Starbird and Palen, 2012], e a mobilização social após o terremoto no Haiti [Starbird and Palen, 2011].

Nesses trabalhos, a rede Twitter foi fundamental como uma plataforma para a difusão de informação e para a organização de atividades relacionadas com essas mobilizações. O tipo de mídia predominantemente utilizado nessas mobilizações é a mídia texto. Os resultados da aplicação da técnica apresentados na Seção 3.6.1 permitem observar que os usuários da rede Facebook, cujos dados de interação foram coletados e analisados na respectiva experimentação, estão propensos a se envolver na mobilização social online contra corrupção no Brasil a partir de publicação de mensagens de texto.

A rede Facebook foi extensamente utilizada para a mobilização política nos EUA, por exemplo, nas eleições parlamentares e para governador no ano de 2006 [Williams, 2007], na eleição presidencial no ano de 2008 [Robertson et al., 2009], e também, para compreender o discurso e a agenda do grupo político conhecido como *Coffee Party* [Mascaro and Goggins, 2011].

Nesses trabalhos relacionados é feita uma análise da propagação de informações na rede em termos do compartilhamento de mídias e da (re) publicação de mensagens, porém, não é realizado o detalhamento de quais são as interações sociais que envolvem os usuários nas mobilizações políticas nos EUA. Em contrapartida, a aplicação da técnica reportada na Seção 3.6.1 permite observar quais interações sociais baseadas em mídias que envolvem os usuários na mobilização social contra corrupção no Brasil realizada na rede Facebook.

3.8 Considerações Finais

Uma técnica para representação, extração e avaliação de interações sociais em Redes Sociais Online foi apresentada neste capítulo. Também foram apresentados resultados da experimentação da técnica que foram reportados na literatura [Gomes and Pimentel, 2011e], [Gomes and Pimentel, 2011c], [Gomes and Pimentel, 2012d] e Gomes and Pimentel [2011f].

Na aplicação da técnica, foram realizadas coletas variadas de dados de interação de mais de 1.000 usuários da rede Facebook no mês de Março de 2011 e de mais de 3.000 usuários entre Setembro e Outubro de 2011. As interações sociais extraídas de forma manual e automática foram avaliadas com as medidas *Confidência*, *Suporte* e *Correlação Cosseno*.

A princípio, a aplicação da técnica é independente do conjunto de dados coletado. A técnica possibilita a representação, a extração e a avaliação de contingências comportamentais que se repetem com regularidade na Rede Social Online. Esse resultado permite observar que a técnica pode ser empregada para detectar padrões que refletem comportamentos coletivos esperados.

A aplicação da técnica permite observar também que interações sociais genéricas podem ser especializadas em interações baseadas em mídia. Por sua vez, interações sociais baseadas em mídia podem ser especializadas em interações sociais baseadas em mídia providas por aplicações disponíveis em *smartphones*. Esse resultado sugere que a técnica pode ser aplicada de forma iterativa e, ainda, mediante sucessivas especializações, a técnica permite que o compartilhamento de mídias e o uso de dispositivos específicos possam ser estudados por meio do detalhamento das interações sociais.

A técnica também possibilita a análise das interações sociais baseadas em mídia associadas com uma movimentação social online e, portanto, pode ser empregada na análise do comportamento coletivo de usuários de Redes Sociais Online. A partir da aplicação da técnica, o comportamento coletivo dos usuários pode ser representado, extraído e avaliado como interações sociais baseadas em mídias.

Os resultados obtidos sugerem que a técnica pode ser aplicada de forma iterativa e ainda, mediante sucessivas especializações, a técnica permite o detalhamento de mídias e aplicações que provém essas mídias na representação, extração e avaliação de interações entre usuários de Redes Sociais Online. A partir da identificação da oportunidade de sistematização da prática de aplicação da técnica, foi elaborado um procedimento de análise de interações sociais baseadas em mídia, e também, foi elaborado um método que guia a aplicação da técnica. O procedimento e o método são apresentados no próximo capítulo.

Um Método Guia para a Aplicação da Técnica

A aplicação da técnica apresentada no capítulo anterior foi bem sucedida por ter permitido a identificação de uma variedade de detalhes envolvidos nas interações entre os usuários de Redes Sociais Online. A fim de fornecer uma orientação para a aplicação da técnica, inicialmente foi proposto um procedimento para análise de interações sociais baseadas em mídia, ainda contemplando a extração manual de contingências comportamentais. Em seguida, por meio do aprimoramento do procedimento de análise, um método guia foi desenvolvido com o propósito de orientar de forma detalhada a aplicação da técnica de representação, extração e avaliação.

O método compreende sete fases: Observação, Captura, Representação e Extração, Avaliação, Validação/Interpretação e Especialização/Reorganização. Os artefatos gerados na saída de uma fase são usados como entrada para outra fase. Esses artefatos são: Rede Social Online, lista de atributos, conjunto de dados brutos (dados crus), conjuntos de dados projetados, conjuntos de regras, regras avaliadas e nova lista de atributos.

O método guia a aplicação da técnica de modo interativo e iterativo. É interativo, pois o usuário do método pode intervir em qualquer fase para efetuar o processo de análise dos dados de interação que desejar. É iterativo, pois o método guia a aplicação da técnica de modo cíclico, se assim o usuário desejar, o que permite, por exemplo, a especialização de interações sociais.

Este capítulo está organizado da seguinte forma: na Seção 4.1 é apresentado o procedimento de análise de interações sociais baseada em mídias baseado na técnica de representação e avaliação; resultados da aplicação do procedimento

Figura 4.1: Visão Geral do Procedimento de Análise de Interações baseadas em Mídias (fonte: Gomes and Pimentel [2012e])

são apresentados na Seção 4.2; na Seção 4.3 é apresentado o método que guia a aplicação da técnica de representação, extração e avaliação; resultados da experimentação do método com a extração manual de interações sociais são apresentados na Seção 4.4; na Seção 4.5 são reportados resultados da aplicação do método com a extração automática de contingências comportamentais, que foram utilizadas na análise da mudança de comportamento de usuários da rede Facebook induzida pelo uso de aplicativos de terceiros; trabalhos relacionados são apresentados na Seção 4.6 e, finalmente, na Seção 4.7 são apresentadas as considerações finais.

4.1 Um Procedimento para Análise de Interações Sociais baseadas em Mídias

Com base na técnica apresentada no capítulo anterior foi especificado um procedimento para analisar as interações sociais baseadas em mídias. O procedimento de análise tem por objetivo orientar a aplicação da técnica na representação e avaliação na análise de interações sociais baseadas em mídias, sem etapa de extração automática. Em outras palavras, a análise é realizada a partir da representação, da aquisição manual e da mensuração de interações baseadas em mídia. A extensão do procedimento que permite a análise de interações soci-

ais genéricas e a especialização em interações sociais baseadas em mídias será apresentada na próxima seção.

Na Figura 4.1 é apresentada uma visão geral do procedimento. O apoio de um especialista em análise de interações sociais é requerido para a aplicação do procedimento. Antes da etapa de coleta de dados, o apoio do especialista é necessário para a observação da Rede Social Online e identificação das interações sociais a serem analisadas.

Os quatro passos principais são os seguintes:

1. **Coleta e Organização de Dados de Interação:** A coleta de dados é realizada automaticamente, por exemplo, usando uma API ou realizando uma análise de *logs* gerados a partir do uso de aplicações sociais. Os dados coletados podem passar por etapas de Transformação, Seleção e Limpeza. Nessa fase, a participação de um especialista em análise de interações é necessária para identificar os dados de interação a serem coletados. Essa fase do procedimento é apresentada na Figura 4.1 1;
2. **Representação das Interações Sociais:** As interações sociais são representados como implicações $B \rightarrow H$ que correspondem a regras no formato *if-then*. A especialização de regras pode ser realizada nessa fase: o apoio de um especialista em análise de interações é necessário para validar as regras extraídas manualmente. Essa fase do procedimento é apresentada na Figura 4.1 2;
3. **Mensuração de Interações Sociais:** As interações sociais são avaliadas utilizando medidas de regras como Confidência (Equação 3.2), Suporte (Equação 3.1) e Correlação Cosseno (Equação 3.3). Essa fase do procedimento é apresentada na Figura 4.1 3;
4. **Interpretação dos Resultados e Melhorias:** Os resultados devem ser interpretados por avaliadores que, contando com o suporte de um especialista em análise de interações, podem sugerir melhorias a serem incorporadas na análise. Essa fase do procedimento é apresentada na Figura 4.1 4.

Na seção seguinte é reportada uma experimentação do procedimento para a análise para identificar um comportamento usual no compartilhamento de mídias entre provedores de conteúdo na web. O procedimento permite tornar explícitos os elementos associados com as atividades dos usuários nas interações sociais representadas como regras *if-then*.

Figura 4.2: Compartilhamento de Vídeo da rede Youtube para a rede Facebook

4.2 Experimentação do Procedimento de Análise

Usuários de redes sociais podem compartilhar mídias dos provedores de conteúdo na Web para Redes Sociais. Por exemplo, usuários podem compartilhar vídeos da rede YouTube para a rede Facebook, como ilustra a Figura 4.2. Na Figura 4.2 Ab (esquerda) pode ser observado um conjunto de ícones que permitem o compartilhamento de vídeos para redes como Facebook, Google+ e Twitter. Por exemplo, um usuário do Youtube pode clicar no ícone da rede *Facebook* e assim compartilhar um vídeo para essa rede. Na Figura 4.2 B é apresentado o resultado da publicação do vídeo da rede Youtube para a rede Facebook.

Além dos ícones específicos para o compartilhamento entre provedor e Rede Social, segunda opção é copiar a URL do vídeo disponível na rede YouTube (Figura 4.2 Aa) e colar essa URL na rede Facebook, o que dá origem a uma nova publicação na Rede Social, como é ilustrado na Figura 4.2 C.

Uma terceira opção é usar o link *Compartilhar* disponível no Facebook para (re) publicar o vídeo na rede. Do ponto de vista da Rede Social, apenas a última opção corresponde a uma resposta para um estímulo social, as outras correspondem a uma *nova* interação social — embora a interação seja, de fato, uma consequência da publicação original.

Se os usuários do Facebook decidirem compartilhar vídeos ou links, eles podem realizar esse compartilhamento utilizando uma das opções de interação mencionadas.

Na experimentação apresentada nesta seção, foi analisada a situação na qual os usuários do Facebook são notificados da publicação de um vídeo ou de um link por um amigo. Observou-se que alguns usuários preferem voltar para o

Tabela 4.1: *Ações Executadas no Compartilhamento de Mídias entre Provedores de Conteúdo (fonte: Gomes and Pimentel [2012e])*

Ação	Descrição
A_4	criar um <i>Post</i> do tipo <i>vídeo</i> com a URL do provedor de conteúdo X em um instante t_1
A_5	criar um <i>Post</i> do tipo <i>vídeo</i> com a URL do provedor de conteúdo X em um instante t_2 ($t_2 > t_1$)
A_6	criar um <i>Post</i> do tipo <i>link</i> com a URL do provedor de conteúdo Y em um instante t_1
A_7	criar um <i>Post</i> do tipo <i>link</i> com a URL do provedor de conteúdo Y em um instante t_2 ($t_2 > t_1$)
A_8	criar um <i>Post</i> do tipo <i>vídeo</i> compartilhado a partir de um amigo da rede Facebook em um instante t_2 ($t_2 > t_1$)
A_9	criar um <i>Post</i> do tipo <i>link</i> compartilhado a partir de um amigo da rede Facebook em um instante t_2 ($t_2 > t_1$)

provedor de conteúdo original da mídia (por exemplo, YouTube), ao invés de usar as visualizações disponíveis no Facebook.

O objetivo do experimento apresentado nesta seção é aplicar o procedimento de análise de interações baseadas em mídia, de forma que situações como as relatadas acima sejam representadas e avaliadas.

4.2.1 Coleta e Organização de Dados de Interação

Com o suporte de um especialista em análise de interações, foram observadas contingências comportamentais associadas com as situações relatadas acima. A partir dessa observação foram identificadas as ações, consequências e usuários (atores das ações). As ações são apresentadas na Tabela 4.1, as consequências e usuários são apresentados na Tabela 4.2. Nessa tabela são representados o usuário a , que é amigo do usuário b , os grupos de usuários k e l .

Utilizando a API da rede Facebook, foram coletados dados de interação de 1.400 usuários, no mês de Agosto de 2011, onde a maioria dos usuários (954) são do Brasil e os demais são de países como Estados Unidos, Canadá, Argentina, Inglaterra, Portugal, Espanha, França, Rússia, Israel, Tailândia, dentre outros.

Os dados coletados foram organizados um conjunto de dados de interação. Na Tabela 4.5 são sumarizadas informações sobre a unidade de observação da rede Facebook utilizada experimentação reportada nesta seção.

4.2.2 Representação das Interações Sociais

A partir da identificação das ações, consequências e usuários (atores das ações), foram representadas as interações sociais apresentadas na Listagem 4.1 com o apoio de um especialista em análise de interações. É importante observar nessas interações que determinadas ações são respostas a ações originais e, portanto,

Tabela 4.2: *Consequências e Usuários no Compartilhamento de Mídias entre Provedores de Conteúdo*(fonte: Gomes and Pimentel [2012e])

Consequências	
C_4	ser notificado da ação A_4
C_5	ser notificado da ação A_5
C_6	ser notificado da ação A_6
C_7	ser notificado da ação A_7
C_8	ser notificado da ação A_8
C_9	ser notificado da ação A_9
Usuários	
a	usuário que executa a ação A_4
b	usuário que executa a ação A_5
l	usuários que executam a ação Curtir
k	usuários que executam a ação Comentar
$\bar{b}$	usuário que não executa a ação A_5
$\bar{l}$	usuários que não executam a ação Curtir
$\bar{k}$	usuários que não executam a ação Comentar

Tabela 4.3: *Informações sobre a Unidade de Observação: Procedimento de Análise*

Seção	Período de Coleta	# de Usuários	# de Contingências
Seção 4.2	Agosto de 2011	1.423	57.156

existem ações que ocorrem em um tempo posterior ao tempo da ação original. Na Tabela 4.1 as ações (A_7 , A_8 e A_9) são respostas a ações originais e ocorrem em um tempo t_2 . O tempo t_2 é posterior ao tempo t_1 .

- R5. $aA_4 \cap bA_5 \rightarrow abklC_4 \cap abklC_5$
R6. $aA_6 \cap bA_7 \rightarrow abklC_6 \cap abklC_7$
R7. $aA_4 \cap \bar{b}A_5 \rightarrow abklC_4$
R8. $aA_6 \cap \bar{b}A_7 \rightarrow abklC_6$
R9. $aA_4 \cap bA_8 \rightarrow abklC_4 \cap abklC_8$
R10. $aA_6 \cap bA_9 \rightarrow abklC_6 \cap abklC_9$

Listagem 4.1: *Contingências Comportamentais Associadas ao Compartilhamento de Mídias entre Provedores de Conteúdo* [Gomes and Pimentel, 2012e]

As regras apresentadas na Listagem 4.1 são descritas como:

R5 *if* o usuário a executa a ação A_4 e o usuário b executa a ação A_5 ,
then usuários a e b e os usuários em grupos k e l são notificados de C_4 e C_5 (usuário a publica um vídeo a partir do provedor X e o usuário b publica novamente o mesmo vídeo depois do usuário a , então todos os usuários são notificados das duas publicações).

R6 *if* o usuário a executa a ação A_6 e o usuário b executa a ação A_7 ,

then usuários a , b e os usuários nos grupos k e l são notificados de C_6 e C_7 (usuário a publica um link a partir do provedor Y e o usuário b publica novamente o mesmo link depois do usuário a , então os usuários são notificados das duas publicações).

R7 *if* o usuário a executa a ação A_4 e o usuário b não executa a ação A_5
then usuários a , b e os usuários nos grupos de k e l são notificados de C_4 (usuário a publica um vídeo a partir do provedor X e o usuário b não publica novamente o mesmo vídeo depois do usuário a , então todos os usuários são (apenas) notificados da publicação original realizada pelo usuário a).

R8 *if* o usuário a executa a ação A_6 e e o usuário b não executa a ação A_7 ,
then usuários a , b e os usuários nos grupos k e l são notificados de C_6 (usuário a publica um link a partir do provedor Y e o usuário b não publica novamente o mesmo link depois do usuário a , então os usuários são notificados das duas publicações).

R9 *if* o usuário a executa a ação A_4 e o usuário b executa a ação A_8 ,
then usuários a , b e os usuários nos grupos k e l são notificados de C_4 e C_8 (usuário a publica um vídeo a partir da publicação de um amigo da rede Facebook e o usuário b compartilha novamente esse vídeo, em seguida, todos os usuários são notificados das duas publicações).

R10 *if* o usuário a executa a ação A_6 e usuário b executa a ação A_9 ,
then usuários a , b e os usuários nos grupos de k e l são notificados de C_6 e C_9 (usuário a publica um link a partir da publicação de um amigo da rede Facebook e o usuário b compartilha novamente esse link, em seguida, todos os usuários são notificados das duas publicações).

4.2.3 Mensuração de Interações Sociais

O conjunto de dados de interação que foi organizado a partir das informações referentes a 27.137 contingências comportamentais. Esse conjunto de dados foi utilizado na avaliação das regras apresentadas na Listagem 4.1 usando as medidas de Suporte, Correlação Cosseno, Sensibilidade e Laplace. Os resultados do computo dessas medidas são apresentados na Tabela 4.4.

As regras $R5$, $R7$ e $R9$ são associadas a interações sociais baseadas em mídias do tipo *vídeo*. As regras $R6$, $R8$ e $R10$ são associadas a interações sociais baseadas em mídias do tipo *link*. Uma ordenação dessas regras do maior valor para o menor valor das medidas de Suporte (Equação 3.1) e de Laplace (Equação 3.5) foi verificada. O resultado dessa ordenação é o mesmo quando observados os valores de Correlação Cosseno (Equação 3.3) e Sensibilidade (Equação 3.4).

Tabela 4.4: Medidas para as Interações Sociais R5 a R10 (fonte: [Gomes and Pimentel, 2012e])

	Suporte	Correlação Cosseno	Sensibilidade	Laplace
R5	3,96%	69,00%	51,52%	50,97 %
R6	0,25%	71,77%	30,83%	50,06 %
R7	8,30%	55,52%	47,61%	52,52 %
R8	12,01%	77,82%	60,56%	54,45 %
R9	0,00%	0,00%	0,00%	50,00%
R10	0,00%	0,00%	0,00%	50,00%

- A ordem das interações sociais baseadas em *vídeo* é $R5 > R7 > R9$.
- A ordem das interações sociais baseadas em *link* é $R8 > R6 > R10$.

4.2.4 Interpretação dos Resultados e Melhorias

Considerando os usuários cujos dados de interação foram coletados, e ainda, mediante a análise do ranqueamento das regras obtido, algumas interpretações podem ser feitas desse resultado.

Analisando o ranqueamento das interações sociais dos usuários que publicam mídias do tipo *vídeo*, pode ser observado que quando os usuários compartilham vídeos publicados por outros usuários, eles primeiro visitam o provedor de conteúdo e, logo após, efetua a publicação da URL do vídeo na rede (R5). Pode ser observado também que os usuários optam por não usar a opção de compartilhamento da rede Facebook (R9) e também que eles decidem por compartilhar vídeos que não são (re) publicados por outros usuários (R7).

No que diz respeito à publicação de links, pode ser observado que quando os usuários compartilham links publicados por outros usuários, eles preferem visitar o provedor de conteúdo e efetua a publicação da URL do link na rede (R6). Pode ser observado também que os usuários optam por não compartilhar links que foram compartilhados por outros usuários (R8) e também optam por não utilizar a opção de compartilhamento do provedor de conteúdo para a rede Facebook (R10).

Essa análise também é importante para a compreensão do comportamento dos usuários cujos dados foram coletados para análise a partir da propagação de conteúdos virais baseados em links e vídeos nas Redes Sociais: ao perceberem um conteúdo frequentemente compartilhado na Rede Social, os usuários preferem visitar o provedor daquele conteúdo e, a partir dessa visita, eles optam por compartilhar novamente o conteúdo do provedor para a rede, ao invés de (re) publicar os conteúdos utilizando os recursos de compartilhamento disponíveis internamente na Rede Social Online.

Designers de aplicações sociais podem usar os resultados da análise de inte-

rações sociais para avaliar a interface de aplicativos que eles projetam. Pode ser o caso, por exemplo, de ferramentas sociais que oferecem muitas opções para que os usuários compartilhem conteúdos como links e vídeos. Na busca de integração com várias redes, designers podem acrescentar várias opções para compartilhamento de conteúdo em suas aplicações, o que pode confundir o usuário.

4.2.5 Síntese das Observações

A experimentação realizada com o procedimento de análise de interações sociais apresenta resultados importantes. Esses resultados foram obtidos a partir de conjuntos de dados coletados da rede Facebook em março de 2011 a partir do perfil de 1.400 usuários. Foi possível observar algumas preferências dos usuários no uso das opções de compartilhamento de mídia disponíveis dentro da Rede Social Online, e também, preferências no uso das opções de compartilhamento de mídias a partir provedores de conteúdo para dentro da rede.

Ainda explorando o compartilhamento de mídias entre provedores de conteúdos e a rede Facebook, em outra experimentação com o procedimento de análise é apresentada no trabalho [Gomes and Pimentel, 2011b]. Nesse trabalho, a aplicação do procedimento de análise permitiu observar uma falha na identificação de mídias compartilhadas da rede Soundcloud para a rede Facebook.

A experimentação do procedimento de análise mostrou a necessidade de uma estruturação mais detalhada das etapas que guiam a aplicação da técnica. A partir dessa constatação, uma expansão do procedimento foi elaborada e organizada no método guia que é apresentado na seção seguinte.

4.3 Um Método para Representação, Extração e Avaliação de Interações Sociais

A partir da sistematização da prática de aplicação da técnica apresentada no capítulo anterior, e também, em uma extensão do procedimento de análise de interações sociais, foi identificada a oportunidade do desenvolvimento de um método guia. Esse método é denominado de método SIREMr (em inglês, *Social Interaction Representation Extraction Measuring rule-based*).

A Figura 4.3 apresenta uma visão geral do método que é especificado como um processo interativo e iterativo com entradas e saídas. A organização do método é inspirada no processo de descoberta de conhecimento (processo KDD) originalmente apresentado por Fayyad et al. [1996].

O método compreende sete fases: Observação, Captura, Preparação, Representação/Extração, Avaliação, Validação/Interpretação e Especialização/Reorganização. Os artefatos gerados na saída de uma fase são usados como entrada para a fase seguinte. Esses artefatos são: Rede Social Online, lista de atributos

Figura 4.3: O Método SIREMr (fonte: Gomes and Pimentel [2012b])

(atributos dos elementos associados as atividades dos usuários), conjunto de dados crus (dados que não passaram por uma seleção ou limpeza), conjuntos de dados (dados tratados), conjuntos de regras, regras avaliadas e nova lista de atributos (atributos dos elementos associados as atividades dos usuários).

As fases, entradas e saídas do método guia são descritas como:

1. **Observação:** O primeiro passo é obter a orientação de psicólogos sociais na observação da Rede Social Online, com o objetivo de identificar quais os elementos que estão relacionados com as interações sociais devem ser capturados, considerando os elementos da linguagem Mechner. Esta etapa produz uma lista de atributos (atributos dos elementos associados as atividades dos usuários) para a próxima fase.
2. **Captura:** A captura de dados acontece automaticamente, por exemplo, quando usando uma API ou via análise de *logs* gerados a partir do uso da rede, levando em conta a lista de atributos gerados na etapa anterior. Podem ser coletadas informações com relacionadas ao conteúdo de uma

publicação, ações executadas pelos usuários, URL's disponíveis pelos servidores que fornecem as mídias compartilhadas na rede, etc. Essa fase produz como saída um conjunto de dados brutos (dados que não passaram pelas fases de Seleção, Limpeza ou Transformação) para a fase seguinte.

3. **Preparação:** Todos os dados brutos coletados podem ser limpos, selecionados e/ou transformados para serem utilizados nas fases subsequentes de representação, de extração e de medição de interações sociais. Nessa fase, a identificação dos elementos da linguagem Mechner (ações, consequências e usuários) é necessária e um conjunto de dados projetado a partir dos dados brutos é construído. Por exemplo, os atributos podem ser o tipo de mídia publicada uma mensagem, as ações do usuário realizadas em torno de uma mídia publicada, as notificações de cada publicação. Essa fase produz um conjunto de dados projetado como saída para a fase seguinte;
4. **Representação e Extração:** As interações sociais são representados como contingências comportamentais na forma de regras *if-then* em situações sociais envolvendo os usuários da rede Facebook que podem ser baseadas em mídia ou ainda providas por dispositivos do tipo *smartphone*, e também em interações sociais que são espalhadas como conteúdo viral, e ainda caracterizam engajamento social contra corrupção no Brasil em Redes Sociais. A extração manual de contingências comportamentais pode ser realizada a partir da observação de situações sociais com o apoio de especialistas que identificam os elementos da linguagem Mechner. A extração automática de contingências comportamentais pode ser realizada a partir de conjuntos de dados projetados com ações, consequências e usuários (atores das ações) coletados de uma Rede Social Online. Essa fase produz um conjunto de regras como saída;
5. **Avaliação:** Cada interação social é descrita como uma regra *if-then* representada na forma geral $B \rightarrow H$ para calcular os valores $bh, b\bar{h}, \bar{b}h, \bar{b}\bar{h}$. Depois disso, é realizada a medição de cada regra usando medidas de Confidência (Equação 3.2, página 45), Suporte (Equação 3.1, página 45) e Correlação Cosseno (Equação 3.3, página 45) são usadas para medir interações sociais baseadas em mídia e também Sensibilidade (Equação 3.4, página 46) e *Laplace* (Equação 3.5, página 46) para medir interações sociais que se espalham como conteúdo viral na rede Facebook. Esta fase produz um conjunto de regras avaliadas como saída;
6. **Interpretação e Validação:** As medições podem ser interpretadas por usuários especializados com interesse em analisar as interações sociais. A va-

validação da representação, extração e avaliação das interações sociais pode ser realizada por usuários especialistas na análise das interações sociais. Por exemplo, após a representação, a extração e a avaliação de interações sociais baseadas no compartilhamento explícito de vídeos ubíquos e na captura de *frames* para a realização de anotações, analistas de interações sociais validaram os conjuntos de regras. Interações sociais genéricas e baseadas em mídias relacionadas com a mobilização social online contra a corrupção no Brasil são representadas, extraídas e avaliadas de maneira automática e, em seguida, validadas por especialistas com interesse em análise dessas interações sociais [Gomes and Pimentel, 2012d] [Gomes and Pimentel, 2012b]. Outro exemplo, designers podem usar os resultados da análise de interações sociais para avaliar a interface de aplicativos que eles projetam [Gomes and Pimentel, 2012e], ou ainda, detectar falhas na identificação de mídias compartilhadas entre provedores de conteúdo e mídias sociais [Gomes and Pimentel, 2011b]. Se os resultados da interpretação e da validação forem suficientes, o processo termina nessa fase. Caso contrário, a especialização das regras ou uma nova organização dos dados pode ser executada e o processo pode ser repetido a partir das fases de Captura ou de Representação. Se o processo não foi terminado, os novos atributos especificados podem ser usados como saída para a fase seguinte;

7. **Especialização/Reorganização:** A especialização de uma regra é realizada por meio do detalhamento dos elementos associados com as atividades dos usuários dentro das regras (por exemplo, aplicações Web e aplicações disponíveis em *smartphones*). Se tais elementos de interesse foram atributos recolhidos, o processo é repetido a partir da fase de Preparação, ou seja, novos conjuntos de dados projetados são gerados, e no próximo passo, regras especializadas são extraídas. Caso contrário, se tais elementos de interesse não faziam parte do conjunto de dados brutos, o processo é repetido a partir da fase de captura. Por exemplo, interações sociais genéricas podem ser especializadas em interações baseadas em mídia, e as interações baseadas em mídias podem ser detalhadas por meio da explicitação de provedores de conteúdo, ou ainda, da explicitação de aplicações Web e aplicações disponíveis em *smartphones*.

A aplicação do método guia permite o detalhamento de ações, de tipos de mídias, de tipos de aplicações e tipos de dispositivos, dentro das contingências comportamentais associadas a interações sociais.

O método apresentado nesta seção foi experimentado com a extração manual e com a extração automática de contingências comportamentais. Na Tabela 4.5

Tabela 4.5: *Informações sobre as Unidades de Observação: Método de Análise*

Seção	Período de Coleta	# de Usuários	# de Contingências
Seção 4.4	entre Maio e Novembro de 2011	1.580	211.390
Seção 4.5	entre Fevereiro e Novembro de 2012	3.685	814.824

são sumarizadas informações sobre as unidades de observação da rede Facebook utilizadas nas experimentações reportadas nas próximas seções. Na seção seguinte é apresentada uma experimentação do método com extração manual de interações sociais.

4.4 Experimentação do Método com Extração Manual de Contingências Comportamentais

Na experimentação apresentada nesta seção, o método foi aplicado na análise de interações sociais baseadas em mídias que, em seguida, foram especializadas em interações sociais baseadas em mídias providas por aplicações Web e por aplicações disponíveis em *smartphones*. Resultados semelhantes obtidos por meio da aplicação da técnica de representação e avaliação foram reportados respectivamente nas Seções 3.5.2 e 3.5.3.

Na avaliação das interações sociais foram utilizadas as medidas Confidência (Equação 3.2, página 45), Suporte (Equação 3.1, página 45), Correlação Cosseno (Equação 3.3, página 45) e Alavancagem (Equação 3.7, página 46).

4.4.1 Observação e Captura

A rede Facebook foi observada com o propósito de identificar os elementos associados com a atividade dos usuários que incluam mídias compartilhadas e aplicações Web e dispositivos móveis utilizados. Foram capturadas informações referentes a 211.390 interações sociais que ocorreram entre Maio de 2011 e Novembro de 2011. Essas informações foram coletadas do perfil de 1.580 usuários onde a maioria (954) são do Brasil e os demais são de países como Estados Unidos, Canadá, Argentina, Inglaterra, Portugal, Espanha, França, Rússia, Israel, Tailândia, dentre outros. Tabela 4.5 são sumarizadas informações sobre a unidades de observação da rede Facebook utilizada na experimentação reportada nesta seção.

As informações coletadas incluem dados sobre o tipo de mídia compartilhada como foto (Figura 2.3, página 22), status (Figura 2.2, página 21), link (Figura 2.6, página 25), vídeo (Figura 2.4, página 23), música (Figura 2.5, página 24)), sobre as atividades dos usuários, tais como o número de *Curtir* de uma publicação, número de *Comentar* de um publicação, os tipos de mídias compartilhados, as aplicações Web e tipos de dispositivos móveis utilizados para compartilhar as

Tabela 4.6: *Ações, Consequências e Usuários da rede Facebook na Experimentação do Método*

Ações	
A_1	ação de criar um Post
A_2	ação de Curtir um Post
A_3	ação de Comentar um Post
Consequências	
C_1	ser notificado de uma ação de criar um Post
C_2	ser notificado de uma ação de Curtir um Post
C_3	ser notificado de uma ação de Comentar um Post
Usuários	
a	usuário que executam a ação de criar um Post
l	usuários que executam a ação de Curtir
k	usuários que executam a ação de Comentar
$\bar{l}$	usuários que não executam a ação de Curtir
$\bar{k}$	usuários que não executam a ação de Comentar

Tabela 4.7: *Especialização da ação A_1 apresentada anteriormente na Tabela 3.3 (fonte: Gomes and Pimentel [2012a])*

Ação	Descrição da Ação
$A_1.status$	ação de criar um Post no Mural do tipo <i>status</i>
$A_1.photo$	ação de criar um Post no Mural do tipo <i>photo</i>
$A_1.video$	ação de criar um Post no Mural do tipo <i>video</i>
$A_1.link$	ação de criar um Post no Mural do tipo <i>link</i>
$A_1.swf$	ação de criar um Post no Mural do tipo <i>swf</i>
$A_1.music$	ação de criar um Post no Mural do tipo <i>music</i>
$A_1.checkin$	ação de criar um Post no Mural do tipo <i>checkin</i>

mídias. A partir de Setembro de 2011, o tipo de mídia *checkin* (Figura 2.7, página 25) ficou disponível para compartilhamento na rede Facebook.

4.4.2 Representação e Extração Manual

Na fase de preparação dos dados foram construídos dois conjuntos a partir das informações coletadas. Entre Maio de 2011 e Agosto de 2011 foram coletados dados de 102.688 contingências comportamentais que foram organizados no conjunto de dados OBC 1. Entre Setembro de 2011 e Novembro de 2011 foram coletados dados de 108.702 contingências comportamentais que foram organizados no conjunto de dados OBC 2. Esses dois conjuntos de dados foram usados na avaliação das interações entre os usuários.

A partir da observação das contingências comportamentais, foram identificados ações, consequências e usuários (atores das ações), que são apresentados na Tabela 4.6.

Na Tabela 4.7 é apresentada a especialização da ação A_1 com os tipos de

Tabela 4.8: *Contingências R4.1 a R4.7 e as Medidas Computadas*

# Regra	Mídia	Suporte		Correlação Cosseno		Alavancagem	
		OBC 1	OBC 2	OBC 1	OBC 2	OBC 1	OBC 2
R4.1	link	2,30%	2,07%	44,19%	41,15%	2,03%	1,82%
R4.2	status	22,62%	22,44%	77,59%	80,37%	14,12%	14,64%
R4.3	video	5,70%	4,77%	55,10%	56,89%	4,95%	4,24%
R4.4	photo	11,29%	11,92%	65,51%	65,40%	7,09%	7,53%
R4.5	swf	0,02%	0,03%	37,80%	60,01%	0,02%	0,03%
R4.6	music	0,02%	0,02%	55,71%	52,34%	0,02%	0,02%
R4.7	checkin		1,77%		65,47%		1,68%

mídia disponíveis na rede Facebook, que foram utilizadas na representação de interações sociais baseadas em mídias.

- R4.1. $aA_1.link \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
R4.2. $aA_1.status \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
R4.3. $aA_1.video \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
R4.4. $aA_1.photo \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
R4.5. $aA_1.swf \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
R4.6. $aA_1.music \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
R4.7. $aA_1.checkin \cap lA_2 \cap kA_3 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$

Listagem 4.2: *Interações Sociais baseadas em Mídia (fonte: Gomes and Pimentel [2012a])*

Na Listagem 4.2 são apresentadas interações baseadas em mídia representadas a partir da identificação de ações, consequências e usuários (atores das ações). Foram representadas interações sociais baseadas em mídias nas quais os usuários executam ações *Curtir* e *Comentar*. Essa representação foi validada por um especialista na análise de interações sociais.

Por exemplo, a descrição da regra R4.1 apresentada na Listagem 4.2 é:

R4.1 *if* usuário a executa a ação de criar um Post do tipo *link* e os usuários l executam a ação A_2 e os usuários k executam a ação A_3 (usuário a fornece o estímulo social que recebe *Curtir* e recebe *Comentar*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem as notificações C_1 e C_2 e C_3 .

4.4.3 Avaliação, Validação e Interpretação

A Tabela 4.8 apresenta os resultados da avaliação das interações sociais apresentadas na Listagem 4.2 com o conjunto de dados OBC 1 e OBC 2. Pode ser observado que a regra R4.7 não é avaliada com o conjunto de dados OBC 1 pois a mídia *checkin* não está presente nesses dados. A medida de Confidência tem valor 100% para cada regra, o que significa que cada regra foi avaliada com todas as instâncias das observações que estão associadas com a regra.

As regras podem ser ranqueadas do maior para o menor valor das medidas de Suporte, Correlação Cosseno e Alavancagem, as quais são computadas a partir do conjunto de dados OBC 1, como $R4.2 > R4.4 > R4.3 > R4.1 > R4.5 > R4.6$. Esse ranqueamento permite observar que os usuários da rede Facebook, para os quais os dados coletados foram analisados nesta seção, estão engajados em interações sociais por meio da execução de *Curtir* e *Comentar* a partir do compartilhamento de mídias tipo *status* ($R4.2$).

As regras também podem ser ordenadas do maior para o menor valor das medidas de Suporte, Correlação Cosseno e Alavancagem, as quais são computadas a partir do conjunto de dados OBC 2, como $R4.2 > R4.4 > R4.3 > R4.1 > R4.7 > R4.5 > R4.6$. Esse ranqueamento permite observar que a mídia *checkin* ($R4.7$) é mais utilizada em interações sociais que as mídias *swf* e *music*, e menos utilizada que os demais tipos de mídia (*link*, *status*, *video* e *photo*).

4.4.4 Especialização

Seguindo com a aplicação do método, uma nova lista de atributos foi elaborada com o propósito de especializar a ação A_1 com informações sobre as aplicações Web (por exemplo, aplicações *Photos* e *Links*) e para dispositivos móveis (por exemplo, Facebook para iPad, iPhone, Blackberry, Android, etc.) que provêm as mídias publicadas nas interações.

4.4.5 Representação e Extração Manual de Novas Interações Sociais

A representação das aplicações Web e disponíveis em *smartphones* nas interações sociais baseadas em mídias é detalhada na Listagem 4.3. Foram representadas interações sociais baseadas em mídias providas por aplicações Web e dispositivos móveis nas quais os usuários executam ações *Curtir* e *Comentar*.

As regras $R4.10$ até $R4.13$ representam interações sociais baseadas em mídias providas por aplicações Web. As regras $R4.14$ até $R4.18$ representam interações sociais baseadas em mídias providas por dispositivos móveis.

$R4.10. aA_1.video.Links \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.11. aA_1.photo.Photos \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.12. aA_1.link.Links \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.13. aA_1.status.Twitter \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.14. aA_1.status.Iphone \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.15. aA_1.photo.Iphone \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.16. aA_1.status.Mobile \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.17. aA_1.status.Blackberry \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$
$R4.18. aA_1.status.Android \cap lA_2 \cap kA_3 \rightarrow alkC_1 \cap alkC_2 \cap alkC_3$ |
|---|

Listagem 4.3: Aplicações Web e de Smartphones nas Interações Sociais baseadas em Mídia
(fonte: [Gomes and Pimentel, 2012a])

Tabela 4.9: Medidas para as Contingências R4.10 to R4.18

# Regra	Mídias, Aplicações	Suporte		Correl. Cosseno		Alavancagem	
		OBC 1	OBC 2	OBC 1	OBC 2	OBC 1	OBC 2
R4.10	video, Links	3,77%	1,07%	63,16%	64,81%	3,42%	1,05%
R4.11	photo, Photos	2,50%	0,97%	58,01%	62,13%	2,31%	0,95%
R4.12	link, Links	0,86%	0,27%	50,00%	50,44%	0,84%	0,27%
R4.13	status, Twitter	1,12%	0,35%	60,37%	60,14%	1,08%	0,35%
R4.14	status, iPhone	9,17%	2,38%	80,60%	83,98%	7,88%	2,30%
R4.15	photo, iPhone	7,48%	2,16%	79,35%	80,31%	6,60%	2,09%
R4.16	status, Mobile	4,51%	1,36%	76,98%	80,10%	4,16%	1,33%
R4.17	status, BlackBerry	1,76%	0,62%	75,94%	82,91%	1,70%	0,62%
R4.18	status, Android	1,49%	0,52%	74,08%	75,95%	1,45%	0,51%

Deve ser observado que são listadas apenas as interações sociais mais frequentes. Uma descrição da regra R4.10 é apresentada em seguida.

R4.10 *if* usuário a executa a ação de criar um Post do tipo *video* por meio da aplicação *Links* e os usuários l executam a ação A_2 e os usuários k executam a ação A_3 (usuário a fornece o estímulo social que recebe *Curtir* e recebe *Comentar*)
then usuário a e grupo de usuários l e k (usuário a e de seus amigos) recebem as notificações C_1 e C_2 e C_3 .

4.4.6 Avaliação e Interpretação de Novas Interações Sociais

Na Tabela 4.9 são apresentados os resultados da avaliação das regras descritas na Listagem 4.3 com o conjunto de dados OBC 1 e OBC 2. A medida de Confidência tem valor 100% para cada regra avaliada, o que significa que cada regra foi avaliada com todas as instâncias das observações que estão associadas com a regra.

As regras R4.10 a R4.13 podem ser ordenadas do maior para o menor valor das medidas de Suporte, Correlação Cosseno e Alavancagem, as quais são computadas a partir dos conjuntos de dados OBC 1 e OBC 2, como $R4.10 > R4.11 > R4.13 > R4.12$.

Considerando os usuários cujos dados de interação foram coletados para a experimentação do método, pode ser observado a partir do ranqueamento das regras R4.10 a R4.13 que as mídias *video* e *photo*, providas respectivamente pelas aplicações Web *Links* e *Photos*¹ (R4.10 e R4.11), são preferencialmente utilizadas em interações sociais nas quais os usuários executam de ações *Curtir* e *Comentar*. Esse resultado é interessante pois a aplicação Web *Links* permite o compar-

¹A rede Facebook se refere a *Photos* como um conjunto de aplicativos utilizados para compartilhamento de fotos, como *Instagram* e *Facebook Câmera*.

tilhamento de qualquer tipo de link Web. Pode ser observado que os usuários cujos dados foram analisados tem preferência por utilizar a aplicação *Links* no compartilhamento de mídia *video* provido pela em interações sociais. Por outro lado, era esperado que houvesse também a preferência por utilizar a aplicação *Photo* no compartilhamento de mídia *photo* em interações sociais.

As regras *R4.14* a *R4.18* podem ser ordenadas do maior para o menor valor das medidas de Suporte, Correlação Cosseno e Alavancagem, as quais são computadas a partir dos conjuntos de dados OBC 1 e OBC 2, como $R4.14 > R4.15 > R4.16 > R4.17 > R4.18$.

Considerando os usuários cujos dados de interação foram coletados para a experimentação do método, pode ser observado a partir da ordenação das regras *R4.14* a *R4.18* que as mídias *status* e *photo*, providas respectivamente pelo dispositivo móvel *iPhone* (*R4.14* e *R4.15*), são preferencialmente utilizadas em interações sociais nas quais os usuários executam de ações *Curtir* e *Comentar*. Esse resultado é consistente com o apresentado na Seção 3.5.3, onde foi reportado que usuários, cujos dados de interação foram analisados, utilizam preferencialmente a aplicação *facebook for iphone* para compartilhar mídias *status* e *photo* em interações sociais.

4.4.7 Síntese das Observações

As experimentações reportadas nesta seção apresentam resultados que foram obtidos a partir de conjuntos de dados coletados da rede Facebook em um período de tempo específico. A aplicação do método com a extração manual de contingências comportamentais permite que observações importantes sejam realizadas. Em particular, foi possível observar o impacto da inclusão de um novo tipo de mídia na rede como opção para interação entre os usuários.

Os resultados da aplicação do método foram consistentes, mostrando que interações sociais baseadas em mídia podem ser especializadas em interações sociais providas por aplicações Web e por aplicações disponíveis em *smartphones*. Também foi possível observar algumas predileções dos usuários cujos dados de interação foram coletados para a experimentação, por exemplo, há uma preferência no uso da aplicação *Links* para o compartilhamento de vídeos nas interações sociais, e também, pelo uso do *smartphone* do tipo *iPhone* no compartilhamento de texto e foto nas interações sociais.

Na próxima seção é apresentada a experimentação do método com a extração automática de contingências comportamentais vinculadas a uma movimentação social que ocorreu na rede Facebook.

4.5 Experimentação do Método na Análise de Comportamento Social

Os resultados reportados nesta seção tratam da análise de mudança do comportamento dos usuários da rede Facebook, mudança essa que foi induzida pela utilização de aplicativos de terceiros no compartilhamento de mídias. Foram analisados dados de interação de 3.000 usuários coletados durante os anos de 2011 e 2012. Esses dados foram utilizados na extração automática e na avaliação de interações sociais baseadas em mídias e, em seguida, na especialização dessas interações em interações sociais baseadas em mídia providas por aplicativos de terceiros.

Figura 4.4: Frequência de Compartilhamento das Mídias (topo), Frequência dos Aplicativos Utilizados no Compartilhamento das Mídias (inferior) (fonte: Gomes and Pimentel [2012c])

4.5.1 Observação da Rede Social e Captura dos Dados

Antes da coleta de dados, foi realizada a observação da Rede Social Facebook com o objetivo de identificar os elementos associados com as interações sociais que podem ser representados na linguagem Mechner.

Foram coletados dados utilizando-se um *crawler* da rede Facebook que foi executado pelo menos uma vez por mês entre Fevereiro de 2011 e Setembro de 2012. Foram capturados dados de mais de 3.000 perfis de usuários, onde a maioria dos usuários (1.267) são do Brasil e os demais são de países como Estados Unidos, Canadá, Argentina, Inglaterra, Portugal, Espanha, França, Rússia, Israel, Tailândia, dentre outros. Tabela 4.5 são sumarizadas informações sobre a unidades de observação da rede Facebook utilizada na experimentação reportada nesta seção.

Esses dados possuem informações sobre 814.824 interações sociais, incluindo dados sobre o tipo de mídia publicada (*checkin*, *photo*, *status*, *video*, *link*, *swf*, *question* e *music*), os aplicativos de terceiros usados no compartilhamento de mídias, informações sobre as atividades do usuário, tais como a adição de *Curtir* ou *Comentar* em uma publicação, e o número de usuários que adicionam tanto *Curtir* e *Comentar* a uma publicação.

A partir dos dados capturados, a frequência de compartilhamento de mídia foi computada (Figura 4.4 (topo)) durante quatro períodos de tempo (1º semestre de 2011, 2º semestre de 2011, 1º semestre de 2012 e 2º semestre de 2012) e, também, a frequência do uso de aplicativos de terceiros no compartilhamento de mídia (Figura 4.4 (inferior)) durante dois períodos de tempo (1º semestre de 2011 e 2º semestre de 2012).

Considerando os usuários para os quais os dados de interação foram coletados, na Figura 4.4 (topo) pode ser observado que há uma mudança de comportamento coletivo usuário no compartilhamento de mídias. Por exemplo, usuários da rede Facebook são mais propensos a compartilhar *status* (texto) (no primeiro semestre de 2011), e um tempo posterior, os usuários estão mais propensos a compartilhar *photo* (no segundo semestre de 2012).

Considerando os usuários para os quais os dados de interação foram coletados, na Figura 4.4 (inferior) pode ser observado que há uma mudança de comportamento coletivo dos usuários no uso de aplicativos de terceiros para o compartilhamento de mídias. Por exemplo, usuários preferem a usar o aplicativo *Photos* e o aplicativo do tipo *Facebook for iPhone* para compartilhar mídia *photo* e o aplicativo do tipo *Facebook for iPhone* para compartilhar mídia *status*.

Nos dados coletados foram identificados 367 aplicativos efetivamente utilizados no compartilhamento de mídias entre um universo de 29.020 aplicações disponíveis no Facebook [Wang, 2012].

Tabela 4.10: *Ações, Consequências e Usuários: Análise do Comportamento Social*

Ações	
A_1	ação de criar um Post
A_2	ação de Curtir um Post
A_3	ação de Comentar um Post
A_4	ação de Compartilhar um Post
Consequências	
C_1	ser notificado de uma ação de criar um Post
C_2	ser notificado de uma ação de Curtir um Post
C_3	ser notificado de uma ação de Comentar um Post
C_4	ser notificado de uma ação de Compartilhar um Post
Usuários	
a	usuário que executam a ação de criar um Post
l	usuários que executam a ação de Curtir
k	usuários que executam a ação de Comentar
s	usuários que executam a ação de Compartilhar
$\bar{l}$	usuários que não executam a ação de Curtir
$\bar{k}$	usuários que não executam a ação de Comentar
$\bar{s}$	usuários que não executam a ação de Compartilhar

4.5.2 Preparação de Dados

Os dados coletados foram organizados em dois conjuntos de dados com ações, consequências e usuários (atores das ações) listados na Tabela 4.10. Os dados coletados no 1º semestre de 2011 foram organizados em um conjunto de dados, o dados coletados no 2º semestre de 2012 foram organizados em outro conjunto de dados. Deve ser observado que na experimentação reportada nesta seção foram incluídas a ação A_4 , a consequência C_4 e os grupos de usuários (atores das ações) s .

Para a extração de interações sociais baseadas em mídias, os tipos de mídia compartilhadas foram especializadas na ação A_1 nos dois conjuntos de dados. Outra especialização da ação A_1 , com mídias e aplicativos de terceiros, foi realizada nos dois conjuntos de dados. Exemplos da especialização da ação A_1 com mídias podem ser observados na Seção 3.5.2 e na Seção 4.4.2. Exemplos da especialização da ação A_1 com mídias, aplicações Web e aplicações disponíveis em *smartphones* podem ser observados na Seção 3.5.3 e na Seção 4.4.5.

4.5.3 Representação e Extração de Interações Sociais

Foi representada a situação na qual ocorre uma interação social baseada em mídia provida via aplicativo de terceiros — o usuário a publica uma *photo* usando o aplicativo *Photos* que recebe *Curtir* e *Comentar*. Essa situação é representada em notação curta como {photo.Photos,L,K}.

A partir dos conjuntos de dados com ação A_1 especializada com os tipos de mídia, foram extraídas as regras apresentadas na Listagem 4.4, usando a implementação do algoritmo descrito na Seção 3.3.

R1: $photo \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R2: $status \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R3: $photo \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2$ R4: $photo \cap lA_2 \cap kA_3 \cap sA_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3 \cap aklsC_4$ R5: $photo \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1$ R6: $link \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1$ R7: $status \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2$ R8: $video \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2$ R9: $checkin \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R10: $video \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R11: $photo \cap lA_2 \cap \bar{k}A_3 \cap sA_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_4$

Listagem 4.4: Interações Sociais baseadas em Mídia na Análise do Comportamento Social [Gomes and Pimentel, 2012c]

A partir dos dois conjuntos de dados com ação A_1 especializada com o tipo de mídia e com o tipo de aplicativo utilizado no provimento da mídia, foram automaticamente extraídas as interações sociais baseadas em mídias providas por aplicativos de terceiros são apresentadas na Listagem 4.5.

R1: $photo.Photos \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R2: $status.FacebookforiPhone \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R3: $photo.Photos \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2$ R4: $photo.FacebookforiPhone \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R5: $photo.Photos \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1$ R6: $photo.Photos \cap lA_2 \cap kA_3 \cap sA_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3 \cap aklsC_4$ R7: $checkin.FacebookforiPhone \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R8: $checkin.FacebookforiPhone \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2$ R9: $status.FacebookforAndroid \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$ R10: $photo.FacebookforAndroid \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklsC_1 \cap aklsC_2 \cap aklsC_3$

Listagem 4.5: Interações Sociais baseadas em Mídia providas via Aplicativos na Análise do Comportamento Social [Gomes and Pimentel, 2012c]

Essas interações sociais foram avaliadas por meio da computação das medidas de Suporte e Confiança.

4.5.4 Avaliação de Interações Sociais

A Figura 4.5 apresenta o valor do Suporte para as interações sociais baseadas em mídia e a Figura 4.6 o valor do Suporte para as interações sociais baseadas em mídias providas por aplicativos. Os valores da medida de Confiança para as regras são 100%, o que significa que cada regra foi avaliada com todas as instâncias das observações que estão associadas com a regra.

Figura 4.5: Suporte das Interações baseadas em Mídia: Análise do Comportamento Social (fonte: Gomes and Pimentel [2012c])

Figura 4.6: Suporte das Interações baseadas em Mídia Providas por Aplicativos de Terceiros: Análise do Comportamento Social (fonte: Gomes and Pimentel [2012c])

4.5.5 Validação e Interpretação

Considerando que os conjuntos de regras apresentados na Seção 3.5.2 e na Seção 3.5.3 foram validados por um especialista em Psicologia Social, e dada a semelhança com os conjuntos de regras apresentados na Listagem 4.4 e na Listagem 4.5, essas regras foram consideradas validadas.

A mudança de comportamento coletivo dos usuários foi analisada a partir da observação dos valores da medida *SupR* apresentados na Figura 4.5 e na Figura 4.6.

Na interpretação desses resultados (Figura 4.5), é possível observar que mudança de comportamento coletivo de usuário no compartilhamento de mídias — de *status* para *photo* (R2 para R1) — é seguida da participação de usuários na interação social através da execução de ações de *Curtir* e *Comentar*. Considerando os usuários para os quais os dados de interação foram coletados, pode ser observado que a mudança de comportamento no compartilhamento de mídias é acompanhado da interação dos usuários por meio da execução de ações *Curtir* e *Comentar*.

Pode ser observada (Figura 4.6) também a mudança de comportamento coletivo de usuário induzida pela utilização de aplicativos no compartilhamento de mídias. O uso do aplicativo *Facebook for iPhone* no compartilhamento de mídia *status* (R2: *status.Facebook for iPhone*, L, K) é mudado para o uso do aplicativo *Photos* no compartilhamento de mídia *photo* (R1: *photo.Photos*, L, K) e essa mudança é seguida pelo envolvimento dos utilizadores na interação social através da execução de ações *Curtir* e *Comentar*.

Considerando os usuários para os quais os dados de interação foram coletados, pode ser observado que a mudança de comportamento no uso de aplicativos para o compartilhamento de mídias é acompanhado da interação dos usuários por meio da execução de ações *Curtir* e *Comentar*.

4.5.6 Síntese das Observações

A experimentação reportada nesta seção apresenta resultados que foram obtidos de conjuntos de dados coletados da rede Facebook, ao longo de dois anos, a partir do perfil de 3.000 usuários. A aplicação do método guia permite observar uma mudança no comportamento dos usuários da rede em termos do compartilhamento de mídias e do uso de aplicativos para compartilhar essas mídias.

4.6 Trabalhos Relacionados

Os trabalhos de pesquisa relacionados com a investigação apresentada neste capítulo incluem métodos de análise do comportamento de usuários de Redes Sociais Online.

Análise do comportamento dos usuários de Redes Sociais têm contribuições valiosas, por exemplo, para o design Web [Burke et al., 2009] e design de interações com dispositivos móveis [Bohmer et al., 2011], marketing viral [Leskovec et al., 2007] [Cha et al., 2012], modelagem do tráfego de informação [Krishnamurthy, 2009] e distribuição de conteúdo [Mislove et al., 2007].

4.6.1 Análise do Comportamento Social dos Usuários da rede Twitter

A análise do comportamento dos usuários a partir do compartilhamento de informações textuais na rede Twitter tem sido amplamente estudada.

Analizando o comportamento dos usuários da rede no encaminhamento de mensagens, Bigonha et al. [2010] propuseram uma técnica para realizar ranqueamento de usuários como evangelistas e detratores. O comportamento individual dos usuários da rede Twitter, estudado a partir do compartilhamento de mensagens textuais, foi analisado no trabalho de Silva et al. [2011], no qual eles conseguiram criar e avaliar um modelo para redes de influencia entre usuários baseados na relação seguidor/seguido.

Os resultados apresentados nesses trabalhos relacionados estão alinhados com o resultado da análise o comportamento coletivo dos usuários, reportado na Seção 4.5. O método apresentado neste capítulo permite que o comportamento coletivo dos usuários de Redes Sociais seja analisado. Entretanto, deve ser observado que nos trabalhos relacionados não é feita uma análise das interações sociais que têm relação com o comportamento desses usuários.

Barbosa et al. [2012] descrevem como os usuários da rede Twitter usam *hashtags* para expressar, classificar, e propagar as suas opiniões e sentimentos. Eles empregam perspectivas diferentes para verificar o uso e a propagação de padrões de *hashtags* que podem apoiar a avaliação e monitoramento de opinião da população online com relação a assuntos como eleições e epidemias. O resultado desse trabalho relacionado está alinhado com o reportado na Seção 4.2 no que se refere a análise da propagação de informações na Rede Social. As avaliações reportadas nesse trabalho relacionado incluem o levantamento de frequências da publicação de mensagens de texto. A ligação entre a publicação desses conteúdos e as interações entre usuários que são estabelecidas a partir dessas publicações não é explicitada. Em contraposição, no caso do resultado reportado na Seção 4.2, são utilizadas diferentes medidas de regras para analisar a propagação de informações baseadas em mídias do tipo vídeo e link.

De Choudhury et al. [2011] propuseram uma técnica de agrupamento iterativo para a seleção de um conjunto de itens sobre um determinado assunto que corresponde a um determinado nível de diversidade. Como resultado, eles discutem implicações para a seleção e avaliação de conjuntos de temas e de itens no

contexto de redes sociais. As diferenças entre usuários especialistas e novatos são exploradas por Liao et al. [2012] para inferir *expertise*. Em seus resultados eles encontraram que informações contextuais fornecem mais subsídios para realizar a inferência da *expertise* dos usuários a partir da informação contida em *tweets*.

Nesses trabalhos relacionados não são explicitamente representadas as ações que os usuários executam a partir do compartilhamento de uma mensagem. Em contraposição, no resultado da aplicação do método na análise da propagação de conteúdos baseados em mídia, reportados na Seção 4.2, são detalhadas as mídias compartilhadas e as ações que os usuários executam a partir desse compartilhamento.

4.6.2 Análise do Comportamento Social dos Usuários da rede Facebook

A análise do comportamento dos usuários da rede Facebook a partir do compartilhamento de diferentes tipos de mídias também é explorada na literatura.

Rabbath et al. [2010] desenvolveram uma abordagem que combina a análise de conteúdos baseados em texto e imagens para automaticamente e semi-automaticamente selecionar fotos de uma história específica. A partir dessa seleção, a abordagem por eles proposta permite a criação automática de álbuns de fotos dos usuários. Kramer [2012] estudou em grande escala o contágio emocional através de um exame de atualizações de status da rede Facebook, indicando que o contágio emocional é possível não apenas através de uma comunicação baseada em texto, mas também que se espalha através de meios indiretos de comunicação de emoção.

Esses trabalhos relacionados reportam resultados da análise de propagação de mídias na rede, respectivamente, foto e texto. Em contraste com o resultado apresentado na Seção 4.2 que trata do compartilhamento de vídeos e links na rede Facebook, esses trabalhos relacionados não reportam uma análise de como usuários interagem a partir do compartilhamento das mídias.

Moore and McElroy [2012] propuseram o uso de um modelo com cinco grandes fatores de personalidade, amplamente utilizado nas áreas de gestão e psicologia, para prever atitudes e comportamentos no Facebook. Seus resultados mostraram que os fatores de personalidade estudados podem ser usados para explicar a natureza das publicações e também o arrependimento na publicação de determinados conteúdos.

Os resultados apresentados nesses trabalhos relacionados estão alinhados com o resultado da aplicação do método no que se refere à análise do comportamento coletivo dos usuários, reportado na Seção 4.5. Entretanto, nesses trabalhos relacionados não são detalhados os elementos associados com as ati-

vidades dos usuários e também não são explicitadas as interações sociais nas quais os usuários do Facebook estão envolvidos, como é realizado nas aplicações do método guia apresentadas neste capítulo.

4.7 Considerações Finais

Neste capítulo foi apresentado um procedimento para análise de interações sociais baseadas em mídias e o método SIREMr. O método é uma extensão do procedimento de análise e também guia a aplicação da técnica apresentada no capítulo anterior. Foram reportadas as contribuições dos trabalhos [Gomes and Pimentel, 2012e], [Gomes and Pimentel, 2011b], [Gomes and Pimentel, 2012b], [Gomes and Pimentel, 2011a], [Gomes and Pimentel, 2012a] e [Gomes and Pimentel, 2012c].

A experimentação realizada com o procedimento de análise de interações sociais permite observar algumas preferências dos usuários no uso das opções de compartilhamento de mídia disponíveis dentro da Rede Social Online, e também, preferências no uso das opções de compartilhamento de mídias a partir de provedores de conteúdo para dentro da rede.

A aplicação do método com a extração manual de contingências comportamentais permite observar o impacto da inclusão de um novo tipo de mídia na rede como opção para interação entre os usuários. A aplicação da técnica e do método guia produziram resultados consistentes com os resultados de experimentações reportadas no capítulo anterior. A aplicação do método com a extração automática de contingências comportamentais permite analisar o comportamento coletivo dos usuários da rede Facebook, que é analisado em termos do compartilhamento de mídias e uso de aplicativos para o compartilhamento dessas mídias. Foi reportado um resultado que mostra a mudança de comportamento dos usuários em termos do compartilhamento de mídias e uso de aplicativos de terceiros.

No próximo capítulo são reportadas duas avaliações da técnica e do método reportados nesta tese. Uma avaliação foi realizada a partir de dados coletados de diferentes sub-redes da rede Facebook. A segunda avaliação é qualitativa e foi realizada com potenciais usuários pesquisadores de área de Computação com o propósito de levantar cenários nos quais a técnica e o método pudessem ser utilizados.

Avaliações da Aplicação da Técnica e do Método

As experimentações da técnica e do método apresentadas nos capítulos anteriores foram realizadas com dados coletados a partir de uma sub-rede da rede Facebook.

Neste capítulo são reportadas duas avaliações da técnica e do método. A primeira avaliação é resultado da experimentação do método com dados coletados a partir de cinco unidades de observação diferentes. Essa avaliação tem o propósito de verificar a consistência dos resultados reportados nos capítulos anteriores com resultados obtidos a partir da coleta dados dessas sub-redes.

A segunda avaliação da aplicação da técnica e do método é qualitativa e foi realizada na forma de entrevista com potenciais usuários pesquisadores da área de Computação. Esses usuários possuem conhecimento especializado para o desenvolvimento de aplicações sociais. Dado o conhecimento especializado desses usuários, foi pedido que eles descrevessem um cenário no qual a técnica e o método pudessem ser utilizados no desenvolvimento de alguma aplicação social.

Este capítulo está organizado da seguinte forma: na Seção 5.2 são relatados os resultados da experimentação da técnica e do método com dados colhidos de diferentes unidades de observação do Facebook; na Seção 5.3 são sumarizados os resultados obtidos da entrevista com engenheiros de software; trabalhos relacionados são apresentados na Seção 5.4 e, finalmente, na Seção 5.5 são apresentadas as considerações finais.

Tabela 5.1: *Informações sobre Perfil dos Potenciais Usuários*

Identificação	Idade	Sexo	# de Amigos na Rede	Formação em Computação
usr 1	28	masculino	137	doutorado incompleto
usr 2	27	masculino	156	mestrado completo
usr 3	25	masculino	325	mestrado incompleto
usr 4	28	feminino	126	doutorado incompleto
usr 5	30	masculino	94	doutorado incompleto

5.1 Recrutamento dos Potenciais Usuários para as Avaliações

O contato inicial com os potenciais usuários foi realizado por e-mail que foi enviado para engenheiros de software, pesquisadores da área de Computação, usuários do Facebook, que possuem conhecimento especializado para o desenvolvimento de aplicações sociais. Esses potenciais usuários foram selecionados do grupo de amigos (e também amigos de amigos) que o autor possui na rede Facebook. A solicitação de participação na avaliação foi respondida por 15 potenciais usuários, dos quais 5 aceitaram disponibilizar o acesso a suas contas para captura de dados, e também, aceitaram participar de uma entrevista.

Na Tabela 5.1 são sumarizadas algumas informações sobre cada voluntário. Pode ser observado que os voluntários têm idade entre 25 e 30 anos. Segundo Holdings [2012], indivíduos na faixa etária de 18 até 34 anos são usuários mais ávidos de mídias sociais, portanto, dados colhidos de indivíduos nessa faixa etária são mais representativos para análise de redes sociais.

Foram realizadas duas avaliações da técnica e do método reportados nesta tese. A primeira avaliação consiste de uma experimentação da técnica e do método realizada a partir de dados colhidos da sub-rede de cada voluntário. O objetivo dessa avaliação é verificar a consistência da aplicação da técnica e do método com resultados reportados nesta tese. A segunda avaliação consiste de uma entrevista com cada potencial usuário após a apresentação dos resultados da aplicação da técnica e do método. O objetivo dessa entrevista é obter a descrição de um cenário no qual a técnica e o método possam ser utilizados. Na próxima seção são reportados os resultados da primeira avaliação.

5.2 Avaliação da Técnica e do Método com Dados Coletados de Diferentes Unidades de Observação do Facebook

Na avaliação do método apresentada nesta seção foram coletados dados de interação a partir de cinco unidades de observação da rede Facebook. Na Tabela 5.1 são sumarizadas informações sobre as unidades de observação utilizadas na experimentação reportada nesta seção.

Tabela 5.2: *Informações sobre as Unidades de Observação: Avaliação da Técnica e do Método*

Identificação	Período de Coleta	# de Usuários	# de Contingências
UOb 1	Março de 2012	137	5.685
UOb 2	Março de 2012	156	6.683
UOb 3	Março de 2012	325	8.194
UOb 4	Março de 2012	126	6.186
UOb 5	Março de 2012	94	4.189
UOb	Março de 2012	838	39.947

Uma implementação do método apresentado no capítulo anterior foi fornecida como um conjunto de scripts em *Python*. As unidades de observação acessadas estão na mesma área geográfica (Brasil). Cada usuário da rede que permitiu a coleta de dados recebeu apoio do autor para executar os scripts em sua máquina pessoal, em um ambiente onde eles se sentiam mais confortáveis.

Os scripts equivalentes a cada etapa do método foram executados com sucesso e, em seguida, os resultados foram preparados para apresentação para cada voluntário em uma sessão de entrevista individual. Os resultados dessas entrevistas são reportados na próxima Seção. O planejamento da avaliação realizada com esses potenciais usuários do método é apresentado no Apêndice B. Depois de apresentar os resultados da aplicação do método, a entrevista foi realizada seguindo o roteiro apresentado no Apêndice C.

5.2.1 Observação e Captura

A observação da rede Facebook para a captura dos dados de interação foi realizada uma única vez no mês de Março de 2012. Foram coletadas informações de um total de 30.947 interações sociais, incluindo, o tipo de mídia publicada (*checkin*, *photo*, *status*, *video*, *link*, *swf*, *question* e *music*), informações sobre as atividades do usuário, tais como a execução da ação de *Curti* e *Comentar* uma publicação, e o número de usuários que executaram ambas ações.

5.2.2 Preparação dos Dados

O mapeamento de ações, consequências e usuários (atores das ações) é o mesmo apresentado anteriormente na Tabela 4.10 da Seção 4.5, que é replicada na Tabela 5.3. Um conjunto de dados foi construído a partir de cada unidade de observação, que foram denominados respectivamente de UOb1, UOb2, UOb3, UOb4 e UOb5. Esses conjuntos de dados foram unidos em um único, denominado de UOb, a partir do qual foram extraídas automaticamente as interações sociais.

Tabela 5.3: *Ações, Consequências e Usuários (replicação da Tabela 4.10)*

Ações	
A_1	ação de criar um Post
A_2	ação de Curtir um Post
A_3	ação de Comentar um Post
A_4	ação de Compartilhar um Post
Consequências	
C_1	ser notificado de uma ação de criar um Post
C_2	ser notificado de uma ação de Curtir um Post
C_3	ser notificado de uma ação de Comentar um Post
C_4	ser notificado de uma ação de Compartilhar um Post
Usuários	
a	usuário que executam a ação de criar um Post
l	usuários que executam a ação de Curtir
k	usuários que executam a ação de Comentar
s	usuários que executam a ação de Compartilhar
$\bar{l}$	usuários que não executam a ação de Curtir
$\bar{k}$	usuários que não executam a ação de Comentar
$\bar{s}$	usuários que não executam a ação de Compartilhar

5.2.3 Representação e Extração de Interações Sociais Genéricas

Utilizando-se uma implementação do algoritmo descrito na Seção 3.3, foram extraídas as interações sociais genéricas apresentadas na Listagem 5.1.

R1.	$A_1 \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow alksC_1$
R2.	$A_1 \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow alksC_1 \cap alksC_2$
R3.	$A_1 \cap \bar{l}A_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow alksC_1 \cap alksC_3$
R4.	$A_1 \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow alksC_1 \cap alksC_2 \cap alksC_3$
R5.	$A_1 \cap \bar{l}A_2 \cap \bar{k}A_3 \cap sA_4 \rightarrow alksC_1 \cap alksC_4$
R6.	$A_1 \cap lA_2 \cap \bar{k}A_3 \cap sA_4 \rightarrow alksC_1 \cap alksC_2 \cap alksC_4$
R7.	$A_1 \cap \bar{l}A_2 \cap kA_3 \cap sA_4 \rightarrow alksC_1 \cap alksC_3 \cap alksC_4$
R8.	$A_1 \cap lA_2 \cap kA_3 \cap sA_4 \rightarrow alksC_1 \cap alksC_2 \cap alksC_3 \cap alksC_4$

Listagem 5.1: *Interações Sociais Genéricas Recorrentes nas Unidades de Observação*

Essas interações sociais podem ser descritas como declarações *if-then* da seguinte forma:

- R1. *if* usuário a executa a ação A_1 e os usuários l não executam a ação A_2 e os usuários k não executam a ação A_3 e os usuários s não executam a ação A_4 (usuário a fornece o estímulo social que não recebe *Curtir*, nem recebe *Comentar* e nem recebe *Compartilhar*)
then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem apenas a notificação C_1 . Em resumo, são interações que envolvem apenas a publicação de conteúdos sem a execução de ações nessas publicações ([P]).

- R2. *if* usuário a executa a ação A_1 e os usuários l executam a ação A_2 e os usuários k não executam a ação A_3 e os usuários s não executam a ação A_4 (usuário a fornece o estímulo social que recebe *Curtir*, não recebe *Comentar* e não recebe *Compartilhar*)
then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 e C_2 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Curtir* apenas ($\{P,L\}$).
- R3. *if* usuário a executa a ação A_1 e os usuários l não executam a ação A_2 e os usuários k executam a ação A_3 e os usuários s não executam a ação A_4 (usuário a fornece o estímulo social que não recebe *Curtir*, recebe *Comentar* e não recebe *Compartilhar*)
then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 e C_3 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Comentar* apenas ($\{P,K\}$).
- R4. *if* usuário a executa a ação A_1 e os usuários l executam a ação A_2 e os usuários k executam a ação A_3 e os usuários s não executam a ação A_4 (usuário a fornece o estímulo social que recebe *Curtir*, recebe *Comentar* e não recebe *Compartilhar*)
then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 e C_2 e C_3 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Curtir* e *Comentar* ($\{P,L,K\}$).
- R5. *if* usuário a executa a ação A_1 e os usuários l não executam a ação A_2 e os usuários k não executam a ação A_3 e os usuários s executam a ação A_4 (usuário a fornece o estímulo social que não recebe *Curtir*, não recebe *Comentar* e recebe *Compartilhar*)
then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 e C_4 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Compartilhar* apenas ($\{P,S\}$).
- R6. *if* usuário a executa a ação A_1 e os usuários l executam a ação A_2 e os usuários k não executam a ação A_3 e os usuários s executam a ação A_4 (usuário a fornece o estímulo social que recebe *Curtir*, não recebe *Comentar* e recebe *Compartilhar*)
then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 , C_2 e C_4 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Curtir* e *Compartilhar* ($\{P, L, S\}$).

Figura 5.1: Suporte das Interações Sociais Genéricas das Unidades de Observação

R7. *if* usuário a executa a ação A_1 e os usuários l não executam a ação A_2 e os usuários k executam a ação A_3 e os usuários s executam a ação A_4 (usuário a fornece o estímulo social que não recebe *Curtir*, recebe *Comentar* e recebe *Compartilhar*)

then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 , C_3 e C_4 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Comentar* e *Compartilhar* ({P, K, S}).

R8. *if* usuário a executa a ação A_1 e os usuários l executam a ação A_2 e os usuários k executam a ação A_3 e os usuários s executam a ação A_4 (usuário a fornece o estímulo social que recebe *Curtir*, recebe *Comentar* e recebe *Compartilhar*)

then usuário a e grupo de usuários l , k e s (usuário a e seus amigos) recebem as notificações C_1 e C_2 , C_3 e C_4 . Em resumo, são interações que envolvem publicação de conteúdos com a execução de ações *Curtir*, *Comentar* e *Compartilhar* ({P, L, K, S}).

As interações sociais apresentadas na Listagem 5.1 foram avaliadas utilizando cada conjunto de dados obtido de cada unidade de observação.

5.2.4 Avaliação, Validação e Interpretação de Interações Sociais Genéricas

As interações sociais genéricas foram por meio do computo das medidas Confidência (Equação 3.2), Suporte (Equação 3.1), Correlação Cosseno (Equação 3.3). A validação das regras extraídas foi realizada de forma indireta, ou seja, a partir de validações anteriormente realizadas por um especialista em Psicologia Social.

O comportamento coletivo dos usuários no engajamento em interações sociais genéricas pode ser analisado a partir dos valores da medida de Suporte apresentados na Figura 5.1. Na Figura 5.1A são apresentados os valores da medida de

Suporte obtidos a partir da avaliação do conjunto de regras da Listagem 5.1 com cada conjunto de dados. Pode ser observado que o comportamento coletivo dos usuários é muito semelhante. Na Figura 5.1B são apresentados os valores da medida de Suporte obtidos a partir da avaliação do conjunto de regras da Listagem 5.1 com o conjunto de dados único contendo dados de todas as unidades de observação. As medidas de Confidência e Correlação tiveram valor 100%.

O comportamento coletivo dos usuários corresponde ao ranking $R4 = \{P, L, K\}$, $R1 = \{P\}$, $R2 = \{P, L\}$, $R3 = \{P, K\}$, $R8 = \{P, L, K, S\}$, $R5 = \{P, S\}$, $R6 = \{P, L, S\}$ e $R7 = \{P, K, S\}$. A partir do ranking dessas regras pode ser interpretado que, primeiro, são mais frequentes as publicações dos usuários que recebem *Curtir* e *Comentar*; em seguida, são frequentes as publicações não recebem nem *Curtir*, *Comentar* e nem *Compartilhar*; em seguida, são frequentes as publicações recebem *Curtir* apenas, e assim por diante.

É interessante observar que os usuários interagem com mais frequência executando as ações *Curtir* e *Comentar* após uma publicação ($Sup(R4) = 39\%$). Pode ser observado também é que uma percentagem substancial das publicações não provocam o engajamento dos usuários por meio da execução de ações ($Sup(R1) = 31\%$). Esse resultado é consistente com resultados obtidos, por exemplo, na análise das interações sociais genéricas vinculadas com a movimentação social contra corrupção no Brasil [Gomes and Pimentel, 2012d], [Gomes and Pimentel, 2012b].

Também é interessante observar que os usuários interagem pouco por meio da execução da ação *Compartilhar* ($Sup(R5)$, $Sup(R6)$, $Sup(R7)$ e $Sup(R8)$ próximos a 0%). Esse resultado é consistente com resultados anteriores que utilizam a técnica de representação e avaliação [Gomes and Pimentel, 2011c].

Na continuidade da aplicação da avaliação do método, as interações sociais genéricas foram especializadas em interações sociais baseadas em mídia.

5.2.5 Especialização

Uma nova lista de atributos foi elaborada com o propósito de especializar a ação A_1 com informações sobre o tipo de mídia compartilhada nas interações sociais. A Figura 5.2 mostra a frequência das mídias compartilhadas e coletadas dos conjuntos de dados UOb1, UOb2, UOb3, UOb4 e UOb5. A partir dessa especialização, a representação e extração de interações sociais baseadas em mídia foram realizadas.

5.2.6 Extração de Interações Sociais Baseadas em Mídia

A ação A_1 foi especializada em cada conjunto de dados e, em seguida, foram extraídas as interações sociais baseadas em mídia, utilizando-se o algoritmo apresentado na Seção 3.3. As interações baseadas em mídias mais frequentes

Figura 5.2: *Frequências das Mídias Compartilhadas obtidas do Conjunto de Dados UOb*

são apresentadas na Listagem 5.2.

- R1.1 $\{status\} : aA_1.status \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1$
 R1.2 $\{status, K\} : aA_1.status \cap \bar{l}A_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_3$
 R1.3 $\{status, L\} : aA_1.status \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2$
 R1.4 $\{status, L, K\} : aA_1.status \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
- R2.1 $\{photo\} : aA_1.photo \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1$
 R2.2 $\{photo, L\} : aA_1.photo \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2$
 R2.3 $\{photo, L, K\} : aA_1.photo \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
 R2.4 $\{photo, K\} : aA_1.photo \cap \bar{l}A_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_3$
 R2.5 $\{photo, L, K, S\} : aA_1.photo \cap lA_2 \cap kA_3 \cap sA_4 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3 \cap aklC_4$
- R3.1 $\{video\} : aA_1.video \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1$
 R3.2 $\{video, L\} : aA_1.video \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2$
 R3.3 $\{video, K\} : aA_1.video \cap \bar{l}A_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_3$
 R3.4 $\{video, L, K\} : aA_1.video \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
 R3.5 $\{video, L, S\} : aA_1.video \cap lA_2 \cap \bar{k}A_3 \cap sA_4 \rightarrow aklC_1 \cap aklC_2 \cap aklC_4$
- R4.1 $\{link\} : aA_1.link \cap \bar{l}A_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1$
 R4.2 $\{link, L\} : aA_1.link \cap lA_2 \cap \bar{k}A_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2$
 R4.3 $\{link, K\} : aA_1.link \cap \bar{l}A_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_3$
 R4.4 $\{link, L, K\} : aA_1.link \cap lA_2 \cap kA_3 \cap \bar{s}A_4 \rightarrow aklC_1 \cap aklC_2 \cap aklC_3$
 R4.5 $\{link, K, S\} : aA_1.link \cap \bar{l}A_2 \cap kA_3 \cap sA_4 \rightarrow aklC_1 \cap aklC_3 \cap aklC_4$

Listagem 5.2: *Interações Sociais Baseada em Mídias para Avaliação com Usuários*

Essas interações sociais baseadas em mídia foram avaliadas utilizando os conjuntos de dados UOb1, UOb2, UOb3, UOb4 e UOb5.

Figura 5.3: Suporte das Interações Baseadas em Mídias obtidas das Unidades de Observação

5.2.7 Avaliação das Interações Sociais baseadas em Mídia

Na avaliação das interações sociais apresentadas na Listagem 5.2 foram computadas as medidas de Suporte, Confidência e Correlação Cosseno. A partir dos maiores valores para os menores valores de cada medida, um ranqueamento das interações sociais baseadas em mídias pode ser elaborado. Na Figura 5.3 são apresentados os valores da medida de Suporte. O valor da medida de Confidência foi 100%.

Na Figura 5.4 são apresentados os valores da medida de Correlação Cosseno. Pode ser observado que o comportamento coletivo dos usuários avaliado pela medida de Suporte é muito semelhante. A avaliação do comportamento coletivo dos usuários pela medida de Correlação Cosseno também é muito semelhante.

5.2.8 Validação e Interpretação das Interações Sociais baseadas em Mídia

A Figura 5.3A mostra os valores da medida de Suporte para interações sociais baseadas em mídia tipo *status* (texto) e, a partir desses valores, é possível estabelecer o ranking $R1.4\{status, L, K\}$, $R1.1\{status\}$, $R1.3\{status, L\}$, $R1.2\{status, K\}$. Esse ranking é o mesmo para interações avaliadas utilizando-se os conjuntos de dados UOb1, UOb2, UOb3, UOb4 e UOb5.

A Figura 5.3B mostra os valores da medida de Suporte para interações sociais baseadas em mídia tipo *photo* e, a partir desses valores, é possível es-

Figura 5.4: Correlação Cosseno das Interações Baseadas em Mídias obtidas das Unidades de Observação

tabelecer o ranking $R1.4\{photo, L, K\}$, $R1.1\{photo\}$, $R1.3\{photo, L\}$, $R1.2\{photo, K\}$ e $R1.2\{photo, L, K, S\}$. Esse ranking é o mesmo para interações avaliadas utilizando-se os conjuntos de dados UOb1, UOb2, UOb3, UOb4 e UOb5.

É interessante observar que o resultado mostrado na Figura 5.3 é consistente com o resultado apresentado na Figura 4.4 (topo) onde pode ser observada a frequência de compartilhamento das mídias no Facebook no 1º semestre de 2012.

Na comparação dessas duas figuras, pode ser percebido que houve um aumento na frequência de compartilhamento de mídia *photo* em contraste com o decréscimo na frequência de compartilhamento de mídia *status*. Esse aumento no compartilhamento de mídia *photo* também é acompanhado do engajamento dos usuários por meio da execução das ações *Curtir* e *Compartilhar*. Esse resultado é consistente com resultados reportados na Seção 4.5.

5.2.9 Síntese das Observações

A experimentação reportada nesta seção apresenta resultados que foram obtidos a partir de conjuntos de dados coletados do Facebook, no mês de Março de 2012, com a finalidade de avaliar a aplicação da técnica e do método a partir de dados coletados de diferentes unidades de observação do Facebook.

Foi possível observar que a aplicação da técnica e do método é consistente

com resultados anteriores, por exemplo, os resultados reportados na Seção 3.5.1 que tratam de interações sociais genéricas, os resultados reportados na Seção 3.5.2 que tratam de interações sociais baseadas em mídia.

5.3 Avaliação Qualitativa com Potenciais Usuários

Os resultados obtidos a partir da avaliação da técnica e do método com dados coletados de diferentes unidades de observação foram apresentados para cada usuário que permitiu a coleta de dados. Após a apresentação dos resultados da avaliação da técnica e do método, foi realizada uma avaliação qualitativa na forma de uma entrevista com cada voluntário. Dado o conhecimento especializado dos voluntários, foi solicitado que eles descrevessem um cenário no qual a técnica e o método pudessem ser utilizados.

A entrevista foi realizada em um ambiente que os voluntários escolheram e se sentiam mais confortáveis, seguindo o roteiro apresentado no Apêndice C. Apesar de suas limitações, as entrevistas são um método valioso para a pesquisa exploratória com usuários [Nielsen, 2010], [Lazar et al., 2010]. A seguir é apresentada uma compilação dos resultados da avaliação.

5.3.1 Compilação dos Resultados das Sessões de Entrevista

Após serem apresentados os dados brutos coletados, os usuários foram consultados se sentiam-se incomodados com o conteúdo dos dados. Todos os entrevistados responderam que não, no entanto, todos ressaltaram que o sigilo na divulgação dessas informações deve ser respeitado.

Foi mostrado como informações coletadas são organizados para que os elementos associados à linguagem de Mechner fossem representados em conjuntos de dados. Em seguida foram extraídas as contingências comportamentais no formato de regras *if-then* de forma automática. Os entrevistados foram consultados se conseguiam ler regras no formato em que foram extraídas e interpretar como as interações sociais ocorriam a partir da leitura das regras. Todos responderam que conseguiam ler e interpretar cada regra.

O usuário *usr4* sugeriu que, além da representação no formato regras *if-then*, uma simplificação dessa representação também fosse elaborada. Por exemplo, a interação *R1* da Listagem 5.1 está representada na forma simplificada como {P}, a interação *R2* está representada na forma simplificada como {P,L}, e assim por diante. Essa sugestão foi acatada e reportada aos demais entrevistados, que concordaram com a representação simplificada.

Em seguida, foi apresentada a especialização das interações genéricas em interações baseadas em mídia. Os resultados da avaliação dessas interações foram apresentados e discutidos com os entrevistados. No passo seguinte, foram

Tabela 5.4: *Pergunta 3 e Respostas de Cada Potencial Usuário*

Pergunta 3	Qual a sua interpretação dos resultados das medidas de avaliação Suporte, Confidência e Correlação Cosseno?
Resposta usr 1	Entendi que a medida de Suporte é utilizada para contar frequências das regras. A medida de Confidência é utilizada para medir o acerto de uma regra. A medida de Correlação Cosseno é parecida com a medida de Suporte.
Resposta usr 2	A medida de Suporte é utilizada para contar frequências das interações. A Confidência mede o acerto de uma regra, portanto, de uma interação. Não consegui interpretar a medida de Correlação Cosseno.
Resposta usr 3	A contagem das frequências das interações é realizada por meio da utilização da medida de Suporte. A medida de Confidência é utilizada para medir o quanto uma regra acerta ou erra no encaixe aos dados. Não consegui interpretar a medida de Correlação Cosseno, pois ela produz um resultado parecido com a medida de Suporte. As medidas de Suporte e Confidência podem ser suficientes para a compreensão dos resultados da avaliação.
Resposta usr 4	Entendo que a medida de Suporte é utilizada para contar frequências das interações sociais. A medida de Confidência pode ser utilizada, por exemplo, para detectar se cada ação tem sua respectiva notificação para o usuário. Se não houve, o resultado é diferente de 100%. A medida de Correlação Cosseno parece não ser necessária para a compreensão dos resultados.
Resposta usr 5	Percebi que a medida de Suporte é utilizada para contar frequências. A medida de Confidência é utilizada para medir o quanto uma regra se encaixa nos dados. Se o resultado dessa medida for diferente de 100%, a Confidência pode ser usada para identificar algum tipo de inconsistência no projeto ou na implementação dos artefatos que permitem as interações sociais. A medida de Correlação Cosseno tem relação com a medida de Suporte. Acredito que as medidas de Suporte e Confidência são suficientes para o entendimento dos resultados da avaliação.

realizadas as perguntas 3 até 8.

As perguntas 3 e 4 e suas respectivas respostas são apresentadas respectivamente na Tabela 5.4 e na Tabela 5.5. As perguntas 5 e 6 e suas respectivas respostas são apresentadas respectivamente na Tabela 5.6 e na Tabela 5.7. As perguntas 7 e 8 e suas respectivas respostas são apresentadas respectivamente na Tabela 5.8 e na Tabela 5.9. Uma síntese dos resultados dessa entrevista é reportada na próxima seção.

A última pergunta da entrevista foi realizada com o propósito de que os usuários descrevessem um cenário no qual a técnica e o método pudessem ser utilizados. A pergunta realizada foi:

Pergunta 9: É possível descrever sucintamente um cenário no qual uma aplicação social possa ser desenvolvida a partir aplicação da técnica e do método?

Tabela 5.5: *Pergunta 4 e Respostas de Cada Potencial Usuário*

Pergunta 4	A representação explícita de informações como tipos de mídias, provedores dessas mídias, e aplicações utilizadas nas interações sociais é útil? Como?
Resposta usr 1	É útil.
Resposta usr 2	Sim. Essa representação é útil para compreender como esses elementos são compartilhados nas interações sociais.
Resposta usr 3	Sim. É útil.
Resposta usr 4	Sim, são úteis e interessantes para entender como as interações sociais ocorrem dentro do ambiente.
Resposta usr 5	Acredito que são úteis e podem ser utilizadas para facilitar a interpretação de como esses elementos são compartilhados nas interações sociais.

Tabela 5.6: *Pergunta 5 e Respostas de Cada Potencial Usuário*

Pergunta 5	O usuário tinha algum tipo de expectativa em relação ao método? Qual?
Resposta de todos potenciais usuários	Não.

Tabela 5.7: *Pergunta 6 e Respostas de Cada Potencial Usuário*

Pergunta 6	O usuário tinha algum tipo de expectativa em relação aos resultados? Qual?
Resposta usr 1	Não, mas os resultados são interessantes.
Resposta usr 2	Não. A representação e a avaliação das interações sociais são bem interessantes.
Resposta usr 3	Não, mas chama a atenção a forma como é feita a representação, a extração e a avaliação.
Resposta usr 4	Não. O processo como um todo é interessante.
Resposta usr 5	Não. Todo o processo de representação, de extração e de avaliação é interessante.

Tabela 5.8: *Pergunta 7 e Respostas de Cada Potencial Usuário*

Pergunta 7	A técnica e o método podem ser utilizados no desenvolvimento de aplicações?
Resposta usr 1	Sim. Por exemplo, a técnica e o método podem ser utilizados para avaliar uma aplicação social.
Resposta usr 2	Sim.
Resposta usr 3	Acredito que sim.
Resposta usr 4	Sim. Por exemplo, a técnica e o método podem ser utilizados em uma etapa de avaliação com um protótipo já desenvolvido ou com a aplicação já pronta.
Resposta usr 5	Sim. A técnica e o método podem ser utilizados, por exemplo, no direcionamento dos esforços voltados para aprimoramento e melhoria de características sociais da aplicação.

Tabela 5.9: *Pergunta 8 e Respostas de Cada Potencial Usuário*

Pergunta 8	Na sua opinião, quais profissionais podem ter interesse no uso desses procedimentos para avaliação de interações sociais?
Resposta usr 1	Desenvolvedores de aplicações.
Resposta usr 2	Engenheiros de software e analistas de sistemas.
Resposta usr 3	Usuários especialistas com interesse em análise de Redes Sociais.
Resposta usr 4	Profissionais da área de Computação e Sociologia que trabalham com Redes Sociais.
Resposta usr 5	Usuários especialistas tais como tomadores de decisão, os analistas de mídias sociais, cientistas sociais e analistas de comportamento em ambientes sociais baseados em computador, podem usar os procedimentos.

Dentre os cinco participantes, apenas três conseguiram elaborar uma descrição de um possível cenário:

Cenário 1. Descrito por urs 1: Os usuários de aplicações sociais são muito egocêntricos. Uma ferramenta que forneça ranqueamento da participação individual de cada usuário em comparação com seus amigos pode ser muito útil. Por exemplo, para cada tipo de interação social, um usuário individual poderia saber quais são seus amigos que participaram com mais frequência em um determinado tipo de interação. Em contrapartida, esse usuário individual poderia saber também quais os tipos de interações sociais que ele mais participa.

Cenário 2. Descrito por usr 4: Uma aplicação que permita a avaliação da sociabilidade, do ponto de vista do engajamento dos usuários nas interações sociais, pode ser construída a partir da técnica e do método. Por exemplo, uma ferramenta poderia oferecer o ranqueamento das interações sociais extraídas de um determinado ambiente social *on the fly*, ou seja, à medida que os usuários usam o ambiente, os valores das medidas são automaticamente atualizados. Essa ferramenta poderia permitir ainda que seus usuários pudessem especializar cada tipo de interação de acordo com um conjunto de possibilidades pré-definidas, por exemplo, as interações sociais genéricas poderiam ser visualizadas a partir da sua especialização com foto, vídeo ou todas as opções de mídia compartilhadas no ambiente; sem prejuízo para a visualização do engajamento dos usuários pela execução de ações a partir da publicação dessas mídias.

Cenário 3. Descrito por usr 5: Usuários especialistas tais como tomadores de decisão, analistas de mídias sociais, cientistas sociais e analistas de comportamento em ambientes sociais baseados em computador poderiam usar uma ferramenta que disponibilizasse o ranqueamento das regras com um tempo de intermitência relativamente curto, permitindo assim que esses usuários possam avaliar estratégias de divulgação de mídias em redes sociais, quais são os usuários ou grupos de usuários mais engajados nas interações sociais a partir da publicação de conteúdos específicos ou até mesmo se um produto lançado recentemente está causando boa impressão entre os seus compradores e potenciais clientes.

A seguir é apresentada uma síntese dos principais resultados obtidos com a entrevista dos potenciais usuários.

5.3.2 Síntese dos Principais Resultados da Entrevista

A entrevista com os potenciais usuários foi importante para que questões relevantes fossem avaliadas com relação a aplicação da técnica e do método apresentados nesta tese.

Uma questão relevante diz respeito ao conteúdo das informações coletadas.

Os usuários não se sentiram incomodados, entretanto, se mostraram preocupados com o sigilo das informações.

Outra questão importante refere-se à interpretação das interações sociais descritas no formato de regras *if-then*. Os usuários participantes da avaliação conseguiram interpretar como as interações sociais ocorrem a partir da leitura das regras. Em seguida, eles conseguiram fazer uma sugestão para uma simplificação da representação das interações sociais.

A interpretação de cada medida utilizada na avaliação das interações sociais também foi avaliada. Os usuários apontaram suas interpretações e também apontaram como cada medida pode ser usada.

Em resumo, os usuários responderam que a medida de Suporte é uma medida de contagem de frequência das interações, portanto, não interpretaram de forma diferente ao realizado na apresentação dos resultados. A medida de Confiança é uma medida do quanto uma regra acerta ou se encaixa nos dados, que pode ser utilizada para detectar algum tipo de inconsistência ou falha no design das interações, quando os resultados dessa medida forem diferentes de 100%. Por exemplo, essa medida pode ser útil para detectar se cada ação tem sua respectiva notificação para o usuário. Com relação a medida de Correlação Cosseno, houve alguma dificuldade para interpretar essa medida. Do ponto de vista dos entrevistados, as medidas de Suporte e Confidência são adequadas para a interpretação e compreensão dos resultados da avaliação apresentada.

Ao final da entrevista, em uma conversa informal, os usuários foram convidados a manifestar suas impressões gerais sobre os resultados da aplicação da técnica e do método. Em resumo, os voluntários manifestaram que:

- A técnica e do método fornecem uma representação legível de como as interações sociais ocorrem, permitindo que o comportamento dos usuários seja determinado a partir da avaliação de regras associadas a interações sociais.
- A partir da contagem de ações e notificações, a técnica e o método permitem a identificação de situações sociais nas quais os usuários estão envolvidos, relacionando ações executadas pelos usuários e notificações do sistema. A explicitação de mídias, por exemplo, permite uma análise detalhada de como as situações sociais são estabelecidas na Rede Social Online;
- No desenvolvimento de aplicações sociais, os resultados podem impactar diretamente sobre o foco dos esforços de trabalho, por exemplo, permitindo que inconsistências no projeto ou na implementação dos artefatos que permitem as interações sociais sejam identificadas;

Novamente os participantes destacaram que a especialização de interações genéricas em interações baseadas em mídia é uma estratégia interessante. Foram confirmados os apontamentos quanto a legibilidade da representação de interações sociais, contagem de situações sociais que envolvem ações dos usuários e notificações do sistema, impacto no design e no desenvolvimento de aplicações sociais.

Três cenários foram apontados pelos participantes da entrevista nos quais aplicações sociais podem ser desenvolvidas a partir da utilização da técnica e do método.

5.4 Trabalhos Relacionados

Os esforços de pesquisa relacionados com a investigação apresentada neste capítulo contemplam a avaliação de aspectos sociais (sociabilidade) em ambientes que permitem a interação entre usuários, e ainda, a participação dos usuários na avaliação de Redes Sociais Online.

5.4.1 Avaliação da Sociabilidade de Aplicações

Ambientes de computação que permitem interações entre os usuários são avaliados usando heurísticas e diretrizes para inspeção sociabilidade. No contexto de avaliações de aplicações de TV sociais, Geerts and De Grooff [2009] propuseram heurísticas e diretrizes de apoio à avaliação de habilidades sociais em sistemas de computador. Questões relacionadas à sociabilidade são também importantes para o design de aplicativos que lidam com as mídias sociais no cenário da TV móvel, como apresenta Chorianopoulos [2010].

Métodos de inspeção de sociabilidade não permitem uma avaliação quantitativa das interações sociais, não fornecem uma representação explícita de situações sociais nas quais os usuários estão envolvidos, e também, não fornecem uma representação explícita de quais ações são executadas pelos usuários após a identificação de uma situação social. A técnica e o método apresentados nesta tese podem contribuir para a superação dessas limitações. Um trabalho futuro envolve a investigação da aplicação da técnica e do método na inspeção da sociabilidade de aplicações sociais.

5.4.2 Participação dos Usuários na Avaliação de Redes Sociais

A participação dos usuários na avaliação de ambientes que permitem interações sociais é reportada em alguns trabalhos.

Benevenuto et al. [2009a] e De Choudhury et al. [2009] mostram que a participação voluntária de usuários em conversações pode ser explicada a partir de fatores intrínsecos (consciência social, características da comunidade e reputação do criador) e fatores extrínsecos (contexto de mídia e interessabilidade de con-

versação) que permitem a compreender a estrutura e a evolução da rede Twitter. Os usuários que participaram das conversações analisadas foram consultados para que pudessem ser determinados os aspectos relacionados a consciência social, características da comunidade e reputação. Lin et al. [2012] apresentaram um novo algoritmo para descobrir estruturas multirrelacionais de fluxos de informações relacionadas a tempo, usuários e ações em Redes Sociais Online. Em um estudo qualitativo com grupos de usuários, foram avaliados os benefícios das estruturas relacionais descobertas.

Em contraposição a esses estudos, que estão concentrados na análise da estrutura e a evolução da Rede Social, o trabalho reportado na Seção 5.2 tem o foco na análise de aspectos relacionados a interação entre usuários. Na Seção 5.3, potenciais usuários mencionaram que a técnica e o método reportados nesta tese podem ser utilizados no desenvolvimento de aplicações sociais.

A utilização de outras técnicas de avaliação com a participação de usuários de Redes Sociais é reportada, por exemplo, no trabalho de Xu et al. [2012] que propuseram um modelo baseado na teoria de gratificações para examinar o uso de Redes Sociais Online a partir de um estudo com foco em grupos. Outras técnicas de avaliação com usuários de aplicações sociais podem ser empregadas em trabalhos futuros para a avaliação da aplicação da técnica e do método reportados nesta tese.

5.5 Considerações Finais

Neste capítulo foram apresentados resultados da experimentação do método *SI-REMr* na análise de dados coletados a partir de diferentes unidades de observação da rede Facebook. Também foram reportados resultados da avaliação da técnica e do método com potenciais usuários por meio de uma entrevista individualizada. As contribuições apresentadas neste capítulo foram reportadas no trabalho [Gomes and Pimentel, 2012f] que está em avaliação.

A consistência com resultados apresentados nos Capítulos 3 e 4 foi verificada a partir dos resultados obtidos na experimentação da técnica e do método com dados de diferentes unidades de observação da rede Facebook.

A entrevista com os potenciais usuários pesquisadores da área de Computação foi importante para que questões relevantes fossem avaliadas com a relação a técnica e o método apresentados nesta tese. Questões importantes foram destacadas em relação a interpretação das interações sociais descritas no formato de regras *if-then*, e ainda, à interpretação de cada medida utilizada na avaliação das interações sociais.

Os potenciais usuários realizaram importantes observações quanto a legibilidade da representação de interações sociais, contagem de situações sociais

que envolvem ações dos usuários e notificações do sistema, impacto no projeto e desenvolvimento de interações sociais em aplicações sociais, e ainda, quais usuários especialistas podem ter interesse no uso da técnica e do método.

Três cenários foram descritos por engenheiros de software que avaliaram qualitativamente a aplicação do método e da técnica, mostrando as potencialidades no uso das contribuições apresentadas nesta tese para o desenvolvimento de aplicações sociais.

Conclusão

As pesquisas científicas que refletem o estado da arte no estudo das atividades dos usuários de Redes Sociais Online tratam de temas como propriedades estruturais das OSNs, interações baseadas em mídias, interações baseadas em aplicativos de terceiros, processamento e análise do comportamento social dos usuários, ligações sociais ativadas pelo uso de dispositivos móveis, técnicas de mineração de dados voltadas para a análise relações sociais, engajamento de usuários em mobilizações cívicas online, e ainda, análise de interações sociais com a avaliação dos usuários.

Na investigação desses temas de pesquisa, um modelo de representação da atividade dos usuários amplamente empregado é baseado na construção de um grafo de interação. A construção de um grafo tem foco na representação de uma interação específica, que é analisada a partir da interpretação de medidas baseadas em grafos e também de medidas estatísticas. Técnicas de mineração de dados podem ser empregadas de forma alternativa e complementar ao modelo baseado em grafos, possibilitando a extração e a avaliação de padrões das atividades dos usuários. Entretanto, tanto na análise baseada em grafos quanto na utilização de técnicas de mineração de dados, não são explicitamente representados elementos associados com a atividade dos usuários como ações executadas, mídias compartilhadas, aplicações e tipos de dispositivos utilizados.

Nesta tese são propostos uma técnica e um método que permitem a representação, a extração e a avaliação das interações entre usuários de Redes Sociais Online. A técnica e o método exploram a representação explícita dos elementos associados com a atividade dos usuários, e também, permitem a extração de

um padrão da atividade dos usuários não somente a partir da análise de uma interação específica.

6.1 Síntese das Contribuições

As principais contribuições desta tese são uma técnica e um método para a representação, a extração e a avaliação de interações sociais nas quais os elementos associados com as atividades dos usuários em Redes Sociais Online são explicitamente tratados. Essas contribuições foram alcançadas mediante sucessivas experimentações com dados coletados do Facebook. Essas contribuições foram desenvolvidas de forma incremental:

- A técnica para representação, extração e avaliação de interações sociais foi proposta por meio da incorporação de uma etapa de extração automática na técnica de representação e avaliação.
- O método que guia a aplicação da técnica foi desenvolvido a partir da extensão de um procedimento de análise de interações sociais baseadas em mídia.

Os ganhos alcançados a partir das contribuições apresentadas estão relacionados ao preenchimento de lacunas do conhecimento científico sistematizado nos vários trabalhos relacionados apresentados nos capítulos anteriores. Entre esses ganhos, podem ser citados:

- A técnica permite que a linguagem de Mechner seja aplicada para representar interações entre usuários de ambientes sociais virtuais como Redes Sociais Online.
- A técnica permite que interações sociais em OSNs representadas na linguagem de Mechner sejam avaliadas quantitativamente. Originalmente, contingências comportamentais representadas na linguagem de Mechner eram avaliadas qualitativamente.
- O método sistematiza a aplicação da técnica de modo cíclico, mediante o detalhamento de elementos associados com a atividade dos usuários, incorporando explicitamente elementos de interesse no processo de análise das interações.
- A aplicação da técnica e do método permite ao pesquisador interessado analisar contingências comportamentais associadas a interações entre usuários de Redes Sociais Online. Nessa análise podem ser observadas interações baseadas em mídia e interações providas via dispositivos móveis e aplicativos de terceiros.

6.2 Síntese dos Resultados

De forma resumida, os resultados de pesquisa que levaram ao alcance das contribuições apresentadas nesta tese podem ser assim organizados:

- Em uma fase anterior à proposição da técnica de representação e a avaliação, foi realizada a computação de contingências comportamentais associadas com o compartilhamento síncrono [Gomes and Pimentel, 2011d] e assíncrono [Gomes et al., 2011] de anotações em vídeo.
- A técnica para a representação e a avaliação de interações sociais [Gomes and Pimentel, 2011f] foi aplicada na identificação de situações frequentes que envolvem os usuários do Facebook em interações sociais [Gomes and Pimentel, 2011e], e também, na análise de interações sociais baseadas em mídias providas por aplicações disponíveis em *smartphones* [Gomes and Pimentel, 2011c]. A técnica foi estendida com a inclusão de uma etapa de extração automática de regras. A técnica expandida com a etapa de extração foi aplicada na análise de uma mobilização cívica online contra corrupção no Brasil [Gomes and Pimentel, 2012d].
- Um procedimento de análise de interações sociais baseadas em mídia foi aplicado na identificação das interações sociais baseadas em mídia mais frequentes, e também, a identificação dos provedores de conteúdo utilizados no compartilhamento dos diferentes tipos de mídia [Gomes and Pimentel, 2012e]. Um método foi desenvolvido a partir da expansão do procedimento de análise [Gomes and Pimentel, 2011b] que foi detalhado em termos da ligação de mídias e outros elementos associados com as ações coletivas executadas por usuários de Redes Sociais Online [Gomes and Pimentel, 2012a].
- O método que guia a aplicação da técnica de representação, de extração e de avaliação de interações sociais foi aprimorado. O método foi aplicado em uma análise mais detalhada da mobilização cívica online contra corrupção no Brasil [Gomes and Pimentel, 2012b], e ainda, na análise de uma mudança de comportamento usuários de aplicativos de terceiros [Gomes and Pimentel, 2012c].
- Uma avaliação qualitativa realizada com potenciais usuários, que tem como objetivo o levantamento de cenários nos quais a técnica e o método possam ser empregados no desenvolvimento de aplicações sociais [Gomes and Pimentel, 2012f].

A experimentação da técnica e do método foi realizada a partir de dados colhidos da rede Facebook. Essa Rede Social Online foi escolhida considerando-se o grande número de usuários conectados, a riqueza de possibilidades de interação entre usuários e a grande variedade de mídias disponíveis para interação. Como a aplicação da técnica e do método é independente da OSN utilizada, a princípio essas contribuições podem ser generalizadas e aplicadas a qualquer Rede Social Online.

6.3 Limitações

O trabalho apresentado nesta tese apresenta algumas limitações, dentre as quais, são destacadas três principais:

- Uma limitação está relacionada a identificação dos elementos que compõem a linguagem de Mechner em Redes Sociais Online. A identificação de ações, usuários e consequências dependem da percepção de quem faz a análise das interações na Rede Social. Outro fator limitador que algumas redes impõem como no caso da rede Facebook, é a exigência de uma autorização prévia dos usuários para que o analista tenha acesso aos dados que podem vir a ser coletados.
- Outra limitação está relacionada ao desempenho do algoritmo *GSP – like* (Seção 3.3), em termos de tempo de execução. Algumas estratégias podem ser implementadas para que o desempenho do algoritmo possa ser melhorado, por exemplo, o armazenamento do conjunto de dados em memória primária.
- A técnica e o método foram experimentados apenas a partir da coleta de dados da rede Facebook. No entanto, considerando-se o grande número de usuários conectados nessa rede, a riqueza de possibilidades de interação entre usuários e a grande variedade de mídias disponíveis para interação, a princípio, as contribuições alcançadas nesta tese podem ser generalizadas e aplicadas a outras Redes Sociais Online.

6.4 Trabalhos Futuros

Dada a relevância e a abrangência do tema abordado nesta tese, o trabalho apresentado abre várias possibilidades e fornece subsídios para investigações futuras:

Extensão da técnica e do método. Algumas extensões para técnica e do método apresentados nesta tese podem ser propostas, por exemplo, a partir do aprofundamento na compreensão das possibilidades de representação de interações sociais na Linguagem de Mechner, interações sociais oportunistas podem

ser representadas, extraídas e avaliadas. Outras extensões são propostas nos trabalhos futuros relatados a seguir.

Aprimoramento da modificação do algoritmo para extração de padrões sequenciais. Além da melhoria do desempenho do algoritmo *GSP – like* para a extração de interações sociais, outros experimentos podem ser realizado no sentido de tornar o algoritmo genérico o suficiente para extrair outros tipos de sequências que não apenas contingências comportamentais.

Estudo e desenvolvimento de um modelo preditivo baseado na técnica e no método. A linguagem de Mechner permite a representação da predição de ações e consequências. A técnica e o método apresentados nesta tese podem estendidos para que um modelo preditivo baseado em regras possa ser construído. Nesse modelo, por exemplo, as ações que podem ser executadas pelos usuários nas interações sociais poderão ser previstas.

Aplicação da técnica e método em outras Redes Sociais Online. A princípio, as contribuições apresentadas nesta tese podem ser generalizadas e aplicadas a outras Redes Sociais Online. Os resultados alcançados a partir da análise de dados do Facebook podem ser verificados em outras Redes Sociais Online populares como Twitter e Google+, por exemplo, na identificação de interações sociais genéricas e suas especializações em interações sociais baseadas em mídias, interações providas via aplicativos e dispositivos móveis, como apresentado no Capítulo 3. A técnica e o método também podem ser aplicados na análise do compartilhamento de mídias entre provedores de conteúdo e a Rede Social Online, como apresentado na Seção 4.1. Outras possibilidades de investigação estão relacionadas a análise do compartilhamento de mídias em movimentações sociais online como apresentado na Seção 3.6.1, e também, na análise do comportamento de usuários de aplicativos de terceiros como reportado na Seção 4.5.

Implementação das ferramentas propostas nos cenários levantados. Os potenciais usuários que participaram da avaliação qualitativa do método e da técnica descreveram três cenários na Seção 5.3.1. Para cada cenário uma ferramenta pode ser construída, preferencialmente em parceria com os engenheiros entrevistados. Outros cenários de aplicação da técnica e do método podem ser levantados com outros perfis de usuários da área de Computação, e as respectivas aplicações que podem ser desenvolvidas.

Avaliação com Pesquisadores de Outras Áreas. Do mesmo modo que foi realizada a entrevista com potenciais usuários pesquisadores área de Computação, entrevistas individuais podem ser realizadas com pesquisadores de outras áreas

interessados em utilizar a técnica e o método apresentados nesta tese. Esses pesquisadores podem ser, por exemplo, das áreas de Ciências Sociais e Psicologia. Outras técnicas de avaliação com potenciais usuários também podem ser aplicadas nesses trabalhos futuros, por exemplo, entrevistas com foco em grupos específicos de usuários.

Aplicação do método e da técnica na análise de interações oportunistas ou latentes. À medida que as APIs das Redes Sociais Online disponibilizarem o registro da atividade dos usuários que possam ser associadas com as interações oportunistas e também com as interações latentes (Seção 2.4.3), o método e a técnica apresentados podem ser utilizados na análise dessas interações.

Aprimoramento do método para avaliação de sociabilidade em aplicações sociais. O método e a técnica apresentados nesta tese têm potencialidades que podem ser exploradas na inspeção da sociabilidade de aplicações sociais. Heurísticas de sociabilidade foram propostas no trabalho de Geerts and De Grooff [2009], assim como métodos de avaliação de IHC que exploram a interação entre usuários na avaliação de aplicações sociais [Souza et al., 2010], [Reis and Prates, 2011]. As contribuições desta tese podem ser alinhadas a iniciativas de avaliação da área de Interface Homem Computador, de tal forma que o método e a técnica apresentados sejam utilizados na avaliação da sociabilidade de aplicações sociais.

Aprimoramento do método e da técnica para a recomendação de ligações e de conteúdos multimídia. A recomendação de ligações sociais e conteúdos multimídia em Redes Sociais Online é um tema explorado em pesquisas recentes [Esmaeili et al., 2011], [Anastopoulos et al., 2011], [Shin et al., 2011], [Yang et al., 2012], [Chaoji et al., 2012]. O método e a técnica apresentados nesta tese podem ser alinhados com essas iniciativas, por exemplo, sugerindo ligações entre usuários que tem um padrão de interação social similar ou que interage com frequência com um amigo em comum, ou ainda, recomendar conteúdos como imagens, vídeos, músicas e localidades entre usuários que têm um mesmo padrão de interação ou têm um grupo de amigos em comum.

6.5 Considerações Finais

Nesta tese foi apresentada uma técnica que permite a representação, a extração e a avaliação de interações entre usuários de Redes Sociais Online, e também foi apresentado um método que guia a aplicação da técnica e permite a análise de interações sociais especializadas. A técnica e o método foram desenvolvidos de modo iterativo a partir do aprimoramento de resultados obtidos em experimentações variadas.

A aplicação da técnica e o método permitem realizar a representação explícita de elementos associados com as atividades de usuários em Redes Sociais Online, e incluem as ações executadas pelos usuários, as mídias compartilhadas, os servidores, os tipos de dispositivos móveis e os tipos de aplicativos de terceiros utilizados para compartilhar essas mídias.

A representação explícita desses elementos pode ser utilizada na extração e na avaliação das interações entre usuários. Foram reportados resultados de experimentos da aplicação da técnica e do método a partir de dados obtidos da rede Facebook. Essa rede social foi escolhida devido ao grande número de usuários por meio dela conectados, a riqueza de possibilidades de interação entre usuários e a variedade de mídias disponíveis para interação.

Como a aplicação da técnica e do método é independente da rede social utilizada, a princípio as contribuições desta tese podem ser generalizadas e aplicadas a qualquer Rede Social Online.

Referências Bibliográficas

- Abrol, S. and Khan, L. (2010). Tweethood: Fuzzy k-closest friends with variable depth for location mining. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing (SOCIALCOM '10)*, pages 153–160. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/SocialCom.2010.30>.
- Agrawal, R. and Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the International Conference on Data Engineering (ICDE '95)*, pages 3–14. IEEE Computer Society. Available from: <http://doi.ieeecomputersociety.org/10.1109/ICDE.1995.380415>.
- Ahn, Y.-Y., Han, S., Kwak, H., Moon, S., and Jeong, H. (2007). Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th International Conference on World Wide Web (WWW '07)*, pages 835–844. ACM. Available from: <http://doi.acm.org/10.1145/1242572.1242685>.
- Amaral, L., Scala, A., Barthelemy, M., and Stanley, H. E. (2000). Classes of small-world networks. In *Proceedings of the National Academy of Science USA*, volume 97, pages 11149–11152. Available from: <http://dx.doi.org/10.1073/pnas.200327197>.
- Anastopoulos, V., Karampelas, P., Kalagiakos, P., and Alhajj, R. (2011). A hybrid framework for building a web-page recommender system. In *Proceedings of the 2011 European Intelligence and Security Informatics Conference (EISIC '11)*, pages 385–390. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/EISIC.2011.40>.
- Anthonyamy, P., Rashid, A., Walkerdine, J., Greenwood, P., and Larkou, G. (2012). Collaborative privacy management for third-party applications in online social networks. In *Proceedings of the 1st Workshop on Privacy and Security in Online Social Media (PSOSM '12)*, pages 5:1–5:4. ACM. Available from: <http://doi.acm.org/10.1145/2185354.2185359>.

- Astudillo, H., Codocedo, V., Canals, G., Torres, D., Diaz, A., Napoli, A., Gomes, A. K., and Pimentel, M. G. C. (2009). Combining knowledge discovery, ontologies, annotations, and semantic wikis. *Webmedia Minicourse Book, 15th Brazilian Symposium on Multimedia and the Web (WebMedia '09)*, pages 82–103. Available from: <http://hal.inria.fr/docs/00/43/56/59/PDF/SemanticWikiWebMedia-FinalVersion-toMarioMeireles.pdf>.
- Barbosa, G. A. R., Silva, I. S., Zaki, M., Meira, Jr., W., Prates, R. O., and Veloso, A. (2012). Characterizing the effectiveness of twitter hashtags to detect and track online population sentiment. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '12)*, pages 2621–2626. ACM. Available from: <http://doi.acm.org/10.1145/2223656.2223846>.
- Barkhuus, L. and Tashiro, J. (2010). Student socialization in the age of Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, pages 133–142. ACM. Available from: <http://doi.acm.org/10.1145/1753326.1753347>.
- Benevenuto, F. (2010). *Uma análise empírica de interações em redes sociais*. Tese de Doutorado em Ciência da Computação., Instituto de Ciências Exatas da Universidade Federal de Minas Gerais. Orientador: Virgílio Augusto Fernandes Almeida. Available from: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/1843/SLSS-85BN96/1/fabriciobenevenutosouza.pdf>.
- Benevenuto, F., Rodrigues, T., Almeida, V., Almeida, J., and Ross, K. (2009a). Video interactions in online video social networks. *ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP)*, 5(4):30:1–30:25. Available from: <http://doi.acm.org/10.1145/1596990.1596994>.
- Benevenuto, F., Rodrigues, T., Cha, M., and Almeida, V. (2009b). Characterizing user behavior in online social networks. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement (IMC'09)*, pages 49–62. Available from: <http://doi.acm.org/10.1145/1644893.1644900>.
- Benevenuto, F., Rodrigues, T., Cha, M., and Almeida, V. (2012). Characterizing user navigation and interactions in online social networks. *Information Sciences: an International Journal*, 195:1–24. Available from: <http://dx.doi.org/10.1016/j.ins.2011.12.009>.
- Bigonha, C. A. S., Cardoso, T. N. C., Moro, M. M., Almeida, V. A. F., and Gonçalves, M. A. (2010). Detecting evangelists and detractors on Twitter. In *Procee-*

dings of the 16th Brazilian Symposium on Multimedia and the Web (WebMedia '10), pages 107–114. SBC.

- Bogdanov, D., Haro, M., Fuhrmann, F., Xambo, A., Gomez, E., and Herrera, P. (2011). A content-based system for music recommendation and visualization of user preferences working on semantic notions. In *Proceedings of the 9th International Workshop on Content-Based Multimedia Indexing (CBMI '11)*, pages 249–252. Available from: <http://mtg.upf.edu/files/publications/Bogdanov-et-al-CBMI2011.pdf>.
- Bohmer, M., Hecht, B., Schoning, J., Kruger, A., and Bauer, G. (2011). Falling asleep with angry birds, Facebook and kindle: a large scale study on mobile application usage. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*, pages 47–56. ACM. Available from: <http://doi.acm.org/10.1145/2037373.2037383>.
- Braghin, S., Ferrari, E., and Trombetta, A. (2011). A rule-based policy language for selective trust propagation in social networks. In *Proceedings of the 2011 ACM SIGMOD/PODS Conference on Databases and Social Networks (DBSocial '11)*, pages 25–30. ACM. Available from: <http://doi.acm.org/10.1145/1996413.1996418>.
- Burke, M., Marlow, C., and Lento, T. (2009). Feed me: motivating newcomer contribution in social network sites. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI '09)*, pages 945–954. ACM. Available from: <http://doi.acm.org/10.1145/1518701.1518847>.
- Burke, M., Marlow, C., and Lento, T. (2010). Social network activity and social well-being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, pages 1909–1912. ACM. Available from: <http://doi.acm.org/10.1145/1753326.1753613>.
- Cattelan, R. G., Teixeira, C., Goularte, R., and Pimentel, M. D. G. C. (2008). Watch-and-comment as a paradigm toward ubiquitous interactive video editing. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 4(4):28:1–28:24. Available from: <http://doi.acm.org/10.1145/1412196.1412201>.
- Cesar, P. and Geerts, D. (2011). Past, present, and future of social TV: A categorization. In *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC 2011)*, pages 347–351. IEEE. Available from: <http://dx.doi.org/10.1109/CCNC.2011.5766487>.

- Cesar, P., Ooi, W. T., Moskowitz, B., Babin, Z., Bulterman, D., Lienhart, R., and Richter, R. (2011). Towards synergy between the open source and the research multimedia communities. In *Proceedings of the ACM International Conference on Multimedia (MM '11)*, pages 597–598. ACM. Available from: <http://doi.acm.org/10.1145/2072298.2072382>.
- Cestnik, B. (1990). Estimating probabilities: A crucial task in machine learning. In *Proceedings of the European Conference on Artificial Intelligence (ECAI'90)*, pages 147–149. Available from: http://www.temida.si/~bojan/_resources/Cestnik_Estimating_probabilities.pdf.
- Cha, M., Benevenuto, F., Ahn, Y.-Y., and Gummadi, K. P. (2012). Delayed information cascades in Flickr: Measurement, analysis, and modeling. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 56(3):1066–1076. Available from: <http://dx.doi.org/10.1016/j.comnet.2011.10.020>.
- Chaoji, V., Ranu, S., Rastogi, R., and Bhatt, R. (2012). Recommendations to boost content spread in social networks. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12)*, pages 529–538. ACM. Available from: <http://doi.acm.org/10.1145/2187836.2187908>.
- Cho, C.-W., Wu, Y.-H., and Chen, A. L. (2009). Effective database transformation and efficient support computation for mining sequential patterns. *Journal of Intelligent Information System*, 32(1):23–51. Available from: <http://dx.doi.org/10.1007/s10844-007-0047-y>.
- Chorianopoulos, K. (2010). Scenarios of use for sociable mobile TV. In Marcus, A., Roibis, A. C., and Sala, R., editors, *Mobile TV: Customizing Content and Experience (Human-Computer Interaction Series)*, pages 243–254. Springer London. Available from: http://dx.doi.org/10.1007/978-1-84882-701-1_18.
- Colombo, G. B., Chorley, M. J., Williams, M. J., Allen, S. M., and Whitaker, R. M. (2012). You are where you eat: Foursquare checkins as indicators of human mobility and behavior. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, volume 0, pages 217–222. IEEE Computer Society. Available from: <http://doi.ieeecomputersociety.org/10.1109/PerComW.2012.6197483>.
- Comarela, G., Crovella, M., Almeida, V., and Benevenuto, F. (2012). Understanding factors that affect response rates in Twitter. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT '12)*, pages 123–132. ACM. Available from: <http://doi.acm.org/10.1145/2309996.2310017>.

- Costa, L. F., Rodrigues, F. A., Travieso, G., and Boas, P. R. V. (2005). Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242. Available from: <http://arxiv.org/abs/cond-mat/0505185>.
- De Choudhury, M., Counts, S., and Czerwinski, M. (2011). Identifying relevant social media content: leveraging information diversity and user cognition. In *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia (HT '11)*, pages 161–170. ACM. Available from: <http://doi.acm.org/10.1145/1995966.1995990>.
- De Choudhury, M. and Sundaram, H. (2011). Why do we converse on social media?: an analysis of intrinsic and extrinsic network factors. In *Proceedings of the 3rd ACM SIGMM International Workshop on Social Media (WSM '11)*, pages 53–58. ACM. Available from: <http://doi.acm.org/10.1145/2072609.2072625>.
- De Choudhury, M., Sundaram, H., John, A., and Seligmann, D. D. (2009). What makes conversations interesting?: themes, participants and consequences of conversations in online social media. In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*, pages 331–340. ACM. Available from: <http://doi.acm.org/10.1145/1526709.1526754>.
- Diakopoulos, N., De Choudhury, M., and Naaman, M. (2012). Finding and assessing social media information sources in the context of journalism. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI '12)*, pages 2451–2460. ACM. Available from: <http://doi.acm.org/10.1145/2208276.2208409>.
- Doerr, C., Blenn, N., Tang, S., and Van Mieghem, P. (2012). Are friends over-rated? a study for the social news aggregator Digg.com. *Computer Communications*, 35(7):796–809. Available from: <http://dx.doi.org/10.1016/j.comcom.2012.02.001>.
- Dong, X., Zheng, Z., Cao, L., Zhao, Y., Zhang, C., Li, J., Wei, W., and Ou, Y. (2011). e-nsp: efficient negative sequential pattern mining based on identified positive patterns without database rescanning. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management (CIKM '11)*, pages 825–830. ACM. Available from: <http://hdl.handle.net/10453/19097>.
- Egele, M., Moser, A., Kruegel, C., and Kirda, E. (2012). Pox: Protecting users from malicious Facebook applications. In *Journal of Computer Communications*, volume 35, pages 1507–1515. Elsevier Science Publishers. Available from: <http://dx.doi.org/10.1016/j.comcom.2012.04.016>.

- Esmaeili, L., Nasiri, M., and Minaei-Bidgoli, B. (2011). Personalizing group recommendation to social network users. In *Proceedings of the 2011 International Conference on Web Information Systems and Mining - Volume Part I, WISM'11*, pages 124–133. Springer-Verlag. Available from: <http://dl.acm.org/citation.cfm?id=2045561.2045581>.
- Fagá, Jr., R., Motti, V. G., Cattelan, R. G., Teixeira, C. A. C., and Pimentel, M. G. C. (2010). A social approach to authoring media annotations. In *Proceedings of the 10th ACM Symposium on Document Engineering (DocEng'10)*, pages 17–26. Available from: <http://doi.acm.org/10.1145/1860559.1860566>.
- Fayyad, U. M., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery: an overview. In Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R., editors, *Advances in Knowledge Discovery and Data Mining*, pages 1–34. American Association for Artificial Intelligence. Available from: <http://dl.acm.org/citation.cfm?id=257938.257942>.
- Fortuna, B., Rodrigues, E. M., and Milic-Frayling, N. (2007). Improving the classification of newsgroup messages through social network analysis. In *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management (CIKM '07)*, pages 877–880. ACM. Available from: <http://doi.acm.org/10.1145/1321440.1321565>.
- Freeman, L. C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*. Empirical Press.
- Fürnkranz, J., Gamberger, D., and Lavrac, N. (2011). *Rule Learning: Essentials of Machine Learning and Relational Data Mining*. Springer, 1st edition.
- Geerts, D. and De Grooff, D. (2009). Supporting the social uses of television: sociability heuristics for social TV. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pages 595–604. ACM. Available from: <http://doi.acm.org/10.1145/1518701.1518793>.
- Gilbert, E. and Karahalios, K. (2009). Predicting tie strength with social media. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI '09)*, pages 211–220. ACM. Available from: <http://doi.acm.org/10.1145/1518701.1518736>.
- Gjoka, M., Sirivianos, M., Markopoulou, A., and Yang, X. (2008). Poking Facebook: characterization of OSNs applications. In *Proceedings of the 1st Workshop on Online Social Networks (WOSN '08)*, pages 31–36. ACM. Available from: <http://doi.acm.org/10.1145/1397735.1397743>.

- Golder, S. A. and Yardi, S. (2010). Structural predictors of tie formation in Twitter: Transitivity and mutuality. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing (SOCIALCOM '10)*, pages 88–95. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/SocialCom.2010.22>.
- Gomes, A. K. (2002). Análise do conhecimento extraído de classificadores simbólicos utilizando medidas de avaliação e de interessabilidade. Dissertação de Mestrado, ICMC-USP. Available from: http://www.teses.usp.br/teses/disponiveis/55/55134/tde-04072002-144610/publico/Diss_Alan.pdf.
- Gomes, A. K. (2007). Small disjuncts grouping by rule coverage and accuracy measures. In *Proceedings of the 7th International Conference on Intelligent Systems Design and Applications (ISDA'07)*, pages 412–415. IEEE Computer Society. Available from: <http://dl.acm.org/citation.cfm?id=1317534.1318271>.
- Gomes, A. K. and Monard, M. C. (2003). Um módulo para avaliação de regras induzidas por algoritmos de aprendizado de máquina. In *Proceedings of the 4th Congress of Logic Applied to Technology (LAPTEC 2003)*, pages 203–211.
- Gomes, A. K., Pedrosa, D. d. C., and Pimentel, M. G. C. (2011). Evaluating asynchronous sharing of links and annotation sessions as social interactions on Internet videos. In *Proceedings of the 2011 IEEE/IPSJ International Symposium on Applications and the Internet (SAINT '11)*, pages 184–189. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/SAINT.2011.33>.
- Gomes, A. K. and Pimentel, M. G. C. (2011a). (in portuguese) Avaliação de interações sociais em ambientes sociais multimídia. In *Workshop de Teses e Dissertações (WTD), Simpósio Brasileiro de Multimedia e Web (WebMedia '11)*, pages 119–126. SBC.
- Gomes, A. K. and Pimentel, M. G. C. (2011b). (in portuguese) Um método para análise de interações sociais na web social como regras se-então. In *Workshop sobre Aspectos da Interação Humano-Computador para a Web Social (WAIHCWS'11) no X Simpósio Brasileiro de Fatores Humanos em Sistemas de Computação (IHC'11)*. Available from: <http://ceur-ws.org/Vol-797/>.
- Gomes, A. K. and Pimentel, M. G. C. (2011c). Measuring media-based social interactions provided by smartphone applications in social networks. In *Proceedings of the 2011 ACM Workshop on Social and Behavioral Networked Media Access (SBNMA '11)*, pages 59–64. ACM. Available from: <http://doi.acm.org/10.1145/2072627.2072644>.

- Gomes, A. K. and Pimentel, M. G. C. (2011d). Measuring synchronous and asynchronous sharing of collaborative annotations sessions on ubi-videos as social interactions. In *Proceedings of the 2011 International Conference on Ubi-Media Computing (U-MEDIA '11)*, pages 122–129. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/U-MEDIA.2011.56>.
- Gomes, A. K. and Pimentel, M. G. C. (2011e). Social interactions representation as users behavioral contingencies and evaluation in social networks. In *Proceedings of the 2011 IEEE International Conference on Semantic Computing (ICSC '11)*, pages 275–278. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/ICSC.2011.103>.
- Gomes, A. K. and Pimentel, M. G. C. (2011f). A technique for human-readable representation and evaluation of media-based social interactions in social networks. In *Proceedings of the 17th Brazilian Symposium on Multimedia and the Web (WebMedia '11)*, pages 119 –126. SBC.
- Gomes, A. K. and Pimentel, M. G. C. (2012a). Evaluation of media-based social interactions: Linking collective actions to media types, applications and devices in social networks– Chapter (under review). In Agarwal, N., Lim, M., and Wigand, R., editors, *Online Collective Action: Dynamics of the Crowd in Social Media – Lecture Notes in Social Networks*. Springer.
- Gomes, A. K. and Pimentel, M. G. C. (2012b). Measuring media-based social interactions in online civic mobilization against corruption in Brazil. In *Proceedings of the 18th Brazilian Symposium on Multimedia and the Web (WebMedia '12)*, pages 173–180. ACM. Available from: <http://doi.acm.org/10.1145/2382636.2382675>.
- Gomes, A. K. and Pimentel, M. G. C. (2012c). Media-based application-induced collective user behavior change (submitted). In Salah, A. A., editor, *IEEE Pervasive Computing*. IEEE.
- Gomes, A. K. and Pimentel, M. G. C. (2012d). Media-based social interaction patterns: a case study in an online civic mobilization. In *Proceedings of the 2012 International Workshop on Socially-Aware Multimedia (IWSAM '12) in conjunction with ACM Multimedia*, pages 37–42. ACM. Available from: <http://doi.acm.org/10.1145/2390876.2390888>.
- Gomes, A. K. and Pimentel, M. G. C. (2012e). A media-based social interactions analysis procedure. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC '12)*, pages 1018–1024. ACM. Available from: <http://doi.acm.org/10.1145/2245276.2245472>.

- Gomes, A. K. and Pimentel, M. G. C. (2012f). Representing, mining and measuring social interactions in social media systems: A rule-based approach (submitted). In de Oliveira, M. C., editor, *Journal of the Brazilian Computer Society*. Springer.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18. Available from: <http://doi.acm.org/10.1145/1656274.1656278>.
- Halvey, M., Vallet, D., Hannah, D., Feng, Y., and Jose, J. M. (2010). An asynchronous collaborative search system for online video search. *Journal of Information Processing Management*, 46:733–748. Available from: <http://dx.doi.org/10.1016/j.ipm.2009.11.007>.
- Han, J. (2005). *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers Inc.
- Hew, K. F. (2011). Students' and teachers' use of Facebook. *Computing Human Behavior*, 27(2):662–676. Available from: <http://dx.doi.org/10.1016/j.chb.2010.11.020>.
- Holdings, N. (2012). State of the social media: Social media report Q3 2012. Available from: <http://blog.nielsen.com/nielsenwire/social/>.
- Jiang, J., Wilson, C., Wang, X., Huang, P., Sha, W., Dai, Y., and Zhao, B. Y. (2010). Understanding latent interactions in online social networks. In *Proceedings of the ACM Conference on Internet Measurement (IMC '10)*, pages 369–382. Available from: <http://doi.acm.org/10.1145/1879141.1879190>.
- Jin, X., Gallagher, A., Cao, L., Luo, J., and Han, J. (2010). The wisdom of social multimedia: using Flickr for prediction and forecast. In *Proceedings of the International Conference on Multimedia (MM '10)*, pages 1235–1244. ACM. Available from: <http://doi.acm.org/10.1145/1873951.1874196>.
- Kajdanowicz, T., Kazienko, P., and Doscocz, P. (2010). Label-dependent feature extraction in social networks for node classification. In *Proceedings of the Second International Conference on Social Informatics (SocInfo'10)*, pages 89–102. Springer-Verlag. Available from: <http://dl.acm.org/citation.cfm?id=1929326.1929333>.
- King, I., Li, J., and Chan, K. T. (2009). A brief survey of computational approaches in social computing. In *Proceedings of the 2009 International Joint*

- Conference on Neural Networks (IJCNN '09)*, pages 2699–2706. IEEE Press. Available from: <http://dl.acm.org/citation.cfm?id=1704555.1704660>.
- Kivran-Swaine, F. and Naaman, M. (2011). Network properties and social sharing of emotions in social awareness streams. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*, pages 379–382. ACM. Available from: <http://doi.acm.org/10.1145/1958824.1958882>.
- Kleinberg, J. (2000). The small-world phenomenon: an algorithm perspective. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC '00)*, pages 163–170. ACM. Available from: <http://doi.acm.org/10.1145/335305.335325>.
- Kramer, A. D. (2012). The spread of emotion via Facebook. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI '12)*, pages 767–770. ACM. Available from: <http://doi.acm.org/10.1145/2207676.2207787>.
- Krishnamurthy, B. (2009). A measure of online social networks. In *Proceedings of the First International Conference on COMMunication Systems And NETWORKS (COMSNETS '09)*, pages 190–199. IEEE Press. Available from: <http://dl.acm.org/citation.cfm?id=1702135.1702163>.
- Lane, N. D., Lymberopoulos, D., Zhao, F., and Campbell, A. T. (2010). Hapori: context-based local search for mobile phones using community behavioral modeling and similarity. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing (Ubicomp '10)*, pages 109–118. ACM. Available from: <http://doi.acm.org/10.1145/1864349.1864366>.
- Lavrac, N., Flach, P. A., and Zupan, B. (1999). Rule evaluation measures: A unifying view. In *International Workshop on Inductive Logic Programming (ILP)*, pages 174–185. Springer.
- Lazar, J., Feng, J., and Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. John Wiley & Sons. Available from: http://books.google.com.br/books?id=H_r6prUFpc4Cem10/01/2013.
- Leskovec, J., Adamic, L. A., and Huberman, B. A. (2007). The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1). Available from: <http://doi.acm.org/10.1145/1232722.1232727>.
- Leung, C. W.-k., Lim, E.-P., Lo, D., and Weng, J. (2010). Mining interesting link formation rules in social networks. In *Proceedings of the 19th ACM Internati-*

- onal Conference on Information and Knowledge Management (CIKM '10)*, pages 209–218. ACM. Available from: <http://doi.acm.org/10.1145/1871437.1871468>.
- Liao, Q. V., Wagner, C., Pirolli, P., and Fu, W.-T. (2012). Understanding experts' and novices' expertise judgment of Twitter users. In *Proceedings of the 2012 ACM annual Conference on Human Factors in Computing Systems (CHI '12)*, pages 2461–2464. ACM. Available from: <http://doi.acm.org/10.1145/2207676.2208410>.
- Lin, Y.-R., Sundaram, H., De Choudhury, M., and Kelliher, A. (2012). Discovering multirelational structure in social media streams. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 8(1):4:1–4:28. Available from: <http://doi.acm.org/10.1145/2071396.2071400>.
- Magnien, C., Latapy, M., and Guillaume, J.-L. (2011). Impact of random failures and attacks on poisson and power-law random networks. *ACM Computing Surveys*, 43(3):13:1–13:31. Available from: <http://doi.acm.org/10.1145/1922649.1922650>.
- Magno, G., Comarela, G., Saez-Trumper, D., Cha, M., and Almeida, V. (2012). New kid on the block: exploring the Google+ social graph. In *Proceedings of the 2012 ACM Conference on Internet Measurement (IMC '12)*, pages 159–170. ACM. Available from: <http://doi.acm.org/10.1145/2398776.2398794>.
- Maia, M., Almeida, J., and Almeida, V. (2008). Identifying user behavior in online social networks. In *Proceedings of the 1st Workshop on Social Network Systems (SocialNets '08)*, pages 1–6. ACM. Available from: <http://doi.acm.org/10.1145/1435497.1435498>.
- Malinen, S. (2010). Photo exhibition or online community? the role of social interaction in Flickr. In *Proceedings of the 2010 5th International Conference on Internet and Web Applications and Services (ICIW '10)*, pages 380–385. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/ICIW.2010.63>.
- Martínez-Ballesteros, M. and Riquelme, J. C. (2011). Analysis of measures of quantitative association rules. In *Proceedings of the 6th International Conference on Hybrid Artificial Intelligent Systems - Volume Part II (HAIS '11)*, pages 319–326. Springer-Verlag. Available from: <http://dl.acm.org/citation.cfm?id=2021503.2021547>.
- Mascaro, C. M. and Goggins, S. P. (2011). Brewing up citizen engagement: the coffee party on Facebook. In *Proceedings of the 5th International Conference on*

- Communities and Technologies (CCT '11)*, pages 11–20. ACM. Available from: <http://doi.acm.org/10.1145/2103354.2103357>.
- Mattaini, M. (1995). Contingency diagrams as teaching tools. *Journal of Behavioral Analyst*, 18:93–98.
- Mechant, P. and Evens, T. (2011). Interaction in web-based communities: a case study of Last.fm. *International Journal Web Based Communities*, 7(2):234–249. Available from: <http://dx.doi.org/10.1504/IJWBC.2011.039513>.
- Mechner, F. (1959). A notation system for the description of behavioral procedures. *Journal of Experimental Analysis of Behavioral*, 2:133–150.
- Mechner, F. (2008). Behavioral contingency analysis. *Behavioral Processes*, 78(2):124–144.
- Mechner, F. (2011). Why behavior analysis needs a formal symbolic language for codifying behavioral contingencies. *European Journal of Behavior Analysis (EJOBA)*, 1:93–104.
- Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., and Bhattacharjee, B. (2007). Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement (IMC '07)*, pages 29–42. ACM. Available from: <http://doi.acm.org/10.1145/1298306.1298311>.
- Moore, K. and McElroy, J. C. (2012). The influence of personality on Facebook usage, wall postings, and regret. *Computing Human Behavior*, 28(1):267–274. Available from: <http://dx.doi.org/10.1016/j.chb.2011.09.009>.
- Moradi, F., Olovsson, T., and Tsigas, P. (2012). Towards modeling legitimate and unsolicited email traffic using social network properties. In *Proceedings of the Fifth Workshop on Social Network Systems (SNS '12)*, pages 9:1–9:6. ACM. Available from: <http://doi.acm.org/10.1145/2181176.2181185>.
- Motti, V. G., Fagá, Jr., R., Catellan, R. G., Pimentel, M. d. G. C., and Teixeira, C. A. (2009). Collaborative synchronous video annotation via the watch-and-comment paradigm. In *Proceedings of the 7th European Conference on Interactive Television (EuroITV '09)*, pages 67–76. Available from: <http://doi.acm.org/10.1145/1542084.1542097>.
- Nazir, A., Raza, S., and Chuah, C.-N. (2008). Unveiling Facebook: a measurement study of social network based applications. In *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement (IMC '08)*, pages 43–56. ACM. Available from: <http://doi.acm.org/10.1145/1452520.1452527>.

- Nazir, A., Raza, S., Gupta, D., Chuah, C.-N., and Krishnamurthy, B. (2009). Network level footprints of Facebook applications. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement (IMC '09)*, pages 63–75. ACM. Available from: <http://doi.acm.org/10.1145/1644893.1644901>.
- Nazir, A., Waagen, A., Vijayaraghavan, V. S., Chuah, C.-N., D'Souza, R. M., and Krishnamurthy, B. (2012). Beyond friendship: modeling user activity graphs on social network-based gifting applications. In *Proceedings of the 2012 ACM Conference on Internet Measurement (IMC '12)*, pages 467–480. ACM. Available from: <http://doi.acm.org/10.1145/2398776.2398826>.
- Negoescu, R.-A., Adams, B., Phung, D., Venkatesh, S., and Gatica-Perez, D. (2009). Flickr hypergroups. In *Proceedings of the 17th ACM International Conference on Multimedia (MM '09)*, pages 813–816. ACM. Available from: <http://doi.acm.org/10.1145/1631272.1631421>.
- Newman, M. E. J. (2003). The structure and function of complex networks. *Review of the Society of Industrial Applied Mathematics (SIAM Review)*, 45:167–256.
- Nielsen, J. (2010). Interviewing users. useit.com: Jakob Nielsen Website – <http://www.useit.com/alertbox/interviews.html>.
- Park, J. Y. and Chung, C.-W. (2012). When daily deal services meet Twitter: understanding Twitter as a daily deal marketing platform. In *Proceedings of the 3rd Annual ACM Web Science Conference (WebSci '12)*, pages 233–242. ACM. Available from: <http://doi.acm.org/10.1145/2380718.2380748>.
- Pei, J., Han, J., Mortazavi-Asl, B., Wang, J., Pinto, H., Chen, Q., Dayal, U., and Hsu, M.-C. (2004). Mining sequential patterns by pattern-growth: the prefixspan approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(11):1424 – 1440. Available from: <http://dx.doi.org/10.1109/TKDE.2004.77>.
- Pesce, J. a. P., Casas, D. L., Rauber, G., and Almeida, V. (2012). Privacy attacks in social media using photo tagging networks: a case study with Facebook. In *Proceedings of the 1st Workshop on Privacy and Security in Online Social Media (PSOSM '12)*, pages 4:1–4:8. ACM. Available from: <http://doi.acm.org/10.1145/2185354.2185358>.
- Quercia, D., Di Lorenzo, G., Calabrese, F., and Ratti, C. (2011). Mobile phones and outdoor advertising: Measurable advertising. *IEEE Pervasive Computing*, 10(2):28 –36.

- Quercia, D., Ellis, J., and Capra, L. (2010). Using mobile phones to nurture social networks. *IEEE Pervasive Computing*, 9(3):12–20.
- Rabbath, M., Sandhaus, P., and Boll, S. (2010). Automatic creation of photo books from stories in social media. In *Proceedings of Second ACM SIGMM Workshop on Social Media (WSM '10)*, pages 15–20. ACM. Available from: <http://doi.acm.org/10.1145/1878151.1878157>.
- Rabbath, M., Sandhaus, P., and Boll, S. (2012). Analyzing Facebook features to support event detection for photo-based Facebook applications. In *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval (ICMR '12)*, pages 11:1–11:8. ACM. Available from: <http://doi.acm.org/10.1145/2324796.2324810>.
- Reis, S. S. and Prates, R. O. (2011). Applicability of the semiotic inspection method: a systematic literature review. In *Proceedings of the 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction (IHC+CLHC '11)*, pages 177–186. Brazilian Computer Society. Available from: <http://dl.acm.org/citation.cfm?id=2254436.2254468>.
- Robertson, S. P., Vatrupu, R. K., and Medina, R. (2009). The social life of social networks: Facebook linkage patterns in the 2008 U.S. presidential election. In *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government (DG.O '09)*, pages 6–15. Digital Government Society of North America. Available from: <http://dl.acm.org/citation.cfm?id=1556176.1556183>.
- Rossi, R. G. and Rezende, S. O. (2011). Building a topic hierarchy using the bag-of-related-words representation. In *Proceedings of the 11th ACM Symposium on Document Engineering (DocEng '11)*, pages 195–204. ACM. Available from: <http://doi.acm.org/10.1145/2034691.2034733>.
- Rummel, R. (1976). Social behavior and interaction. In *Understanding Conflict and War – The Conflict*, chapter 9. Sage Publications. Available from: <http://www.hawaii.edu/powerkills/TCH.CHAP9.HTMvisitadoem10/01/2013>.
- Savage, N. (2011). Twitter as medium and message. *Communications of the ACM*, 54(3):18–20. Available from: <http://doi.acm.org/10.1145/1897852.1897860>.
- Scott, J. P. (2000). *Social Network Analysis: A Handbook*. SAGE Publications, 2 edition.

- Shamma, D. A., Kennedy, L., and Churchill, E. F. (2011). Peaks and persistence: modeling the shape of micro blog conversations. In *Proceedings of the 2011 ACM Conference on Computer Supported Cooperative Work (CSCW '11)*, pages 355–358. ACM. Available from: <http://doi.acm.org/10.1145/1958824.1958878>.
- Shin, S., Jang, S.-J., and Lee, S.-P. (2011). The user-group based recommendation for the diverse multimedia contents in the social network environments. In *Proceedings of the 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing (DASC '11)*, pages 202–206. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/DASC.2011.206>.
- Sidman, M. (2000). Equivalent relations and the reinforcement contingency. *Journal of Experimental Analysis of Behavioral*, 74:127–146. Available from: <http://seab.envmed.rochester.edu/jeab/articles/2000/jeab-74-01-0127.pdf>.
- Silva, A., Meira, Jr., W., and Zaki, M. J. (2012). Mining attribute-structure correlated patterns in large attributed graphs. *VLDB Endowment*, 5(5):466–477. Available from: <http://dl.acm.org/citation.cfm?id=2140436.2140443>.
- Silva, A., Valiati, H., Guimarães, S., and Wagner, Jr., M. (2011). From individual behavior to influence networks: A case study on Twitter. In *Proceedings of the 17th Brazilian Symposium on Multimedia and the Web (WebMedia '11)*, pages 135–142. SBC.
- Skinner, B. F. (1953). *Science and Human Behavior*. New York Press.
- Souza, C. S., Leitão, C. F., Prates, R. O., Amélia Bim, S., and da Silva, E. J. (2010). Can inspection methods generate valid new knowledge in HCI? the case of semiotic inspection. *International Journal of Human-Computer Studies*, 68(1-2):22–40. Available from: <http://dx.doi.org/10.1016/j.ijhcs.2009.08.006>.
- Starbird, K. and Palen, L. (2011). "Voluntweeters": Self-organizing by digital volunteers in times of crisis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, pages 1071–1080. ACM. Available from: <http://doi.acm.org/10.1145/1978942.1979102>.
- Starbird, K. and Palen, L. (2012). (How) Will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12)*, pages 7–16. ACM. Available from: <http://doi.acm.org/10.1145/2145204.2145212>.
- Stroele, V., Oliveira, J., Zimbrão, G., and Souza, J. M. (2009). Mining and analyzing multirelational social networks. In *Proceedings of the 2009 International*

- Conference on Computational Science and Engineering (CSE '09)*, pages 711–716. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/CSE.2009.69>.
- Sun, B. and Ng, V. T. (2012). Identifying influential users by their postings in social networks. In *Proceedings of the 3rd International Workshop on Modeling Social Media (MSM '12)*, pages 1–8. ACM. Available from: <http://doi.acm.org/10.1145/2310057.2310059>.
- Susarla, A., Oh, J.-H., and Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from Youtube. *Journal of Information System Research*, 23(1):23–41. Available from: <http://dx.doi.org/10.1287/isre.1100.0339>.
- Tan, P.-N., Steinbach, M., and Kumar, V. (2005). *Introduction to Data Mining*. Addison Wesley.
- Tilson, D., Lyytinen, K., and Sørensen, C. (2010). Research commentary – digital infrastructures: The missing is research agenda. *Information System Research*, 21(4):748–759. Available from: <http://dx.doi.org/10.1287/isre.1100.0318>.
- Tsai, D., Flagg, M., Nakazawa, A., and Rehg, J. M. (2012). Motion coherent tracking using multi-label mrf optimization. In *International Journal of Computer Vision*, volume 100, pages 190–202. Kluwer Academic Publishers. Available from: <http://dx.doi.org/10.1007/s11263-011-0512-5>.
- Valafar, M., Rejaie, R., and Willinger, W. (2009). Beyond friendship graphs: a study of user interactions in flickr. In *Proceedings of the 2nd ACM Workshop on Online Social Networks (WOSN '09)*, pages 25–30. ACM. Available from: <http://doi.acm.org/10.1145/1592665.1592672>.
- Vasconcelos, M. A., Ricci, S., Almeida, J., Benevenuto, F., and Almeida, V. (2012). Tips, dones and todos: uncovering user profiles in Foursquare. In *Proceedings of the International Conference on Web Search and Data Mining (WSDM '12)*, pages 653–662. ACM. Available from: <http://doi.acm.org/10.1145/2124295.2124372>.
- Vinciarelli, A., Pantic, M., and Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Journal of Image and Vision Computing*, 27:1743–1759. Available from: <http://dx.doi.org/10.1016/j.imavis.2008.11.007>.

- Viswanath, B., Mislove, A., Cha, M., and Gummadi, K. P. (2009). On the evolution of user interaction in Facebook. In *Proceedings of the ACM Workshop on Online Systems Network (WOSN '09)*, pages 37–42. Available from: <http://doi.acm.org/10.1145/1592665.1592675>.
- Wang, H. and Chin, A. (2010). Evolution analysis of a mobile social network. In *Proceedings of the 6th International Conference on Advanced Data Mining and Applications: Part I (ADMA'10)*, pages 310–321. Springer-Verlag. Available from: <http://dl.acm.org/citation.cfm?id=1947599.1947632>.
- Wang, N. (2012). Third-party applications' data practices on Facebook. In *Extended Abstracts of the ACM Annual Conference on Human Factors in Computing Systems (CHI EA '12)*, pages 1399–1404. ACM. Available from: <http://doi.acm.org/10.1145/2212776.2212462>.
- Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., and Cranor, L. F. (2011). I regretted the minute I pressed share: a qualitative study of regrets on Facebook. In *Proceedings of the 7th Symposium on Usable Privacy and Security (SOUPS '11)*, pages 10:1–10:16. Available from: <http://doi.acm.org/10.1145/2078827.2078841>.
- Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge U. Press.
- Watts, D. J. (1999). *Small worlds: the dynamics of networks between order and randomness*. Princeton University Press.
- Weingarten, K. and Mechner, F. (1966). The contingency as an independent variable of social interaction. *Readings of Experimental Analysis of Behavior*, pages 447–459. Available from: http://mechnerfoundation.org/pdf_downloads/contingency.pdf.
- Wilkinson, D. and Thelwall, M. (2011). Researching personal information on the public web. *Social Science Computer Review*, 29(4):387–401. Available from: <http://ssc.sagepub.com/content/29/4/387.abstract>.
- Williams, C. B. (2007). Social networks in political campaigns: Facebook and the 2006 midterm elections. *Journal of the Meeting of the American Political Science Association*, 1:49 –62. Available from: <http://blogs.bentley.edu/politechmedia/wp-content/uploads/2007/09/apsa-2007-{F}acebook1.pdf>.
- Wilson, C., Boe, B., Sala, A., Puttaswamy, K. P., and Zhao, B. Y. (2009). User interactions in social networks and their implications. In *Proceedings of the 4th*

- ACM *European Conference on Computer Systems (EuroSys '09)*, pages 205–218. ACM. Available from: <http://doi.acm.org/10.1145/1519065.1519089>.
- Wilson, C., Sala, A., Puttaswamy, K. P. N., and Zhao, B. Y. (2012). Beyond social graphs: User interactions in online social networks and their implications. *ACM Transactions on the Web (TWEB)*, 6(4):17:1–17:31. Available from: <http://doi.acm.org/10.1145/2382616.2382620>.
- Xu, C., Ryan, S., Prybutok, V., and Wen, C. (2012). It is not for fun: An examination of social network site usage. *Information & Management*, 49(5):210–217. Available from: <http://dx.doi.org/10.1016/j.im.2012.05.001>.
- Yang, X., Steck, H., Guo, Y., and Liu, Y. (2012). On top-k recommendation using social networks. In *Proceedings of the Sixth ACM Conference on Recommender Systems (RecSys '12)*, pages 67–74. ACM. Available from: <http://doi.acm.org/10.1145/2365952.2365969>.
- Zaki, M. J. (2001). Spade: An efficient algorithm for mining frequent sequences. *Journal of Machine Learning*, 42(1-2):31–60. Available from: <http://dx.doi.org/10.1023/A:1007652502315>.
- Zhai, G., Fox, G. C., Pierce, M., Wu, W., and Bulut, H. (2005). esports: Collaborative and synchronous video annotation system in grid computing environment. In *Proceedings of the Seventh IEEE International Symposium on Multimedia (ISM '05)*, pages 95–103. IEEE Computer Society. Available from: <http://dx.doi.org/10.1109/ISM.2005.55>.
- Zhou, Z., Bandari, R., Kong, J., Qian, H., and Roychowdhury, V. (2010). Information resonance on Twitter: watching Iran. In *Proceedings of the 1st Workshop on Social Media Analytics (SOMA '10)*, pages 123–131. ACM. Available from: <http://doi.acm.org/10.1145/1964858.1964875>.

Experimentação da Técnica com Dados de Interação da *CWaCTool*

A ferramenta *Collaborative Watch-and-Comment Tool (CWaCTool)* Motti et al. [2009] é uma ferramenta multimídia de Computação Ubíqua que dá permissão para que os usuários compartilhem links e fazerem download de vídeos do YouTube, além de permitir a captura de frames dos vídeos e a execução de anotações de texto, de voz e de tinta sob os vídeos baixados. As primeiras experimentações da técnica aqui apresentada foram realizadas em um estudo envolvendo grupos de usuários da *CWaCTool* Fagá et al. [2010]. Foram utilizados três computadores localizados em uma mesma sala. Os computadores possuíam Windows XP instalado, com acesso à Internet e a mesma configuração de hardware. A *CWaCTool* foi instalada em as três máquinas. Foram organizados quatro grupos (G1, G2, G3 e G4) com três usuários (a, b, c) cada grupo. A interação com a ferramenta foi realizada usando-se mouse e teclado. Cada usuário selecionou aleatoriamente uma máquina e os artefatos resultantes da interação (anotações, vídeos baixados, frames capturados, etc) foram mantidos nas máquinas de um grupo para outro. Todas as sessões foram registradas em vídeos dos quais, em uma análise posterior, foram extraídas manualmente as interações sociais na representação de Mechner.

A.1 A Ferramenta *CWaCTool* em Uso

A versão mais recente da ferramenta *CWaCTool* em uso é apresentada na Figura A.1. A Figura A.1 A mostra a tela na qual um vídeo baixado do YouTube é apresentado. A Figura A.1 B mostra as abas *Search* e *Home*. Usando a aba *Se-*

Figura A.1: *Iniciando a CWaCTool.*

A: tela para exibição do vídeo, B: abas *Search* e *Home*, C: vídeo em download, D: ferramenta de bate-papo, E: informações sobre o vídeo, menu e bate-papo.

arch o usuário pode procurar por vídeos do YouTube usando palavras-chave ou um identificador do vídeo. Quando um usuário selecionar com duplo clique um vídeo encontrado na busca, é dado início ao download e a miniatura do vídeo fica destacada na cor amarela (Figura A.1 C). Além disso, o usuário pode conversar com outros usuários conectados no bate-papo (Figura A.1 D), ou ainda usar o menu para ver informações de vídeo, ler e inserir comentários no Youtube, ver tags (texto) associadas ao vídeo ou a grupos de vídeo (Figura A.1 E).

A *CWaCTool* ao ser iniciada, apresenta vídeos baixados armazenados em cache, frames e anotações de sessões anteriores. O compartilhamento assíncrono de links de vídeo e sessões de anotação é originado a partir desses artefatos.

No que se refere aos recursos que permitem a anotação em vídeo, um usuário ativa essa função de anotação clicando no vídeo exibido na janela principal, como resultado é exibida uma janela *pop-up* com três opções: uma para digitar a *anotação de texto*, uma para fazer a gravação de *anotação de áudio* e outra para iniciar uma *anotação de tinta*.

Se uma *anotação de tinta* for selecionada, o vídeo continua em reprodução numa janela minimizada (Figura A.2 A) e uma sessão de anotação é iniciada. No momento que é realizado o clique (seleção), é capturado o frame do vídeo, e esse frame é apresentado na tela principal (Figura A.2 B). As funções de anotação de tinta são exibidas na base da tela (entre Figura A.2 C e A.2 D).

A *CWaCTool* tem importantes recursos de colaboração como o compartilhamento de links de vídeo e de sessões de anotação. Para compartilhar as anotações, os usuários têm que explicitamente convidar outros usuários conectados na ferramenta para compartilhar uma sessão. O convite é realizado a partir de

Figura A.2: CWaCTool em uso [Gomes and Pimentel, 2011d].

A: vídeo em playback minimizado, B: frame de vídeo capturado no painel, C: funções de anotação de tinta, D: botão de compartilhamento de sessão de anotação, E: notificação de aceitação de compartilhamento de sessão, F: compartilhamento explícito de link de vídeo, G: início de download de vídeo na *Search*, H: vídeo apresentado na aba *Home*.

um clique no botão mostrado na Figura A.2 D, que faz com que outros demais usuários conectados recebam um convite para participar de um compartilhamento na janela de bate-papo (Figura A.2 E).

Quando um usuário aceita participar de uma sessão de anotação compartilhada, se ele não tiver baixado o vídeo do qual o frame compartilhado foi capturado, imediatamente é iniciado o download do vídeo. Nessa situação é realizado um compartilhamento implícito de links de vídeo (Figura A.2 H).

O compartilhamento explícito de vídeo é executado quando um usuário clica com o botão esquerdo do mouse sobre uma miniatura do vídeo na aba *Home* e, em seguida, escolhe a opção de compartilhar no bate-papo. Como resultado, um link de vídeo em cor azul é mostrado na janela de bate-papo (Figura A.2 F.). Um usuário aceita o compartilhamento do vídeo clicando nesse link e, se o usuário não tinha baixado esse vídeo antes, o download é iniciado (Figura A.2 G.).

A.2 Representando e Avaliando Interações Sociais da CWaCTool

A técnica apresentada neste capítulo foi experimentada com o objetivo de medir contingências comportamentais que envolvem usuários da ferramenta CWaCTool em interações sociais. Nessas experimentações, a extração de regras que representam interações sociais foi realizada manualmente.

No trabalho Gomes et al. [2011] foram representadas contingências comportamentais que envolvem o compartilhamento assíncrono de links de vídeo da rede YouTube e sessões de anotação. As seguintes ações e respectivas con-

sequências foram identificadas como relevantes para essas interações sociais disponíveis pela *CWaCTool*:

- A_1 = compartilhar uma sessão de anotação (usando um frame capturado de um usuário não está conectado)
- A_2 = compartilhar um link de vídeo no bate-papo (usando um vídeo baixado a partir de um usuário não está conectado)
- A_3 = aceitar um convite (link vídeo clicando compartilhada em bate-papo ou aceitação de uma sessão de partilha de anotação)
- C_1 = efetivamente fazer anotações em uma sessão compartilhada (anotação colaborativa)
- C_2 = capturar algum frame de vídeo e fazer anotações

As contingências comportamentais representadas a partir dessas ações e consequências, com suas respectivas declarações *if-then*, são:

- R1** $aA_1 \cap bcA_3 \rightarrow abcC_1$; *if* um usuário convida para uma sessão de anotação e os outros dois aceitam *then* os três usuários efetivamente fazem anotações (anotação colaborativa)
- R2** $aA_1 \cap bcA_3 \rightarrow ab\bar{c}C_1$; *if* um usuário convida para uma sessão de anotação e os outros dois aceitam *then* (apenas) dois usuários (efetivamente) colaboram na sessão de anotação (um dos usuários não anota)
- R3** $aA_1 \cap b\bar{c}A_3 \rightarrow ab\bar{c}C_1$; *if* um usuário convida para uma sessão de anotação e um dos dois não aceitam *then* (apenas) um usuário anota (os dois usuários convidados não anotam)
- R4** $aA_1 \cap \bar{b}cA_3 \rightarrow ab\bar{c}C_1$; *if* um usuário convida para uma sessão de anotação e apenas um usuário aceita (o outro usuário não aceita) *then* dois usuários (efetivamente) colaboram na sessão de anotação (o mesmo usuário que não aceita o compartilhamento também não anota)
- R5** $aA_2 \cap bcA_3 \rightarrow abcC_2$; *if* um usuário faz um compartilhamento de link de vídeo e os outros dois aceitam *then* cada um dos três usuários faz a captura de algum frame de vídeo e faz anotações no frame capturado
- R6** $aA_2 \cap b\bar{c}A_3 \rightarrow ab\bar{c}C_2$; *if* um usuário faz um compartilhamento de link de vídeo e um dos usuário aceita (o outro usuário não aceita) *then* dois usuários capturam algum frame de vídeo e faz anotações no frame capturado (o mesmo usuário que não aceita o compartilhamento também não captura e não anota)

R7 $aA_2 \cap \overline{bc}A_3 \rightarrow \overline{abc}C_2$; *if* um usuário faz um compartilhamento de link de vídeo e os outros dois não aceitam *then* apenas um usuário captura algum frame de vídeo e faz anotações no frame capturado (os outros dois usuários que não aceitam o compartilhamento também não capturam e não anotam).

Tabela A.1: Medidas para grupo G_2 e grupo G_3 [Gomes et al., 2011]

	ConR	SupR	wSupR	ConR	SupR	wSupR
$R1$	—	0	0	—	0	0
$R2$	—	0.5	3	—	0	0
$R3$	—	0	0	—	0	0
$R4$	—	0	0	—	0.5	2.125
$R5$	0	0	0	—	0	0
$R6$	1	0.5	2.125	—	0	0
$R7$	—	0	0	1	0.5	1.375

Tabela A.2: Medidas para grupo G_4 e todos Grupos [Gomes et al., 2011]

	ConR	SupR	wSupR	ConR	SupR	wSupR
$R1$	—	0	0	0	0	0
$R2$	—	0	0	1	0.1428	0.8571
$R3$	1	0.33	1.167	0.5	0.1428	0.5
$R4$	1	0.33	1.417	1	0.2857	1.2142
$R5$	—	0	0	0	0	0
$R6$	—	0.333	1.417	1	0.2857	1.2142
$R7$	—	0	0	1	0.1428	0.3928

As contingências comportamentais de $R1$ até $R4$ estão associadas ao compartilhamento assíncrono de sessões de anotação. As contingências comportamentais de $R5$ até $R7$ estão associadas ao compartilhamento assíncrono de links de vídeo. As contingências de $R1$ até $R7$ foram avaliadas usando as medidas de Suporte (Equação 3.1), Confidência (Equação 3.2) e Utilidade (Equação 3.6). Os resultados dessa avaliação apresentados nas Tabelas A.1 e A.2, foram usados para criar um ranqueamento entre as regras. O grupo de usuários $G1$ não se envolveu nas interações. Foi possível observar que os usuários estavam mais frequentemente envolvidos nas situações descritas em $R4$ (compartilhamento de sessão de anotação) e $R6$ (compartilhamento de links de vídeo).

No trabalho Gomes and Pimentel [2011d] foram representadas e avaliadas contingências comportamentais que envolvem o compartilhamento síncrono e assíncrono de sessões de anotação. No compartilhamento síncrono uma sessão de anotação é usando um frame capturado por usuário que está conectado a *CWaCTool*. Foram representadas 9 contingências comportamentais, que também foram avaliadas usando-se as medidas de Suporte, Confidência e Utilidade.

Contingências de $R1$ até $R6$ foram codificadas para representar situações que envolvem os usuários no compartilhamento síncrono de sessões de anotação. Contingências de $R7$ até $R9$ foram codificadas para representar situações que envolvem os usuários no compartilhamento assíncrono de sessões de anotação. O grupo de usuários $G1$ também não se envolveu nas interações sociais representadas nesse trabalho. Foi possível observar que os usuários estavam mais frequentemente envolvidos na situação descrita como:

- $aA_1 \cap bcA_3 \rightarrow abcC$. *If* um usuário compartilha uma sessão de anotação (usando um frame capturado por usuário que está conectado) e os outros dois usuários aceitam *then* os três usuários (efetivamente) colaborativamente fazem anotações.

Foram representadas 35 contingências comportamentais que envolvem, além das interações representadas e avaliadas nos trabalhos Gomes et al. [2011] e Gomes and Pimentel [2011d], situações nas quais os usuários fazem:

- compartilhamento explícito de link de vídeo e captura de um frame do vídeo compartilhado
- compartilhamento explícito de link de vídeo e comentário do vídeo compartilhado
- compartilhamento implícito de link de vídeo e captura de um frame do vídeo compartilhado
- compartilhamento implícito de link de vídeo e comentário do vídeo compartilhado

As 35 contingências comportamentais foram avaliadas usando as medidas de Suporte, Confidência, e ainda Correlação Cosseno (Equação 3.3) e Alavanca-gem 3.7. As medidas foram usadas nessa ordem para que o ranqueamento fosse realizado e, a partir desse ranqueamento, foi possível observar que os usuários estavam mais frequentemente envolvidos na situação descrita como:

- $aA_1 \cap b\bar{c}A_2 \rightarrow ab\bar{c}C_1$. *If* um usuário compartilha uma sessão de anotação (usando um frame capturado por usuário que está conectado) e um usuário aceitar (o outro não) *then* dois usuários (efetivamente) fazem anotações colaborativas (o mesmo usuário que não concordar em participar do compartilhamento também não fazem anotações).

As experimentações apresentadas nesta seção permitiram fazer observações iniciais sobre a utilização da linguagem de Mechner na representação de contingências comportamentais que ocorrem a partir do uso da ferramenta. Também

foi possível realizar observações iniciais sobre a utilização de medidas de regras na avaliação dessas contingências comportamentais.

Contingências comportamentais associadas a interações sociais que ocorrem a partir do uso da *CWaCTool*, foram codificadas utilizando-se a linguagem de Mechner. Foi possível detectar situações mais frequentes que envolveram os usuários da *CWaCTool* em interações sociais foi realizada usando-se a medida de Suporte. Foi possível ainda observar que algumas das situações representadas como contingências comportamentais não foram efetivamente executadas pelos usuários, pois essas contingências foram avaliadas com $SupR = 0$.

Foi observado o uso de diferentes medidas de regra para o ranqueamento de contingências comportamentais. Medidas que são derivadas de outras medidas podem não contribuir para que seja realizado um desempate no ranqueamento de contingências. Por exemplo, as regras $R1$ e $R6$ mencionadas acima tem os mesmos valores de $SupR$ e também os mesmos valores de $wSupR$.

A partir da representação de avaliação das interações sociais, foram sugeridas novas funcionalidades de interação para a *CWaCTool*, tais como:

- uma indicação que o item compartilhado não foi capturado pelo usuário que aceitou um compartilhamento,
- uma indicação de que o usuário aceitou assistir à sessão de anotação e que ele não tem (ou já têm) o vídeo, de onde o frame foi capturado, em sua máquina,
- uma indicação de que o usuário aceitou a assistir à sessão de anotação e ele tem (ou não) o quadro compartilhado em sua máquina.

A adição desses recursos pode contribuir para a identificação de novas situações nas quais os usuários serão envolvidos em interações sociais. Em seguida, a técnica aqui apresentada poderá ser aplicada para a análise dessas novas interações sociais.

A.3 Trabalhos Relacionados

Os esforços de pesquisa relacionados com a investigação apresentada neste apêndice contemplam o suporte para atividades colaborativas com conteúdo multimídia.

O apoio para atividades colaborativas com conteúdo multimídia para viabilizar a interação entre usuários é explorado em várias contribuições. Zhai et al. [2005] apresentam um sistema de anotação de vídeo colaborativo chamado de *eSports*, que utiliza uma arquitetura cliente-servidor para compartilhar anotações e dar suporte para a criação de conteúdo multimídia. Halvey et al. [2010]

descrevem um sistema para busca de vídeos on-line, que fornece mecanismos para grupos de usuários colaborarem entre si, tanto de forma síncrona e assíncrona, presencial ou a distância, em tarefas de busca de vídeo. O paradigma *Watch-and-Comment* [Cattelan et al., 2008] foi explorado em uma perspectiva social, por meio da ferramenta *CWaCTool* apresentada por Motti et al. [2009], que dá permissão para os usuários colaborem por meio de anotações de texto, voz e tinta sobre vídeos da rede YouTube. No trabalho aqui apresentado, atividades colaborativas com conteúdo multimídia são representadas e avaliadas como interações sociais entre usuários. A técnica aqui apresentada pode ser usada, por exemplo, como complementação para a avaliação de usabilidade dessas ferramentas.

Planejamento da Avaliação com Engenheiros de Software

1 Objetivo: por que avaliar?

- a) Entender quais são as necessidades e expectativas dos potenciais usuários
- b) Verificar se técnica e o método tem alguma utilidade para os usuários potenciais
- c) Obter opiniões técnicas dos usuários potenciais sobre a utilização da técnica e do método

2 Escopo: o que avaliar?

- a) Aspectos do sistema
 - Funcionalidade
Os procedimentos implementam tarefas que os usuários entendem como úteis?
- b) Aspectos no usuário
 - Aprendizado
O usuário compreendeu como é computada de cada etapa da técnica e do método?
 - Satisfação
Os resultados apresentados estão dentro das expectativas dos usuários?
O usuário está satisfeito com a forma com que os resultados foram apresentados?

3 Quando avaliar?

- a) Avaliação Formativa:

- Buscar informações que podem ser utilizadas para melhorar a técnica e o método apresentados
- Número pequeno de pessoas participam (potenciais usuários)
- Coletar opiniões antes da elaboração de um produto final (software)

4 Onde avaliar?

- a) Em ambiente escolhido pelo usuário

5 Participantes: Quem avaliar?

- a) Potenciais usuários – usuários do Facebook com conhecimento técnico para desenvolver aplicações de Computação Social
- b) Interesse na participação – voluntario

6 Foco da avaliação:

- a) Observar a repetição dos procedimentos
- b) Avaliar a utilidade da técnica e do método

7 Técnicas de avaliação

- a) Pesquisa de opinião: entrevista

Listagem B.1: *Planejamento da Avaliação com Engenheiros de Software*

Roteiro da Entrevista com Engenheiros de Software

Antes da entrevista , com a participação do usuário

- a) apresentar a técnica empregada no método (representação , extração e avaliação de interações sociais)
- b) apresentar o método e suas fases (descrição das fases e exemplos de arquivos de entrada e saída de cada fase)
- c) explicar o que será apresentado nos resultados (interações sociais genéricas e baseadas em mídia)

Iniciando a entrevista :

- a) mostrar os dados coletados
- b) explicar o que significa cada informação coletada
- c) explicar a limitações do que pode ser coletado

Pergunta 1:

O usuário se sente incomodado ao saber que o conteúdo de alguma informação coletada?

Executando o método:

- a) executar os scripts mostrando a saída de cada etapa
- b) interpretar com o usuário os resultados das medições das interações sociais genéricas , mostrando as regras e os gráficos elaborados a partir dos resultados das avaliações

Pergunta 2:

O usuário conseguiu interpretar o que está escrito nas regras?

Pergunta 3:

Qual a sua interpretação dos resultados das medidas de avaliação Suporte, Confidência e Correlação Cosseno?

Continuando a execução do método:

- a) mostrar como é realizada a especialização das interações sociais genéricas em interações sociais baseadas em mídia,
- b) executar o script que faz especialização da regras
- c) mostrar que as regras podem ser especializadas com tipos de mídias, provedores de conteúdo, e aplicações dos usuários
- d) interpretar com o usuário os resultados das medições das interações sociais genéricas, mostrando as regras e os gráficos elaborados a partir dos resultados das medições

Repetir perguntas 2 e 3 considerando interações sociais baseadas em mídia

Mostrar outras possibilidades de especialização das regras (explicitação do servidor que provem a mídia e explicitação das aplicações usadas no compartilhamento das mídias, por exemplo).

Pergunta 4:

A representação explícita de informações como tipos de mídias, provedores dessas mídias, e aplicações utilizadas nas interações sociais é útil? Como?

Pergunta 5:

O usuário tinha algum tipo de expectativa em relação ao método? Qual?

Pergunta 6:

O usuário tinha algum tipo de expectativa em relação aos resultados? Qual?

Pergunta 7:

A técnica e o método podem ser utilizados no desenvolvimento de aplicações?

Pergunta 8:

Na sua opinião, quais profissionais podem ter interesse no uso desses procedimentos para avaliação de interações sociais?

Pergunta 9:

É possível descrever sucintamente um cenário no qual uma aplicação social possa ser desenvolvida a partir aplicação da técnica e do método?

Listagem C.1: *Questionário aplicado na Avaliação com Engenheiros de Software*