

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

GUILHERME AUGUSTO MALDONADO DA CRUZ

**Determinação da confiança no desenvolvimento global de
software utilizando análise de sentimentos**

Maringá

2016

GUILHERME AUGUSTO MALDONADO DA CRUZ

**Determinação da confiança no desenvolvimento global de
software utilizando análise de sentimentos**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Informática, Centro de Tecnologia da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Orientadora: Prof^a. Dr^a. Valéria Delisandra Feltrim

Maringá
2016

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá – PR., Brasil)

C957d Cruz, Guilherme Augusto Maldonado da
Determinação da confiança no desenvolvimento
global de software utilizando análise de
sentimentos/ . -- Maringá, 2016.
128 f., il., color., figs., tabs.

Orientadora: Prof.a. Dr.a. Valéria Delisandra
Feltrim.

Dissertação (mestrado) - Universidade Estadual de
Maringá, Centro de Tecnologia, Programa de Pós-
graduação em Ciência da Computação, 2016.

1. Determinação de confiança - Equipe Virtual. 2.
Desenvolvimento global de Software. 3. Analise de
sentimento. I. Feltrim, Valéria Delisandra, orient.
II. Universidade Estadual de Maringá. Centro de
Tecnologia. Programa de Pós-Graduação em Ciência da
Computação. III. Título.

CDD 22. ED.005.1

JLM-000189

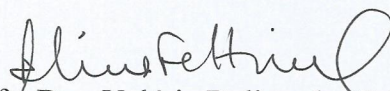
FOLHA DE APROVAÇÃO

GUILHERME AUGUSTO MALDONADO DA CRUZ

Determinação da confiança no desenvolvimento global de software utilizando análise de sentimentos

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Informática, Centro de Tecnologia da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Ciência da Computação pela Banca Examinadora composta pelos membros:

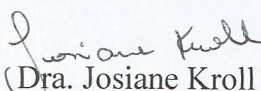
BANCA EXAMINADORA



Profa. Dra. Valéria Delisandra Feltrim
Universidade Estadual de Maringá – DIN/UEM



Profa. Dra. Elisa Hatsue Moriya Huzita
Universidade Estadual de Maringá – PCC/UEM



(Dra. Josiane Kroll
Pontifícia Universidade Católica do Rio Grande do Sul – PPGCC/PUCR

Aprovada em: 28 de março de 2016.

Local da defesa: Sala 101, Bloco C56, *campus* da Universidade Estadual de Maringá.

AGRADECIMENTOS

Agradeço a Deus por ter me abençoado nesta caminhada e as pessoas que contribuíram na realização deste trabalho. Em especial:

A minha família pelo carinho, apoio e incentivo.

A minha orientadora professora Dr^a. Valéria Delisandra Feltrim e a professora Dr^a. Elisa Hatsue Moriya Huzita pelo apoio, comentários e sugestões no desenvolvimento deste projeto.

Aos demais professores das disciplinas que cursei e aos colegas ...

E a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro concedido a este trabalho.

Determinação da confiança no desenvolvimento global de software utilizando análise de sentimentos

RESUMO

Com o avanço da tecnologia foi possível o surgimento de uma nova estrutura de trabalho, as equipes virtuais. Nessas equipes as pessoas se encontram em diferentes locais e possivelmente em fusos horários diferentes, fazendo uso de comunicação mediada por computador. Tal estrutura fornece alguns benefícios, tais como: maior flexibilidade para a realização de atividades e, também, menor dificuldade para encontrar mão de obra qualificada. Contudo, a distância gera alguns desafios, principalmente, por causa do meio de comunicação utilizado. O meio de comunicação, por sua vez, pode acarretar em maior dificuldade de se estabelecer confiança, essencial para a eficiência dessas equipes. Com o intuito de garantir maior eficiência destas equipes, pode-se utilizar a confiança estabelecida durante projetos anteriores no processo de alocação dos membros para uma nova equipe e/ou, durante a execução do projeto para monitorar o relacionamento entre os membros. Tendo como foco equipes de desenvolvimento global de software, a determinação da confiança pode ser realizada utilizando a técnica de análise de sentimentos sobre comentários em sistemas de versionamento, um sistema utilizado para o desenvolvimento de software, assim como outras características dos sistemas de versionamento. Esta dissertação apresenta um *framework* automático, denominado ARSENAL-GSD, para estimar a existência de confiança entre membros de equipes de desenvolvimento global de software utilizando análise de sentimentos desenvolvido a partir dos indícios de confiança e modelos de confiança encontrados na literatura. O *framework* permite aos gerentes de equipes de desenvolvimento global de software avaliar a interação entre os membros e o nível de confiança entre eles. Essas informações auxiliam os gerentes a escolher membros para novas equipes e permite o monitoramento da confiança entre os membros, garantindo maior eficiência as equipes.

Palavras-chave: Confiança. Desenvolvimento Global de Software. Análise de Sentimentos. Determinação de Confiança, Equipe Virtual.

Trust estimation in global software development using sentiment analysis

ABSTRACT

The advance in technology has enabled the emergence of a new work structure, virtual teams. In these teams, people are in different places and possibly over different time zone, making use of computer mediated communication. This structure provides great flexibility and some benefits such as to find skilled labor easier. However, distance creates some challenges, especially, regarding to means of communication. The means of communication in turn can bring greater difficulty in developing trust, essential for efficiency in these teams. In order to ensure greater efficiency of these teams, we can use the trust established during previous projects in the process of members allocation to a new team and/or, during the project execution in order to monitor the relationship between members. Focusing on global software development teams, trust estimation can be performed using sentiment analysis of versioning systems comments, a kind of system used for software development, as well as other of its features. This master thesis presents an automatic framework, called ARSENAL-GSD, for trust estimation between members of global software development teams using sentiment analysis developed from trust evidences and trust models found in literature. The framework allows global software development team managers to evaluate interaction between members and their trust level. This information assists managers to select members to new teams and allows trust monitoring between members, ensuring more efficiency to the teams.

Keywords: Trust. Global Software Development. Sentiment Analysis. Trust Estimation. Virtual team.

LISTA DE FIGURAS

2.1	Processo de formação da confiança interpessoal.	24
2.2	Antecedentes de Confiabilidade.	25
2.3	<i>Framework</i> para estimar confiança.	34
2.4	Modelo para estimar confiança utilizando questionários como fonte dos dados.	36
2.5	Conversação e Propagação.	37
2.6	Esquema de funcionamento das <i>pull requests</i>	47
3.1	Metodologia empregada	49
3.2	Representação do <i>framework</i> ARSENAL-GSD e seu modelo conceitual	52
3.3	Grafo de relações inicial.	55
3.4	Arestas entre nós em um grafo de relacionamentos.	55
3.5	Diagrama de casos de uso do <i>framework</i> ARSENAL-GSD.	57
3.6	Diagrama de componentes para o <i>framework</i> ARSENAL-GSD.	63
3.7	Diagrama de pacotes do <i>framework</i> ARSENAL-GSD.	64
3.8	Diagrama de classes do <i>framework</i> ARSENAL-GSD.	66
3.9	Diagrama de sequência - Extrair Dados do Projeto	70
3.10	Diagrama de sequência - Gerar Grafo de Confiança	71
4.1	Frequência para as respostas de caracterização dos participantes.	84
4.2	Frequência para as respostas de validação dos indícios.	85
4.3	Frequência para as respostas de validação dos antecedentes.	86
4.4	Frequência para as respostas de validação das fórmulas para os indícios.	87
4.5	Frequência para as respostas de validação das fórmulas para os antecedentes.	88
4.6	Frequência para as respostas de validação da fórmula de confiança.	88
4.7	<i>Framework</i> ARSENAL-GSD com os pesos ajustados.	89
5.1	Tela da aplicação de exemplo.	97
5.2	Grafo de relacionamentos inicial.	98
5.3	Grafo de confiança.	98
5.4	<i>Pull request</i> 135.	100
5.5	<i>Pull request</i> 138.	100
5.6	Arestas entre cdzombak e dzlobin no grafo de relacionamentos.	101

1.1	<i>Framework</i> proposto para estimar a confiança.	115
1.2	Fórmulas utilizadas para extração dos indícios no <i>framework</i> . . .	118
2.1	Questionário - Parte 1	122
2.2	Questionário - Parte 2	123
2.3	Questionário - Parte 3	124
2.4	Questionário - Parte 4	125
2.5	Questionário - Parte 5	126
2.6	Questionário - Parte 6	127
2.7	Questionário - Parte 7	128

LISTA DE TABELAS

2.1	Trabalhos aceitos	28
2.2	Síntese dos Modelos de Confiança	42
2.3	Taxa de acerto das ferramentas de análise de sentimentos sobre os comentários anotados.	45
3.1	Características do ARSENAL-GSD	56
4.1	Sete etapas para realização de uma <i>survey</i> (Kasunic, 2005).	80
4.2	Estatística descritiva indícios e antecedentes da confiabilidade.	90
4.3	Resultado teste Shapiro-Wilk para normalidade.	91
4.4	Resultado testes.	92
4.5	Estatística descritiva validade das fórmulas para indícios e antecedentes.	92
4.6	Distribuição das respostas agrupando por conhecimento em DGS e Aceito/Rejeitado juntamente com o resultado do teste exato de Fisher.	93

LISTA DE SIGLAS E ABREVIATURAS

API: *Application Programming Interface*

ARSENAL-GSD: *Automatic tRust eStimator based on sENTiment anALysis for Global Software Development*

CMC: Comunicação Mediada por Computador

DGS: Desenvolvimento Global de Software

RDF: *Resource Description Framework*

TI: Tecnologia da Informação

WOT: *Web of Trust*

SUMÁRIO

1	Introdução	12
1.1	Motivação e Justificativa	13
1.2	Objetivos	15
1.3	Estrutura do Trabalho	16
2	Revisão de literatura	17
2.1	Considerações Iniciais	17
2.2	Equipes Virtuais	17
2.3	Desenvolvimento Global de Software	19
2.4	Confiança	21
2.4.1	Confiança Interpessoal	23
2.4.2	Confiança em DGS	24
2.4.3	Indícios de Confiança	26
2.4.4	Modelos de Confiança	33
2.5	Análise de Sentimentos	43
2.5.1	SentiStrength	44
2.6	GitHub	46
2.7	Considerações finais	47
3	ARSENAL-GSD	48
3.1	Considerações Iniciais	48
3.2	Metodologia	48
3.3	Características	53
3.3.1	Trabalhos relacionados	55
3.4	Projeto e Implementação	56
3.4.1	Técnicas de extração de indícios	72
3.4.2	Funcionamento	75
3.5	Considerações Finais	76
4	Validação	78
4.1	Considerações Iniciais	78
4.2	A <i>Survey</i>	79
4.2.1	Identificar os objetivos de pesquisa	79
4.2.2	Identificar e caracterizar o público alvo	80
4.2.3	Projetar o plano de amostragem	81

4.2.4	Projetar e escrever o questionário	81
4.2.5	Realização do teste piloto	82
4.2.6	Analisar os resultados e escrever o relatório	82
4.2.7	Ameaças a Validade	94
4.3	Considerações Finais	95
5	Utilizando o ARSENAL-GSD	96
5.1	Considerações Iniciais	96
5.2	Instanciando o ARSENAL-GSD	96
5.3	Um Exemplo de Uso	97
5.3.1	Arestas Parciais e Finais	98
5.3.2	Utilizando as Saídas	99
5.4	Considerações Finais	101
6	Conclusão	103
6.1	Limitações	104
6.2	Contribuições	104
6.3	Trabalhos Futuros	105
	REFERÊNCIAS	107
A	Documento para Realização da <i>Survey</i>	115
A.1	Visão Geral	115
A.2	Fórmulas	117
B	Questionário - <i>Survey</i>	121

Introdução

A comunicação é essencial para os seres humanos. Desde o momento em que uma criança nasce ela tenta se comunicar com seus pais por meio do choro. Conforme o tempo passa, ela aprende a falar e passa a se comunicar com as pessoas ao redor, permitindo que ela desenvolva relacionamentos de todos os tipos. Em uma empresa a comunicação é muito importante, pois é por meio da comunicação que os funcionários conseguem colaborar e alcançar os objetivos da empresa. Assim, cada setor deve se comunicar com os outros de forma eficiente para que a empresa prospere.

Quando as pessoas estão geograficamente distantes, a comunicação se torna mais difícil. Desde os primórdios da humanidade o homem tenta se comunicar a distância, por meio dos sinais de fumaça, passando pelos tambores, pombos correio, telegrafo, telefone até, mais recentemente com a evolução das tecnologias de comunicação e informação, o e-mail, fóruns, aplicativos de mensagem instantânea e telepresença (Khan, 2012). O avanço dessas tecnologias facilitaram a comunicação entre pessoas distribuídas geograficamente permitindo a aumento de equipes virtuais.

Paralelamente a isso, no final da década de 80, com o aumento da demanda por produtos de software, houve uma grande quantidade de fusões e aquisições entre as empresas de tecnologia da informação (TI). Essas empresas passaram a focar em uma área de expertise e a terceirizar o restante das atividades para outras empresas especializadas (Sengupta et al., 2006) por meio da utilização de equipes virtuais. Com a globalização dos negócios, as equipes virtuais tornaram-se indispensáveis para as organizações globais (Berry, 2011; Paul e He, 2012). Nas equipes virtuais os membros devem trabalhar em conjunto por um curto período, aprender uns com os outros e alcançar os objetivos. Entretanto, coordenar

o trabalho desses membros que não se conhecem utilizando comunicação mediada por computador (CMC) é um desafio (Iacono e Weisband, 1997).

O desenvolvimento de software utilizando equipes virtuais caracteriza o desenvolvimento global de software (DGS) que tem como objetivo proporcionar algumas vantagens, mas a distribuição geográfica acrescenta alguns desafios. Dentre as vantagens proporcionadas por essa estrutura podemos citar: facilidade em encontrar mão de obra especializada, aumento da jornada de trabalho e diminuição dos custos (O’Conchuir et al., 2006). Mas a distribuição geográfica e as diferenças culturais também trazem seus desafios, principalmente no que se refere a comunicação, que é dependente, em sua maioria, de CMC (Herbsleb e Moitra, 2001; Sengupta et al., 2006).

Um dos desafios enfrentados pelas equipes virtuais e, portanto, no desenvolvimento global de software é a confiança entre os membros da equipe, em parte ocasionado pelo tipo de comunicação utilizado. Existem vários trabalhos que mostram a importância da confiança para as equipes de DGS (Al-Ani et al., 2013, 2011; Kuo e Yu, 2009; Pangil e Chan, 2014; Siakas e Siakas, 2008). O compartilhamento de informações e conhecimento, necessários para a realização das tarefas, depende da confiança uma vez que as pessoas tendem a buscar e aceitar o conhecimento das pessoas em quem elas confiam. A confiança afeta, ainda, a cooperação (McNab et al., 2012; Striukova e Rayna, 2008), a coordenação (Peñarroja et al., 2013), a comunicação (Siakas e Siakas, 2008) e o estabelecimento da coesão (Yusof e Zakaria, 2012) entre os membros da equipe.

1.1 Motivação e Justificativa

Informações relativas à confiança têm sido utilizadas por algoritmos de recomendação de pessoas, serviços, itens e vendedores como forma de melhorar a qualidade da recomendação (Cheng et al., 2012; Li et al., 2010a,c; Nepal et al., 2013). Da mesma forma, as informações de confiança entre os membros de uma equipe de desenvolvimento podem ser utilizadas para recomendação de equipes (Awal e Bharadwaj, 2014) e também, para monitorar o relacionamento entre os membros (Fan et al., 2011).

Dada a importância da confiança para o desenvolvimento global de software, estimar a confiança entre as pessoas envolvidas no DGS é um passo importante para garantir a eficiência da equipe. Segundo Kuo e Yu (2009) equipes de alto desempenho requerem alto nível de confiança, uma vez que equipes de baixa confiança precisam se esforçar mais do que equipes de alta confiança para produzir com a mesma qualidade (Jarvenpaa et al., 2004). Para estimar a confiança podemos utilizar um modelo de confiança, como em Awal e Bharadwaj (2014). Vários modelos de confiança (Cheng et al., 2012; Huynh et al., 2006;

Lavrac et al., 2007; Skopik et al., 2009) podem ser encontrados na literatura, geralmente incorporados em algoritmos para recomendação de pessoas, produtos e serviços.

As informações utilizadas por esses modelos podem ser extraídas, por exemplo, a partir de redes sociais (Li et al., 2010c). Em geral as informações referem-se à quantidade de interações, avaliação dessas interações e o seu sucesso. Contudo, em um ambiente de trabalho as pessoas podem não se sentir à vontade para fornecer uma avaliação ao final da interação com um colega de trabalho. Ainda, se houverem grandes quantidades de interações, as pessoas podem passar a fornecer avaliações incorretas o que impede o cálculo correto da confiança. A estimativa do sucesso de uma interação, por sua vez, vai depender do tipo da interação. Em Skopik et al. (2009) um conjunto de métricas foi desenvolvido para analisar o sucesso de uma interação. No entanto, essas métricas são dependentes do domínio e, ainda, ignoram a subjetividade da avaliação, usam uma concepção de sucesso que é igual para todas as pessoas, que não é verdade. O que é considerado bom para uma pessoa, pode não ser bom para outra.

Por meio de um mapeamento sistemático buscamos formas de se determinar automaticamente a presença de confiança entre os membros de uma equipe. Não encontramos na literatura um meio de determiná-la exatamente, mas encontramos um conjunto de comportamentos e características exibidas por equipes em que existia confiança entre os membros. Chamamos esses comportamentos/características de indícios de confiança. Exemplos desses indícios são o tom positivo, a comunicação intensa e a mímica de vocabulário. Parte desses indícios podem ser capturados sem intervenção humana. Dessa forma, incrementando os modelos de confiança com os indícios encontrados, podemos melhorar as estimativas de confiança entre membros de uma equipe.

Como o foco deste trabalho é o DGS, uma fonte de dados que pode ser utilizada a fim de extrair as informações para se estimar confiança (os indícios de confiança) são os sistemas de versionamento. Neste trabalho o interesse está nos sistemas de versionamento que permitem aos usuários submeter *commits* e comentar o *commit* dos colegas, pois a partir desses comentários podem ser extraídos alguns dos indícios de confiança considerados, como o tom positivo dos comentários. Exemplos desses sistemas são o GitHub¹ e o BitBucket². Neste trabalho utilizamos o GitHub como estudo de caso, focando nas *pull requests*, nas quais um usuário submete um *commit* ao repositório e outros usuários podem interagir fornecendo comentários até que o *commit* seja aceito ou não.

Uma vez que a confiança é um aspecto subjetivo e próprio de cada indivíduo, seus indícios também envolvem questões afetivas e subjetivas. A área de pesquisa que busca

¹<https://github.com/>

²<https://bitbucket.org/>

capturar automaticamente aspectos afetivos a partir da comunicação em diferentes mídias é a análise de sentimentos. Entre os problemas tratados pela análise de sentimentos, o mais comum e difundido é a análise da polaridade, que busca classificar, para um determinado texto, a polaridade do sentimento expresso: se positiva, negativa ou neutra. Dessa forma, ferramentas de análise de sentimentos podem ser utilizadas para determinar se um comentário tem um tom positivo ou não. Além disso, outros indícios de confiança podem ser capturados a partir dessa análise, como a benevolência e, neste caso específico, a habilidade de um usuário, uma vez que, o *commit* é efetuado como resultado da execução de atividades (incluindo-se alterações e/ou atualizações) realizadas usando as habilidades que uma pessoa possui. Tanto a benevolência quanto a habilidade impactam na confiabilidade de uma pessoa, que por sua vez é um indício de confiança.

Observe que os indícios de confiança tom positivo e mímica de vocabulário são particulares de cada membro. E ao utilizá-los conseguimos capturar ao menos em parte a subjetividade inerente da confiança que é desprezada com o uso de métricas. Dessa forma, a incorporação desses indícios aos modelos de confiança permite uma estimativa de confiança mais específica para cada membro da equipe, eliminando a necessidade de avaliação do sucesso das interações por parte dos membros e reduzindo a dependência do domínio.

Nesse contexto, a questão de pesquisa que guiou este trabalho é:

- Como podemos estimar a existência de confiança entre membros de uma equipe DGS de forma automática (dispensa a necessidade dos membros fornecerem informações) com o auxílio da análise de sentimentos?

1.2 Objetivos

Este trabalho tem como objetivo desenvolver um meio automático de estimar a confiança entre os membros de uma equipe de DGS a partir de indícios de confiança observados nas interações em sistemas de versionamento, com o auxílio da análise de sentimentos. Para alcançar esse objetivo geral, foram traçados os seguintes objetivos específicos:

- Identificar os indícios de confiança relevantes para DGS;
- Identificar modelos para estimar confiança existente em DGS e equipes virtuais;
- Mapear os indícios e elementos dos modelos, que podem ser obtidos com a análise de sentimentos;

- Identificar os indícios que podem ser obtidos a partir das interações em sistemas de versionamento; e
- Projetar um *framework* para estimar a confiança entre membros de equipes DGS.

1.3 Estrutura do Trabalho

Na Seção 2 apresentamos os conceitos utilizados no desenvolvimento do trabalho, sendo eles equipes virtuais, DGS, confiança e análise de sentimentos. A Seção 3 apresenta o *framework* ARSENAL-GSD. A validação dos indícios e fórmulas utilizadas no *framework* estão na Seção 4. Um exemplo de utilização do *framework* é apresentado na Seção 5. Por fim, a Seção 6 apresenta as conclusões e direcionamentos para trabalhos futuros.

Revisão de literatura

2.1 Considerações Iniciais

Esta Seção apresenta uma revisão da literatura sobre os assuntos abordados nesta dissertação. A Seção 2.2 define equipes virtuais e apresenta as vantagens e desafios dessas equipes. O desenvolvimento global de software é abordado na Seção 2.3. A Seção 2.4 trata de confiança, um dos desafios presentes em equipes virtuais. Um conjunto de indícios da confiança é apresentado na Seção 2.4.3 e os trabalhos relacionados são apresentados na Seção 2.4.4. A Seção 2.5 apresenta brevemente a análise de sentimentos, uma área ligada ao processamento de linguagem natural e a ferramenta para análise de sentimentos SentiStrength. Na Seção 2.6 apresentamos o GitHub, um sistema de versionamento para o qual implementamos uma instância do *framework* e, por fim, na Seção 2.7 são feitas as considerações finais desta seção.

2.2 Equipes Virtuais

Segundo Mitchell e Zigurs (2009, p.62, tradução nossa) uma equipe virtual pode ser definida como: “[...]um grupo de indivíduos que trabalham de forma interdependente considerando espaço, tempo, fronteiras organizacionais e/ou práticas de trabalho por meio do uso de tecnologias colaborativas para alcançar objetivos em comum”. Khan (2012) acrescenta que essas equipes têm uma natureza temporária (podem nunca ter trabalhado e nem trabalhar juntos novamente). Essa natureza temporária dificulta o desenvolvimento da confiança entre os membros da equipe (Pangil e Chan, 2014).

O acesso à reputação dos membros pode facilitar o desenvolvimento de confiança em equipes virtuais, mas quando a equipe envolve mais de uma organização, os membros não têm como acessar a reputação dos membros de outra empresa (Rusman et al., 2010).

Dentre os benefícios que levam as empresas a adotar equipes virtuais podem ser citados (Pangil e Chan, 2014):

- Contratação de mão de obra qualificada: as equipes virtuais permitem a contratação de pessoas que se encontram em locais diferentes aumentando a oferta de pessoal qualificado;
- Aumento da jornada de trabalho para até 24h: as equipes podem explorar a diferença de fusos horários criando uma equipe virtual que, trabalha até 24h por dia, pois quando a jornada de trabalho acaba em um local, ela começa em outro;
- Maior flexibilidade permitindo uma maior competitividade e responsividade ao mercado. As empresas podem criar equipes em locais diferentes de acordo com as oportunidades de negócio.

Junto das vantagens temos algumas desafios que dificultam o estabelecimento/manutenção da comunicação e da confiança entre os membros de equipes virtuais (Nunamaker et al., 2009 apud Pangil e Chan, 2014):

- Perda de pistas não-verbais: a utilização de CMC causa perda de pistas não verbais como a expressão facial do ouvinte;
- Redução dos mecanismos de conversas informais: como os membros das equipes não estão próximos, as oportunidades de conversas informais diminuem. Por exemplo, as conversas de corredor não são possíveis com alguém que está em outro continente.
- Redução das oportunidades de se construir amizades: a distância entre as pessoas dificulta a interação entre elas, que por sua vez dificulta o desenvolvimento de uma amizade;
- Fusos horários diferentes: dificulta o agendamento de reuniões e a comunicação síncrona;
- Tecnologia de difícil utilização e não confiável: alguns membros não compreendem como utilizar as tecnologias de um ambiente virtual. Mesmo em uma equipe DGS, nem todos os envolvidos são da área de TI;

- Dificuldade de construir um consenso: nem sempre os membros vão concordar em tudo, principalmente se tiverem culturas diferentes;
- Culturas diferentes: a diferença cultural impacta em como as pessoas reagem as situações e o que é considerado “normal” em uma cultura, pode ser ofensivo em outra;
- Diferentes processos de trabalho: os membros de uma equipe virtual podem utilizar processos de trabalho diferentes e essa diferença pode ocasionar conflitos.

2.3 Desenvolvimento Global de Software

Buscando a redução dos custos, agilidade no desenvolvimento e acesso mão de obra qualificada, as organizações passaram a utilizar equipes virtuais aliadas à terceirização, o chamado desenvolvimento global de software (DGS) (Jiménez e Piattini, 2009). Dentre os fatores que aceleraram o aparecimento dessas equipes globais, além dos benefícios fornecidos pelas equipes virtuais, temos (Herbsleb e Moitra, 2001):

- Necessidade de capitalizar o conjunto de recursos globais para permitir sua utilização independentemente da localização;
- Vantagens da proximidade de mercado, como conhecimento do cliente e situação local;
- Formação rápida de equipes para aproveitar as oportunidades do mercado;
- Pressão para entrega mais rápida dos produtos;
- Flexibilidade para capitalizar oportunidades independentemente de onde elas se encontrem.

O DGS é uma atividade colaborativa, que pode ser caracterizada por ter: membros de diferentes culturas e organizações, separados pelo tempo e espaço que usam CMC para colaborar (O’Conchuir et al., 2006; Sengupta et al., 2006). Essa organização das equipes tem como objetivo proporcionar alguns benefícios:

- Redução do custo de desenvolvimento: alguns locais oferecem mão de obra a um custo menor do que onde a empresa se encontra;

- Desenvolvimento *follow-the-sun*: explora-se a diferença de fuso horário de forma que quando uma equipe termina sua jornada de trabalho a outra inicia, a fim de agilizar o desenvolvimento;
- Modularização do trabalho entre os locais de desenvolvimento: o projeto pode ser modularizado de forma que cada local de desenvolvimento seja responsável pelo desenvolvimento de um módulo;
- Acesso a mão de obra qualificada: as empresas podem contratar pessoas em qualquer local do mundo;
- Inovação e compartilhamento de práticas: as equipes podem utilizar práticas diferentes, que por sua vez podem ser compartilhadas entre os sítios. As diferenças entre os membros podem favorecer inovações.
- Proximidade do mercado: é possível ter uma equipe próxima ao cliente que possua a mesma cultura e um entendimento melhor da situação do cliente.

Contudo, O'Conchuir et al. (2006) apontam que nem sempre esses benefícios são garantidos. Enviar algumas tarefas para países com baixo custo de produção pode acrescentar problemas se não houver planejamento ou se a equipe não tiver conhecimento suficiente para sua realização. O desenvolvimento *follow-the-sun* aumenta a quantidade de horas de trabalho, mas se não houver sobreposição de horários, a comunicação entre as equipes é afetada. Dessa forma, atividades de desenvolvimento deveriam evitar o *follow-the-sun*, enquanto que atividades de resolução de defeitos e suporte podem se dar bem com o *follow-the-sun*. A modularização do trabalho permite desenvolvê-lo paralelamente, mas pode gerar problemas de integração e, dependendo do tamanho das tarefas, pode não trazer aumento de eficiência. A possibilidade de encontrar mão de obra em outros locais é um grande atrativo, mas problemas devidos às diferenças culturais podem surgir. Devido às diferenças de perfil, existe uma maior possibilidade de inovação e compartilhamento de práticas, mas com a falta de comunicação informal isso não acontece muito. Ao mesmo tempo que se tem uma equipe perto do cliente (proximidade do mercado) e que tem a mesma cultura do cliente, os problemas de diferença cultural podem aparecer entre as equipes.

Assim, além de benefícios, o DGS também traz alguns desafios que se somam aos desafios das equipes virtuais (Herbsleb e Moitra, 2001; Sengupta et al., 2006):

- Problemas estratégicos: nessa categoria podemos citar como será realizada a divisão do trabalho entre as equipes e a privacidade dos dados transmitidos entre essas

equipes. A resistência organizacional, devido a funcionários que são contra a adoção do DGS na empresa, é outro problema estratégico que pode ocorrer.

- Comunicação inadequada: Comunicados formais devem seguir uma interface clara para que todas as equipes possam entender.
- Gerência do conhecimento: importante quando uma nova equipe é adicionada ao projeto. O compartilhamento de conhecimento e informações deve ser realizado de forma adequada para que os membros saibam o que é crítico ou não, e a documentação em DGS deve estar sempre atualizada para que membros possam acessar sempre a versão atual.
- Problemas técnicos: incluem formatos e ferramentas diferentes, problemas de rede e segurança de transmissão de dados.

2.4 Confiança

A confiança é estudada em diversas áreas do conhecimento, como: a psicologia, a filosofia e a economia. Com base nas definições de diversas áreas, Rusman et al. (2010, p.836, tradução nossa) definiram confiança como:

Um estado psicológico positivo (cognitivo e emocional) do confiante em relação ao confiado, que compreende as expectativas positivas do confiante a respeito das intenções e comportamentos futuros do confiado, levando a uma vontade de expressar um comportamento de confiança em um contexto específico.

Essa definição apresenta uma das características da confiança, que é a sensibilidade ao contexto. Outra definição que foi identificada como a mais utilizada pela comunidade na revisão de Mitchell e Zigurs (2009) é:

[...]a vontade de uma das partes ser vulnerável às ações de outra parte com base na expectativa de que a outra parte irá executar uma determinada ação importante para o confiante, independentemente da capacidade de monitorar ou controlar a outra parte (Mayer et al., 1995 apud Mitchell e Zigurs, 2009, p.62, tradução nossa).

Essa definição menciona a independência da capacidade de monitorar ou controlar a outra parte, que é justamente o motivo pelo qual as equipes virtuais necessitam da

confiança. Devido a distância entre os membros, estes não são capazes de monitorar o que os outros membros da equipe estão fazendo.

Sherchan et al. (2013) apresentam algumas propriedades da confiança:

- **Específica ao contexto:** Tem-se confiança em um médico para dar um diagnóstico e em um mecânico para consertar o carro, mas não o contrário.
- **Dinâmica:** A confiança muda com o tempo, podendo aumentar ou diminuir com base nas experiências do confiante.
- **Não-transitiva:** A confiança não é transitiva. Se A confia em B e B confia em C , não necessariamente A confia em C .
- **Propagativa:** Apesar de não ser transitiva, se A confia em B e B confia em C , A pode derivar um nível de confiança em C .
- **Composta:** Se não existem interações com uma pessoa, a confiança nela pode ser obtida agregando recomendações de outros a seu respeito.
- **Subjetiva:** Considerando revisões sobre filmes, se duas pessoas têm o mesmo gosto por filmes, provavelmente uma irá confiar na revisão da outra, mas se os gostos são diferentes, não.
- **Assimétrica:** Se você confia em alguém não necessariamente essa pessoa confia em você.
- **Auto-reforço:** As pessoas tratam bem quem elas confiam, aumentando a confiança. Da mesma forma, elas não se relacionam com quem elas não confiam, impedindo o desenvolvimento de confiança.
- **Sensível a eventos:** A confiança requer tempo para ser desenvolvida, mas um evento negativo pode destruí-la rapidamente.

As três principais categorias de confiança são (McKnight e Chervany, 1996 apud Mumbi, 2007) (i) confiança interpessoal, que é a confiança que se estabelece entre pessoas, seja individualmente ou como grupo; (ii) confiança institucional, que é a confiança relacionada à organização, imposta pela sua estrutura e (iii) disposição à confiança, que se refere a quão disposta uma pessoa é a confiar em alguém. Este trabalho foca na confiança interpessoal e por isso ela é abordada na próxima subseção.

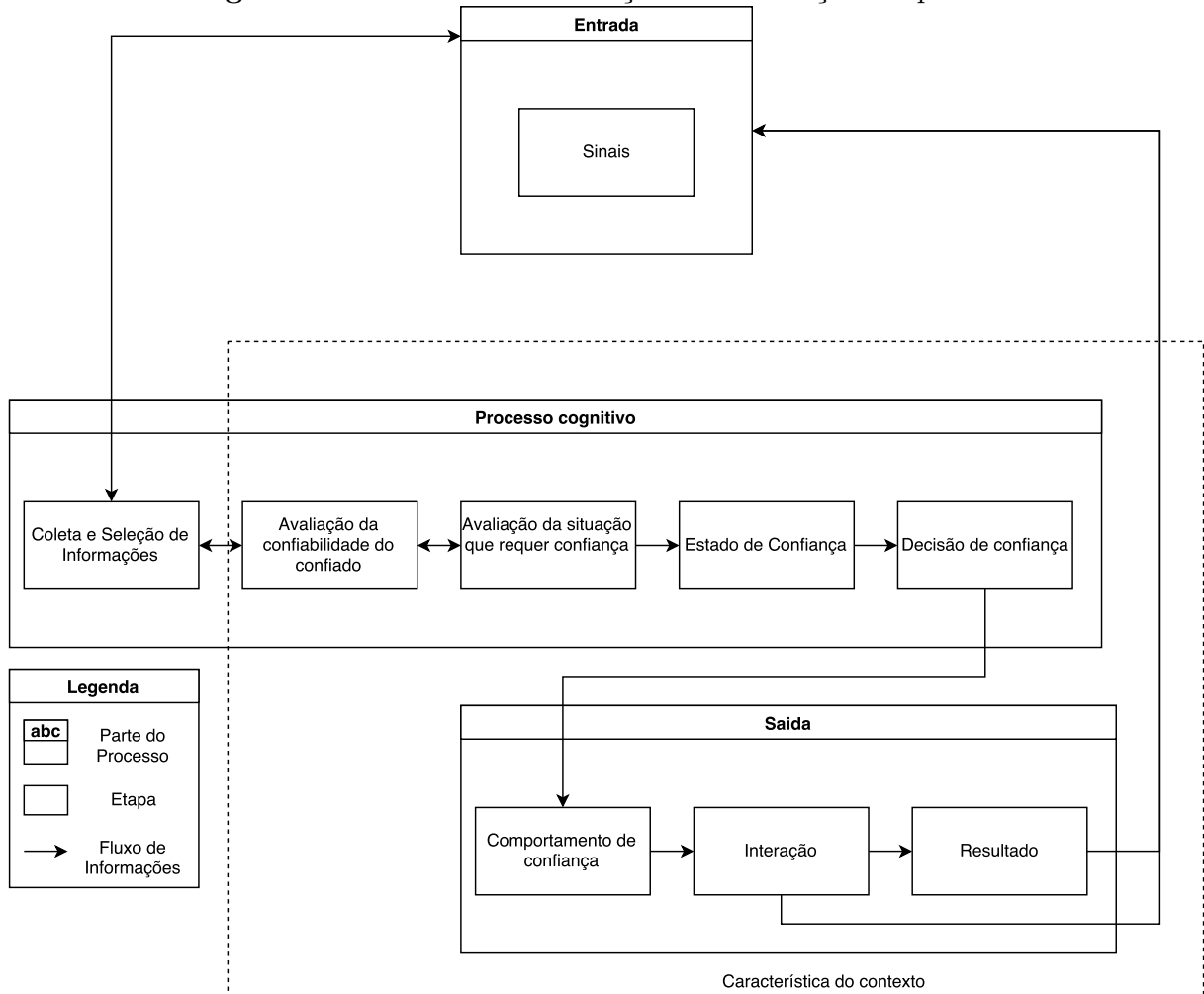
2.4.1 Confiança Interpessoal

Antes de tomar a decisão de confiar, uma pessoa avalia a confiabilidade da pessoa a ser confiada e o risco. Confiança e confiabilidade não são a mesma coisa, “a confiança compreende a confiabilidade observada do confiado” enquanto “a confiabilidade é a avaliação do confiante de quanto e pra que tipo de desempenho um confiado (pessoa) pode ser confiado” (Rusman et al., 2010, tradução nossa). Dessa forma que se a pessoa optar por confiar estará vulnerável ao confiado positiva e negativamente, assumindo um risco (Rusman et al., 2010). Dessa forma, a disponibilidade de informações pessoais dos membros da equipe influencia positivamente o nível de confiança interpessoal da equipe, colaborando para que os membros avaliem a confiabilidade de um membro.

Segundo Rusman et al. (2010) a formação de confiança interpessoal segue um processo cognitivo que possui entradas e gera uma saída (Figura 2.1). As entradas são os sinais observados por uma pessoa. Como nas equipes virtuais esses sinais estão presentes em menor quantidade devido à CMC, a formação de confiança é dificultada. Nesse processo o confiante coleta e seleciona as informações, obtidas por meio dos sinais de entrada, que considera relevantes para avaliar a confiabilidade da pessoa. A confiabilidade é avaliada com base em um conjunto de propriedades chamadas de antecedentes de confiabilidade (Figura 2.2). A partir da análise da confiança no confiado e da situação que requer confiança, o confiante opta por confiar no confiado ou não. Caso o confiante opte por confiar, ele passa a exibir um comportamento de confiança. Para equipes virtuais, o comportamento de confiança pode ser caracterizado pelo compartilhamento de recursos, delegação de tarefas, entre outros. Ao final da interação entre o confiante e confiado o resultado da interação realimenta o processo.

Conforme pode ser visto no processo da Figura 2.1, o estabelecimento da confiança depende em grande parte da confiabilidade das pessoas envolvidas (Rusman et al., 2010). Os antecedentes de confiabilidade são atributos utilizados para se avaliar a confiabilidade de uma pessoa. A Figura 2.2 apresenta os antecedentes de confiabilidade levantados por Rusman et al. (2010) divididos em cinco categorias: (i) Comunalidade, que diz respeito a quão similares as pessoas são, (ii) Capacidade, que se refere às capacidades intelectuais da pessoa, (iii) Benevolência, que se refere ao nível de cortesia e positividade que o confiado exibe em relação ao confiante, (iv) Normas internalizadas, se referem às normas morais do confiado e (v) Dependabilidade, que se refere ao quanto se pode depender de uma pessoa. Em Rusman et al. (2013) foi constatado que no início da colaboração as pessoas têm preferência por informações que digam respeito à competência, ao comprometimento, a responsabilidade, a disponibilidade e a comunalidade.

Figura 2.1: Processo de formação da confiança interpessoal.

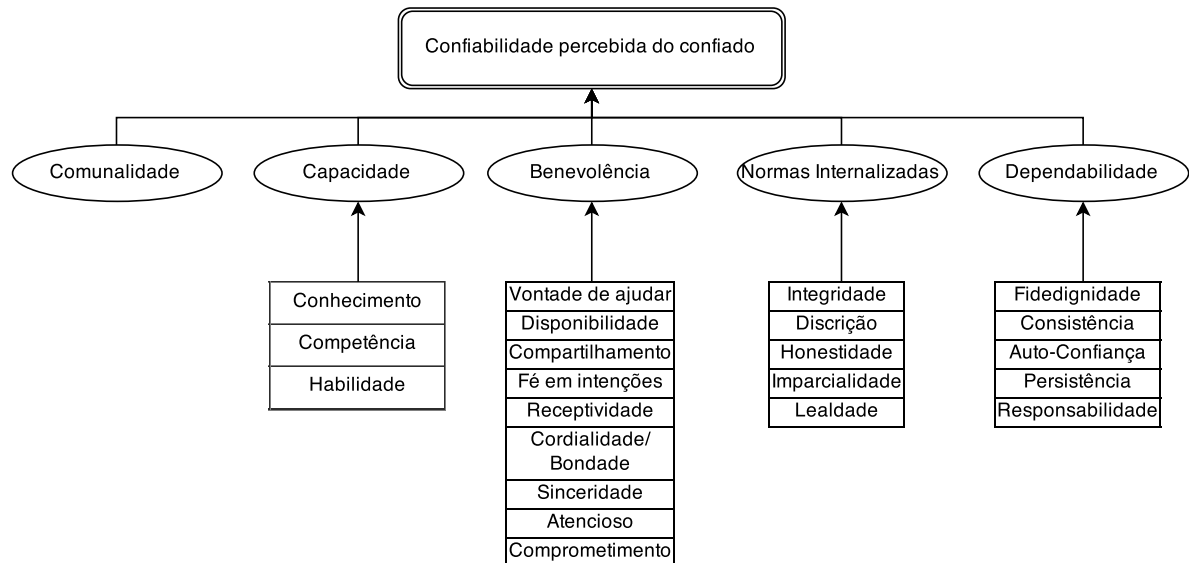


Fonte: traduzido de Rusman et al. (2010).

2.4.2 Confiança em DGS

Um dos desafios presentes em equipes virtuais e, portanto, em DGS é a criação e manutenção da confiança (Al-Ani et al., 2013; Gomes e Marczak, 2012; Khan, 2012). Isso se deve, a uma serie de fatores já mencionados anteriormente, tais como à falta de informações passadas pela CMC (Al-Ani et al., 2014; Peñarroja et al., 2013) e ao fato de que as vezes essas equipes têm baixa coesão (Jarvenpaa e Leidner, 1998).

Nesse contexto, a confiança é importante nas equipes DGS devido à incapacidade dos membros em checar o que os outros membros estão fazendo simplesmente observando o que eles estão fazendo (Jarvenpaa et al., 1998). Monitorar os membros dessas equipes torna-se caro e a diferença de informação entre os membros aumenta a chance de risco moral e de ocorrer o problema do principal-agente. Risco moral refere-se a incapacidade

Figura 2.2: Antecedentes de Confiabilidade.

Fonte: adaptado de Rusman et al. (2010).

de monitorar as ações de uma das partes após iniciada a interação. A confiança diminui esses riscos bem como o custo de monitoramento (Al-Ani et al., 2013; Striukova e Rayna, 2008). Segundo Kuo e Yu (2009) equipes de alto desempenho requerem alto nível de confiança. As equipes de baixa confiança devem se esforçar mais do que equipes de alta confiança para produzir com a mesma qualidade (Jarvenpaa et al., 2004).

De acordo com Khan (2012) a confiança é a precursora das relações, impactando no compartilhamento de conhecimento (Al-Ani et al., 2013; Pangil e Chan, 2014; Striukova e Rayna, 2008) e informações que são importantes para o sucesso de equipes virtuais (Filstad e Gottschalk, 2010). Quando novas informações são compartilhadas a qualidade da solução aumenta (Paul e He, 2012). Entretanto, para existir esse compartilhamento é necessário que exista confiança, principalmente quando as informações compartilhadas são confidenciais.

Além do impacto no compartilhamento de conhecimento e informações, a confiança também afeta de forma positiva a coesão (Yusof e Zakaria, 2012), a cooperação (McNab et al., 2012; Striukova e Rayna, 2008) e a coordenação (Peñarroja et al., 2013), atributos essenciais para as equipes virtuais. Afeta também como as pessoas reagem às situações. Como exemplo, considere que uma pessoa A envie um e-mail para uma pessoa B e esta demora a responder. Se A confia em B, A pode atribuir a demora a problemas técnicos, caso contrário, A pode assumir que B simplesmente não quis responder (Jarvenpaa et al., 2004).

Siakas e Siakas (2008) listam algumas vantagens provenientes do estabelecimento de confiança nas relações de terceirização:

- Mitiga o comportamento oportunista: as empresas envolvidas (cliente e vendedor) trabalham em conjunto mas buscam seus próprios objetivos, e o estabelecimento de confiança entre eles impede que um se aproveite do outro;
- Melhora a comunicação: quanto maior a confiança entre os membros, melhor é a comunicação entre eles;
- Desenvolve a coesão: a confiança cria expectativas compartilhadas entre os membros da equipe como um todo;
- Suporta a criação de objetivos mútuos: os membros passam a buscar os objetivos da equipe em vez de seus próprios objetivos;
- Facilita a transferência de conhecimento e recursos: a confiança promove o contato interpessoal permitindo o compartilhamento e cooperação. Ainda, segundo Al-Ani et al. (2011), os membros preferem buscar conhecimento daqueles em quem eles têm confiança. A aceitação do conhecimento ocorre da mesma forma;
- Aumenta a eficiência da equipe e melhora a qualidade das saídas (Khan, 2012): a confiança aumenta o esforço da equipe, diminui os custos de monitoramento, favorece colaboração, facilita o compartilhamento de informações e recursos; e, conseqüentemente melhora a qualidade da saída.
- Facilita alcançar os objetivos: a presença de confiança entre os membros da equipe ajuda a alcançar os objetivos da equipe.

2.4.3 Índícios de Confiança

Como parte deste trabalho, conduzimos um mapeamento sistemático da literatura entre setembro de 2014 e janeiro de 2015 em para responder: "Como determinar a existência de confiança entre membros de uma equipe DGS?". O mapeamento utilizou a *string* de busca apresentada no Quadro 2.1.

Quadro 2.1: *String* de busca.

```
((‘Global software development’ OR ‘Distributed software
development’ OR ‘virtual team’ OR ‘virtual teams’) AND
trust) OR ((‘Global software development’ OR
‘Distributed software development’ OR ‘virtual teams’)
AND (‘text mining’ OR ‘opinion mining’ OR ‘sentiment
analysis’))
```

Essa *string* de busca foi submetida a seis bases, das quais tínhamos acesso: ACM, *Engineering Village*, IEEE, *Science Direct*, *Scopus* e *SpringerLink*, sem nenhuma restrição de tempo, retornando um total de 996 trabalhos dos quais 444 eram duplicados. Dos 552 trabalhos únicos. Para seleção dos trabalhos foram utilizados os seguintes critérios de inclusão:

- O título e/ou resumo apresentam as palavras chaves
- Aborda confiança no DGS ou equipes virtuais
- Aborda a confiança interpessoal
- Determina a confiança entre os membros ou apresenta um comportamento/característica apresentada por essas equipes na presença de confiança

e de exclusão:

- Trabalho não é acessível
- Descrição de eventos/*call for papers*

Após a leitura do resumo, introdução e conclusão 532 foram excluídos. Após a leitura total, dos 30 trabalhos restantes 11 foram aceitos (Tabela 2.1), sendo esses estudos publicados entre 1997 e 2012. Ao término do mapeamento, não encontramos uma forma exata de determinar se existe ou não confiança entre os membros de uma equipe, mas alguns trabalhos que são descritos a seguir, apontam comportamentos e características que servem como indícios de sua existência. Chamamos esses comportamentos e características de indícios de confiança.

Jarvenpaa et al. (1998) realizaram um estudo envolvendo 385 estudantes de pós-graduação de 28 universidades com o objetivo de analisar a relação entre os antecedentes de confiança (habilidade, benevolência, integridade e propensão a confiar), exercícios de formação de equipes e a confiança em uma equipe virtual. Para a realização do estudo foram formadas equipes de até quatro alunos, sendo que os alunos tinham a opção

Tabela 2.1: Trabalhos aceitos

Título	Ano	Autor
Developing trust in virtual teams	1997	Iacono e Weisband
Is Anybody out There? Antecedents of Trust in Global Virtual Teams	1998	Jarvenpaa et al.
Linguistic Mimicry and Trust in Text-based CMC	2008	Scissors et al.
Fostering trust in virtual project teams: Towards a design framework grounded in a TrustWorthiness ANtecedents (TWAN) schema	2010	Rusman et al.
The joint relationships of communication behaviors and task interdependence on trust building and change in virtual project teams	2009	Rico et al.
An Exploratory Study of Trust Dynamics in Work-Oriented Virtual Teams	2009	Kuo e Yu
Trust in Virtual Teams: Solved or Still a Mystery?	2009	Mitchell e Zigurs
Learning About the Quality of Teamwork from Wikiteams	2010	Turek et al.
An Understanding of the Role of Trust in Knowledge Seeking and Acceptance Practices in Distributed Development Teams	2011	Al-Ani et al.
Time Pressure, Cultural Diversity, Psychological Factors, and Information Sharing in Short Duration Virtual Teams	2012	Paul e He
Role of trust and relationships in geographically distributed teams: exploratory study on development sector	2012	Khan

de participar ou não de dois exercícios de formação de equipes e, em seguida, deviam responder a um questionário para avaliar a confiança e seus antecedentes. Respondido o questionário, os alunos realizaram um projeto obrigatório que consistia em desenvolver um *website*, respondendo ao questionário novamente ao final do projeto. Os participantes podiam se comunicar somente por meios eletrônicos. Nesse trabalho observou-se que: (i) a confiabilidade é altamente correlacionada com a confiança, (ii) os exercícios de formação não influenciam diretamente a confiança, mas indiretamente por meio dos antecedentes, (iii) habilidade, integridade e propensão a confiar são bons preditores de confiança em um primeiro momento (antes do início do projeto), mas em um segundo momento (ao final do projeto) benevolência, integridade e propensão a confiar passam a ser bons preditores. Os autores também realizaram um estudo qualitativo considerando as três equipes com alto nível de confiança e três equipes com baixo nível de confiança. Nesse estudo observou-se características comuns às equipes de alta confiança que não apareciam nas de baixa confiança. As características encontradas foram: proatividade, comunicação orientada à tarefas, tom positivo nas comunicações, liderança rotativa, clareza de objetivos, divisão de papéis, gerenciamento de tempo, *feedback* e comunicação intensa.

Em outro estudo, Khan (2012) realizou entrevistas com 12 participantes de equipes globalmente distribuídas do setor de desenvolvimento para entender qual o papel dos relacionamentos e da confiança nessas equipes. O conteúdo das entrevistas foi codificado e analisado levando em consideração cinco categorias: (i) definição e relevância da confiança e relacionamentos, (ii) relacionamentos informais, (iii) papel do líder em desenvolver confiança e relacionamentos, (iv) o uso da tecnologia para construir confiança e relacionamentos e (v) desempenho da equipe e relacionamentos confiáveis. Nesse trabalho foi constatado que a qualidade das saídas, o entusiasmo, a delegação de autoridade e a quantidade de trabalho são sinais de confiança.

A relação entre a mímica de vocabulário, repetição de palavras ou frases, em conversas por mensagens instantâneas, foi investigada em Scissors et al. (2008). Nesse estudo, 52 participantes jogaram em duplas um jogo de investimentos, que necessitava de confiança entre a dupla para alcançar o sucesso. Cada participante jogou 28 *rounds* podendo conversar somente via mensagens instantâneas por 5 minutos a cada 5 *rounds*. Observou-se que as mímicas do tipo léxica, que seria a repetição de palavras e abreviações estão relacionadas com a confiança entre as pessoas, de forma que o compartilhamento do mesmo vocabulário seria um indicativo de confiança.

Al-Ani et al. (2011) investigaram o papel da confiança na busca e aceitação de conhecimento. Nesse estudo, 43 funcionários de uma grande multinacional, líder no desenvolvimento de tecnologias baseadas em computador, foram entrevistados. A entrevista consistiu de três fases: na primeira as perguntas eram relacionadas ao *background* do entrevistado; na segunda fase as questões se referiam a um projeto recente que eles haviam participado e na terceira fase as perguntas eram relacionadas a cenários nos quais eles deviam buscar por conhecimento em outras pessoas. Analisando as entrevistas percebeu-se que os entrevistados preferiam buscar conhecimento com aqueles em quem eles confiavam e tendiam a aceitar esse conhecimento se eles confiavam na fonte.

Em (Iacono e Weisband, 1997) realizou um estudo com alunos de graduação e pós-graduação de três universidades dos Estados Unidos. Nesse estudo os alunos foram divididos em equipes de três alunos, preferencialmente um de cada universidade, que deviam desenvolver um trabalho escolar. Os dados coletados durante o estudo foram resposta de um questionário de caracterização, os e-mails trocados e as notas atribuídas aos trabalhos. Os e-mails, foram codificados em iniciações e repostas, sendo a quantidade de iniciações e repostas considerada uma medida de confiança. Observou-se que equipes de alto desempenho (nota acima de 90%) possuíam grande quantidade de iniciações e repostas.

Rusman et al. (2010) realizaram uma síntese a literatura de várias disciplinas a respeito da confiança. Com as informações obtidas propuseram um modelo de formação de confiança interpessoal. Analisando esse modelo de formação percebeu-se que uma das maneiras de facilitar o desenvolvimento de confiança em equipes virtuais é facilitar a avaliação da confiabilidade de um membro, fornecendo informações sobre os outros membros. Sendo assim eles propuseram o *TrustWorthiness ANtecedentes (TWAN) schema*, que auxilia a determinar quais informações dos membros devem ser disponibilizadas para facilitar a avaliação da confiabilidade pelos outros membros, facilitando por sua vez do desenvolvimento da confiança.

Rico et al. (2009) coletaram dados de 53 equipes virtuais (187 participantes) de uma empresa de desenvolvimento. Dentre esses dados temos confiança, confiabilidade, orientação das comunicações (social ou a tarefas), entusiasmo, virtualidade. Uma das conclusões obtidas ao avaliar os dados é que a comunicação orientada a tarefas é fortemente relacionada a confiança.

Mitchell e Zigurs (2009) analisaram os trabalhos a respeito da confiança presentes na literatura a fim de analisar os desafios que já foram resolvidos e o que ainda necessita de estudo. Ao todo foram analisados 42 trabalhos, publicados de 1997 até 2007 e que envolviam diversos temas relacionados a confiança. A partir dos trabalhos analisados os autores desenvolveram um *framework* para pesquisa sobre confiança em equipes virtuais e apontaram algumas áreas que merecem estudos como: teorias para sistemas de informação, utilização de novos métodos para obtenção de dados, utilização de profissionais da área em estudos.

Turek et al. (2010) apresentaram uma abordagem para avaliar equipes de colaboradores. A abordagem consiste em gerar uma rede social com base nos históricos de edição de artigos dos colaboradores. Essa rede é composta de quatro dimensões: confiança, desconfiança, conhecimento e afinidade. A dimensão de confiança foi derivada do número de palavras escritas por um colaborador que foram copiadas/movidas/restauradas por outro colaborador.

Paul e He (2012) analisaram o impacto da pressão temporal e diversidade cultural em fatores psicológicos (confiança e motivação) de uma equipe virtual e se esses fatores psicológicos afetam o compartilhamento de conhecimento e este por sua vez a qualidade das saídas. Para isso, 24 equipes, de três membros cada, formadas por estudantes de graduação e pós-graduação em negócios, computação e engenharia foram estudadas. As variáveis motivação e confiança foram obtidas por meio de questionários, diversidade cultural com base nos dados demográficos informados no início do estudo, para avaliar o compartilhamento de conhecimento um estudante de fora do estudo contabilizou as

informações compartilhadas entre os membros e por fim, a qualidade das saídas foi obtida ao corrigir o diagrama de entidade e relacionamentos gerado pelas equipes. Analisados os dados, os autores concluíram que: pressão temporal melhora a motivação, mas não a confiança. A diversidade cultural não afeta os fatores psicológicos. Diversidade cultural e os fatores psicológicos influenciam o compartilhamento de informações. Motivação e confiança são altamente correlacionadas. Dependendo do tipo de informação compartilhada, o compartilhamento de informações melhora a qualidade das saídas.

Kuo e Yu (2009) estudaram o desenvolvimento e manutenção da confiança equipes virtuais orientadas a trabalho, e como essa confiança afeta a comunicação e a coesão. Foram acompanhadas 25 equipes virtuais (158 alunos) de um curso universitário, cada equipe possuindo de 6 a 8 membros. Dentre os resultados obtidos, os autores observaram que a confiança aumenta a comunicação e o sentimento de coesividade.

A lista abaixo sumariza os indícios encontrados que podem ser utilizados para se determinar a existência de confiança interpessoal:

- Iniciações e respostas: Segundo Iacono e Weisband (1997), iniciações são definidas como perguntas ou propostas que levem o receptor a fornecer uma resposta relevante. A quantidade de iniciações e respostas foi utilizada pelos autores para medir confiança, uma vez que equipes de alto desempenho exibiram maior confiança (quantidade de iniciações e respostas) que as de baixo desempenho.
- Motivação: Paul e He (2012) mediram confiança, motivação, compartilhamento de informações e qualidade da solução, em 24 equipes que realizaram uma tarefa de solução de problemas. Por meio do uso da regressão “*ridge*” para identificar se motivação e confiança impactam no compartilhamento de informações, constataram que motivação e confiança impactam no compartilhamento de informações e eles (motivação e confiança) são altamente correlacionados.
- Compartilhamento de conhecimento: o experimento realizado em Paul e He (2012) também mostrou que quanto maior a confiança entre os participantes, maior é o compartilhamento de informações entre eles. Mitchell e Zigurs (2009) também apontam o impacto da confiança na transferência de conhecimento.
- Aceitação de conhecimento: as pessoas tendem a aceitar conhecimento de quem elas confiam (Al-Ani et al., 2011). Em Turek et al. (2010) a movimentação do texto de um autor por outro em artigos da Wikipédia foi considerada um sinal de confiança, pois expressa a tendência de um autor em acreditar na credibilidade do outro.

- Confiabilidade: confiança e confiabilidade estão altamente correlacionadas, quando a confiança em alguém é alta a confiabilidade também é (Jarvenpaa et al., 1998).
- Proatividade: em equipes com alta confiança os membros se voluntariam para os papéis e apresentam iniciativa (Jarvenpaa et al., 1998).
- Comunicação orientada à tarefas: as conversas em equipes de alta confiança concentram-se em torno das tarefas que devem ser realizadas. Jarvenpaa et al. (1998); Rico et al. (2009) notaram em seus trabalhos que raramente ocorrem conversas relacionadas a temas sociais nessas equipes.
- Tom positivo: em equipes de alta confiança os membros mostram entusiasmo nas conversas, elogiam e encorajam os outros. Jarvenpaa et al. (1998) observaram que as discussões que ocorriam nessas equipes eram resolvidas tão gentilmente que quase não eram percebidas.
- Clareza de objetivos: equipes de alta confiança discutem mais os objetivos e, em caso de dúvida, os membros buscam esclarecer com os coordenadores em vez de criar suposições (Jarvenpaa et al., 1998).
- Liderança rotativa: vários membros apresentam traços de liderança e, conforme a necessidade, eles assumem a liderança da equipe (Jarvenpaa et al., 1998).
- Divisão de papéis: os membros assumem papéis na equipe, o que não significa que sejam totalmente independentes, pois o responsável por um papel mostra o resultado do trabalho e os outros fornecem *feedback* (Jarvenpaa et al., 1998).
- Gerenciamento do tempo: em equipes de alta confiança os membros discutem os prazos, estabelecem marcos e cuidam para cumpri-los (Jarvenpaa et al., 1998).
- *Feedback* e comunicação intensa: equipes de alta confiança exibem grande quantidade de comunicação e fornecem *feedback* sobre o trabalho dos colegas de equipe (Jarvenpaa et al., 1998; Kuo e Yu, 2009; Mitchell e Zigurs, 2009; Rusman et al., 2010).
- Alto desempenho: a confiança está relacionada positivamente com a coesão, o comprometimento, a satisfação, o desempenho e o entusiasmo (Mitchell e Zigurs, 2009; Rusman et al., 2010) (Kuo e Yu (2009) utilizaram análise de correlação sobre as variáveis confiança e coesividade).

- Qualidade das saídas: Khan (2012) realizou um estudo qualitativo para avaliar o papel dos relacionamentos e da confiança em equipes geograficamente distribuídas no setor de desenvolvimento. Constataram que a qualidade das saídas é um indicador de confiança e bom relacionamento.
- Mímica de vocabulário: quando existe confiança entre as pessoas, elas tendem a compartilhar um vocabulário comum em CMC (Scissors et al., 2008).
- Delegação: a delegação de autoridade e tarefas foram considerados indícios de confiança em Khan (2012) e Rusman et al. (2010).

2.4.4 Modelos de Confiança

Por meio de um segundo mapeamento sistemático da literatura nós buscamos responder a seguinte pergunta: “Como estimar a confiança entre membros de equipes DGS?”. Para esse mapeamento utilizamos a *string* de busca do Quadro 2.2. Essa *string* tinha por objetivo encontrar modelos e *frameworks* de confiança para equipes DGS.

Quadro 2.2: *String* de busca.

<pre>("Trust Framework" OR "Trust Model") AND ("Social Network" OR "Virtual Team")</pre>
--

Nesse mapeamento consideramos as mesmas condições do mapeamento anterior, exceto pelos critérios de inclusão:

- Título e/ou resumo apresentam as palavras chaves
- Descrevem um modelo/*framework*
- Estimam a confiança interpessoal

e de exclusão:

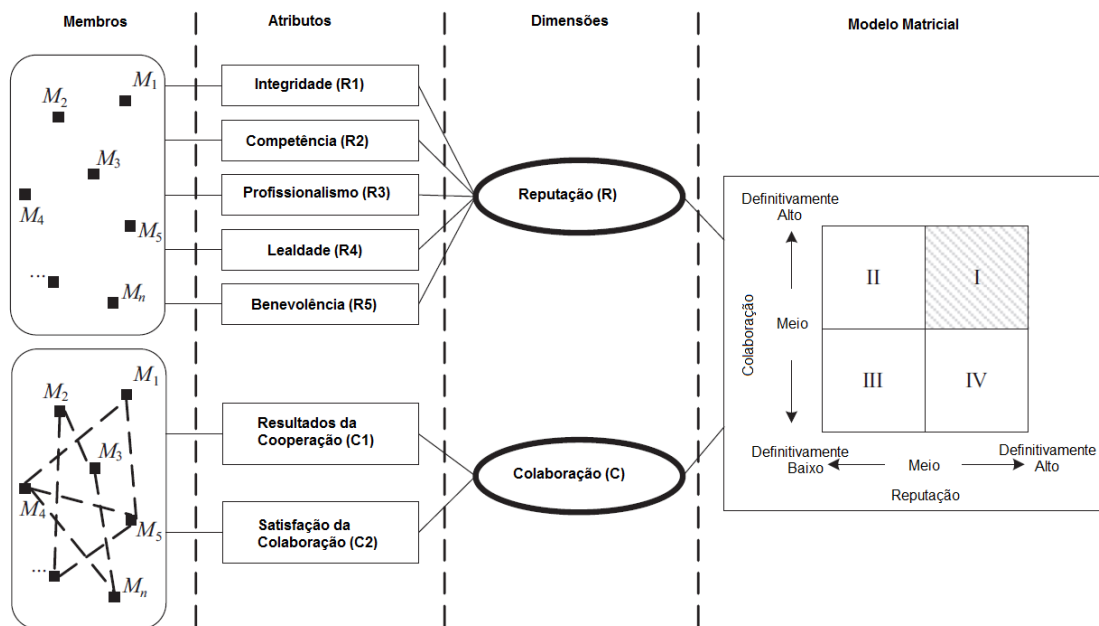
- Trabalho não é acessível
- Descrição de eventos/*call for papers*
- Modelos/*frameworks* que gerenciam a confiança

Após submeter a *string* de busca as bases, obtivemos 413 trabalhos, desses 108 eram duplicados, restando 305 trabalhos. Dos 305 realizada a leitura do resumo, 158 foram excluídos. Com a leitura da introdução e conclusão mais 97 foram excluídos, restando 50

trabalhos que foram lidos integralmente. Por fim 13 trabalhos foram aceitos, sendo esses publicados de 2006 a 2013, e são descritos a seguir.

Fan et al. (2011) desenvolveram um *framework*, que utiliza a reputação (como a confiabilidade dos membros) e a colaboração (representando a cooperação entre os membros) para estimar a confiança. A Figura 2.3 apresenta o *framework* desenvolvido com duas dimensões: reputação e colaboração. Cada uma delas é calculada utilizando uma análise de decisão multiatributo *fuzzy*, sobre os valores fornecidos pelos membros da equipe para cada dimensão na forma de termos *fuzzy*. Uma vez computado os valores, eles formam um ponto em uma matriz na qual cada quadrante representa um nível de confiança. Pontos no quadrante *I* são considerados ideais. Pontos em outros quadrantes indicam que providências devem ser tomadas para melhorar a dimensão com valor baixo. Apesar das entradas serem fornecidas por cada membro da equipe, o *framework* não calcula a confiança entre os membros, mas fornece um nível de confiança à um membro agregando as informações fornecidas por todos os membros.

Figura 2.3: *Framework* para estimar confiança.



Fonte: traduzido de Fan et al. (2011).

Li et al. (2010b) propuseram um modelo de confiança pensando em como as pessoas resolvem problemas: ao receber um problema, a pessoa tenta resolvê-lo com seu conhecimento; se ela não consegue, pede ajuda a seus conhecidos e assim sucessivamente. Esse método de resolução se assemelha a um sistema multiagente sobre o qual foi criada uma rede para cálculo de confiança. Nessa rede, a confiança é calculada com base nas

habilidades de um agente e no relacionamento entre os agentes (*Cap* e *Rel* na Equação 2.1, respectivamente) ponderados por pesos α referentes à importância de cada componente. A confiança entre o agente a e b calculada por meio da Equação 2.1 é utilizada para escolher a qual agente recorrer quando um determinado problema aparecer.

$$Trust_{ab} = \alpha_1 Cap(b) + \alpha_2 Rel(a, b) \quad (2.1)$$

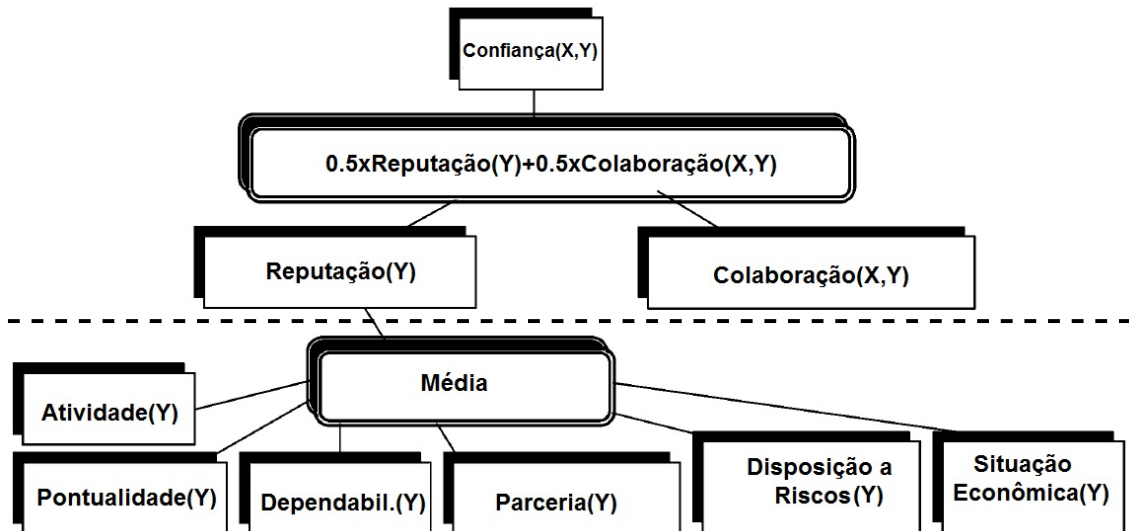
O *framework* proposto em Skopik et al. (2009) tem como objetivo determinar a confiança entre pessoas e serviços de forma automática, sem a necessidade de *feedbacks*. O *framework* gera um grafo no qual os nós representam serviços e pessoas e as arestas representam o valor da confiança entre os nós. A confiança é derivada da quantidade de interações bem-sucedidas em relação ao total de interações. O sucesso de uma interação é calculado por um conjunto de métricas, tais como a ocorrência de erros em um serviço. Uma vez que o grafo foi gerado é possível recomendar equipes, recomendar parceiros/serviços, e dar suporte ao acesso de recursos confiáveis.

Contudo, ao utilizar métricas para calcular o sucesso de uma interação, o trabalho de Skopik et al. (2009) ignora a subjetividade, uma das características da confiança, ao manter a mesma noção de sucesso para todas as pessoas. Por exemplo, para uma pessoa, se um serviço demorar até 30 segundos para enviar uma resposta, a interação é considerada um sucesso, enquanto que, para outra pessoa, se passar de 10 segundos é um fracasso. Esse tipo de métrica não consegue capturar a subjetividade inerente a expectativa de cada indivíduo.

O trabalho de Lavrac et al. (2007) apresenta uma abordagem para a modelagem de confiança em redes colaborativas de organizações e redes de pessoas utilizando suporte de decisão hierárquico multiatributo. O modelo proposto utiliza a reputação e a colaboração para determinar confiança, que por sua vez podem ser calculadas de formas diferentes, dependendo do domínio de aplicação: (i) na forma de questionários (Figura 2.4) nos quais os membros da equipe informam suas opiniões a respeito de outros membros; ou (ii) utilizando informações extraídas da web. Por exemplo, no domínio de pesquisa científica, a confiança entre autores pode ser estimada a partir de dados da *Web of Science*, do *CiteSeer* e do Google. No caso de domínios em que se opte pelo uso de questionários, a opinião é fornecida atribuindo-se valores de 1 a 6 (muito ruim a muito bom) para os atributos atividade, pontualidade, dependabilidade, parceria, disposição a riscos e situação econômica. Já para o atributo colaboração os valores variam de 0 a 3. No caso de domínios como o da pesquisa científica, os dados do *Web of Science* e *CiteSeer* fornecem valores de reputação baseados nos registros de publicações, enquanto *CiteSeer* e

Google fornecem o valores referentes à colaboração. Na Figura 2.4 a parte acima da linha pontilhada é o modelo geral, já a parte abaixo da linha pontilhada refere-se ao meio de obtenção dos dados por meio de questionários.

Figura 2.4: Modelo para estimar confiança utilizando questionários como fonte dos dados.



Fonte: adaptado de Lavrac et al. (2007)

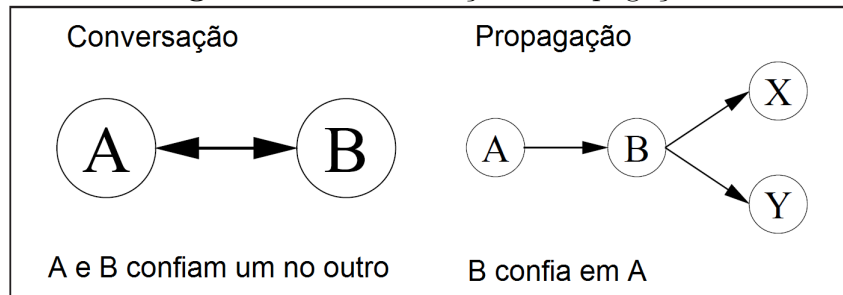
Modelos de Confiança para Redes Sociais

Uma rede social é um grafo, no qual os vértices representam pessoas ou organizações e as arestas representam o relacionamento entre elas (Turek et al., 2010). Como o GitHub, ferramenta utilizada na realização deste trabalho, pode ser vista como uma rede social e as próprias equipes virtuais podem ser vistas como redes sociais nas quais as relações são a troca de informações e trabalho (Khan, 2012), é importante analisar os modelos de confiança para redes sociais. Apesar do domínio ser diferente do deste trabalho, acredita-se que esses modelos possam ter características em comum com o domínio de equipes de DGS.

Adali et al. (2010) apresentam um modelo para o cálculo de confiança em redes sociais baseado no comportamento das pessoas. Considerou-se que quando as pessoas confiam umas nas outras elas tendem a conversar e que se B propaga o que A disse para outras pessoas, então B confia em A . A Figura 2.5 mostra as relações de propagação e conversação previstas no modelo.

O'Donovan et al. (2007) analisaram os comentários dos *feedbacks* do Ebay para medir a confiança dos vendedores, visando ajudar os compradores a escolher os melhores

Figura 2.5: Conversação e Propagação.



Fonte: adaptado de Adali et al. (2010).

vendedores, dado que uma grande quantidade de vendedores tem alta reputação (99%). Um conjunto de características, tais como custo e produto; foi utilizado na análise dos comentários. Quando uma palavra relacionada a uma das características era encontrada procurou-se por adjetivos associados. O adjetivo, por sua vez, era procurado em duas listas (uma positiva e outra negativa) com 20 adjetivos cada. Se o adjetivo fosse encontrado em uma das listas a polaridade da característica era armazenada. Caso o adjetivo não estivesse em nenhuma das listas, recuperava-se 20 sinônimos e 20 antônimos a partir de um tesouro. Os sinônimos e antônimos eram então procurados nas listas de adjetivos. A partir da polaridade registrada gerava-se a confiança granular (um valor de confiança para cada característica) e a confiança interpessoal. A confiança granular era gerada a partir da polaridade registrada em cada *feedback* de um vendedor. A confiança interpessoal era gerada somando-se os *feedbacks* de um comprador para um vendedor.

Em Huang e Arasan (2010) foi desenvolvido um modelo para se avaliar a confiança entre usuários de e-mail. Os autores consideram que quanto mais os usuários trocam e-mails, maior é a confiança entre eles. Dessa forma, a confiança é calculada com as seguintes equações:

$$\gamma_{ij} = \left(\frac{2}{\frac{L_{ij}}{L_{ji}} + \frac{L_{ji}}{L_{ij}}} \right) \frac{L_{ij}}{L_{ij} + L_{ji}} \beta \quad (2.2)$$

e

$$T_{ij} = [\log_2(N_{ij} + 2)]^{\gamma_{ij}} \quad (2.3)$$

onde L_{ij} é a quantidade de e-mails enviados de i para j , $\beta = 1/0.553$ é uma constante para escalar o valor da função γ_{ij} até 1, N_{ij} é o total de e-mails trocados entre i e j e T_{ij} é a confiança de i em j .

Previendo o surgimento de uma rede que una as redes sociais e as redes de distribuição de conteúdo, Li et al. (2010c) propuseram um modelo de confiança para aumentar a

segurança dos usuários dessas redes, de forma que bons usuários sejam recompensados e os maus sejam punidos. Nesse trabalho, a confiança é calculada pelas equações 2.4 e 2.5.

$$T = \sum Transacoes_{ab} * D \quad (2.4)$$

$$D = \frac{P_{ab} - N_{ab}}{\alpha} + \beta \quad (2.5)$$

Na Equação 2.4, T é a confiabilidade de a em b e $Transacoes_{ab}$ é o valor dado a uma transação entre a e b . Transações positivas, como compartilhar um bom arquivo, têm valor positivo, enquanto as negativas, como compartilhar um arquivo ruim, têm valor negativo. Na Equação 2.5, P_{ab} é a quantidade de *feedbacks* (avaliações sobre as interações) positivos entre a e b , N_{ab} é o número de *feedbacks* negativos, α é um limiar para determinar o número mínimo de transações necessárias para a relação ser considerada confiável e β é um fator de ajuste que indica a regressão natural da confiança, isto é, a diminuição natural da confiança com o passar do tempo.

O modelo de confiança proposto em Li et al. (2010a) tem como objetivo auxiliar os usuários de *E-commerce* a escolher os melhores vendedores. Nesse trabalho a confiança é calculada baseada nas interações entre os usuários. Na falta de interações, é utilizada a similaridade entre as avaliações fornecidas por eles, se o valor da similaridade for maior que um limiar Θ (definido como 0,3, pois diminuiu o erro durante o experimento). Caso contrário, a confiança é considerada inexistente. Esse calculo é descrito pela Equação 2.6, cujos resultados são utilizados como pesos das arestas em um grafo de usuários representando a confiança entre eles.

$$T = \begin{cases} DT_{a \rightarrow b} & \text{se existe uma aresta direta} \\ SIM(u_a, u_b) & \text{se } SIM(u_a, u_b) > \Theta \\ 0 & \text{senão} \end{cases} \quad (2.6)$$

A similaridade ($SIM(u_a, u_b)$) é calculada utilizando o coeficiente de *Spearman* e a confiança direta ($DT_{a \rightarrow b}$) é calculada pela Equação 2.7.

$$DT_{a \rightarrow b} = Avaliacao_{a \rightarrow b} \times e^{-[tempo(n) - tempo(i)]/D} \quad (2.7)$$

na qual $tempo(n)$ refere-se ao tempo atual e $tempo(i)$ ao tempo em que a avaliação foi fornecida, D é o valor de decaimento, que é aplicado com base na data da avaliação. Quanto mais recente, mais importante a avaliação.

O METrust é um modelo de confiança que utiliza avaliação mútua ao final de uma transação (ambos avaliam a transação), proposto por Cheng et al. (2012) com o objetivo de proteger usuários de redes P2P de usuários maliciosos. O cálculo da confiança é baseado na confiança direta e nas recomendações dos usuários. Para diminuir o efeito de recomendações falsas, a similaridade entre quem recomenda e quem recebe a recomendação é calculada e utilizada como peso para a recomendação. As Equações 2.8, 2.9 e 2.10 calculam a confiança direta, de recomendação e final de um usuário i em relação a um usuário j , respectivamente.

$$D_{ij} = \frac{\sum_{t_k=t_1}^{t_n} f(t_k) d_{ij}^{t_k}}{\sum_{t_k=t_1}^{t_n} f(t_k)} \quad (2.8)$$

$$R_{ij} = \sum_{k \in Q} w_k \times f_{kj} \quad (2.9)$$

$$T_{ij} = \delta D_{ij} + (1 - \delta) R_{ij} \quad (2.10)$$

Na Equação 2.8, t_n é a n -ésima transação, $f(t_k)$ é a função de decaimento em função do tempo, $d_{ij}^{t_k}$ é o valor da k -ésima avaliação de i em relação a j . Na Equação 2.9, Q é o conjunto de amigos de i , w_k é o peso da recomendação f_{kj} , calculado utilizando a similaridade das avaliações de k e i . Na Equação 2.10, δ é o peso que define a importância da confiança direta e de recomendação no valor de confiança final.

O modelo STrust de Nepal et al. (2013) foi desenvolvido para recomendação de amigos em uma rede social. O STrust utiliza confiança social derivada do comportamento dos usuários nas redes sociais. A confiança social possui dois componentes: a confiança da popularidade (*PopTrust*) e a de compromisso (*EngTrust*). A confiança final é calculada com a fórmula a seguir:

$$SocialTrust(m_i, m_j, k, d) = \alpha.PopTrust(m_i, m_j, k, d) + (1 - \alpha)EngTrust(m_i, m_j, k, d) \quad (2.11)$$

Onde m é um nó da rede, k é um contexto e d é uma atividade. O valor α define a importância dos dois componentes, que são calculados como se segue:

$$PopTrust(m_i, m_j, k, d) = \frac{|PT_{ij}^{kd+}| + 1}{|PT_{ij}^{kd+}| + |PT_{ij}^{kd-}| + 2} \quad (2.12)$$

e

$$EngTrust(m_i, m_j, k, d) = \frac{|ET_{ji}^{kd+}| + 1}{|PT_{ji}^{kd+}| + |PT_{ji}^{kd-}| + 2} \quad (2.13)$$

onde $|PT_{ji}^{kd+}|$ representa as interações positivas que m_i teve com m_j em uma atividade d no contexto k . $|ET_{ji}^{kd+}|$ por sua vez refere-se as interações positivas que m_j teve em relação a m_i .

Yin et al. (2012) desenvolveram um modelo chamado AUTrust para calcular a confiança entre usuários adjacentes em redes sociais (amigos, por exemplo). Nesse modelo, a confiança é composta de três componentes: similaridade entre usuários, familiaridade e reputação. O valor de confiança é calculado usando a Equação 2.14.

$$Tr(u \Rightarrow v) = \alpha \cdot STr(u, v) + \beta \cdot FTr(u, v) + \gamma \cdot RTr(v) \quad (2.14)$$

Nessa equação, $Tr(u \Rightarrow v)$ é a confiança de u em v e $STr(u, v)$, $FTr(u, v)$, $RTr(v)$ representam os componentes similaridade, familiaridade e reputação, respectivamente. α , β e γ são os pesos desses componentes. $STr(u, v)$ é a similaridade entre os atributos de u e v e pode ser calculada utilizando fórmulas para o cálculo de similaridade, como a similaridade de cosseno ou coeficiente de Jaccard. $FTr(u, v)$ é a familiaridade entre u e v , e é baseada no número de vizinhos em comum e na quantidade de interações com v em relação ao total de interações com os vizinhos. $RTr(v)$ é a reputação de u , que dever ser calculada de acordo com cada rede social, uma vez que o cálculo pode ser diferente para cada rede. Nesse trabalho os dados utilizados no experimento foram coletados de um microblog e, portanto, $RTr(v)$ foi calculada com base na quantidade de seguidores, na taxa seguidores/seguidos e quantidade de *tweets* comentados e *retweets*.

Em Huynh et al. (2006) foi proposto o modelo FIRE para sistemas multiagente abertos. A confiança no modelo FIRE possui quatro componentes: confiança baseada nas interações, confiança baseada em papéis, reputação das testemunhas e reputação certificada. O valor final da confiança é dado pela Equação 2.15, na qual K refere-se aos componentes, a e b são agentes, c é um termo (qualidade, honestidade, ...) e w_K é o peso da componente.

A confiança baseada em interações e a reputação das testemunhas são como descritas nos modelos apresentados anteriormente. Já a confiança baseada em papéis é constituída de um conjunto de regras que cada agente possui. As regras têm o formato $(papel_a, papel_b, c, e, v)$, na qual e é o nível de influência da regra e v é o desempenho esperado em relação ao termo c . Essas regras são dependentes do domínio e, geralmente, são especificadas pelo projetista dos agentes.

A reputação certificada é fornecida pelo próprio agente e constitui-se de um conjunto de avaliações fornecidas pelos agentes com quem o agente em questão interagiu. Vale notar que essas avaliações não são necessariamente as avaliações mais recentes.

$$\tau(a, b, c) = \frac{\sum_{K \in I, R, W, C} w_K \cdot \tau_K(a, b, c)}{\sum_{K \in I, R, W, C} w_K} \quad (2.15)$$

Cada componente é calculada utilizando a Equação 2.16:

$$\tau_K(a, b, c) = \frac{\sum_{r_i \in R_K(a, b, c)} \omega_K(r_i) \cdot v_i}{\sum_{r_i \in R_K(a, b, c)} \omega_K(r_i)} \quad (2.16)$$

na qual, r_i é uma avaliação no formato $r = (a, b, c, i, v)$, i é a interação e v é o valor da avaliação ($[-1,1]$). R_K é o conjunto de avaliações de uma componente. $\omega_K(r_i)$ é a função peso. Para as componentes baseada em interação, reputação das testemunhas e reputação certificada é uma função de decaimento do tempo, que retorna um valor entre 0 e 1 dependendo de quão recente é o valor da componente. Para a componente baseada em papéis, a função peso é o nível de influência da regra.

Síntese dos Modelos de Confiança

A 2.2 apresenta uma síntese dos modelos de confiança descritos nessa seção. A coluna **Automatizado** indica se é necessário que o usuário forneça os dados manualmente ou não. A coluna **Informações Consideradas** aponta quais são as informações consideradas para calcular a confiança. A coluna **Domínio** refere-se ao domínio para o qual o modelo foi criado. A coluna **Subjetivo** tem o valor “Sim” para modelos que consideram a subjetividade da confiança. Por fim, a coluna **Local** indica quais são os modelos que calculam a confiança entre duas pessoas.

Tabela 2.2: Síntese dos Modelos de Confiança

Trabalho	Auto- mati- zado	Informações Considera- das	Domínio	Subjeti- vo	Local
Fan et al. (2011)	Não	Reputação e Cola- boração	Equipes Virtuais	Sim	Não
Li et al. (2010b)	Não	Habilidades e Relacio- namentos	Resolução de Problemas	Sim	Sim
Skopik et al. (2009)	Sim	Sucesso das Interações	Equipes Virtuais	Não	Sim
Lavrac et al. (2007)	Ambos	Reputação e Cola- boração	Redes Colabora- tivas e de Pessoas	Sim (quando manual)	Sim
Adali et al. (2010)	Sim	Conversação e Propagação	Redes So- ciais	Não	Sim
O'Donovan et al. (2007)	Sim	Análise de Adjetivos	E-bay	Sim	Sim
Huang e Arasan (2010)	Sim	Troca de E-mails	E-mails	Não	Sim
Li et al. (2010c)	Não	Sucesso das Interações	Redes So- ciais	Sim	Sim
Li et al. (2010a)	Não	Sucesso das Interações/ Similari- dade	<i>E-com- merce</i>	Sim	Sim
Cheng et al. (2012)	Não	Interações	Redes P2P	Sim	Sim
Nepal et al. (2013)	Não	Comporta- mento	Recomen- dação de Amigos	Sim	Sim
Yin et al. (2012)	Sim	Similaridade, Familiari- dade e Reputação	Redes So- ciais	Não	Sim
Huynh et al. (2006)	Não	Interações, papéis, reputação das teste- munhas e certificada	Sistemas Multi- agente Abertos	Sim	Sim

2.5 Análise de Sentimentos

A análise de sentimentos ou mineração de opiniões é uma área de pesquisa ampla e interdisciplinar que envolve o estudo das opiniões, atitudes e emoções. A área tem ganhado muita atenção da comunidade acadêmica na última década e encontrando aplicações em praticamente todos os domínios de negócio e social (Liu, 2012).

A análise de sentimento também tem sido aplicada no contexto de pesquisas em desenvolvimento de software. Guzman (2013) utilizou análise de sentimentos para capturar emoções durante as diversas fases do desenvolvimento de software e fornecer conhecimento do clima emocional dos membros da equipe.

Borbora et al. (2013) analisaram características que expressassem confiança em redes sociais de diferentes domínios. Um dos domínios analisados foi uma rede formada por e-mails de uma empresa de tecnologia da informação (TI). Um grafo foi gerado para representar a rede na qual os nós eram os usuários e as arestas representavam a comunicação entre eles. A análise de sentimentos foi utilizada para determinar a polaridade das mensagens, que por sua vez foi utilizada para rotular as arestas, uma vez que o sentimento expresso na comunicação é um indicador da presença/ausência de confiança entre os envolvidos (Pang e Lee, 2008 apud Borbora et al., 2013).

Além disso, como foi possível observar na Seção 2.4.3, um dos indícios de confiança é o tom positivo da comunicação, que pode ser capturada por meio da análise de sentimentos. Em Zhang et al. (2009) foi proposto um modelo de confiança para adicionar informações de contexto à *Web of Trust* (WOT), redes em que os *links* representam a existência de confiança, no caso Epinions¹. Como direções para trabalhos futuros, Zhang et al. (2009) sugeriram a utilização de análise de sentimentos no conteúdo gerado pelos usuários para se obter uma compreensão mais profunda da confiança entre os usuários. Sendo assim, a análise de sentimentos se apresenta como uma técnica que pode ajudar a estimar a existência a confiança entre membros de uma equipe DGS.

Uma das tarefas realizadas por sistemas de análise de sentimentos é determinar para um dado texto a polaridade do sentimento expresso: se positivo, neutro ou negativo. A unidade utilizada na análise determina o seu nível e geralmente se enquadra em uma das três (Liu, 2012):

- **Documento:** Nesse nível de análise o texto é visto como um todo, e se considera que o mesmo expressa um único sentimento, mesmo que várias entidades e/ou aspectos sejam citados no texto.

¹www.epinions.com

- **Sentença:** É um nível de análise mais refinado que o nível de documento. Nesse nível, o texto é separado em sentenças e considera-se que cada sentença expressa um sentimento. De modo similar ao nível de documento, desconsidera-se a possibilidade de que existam varias entidades e/ou aspectos com sentimentos diferentes em uma mesma sentença.
- **Entidade e Aspecto:** Nesse nível busca-se identificar quais são as entidades mencionadas no texto e quais são os aspectos dessas entidades que são alvos de uma opinião. Dessa forma, uma entidade pode possuir opiniões com um sentimento positivo em um aspecto e negativo em outro. É um nível de análise mais refinado, que os anteriores, que busca identificar, dado um documento d , as quintuplas de opinião $(e_i, a_{ij}, s_{ijkb}, h_k, t_l)$ em d , nas quais e_i é uma entidade, a_{ij} um aspecto da entidade, s_{ijkb} o sentimento sobre o aspecto, h_k a pessoa que forneceu a opinião e t_l quando a opinião foi expressa.

Neste trabalho utilizamos a análise de sentimentos em nível de documento sobre os comentários do sistema de versionamento, uma vez que consideramos que a polaridade encontrada em um comentário é dirigida ao criador de uma *pull request* e aos usuários mencionados no comentário.

2.5.1 SentiStrength

Jongeling et al. (2015) avaliaram o desempenho de ferramentas de análise de sentimentos no domínio das pesquisas de engenharia de software. As ferramentas SentiStrength (Thelwall et al., 2010b), NLTK (Bird et al., 2009), Alchemy API ² e Stanford (Socher et al., 2013) foram avaliadas usando um conjunto de 392 comentários anotados. As ferramentas avaliadas não obtiveram um bom desempenho, mas dentre elas as ferramentas NLTK e SentiStrength obtiveram os melhores resultados. Os autores também observaram que a escolha de uma ferramenta de análise de sentimentos pode levar a resultados contraditórios, uma vez que essas ferramentas além de um baixo desempenho no domínio de engenharia de software, elas não concordam entre si.

Antes de escolher uma ferramenta de análise de sentimentos para utilização em neste trabalho, classificamos manualmente um conjunto de 228 comentários de *pull requests* do GitHub de acordo com sua polaridade: se positiva, negativa ou neutra. Analisamos a taxa de acerto das ferramentas SentiStrength, StanfordNLP e Alchemy API.

²<http://www.alchemyapi.com/products/alchemylanguage>

Tabela 2.3: Taxa de acerto das ferramentas de análise de sentimentos sobre os comentários anotados.

Comentários/Taxa de Acerto	SentiStrength	Alchemy	StandfordNLP
Positivo	0,573	0,640	0,333
Neutro	0,465	0,209	0,232
Negativo	0,626	0,805	0,567
Geral	0,548	0,526	0,364

A 2.3 apresenta a taxa de acerto das ferramentas para o conjunto de comentários. As linhas Positivo, Neutro e Negativo apresentam as taxas de acerto para essas categorias. A linha geral apresenta o resultado considerando todas as categorias, isto é, a taxa de acerto da ferramenta como um todo. Como podemos observar na 2.3, a ferramenta com melhores resultados, foi a SentiStrength, com aproximadamente 54,8% de acertos. Sendo assim, nosso teste corrobora com o trabalho de Jongeling et al. (2015).

A ferramenta SentiStrength foi criada a partir de um conjunto inicial de 2600 comentários extraídos da rede social MySpace e classificados manualmente. A ferramenta realiza a análise de polaridade em nível de sentença e informa um valor de polaridade usando uma de 5 pontos negativos e 5 positivos. Entre os principais aspectos referentes a implementação da ferramenta, estão:

- Uma lista de 298 palavras anotadas positivamente e 465 negativamente. A força do sentimento de cada palavra é dada em uma escala de 2 a 5, inicialmente atribuída por meio do julgamento humano e refinada automaticamente com um algoritmo de treinamento. Além das palavras anotadas, a lista conta com *wild cards* e palavras específicas ao contexto do MySpace, como, por exemplo, *lol*.
- Um algoritmo de treinamento para otimizar a força do sentimento de cada palavra.
- A palavra *miss* tem sentimento positivo e negativo com força 2, pois a frase “*I miss you*” sugere tristeza e amor.
- Um algoritmo de correção ortográfica usado devido a inclusão de letras repetidas como uma forma de expressão.
- Uma lista de palavras que amplificam ou diminuem o sentimento expresso pela palavra seguinte, por exemplo, muito ou algum.
- Uma lista de palavras de negação. As palavras de negação são palavras que invertem o sentimento expresso pelas palavras que seguem.

- Letras repetidas aumentam a força do sentimento expresso, desde que se repitam mais de 2 vezes.
- Uma lista de *emoticons* anotada com sentimentos.
- Sentenças terminadas com pontos de exclamação são consideradas como tendo pelo menos um sentimento de valor 2 positivo.
- Repetição da pontuação incluindo pelo menos um ponto de exclamação aumenta a força do sentimento da palavra/sentença anterior em 1.
- Sentimento negativo em perguntas é ignorado, para que perguntas como “*Are you angry?*” sejam classificadas como sem sentimento.

A ferramenta SentiStrength está disponível em duas versões: uma versão gratuita para Windows e uma versão comercial em Java, que pode ser obtida gratuitamente para fins de pesquisa. A versão comercial tem todas as funcionalidades da versão gratuita e permite fazer classificação binária (+/-), ternária (+/0/-) e de escala única (-4 a 4).

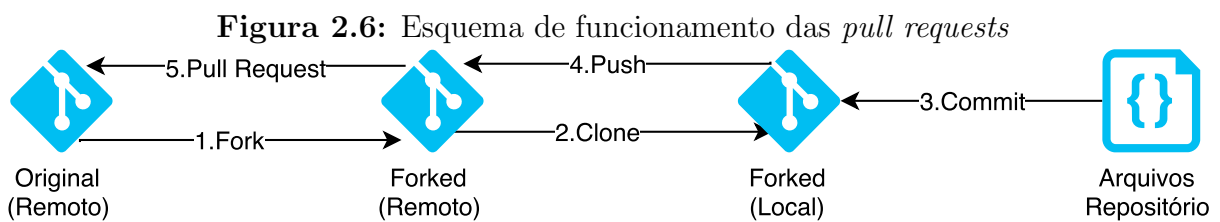
2.6 GitHub

O GitHub é uma ferramenta colaborativa construída sobre o sistema de versionamento Git³. O GitHub fornece aos seus usuários repositórios remotos de dois tipos, públicos e privados, sendo que as contas gratuitas só podem fazer uso de repositórios públicos. Cada repositório pode ter vários colaboradores, esses podem colaborar com o projeto por meio de *commits* ou *pull requests*. Para gerenciar problemas, tarefas e aprimoramentos o GitHub permite que sejam criadas *issues*, onde a atividade em questão é descrita e os interessados podem colaborar para realizá-las. Além dessas funções o GitHub provê: acompanhamento integrado de problemas, revisão de código colaborativa, gerenciamento de equipes dentro de uma organização e *syntax highlight*. Essas funções permitem a colaboração no desenvolvimento de um projeto.

Neste trabalho, a função que nos interessa é a *pull request*, descrita na Figura 2.6. Os arquivos de um projeto estão armazenados em repositórios, somente pessoas com privilegio de escrita (fornecido pelo proprietário) podem alterar o projeto. Contudo, se algum usuário do GitHub tiver interesse em contribuir para aquele projeto, ele pode fazer um *fork* do projeto (1). Ao fazer um *fork* o usuário passa a ter um repositório idêntico ao original com permissão de escrita. Em seguida o usuário clona o projeto para

³<http://git-scm.com/>

a sua máquina (2), criando uma cópia dos arquivos do repositório remoto no repositório local. Após modificar os arquivos necessários o usuário realiza o *commit* (3), efetuando as alterações no repositório local. Um *push* (4) deve ser efetuado para transferir as modificações do repositório local para o remoto. Como o usuário não tem permissões de escrita para enviar suas modificações para o repositório original, ele faz uma *pull request*, solicitando ao proprietário do projeto original que adicione suas modificações ao repositório original. Uma vez criada a *pull request* outros usuários podem comentar as modificações, até que se decida por aceitar ou não a *pull request*.



2.7 Considerações finais

Esta seção apresentou os desafios enfrentados pelas equipes de DGS e porque a confiança é importante para essas equipes. Também foram apresentados conceitos relacionados à confiança, modelos de confiança existentes, um levantamento dos indícios da confiança e a técnica de análise de sentimentos. Observando os modelos de confiança apresentados, percebemos que a maioria não é automática. Os modelos automáticos que foram apresentados ignoram a subjetividade, que é inerente a confiança, exceto no caso de O'Donovan et al. (2007), mas, assim como a maioria dos modelos, os autores não consideram o domínio de equipes virtuais e nem DGS. Dadas as características da análise de sentimentos, a técnica se apresenta como uma possível solução para incorporar a subjetividade aos modelos automáticos de confiança. A próxima seção apresenta o *framework* ARSENAL-GSD, que utiliza os indícios de confiança e análise de sentimentos para estimar a confiança entre membros de equipes de DGS.

ARSENAL-GSD

3.1 Considerações Iniciais

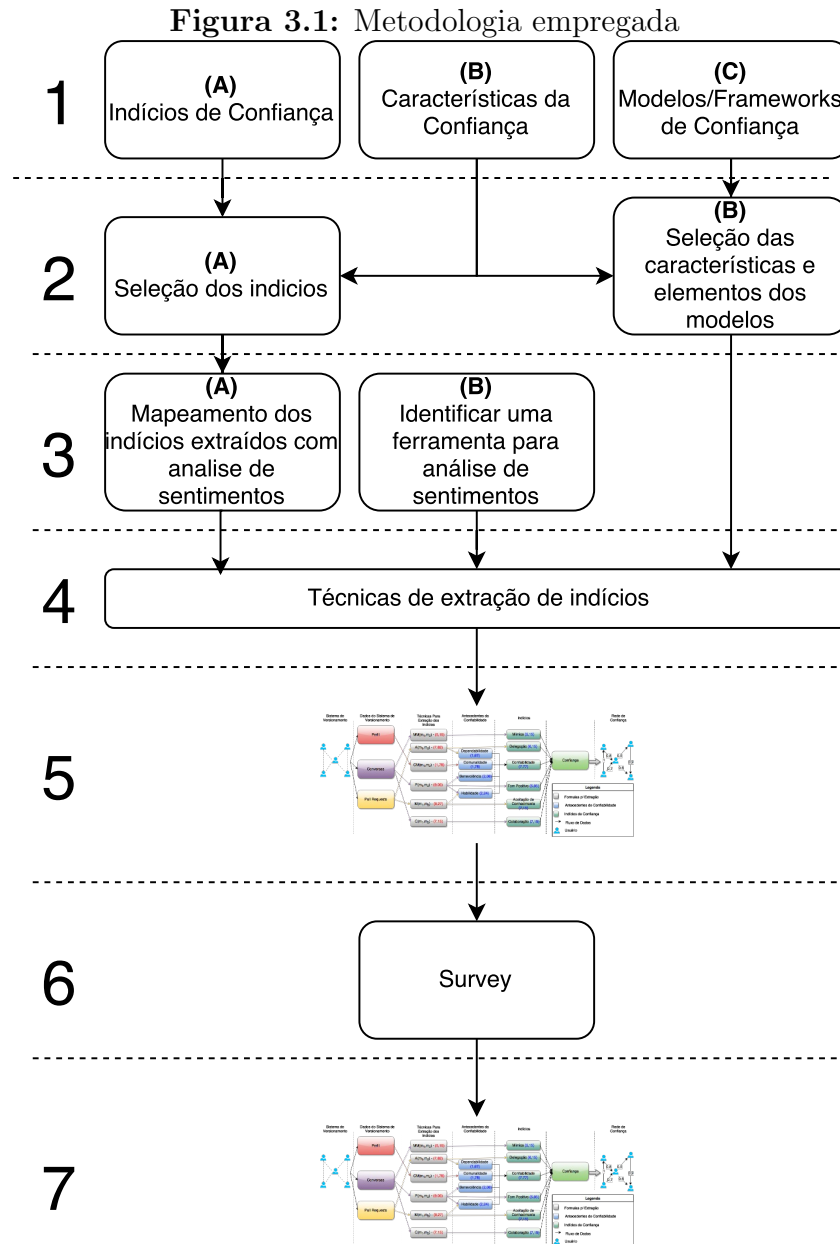
Equipes virtuais necessitam de confiança entre os membros a fim de alcançar seus objetivos, uma vez que a confiança é relacionada à eficiência de uma equipe (Pangil e Chan, 2014).

Modelos de confiança podem ser utilizados para avaliar a confiança entre os membros. Contudo, alguns modelos requerem que os usuários forneçam avaliações dos outros, ou que exista um meio de informar se uma interação foi positiva ou não. O problema dessa abordagem é que, no contexto de DGS, os membros da equipe podem não se sentir à vontade para avaliar um colega de trabalho. Mesmo que isso não seja visto como um problema, existiriam muitas interações e os membros poderiam se cansar de avaliar cada uma delas, fornecendo avaliações sem sentido que comprometem a saída dos modelos de confiança (Li et al., 2010a).

Nesse contexto, nós propomos o *framework* ARSENAL-GSD (*Automatic tRust eStimator based on sENTiment anALysis for Global Software Development*), para estimar a confiança entre membros de equipes DGS de forma automática. Esta seção apresenta a metodologia utilizada para desenvolver o ARSENAL-GSD, sua descrição em termos de características, projeto e implementação, e seu funcionamento.

3.2 Metodologia

A metodologia utilizada para desenvolver o ARSENAL-GSD é apresentada na Figura 3.1 e consiste de 7 etapas:



1^a etapa: Nessa etapa realizamos uma revisão da literatura para entender como a confiança se manifesta nessas equipes e como podemos estimá-la.

A Realizamos um levantamento dos indícios de confiança apresentados por equipes virtuais/DGS por meio do mapeamento sistemático da literatura descrito na Seção 2.4.3.

B Dentre as características de confiança apresentadas na Seção 2.4 optamos por considerar dinamicidade, por meio da janela temporal, subjetividade, com o

uso da análise de sentimentos e mímica de vocabulário, assimetria, o valor de confiança A para B é calculado de forma independente do valor de B para A. Não-transitiva, propagativa e composta não foram consideradas uma vez que nosso interesse se encontra somente na confiança direta (quando existe interação entre as pessoas). Auto-reforço e sensibilidade a eventos então embutidas nas próprias interações. Quanto a especificidade ao contexto, consideramos o contexto de DGS, não consideramos contextos mais específicos.

C Realizamos um levantamento dos modelos de confiança/*frameworks* utilizados para estimar a confiança em equipes virtuais/DGS por meio do mapeamento sistemático da literatura descrito na Seção 2.4.4.

2² etapa: Nessa etapa identificamos quais dos elementos serão utilizados no ARSENAL-GSD levando em conta as características da confiança.

A Dentre os indícios apresentados na Seção 2.4.3, escolhemos aqueles que julgamos possíveis de serem extraídos automaticamente a partir do GitHub, o sistema de versionamento que utilizamos. Outros sistemas de versionamento que podem conter dados diferentes, o que pode possibilitar a extração de um número maior ou menor de indícios. Iniciações e respostas, *feedback* e comunicação, intensa foram agrupados como colaboração no ARSENAL-GSD, uma vez que a colaboração é um aspecto presente também em alguns modelos de confiança. Embora, motivação, compartilhamento de conhecimento, proatividade, comunicação orientada a tarefas, clareza de objetivos, liderança rotativa, divisão de papéis, gerencia de tempo, alto desempenho, qualidade das saídas, grande quantidade de trabalho, delegação de autoridade, entusiasmo, compartilhamento de recursos e divisão de tarefas sejam indícios válidos, não podem ser extraídos automaticamente ou necessitam de mais dados do que os fornecidos pelos sistemas de versionamento para suas estimativas.

B A habilidade dos membros e colaboração entre eles foi utilizada em vários dos modelos apresentados (Fan et al., 2011; Lavrac et al., 2007; Li et al., 2010b), assim como a utilização do tempo (Huang e Arasan, 2010; Li et al., 2010a) para em sua estrutura uma vez que a confiança se altera com o tempo (dinâmica). Nós consideramos esses elementos no ARSENAL-GSD e utilizamos uma estrutura similar aos trabalhos de Fan et al. (2011), Li et al. (2010b) e Lavrac et al. (2007).

3^a etapa: Nessa etapa identificamos quais indícios podem ser extraídos com análise de sentimentos e escolhemos qual ferramenta de análise de sentimentos utilizar.

A Dos indícios e elementos selecionados no passo anterior identificamos quais destes podem ser extraídos com a utilização da análise de sentimentos. Os indícios identificados foram o tom positivo, a benevolência (nível de cortesia e positividade) e nesse caso específico a habilidade uma vez que o *commit* é resultado das habilidades de um usuário.

B Para extrair a polaridade dos comentários do GitHub utilizamos uma ferramenta de análise de sentimentos. Para escolher qual ferramenta utilizar comparamos três ferramentas: SentStrength, AlchemyAPI e StanfordNLP (Seção 2.5.1). A escolhida foi a SentiStrength.

4^a etapa: Uma vez identificados os dados que devem ser extraídos, desenvolvemos técnicas de extração de indícios, que são formulas utilizadas para estimar os indícios, e uma formula para estimar a confiança. Ao todo foram desenvolvidas 6 técnicas de extração de indícios que serão descritas na Seção 3.4.1.

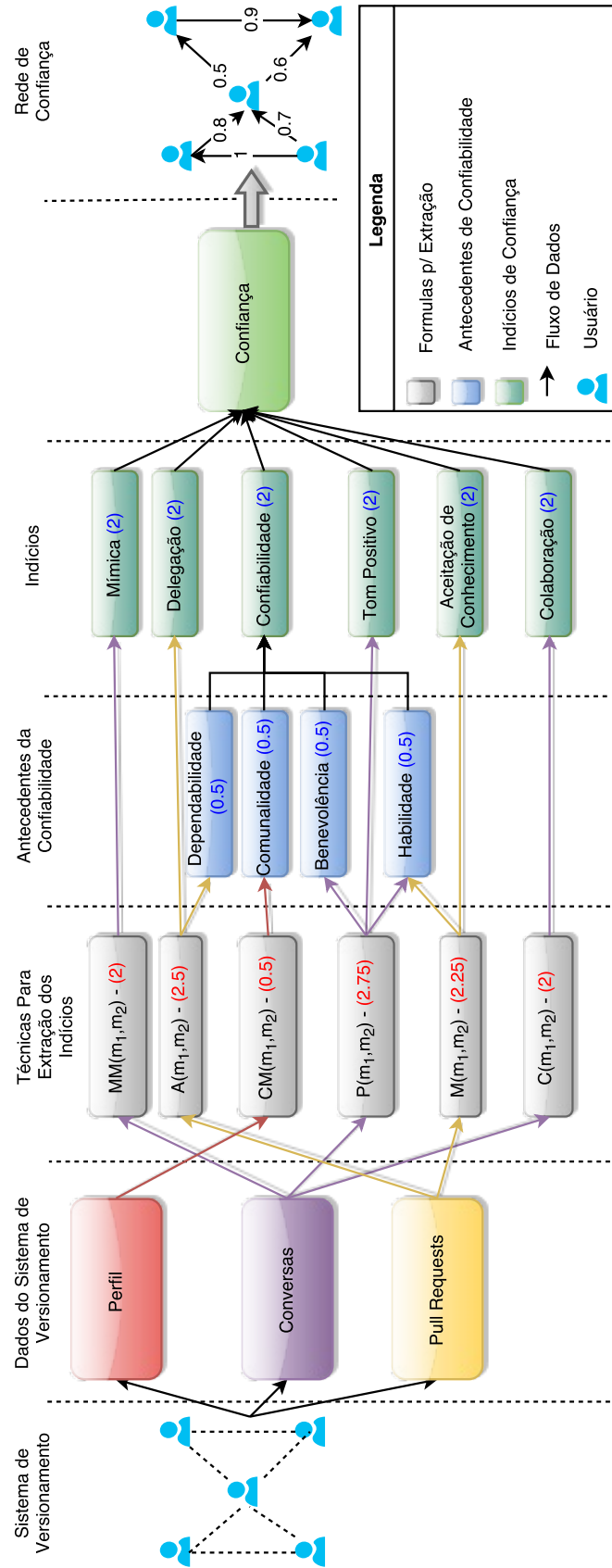
5^a etapa: Com os resultados obtidos nas etapas anteriores chegamos a uma representação do *framework* ARSENAL-GSD. A Figura 3.2 apresenta uma representação do ARSENAL-GSD e seu modelo conceitual proposto em termos de entradas, técnicas utilizadas, indícios de confiança considerados e saída. Uma instancia do ARSENAL-GSD foi implementada em Java para o GitHub.

6^a etapa: Devido a falta de uma base de dados anotada para avaliar o *framework* realizamos uma *survey* para avaliar a importância de cada indicio para o calculo da confiança e a validade das técnicas de extração de indícios.

7^a etapa: Com a importância dos indícios encontrada por meio da *survey* recomendamos os pesos para cada uma das técnicas de extração de indícios no calculo da confiança chegando a ultima versão do ARSENAL-GSD

As seções que seguem descrevem detalhadamente o ARSENAL-GSD.

Figura 3.2: Representação do *framework* ARSENAL-GSD e seu modelo conceitual



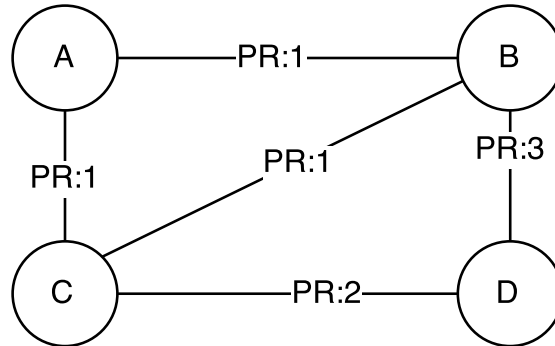
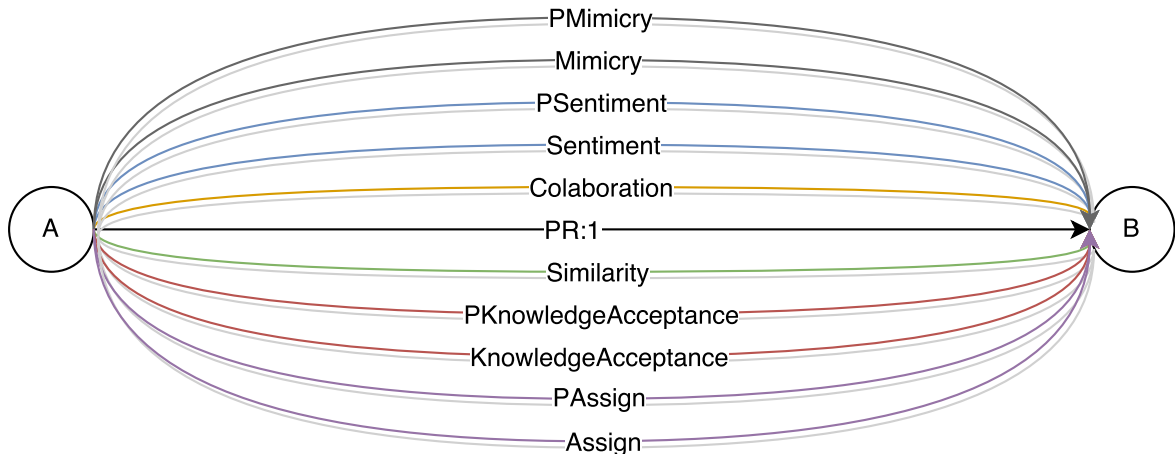
3.3 Características

As principais características do *framework* ARSENAL-GSD são as seguintes:

- a) Utiliza sistemas de versionamento como fonte de dados. Sistemas de versionamento são ferramentas amplamente utilizadas em desenvolvimento de software e em DGS (Robbes e Lanza, 2005) e tem servido de motivação para diversos trabalhos. Yu et al. (2014) desenvolveram um recomendador de revisores que sugere revisores para uma determinada *pull request*. Izquierdo et al. (2015) desenvolveram uma ferramenta para gerar visualizações baseadas em rótulos atribuídos as *issues*. As visualizações geradas podem ajudar a identificar os membros mais ativos e a evolução da *issue*. Particularmente, os sistemas de versionamento que nos interessam são aqueles que permitem aos usuários realizar *commits* e comentar os *commits* dos outros.
- b) Utiliza indícios de confiança extraídos dos dados do sistema de versionamento (comentários, estado do *commit* e perfil do usuário), para estimar a confiança entre membros de um projeto. Os indícios utilizados no *framework* foram mímica de vocabulário, tom positivo, confiabilidade, aceitação de conhecimento, delegação e colaboração.
- c) É automático. A configuração do *framework*, requer que seja informado o projeto alvo, as técnicas de extração de indícios que serão utilizadas, o mecanismo de extração de dados, o tamanho da janela temporal e a frequência de atualização. Feito isso, o *framework* recupera os dados sem intervenção humana, estima a existência de confiança de acordo com a janela temporal e atualiza os valores estimados de acordo com a frequência de atualização. Isso garante a característica de atualização automática das estimativas de confiança, sem a necessidade de intervenção humana para fornecer os dados.
- d) Preserva a subjetividade inerente à confiança utilizando análise de sentimentos e considerando a mímica de vocabulário como um indício de confiança. Os valores provenientes da análise de sentimentos e mímica são extraídos de como os membros escrevem seus comentários, e a expressão escrita é pessoal para cada membro. Assim, o *framework* leva em consideração a subjetividade, uma vez que ele infere indícios de confiança a partir de dados pessoais. Com a análise de sentimentos nós podemos inferir tom positivo da comunicação diretamente. A benevolência pode ser inferida a partir da análise de sentimentos, uma vez que foi definida em Rusman et al. (2010) como a atitude positiva e cortesia apresentada pelo confiado. A habilidade também

pode ser inferida a partir da análise de sentimentos, uma vez que a polaridade de cada comentário pode ser vista como uma avaliação a respeito do *commit* e o *commit*, por sua vez, é o resultado das habilidades de um membro para resolver um problema. Assim, nos utilizamos a polaridade dos comentários como a avaliação das habilidades de um membro.

- e) Atualiza os valores de indícios e de confiança com o tempo. O *framework* considera uma janela de tempo para realizar a extração dos dados e, de tempos em tempos, ele move essa janela, descartando dados antigos, capturando dados novos e atualizando os valores dos indícios de confiança de acordo com os dados da janela de tempo atual.
- f) Gera um grafo de relacionamentos inicial, como exemplificado na Figura 3.3. Esse grafo informa em qual *pull request* os membros interagiram. O grafo inicial tem uma aresta **PR:*** entre os pares de membros que interagem em uma mesma *pull request*. Durante a execução são adicionadas arestas parciais e finais ao grafo de relacionamentos inicial a partir da análise dos dados provenientes do sistema de versionamento. Arestas parciais mantêm valores parciais/intermediários utilizados para calcular as arestas finais. Existe uma aresta final para cada técnica de extração de indícios. Por exemplo, na Figura 3.4 a aresta **PSentiment** mantêm o valor de polaridade de um comentário e a aresta **Sentiment** mantêm a taxa de comentários positivos. A partir do momento em que são adicionadas arestas parciais/finais ao grafo de relacionamentos inicial, ele passa a ser chamado simplesmente de grafo de relacionamentos.
- g) Gera um grafo de confiança com uma estimativa da existência de confiança entre cada par de membros que interagiram em pelo menos uma *pull request*. Nesse grafo, os nós representam membros de um projeto e as arestas representam a existência de confiança entre eles. O peso da aresta varia entre 0 e 1 e representa a probabilidade de existir confiança entre dois membros. Quanto mais próximo o valor das aresta é de 1, maior é a chance de existir confiança entre dois membros.
- h) Foco na confiança interpessoal, mais especificamente, entre duas pessoas. O *framework* estima a confiança entre cada par de pessoas que interagiram no projeto baseado nos indícios encontrados entre elas.
- i) Estima somente a confiança direta entre duas pessoas. O *framework* não utiliza a recomendação dos outros membros para estimar a confiança quando não existem interações (características propagativa e composta).

Figura 3.3: Grafo de relações inicial.**Figura 3.4:** Arestas entre nós em um grafo de relacionamentos.

3.3.1 Trabalhos relacionados

A principal diferença entre o ARSENAL-GSD e os trabalhos apresentados na Subseção 2.4.4 é que este é um *framework* automático, mas que ainda assim considera a subjetividade. A estrutura multi-atributos utilizada é similar a dos trabalhos de Fan et al. (2011) e Lavrac et al. (2007). Não consideramos reputação no ARSENAL-GSD, uma vez que nos interessa somente a confiança direta entre duas pessoas, mas algoritmos de inferência de confiança podem ser utilizados sobre a saída do ARSENAL-GSD a fim de estimar a confiança entre membros que não interagiram. O ARSENAL-GSD utiliza dados provenientes das interações de forma similar a vários dos trabalhos analisados (Cheng et al., 2012; Huynh et al., 2006; Li et al., 2010a,c; Skopik et al., 2009), mas removemos a necessidade de avaliação explícita empregando a análise de sentimentos. Essa abordagem

é similar ao que foi feito em O'Donovan et al. (2007), o que preservou a subjetividade juntamente com o emprego do índice mímica de vocabulário. Assim como os trabalhos de Huang e Arasan (2010), Li et al. (2010a), Cheng et al. (2012) e Huynh et al. (2006) mantivemos o fator de tempo utilizando uma janela temporal para a extração dos dados e a atualização dos valores utilizados nas estimativas, garantindo uma característica de atualização automática dos valores de confiança.

A Tabela 3.1 apresenta ainda as características consideradas na Tabela 2.2, para o ARSENAL-GSD a fim de permitir a comparação com os outros trabalhos.

Tabela 3.1: Características do ARSENAL-GSD

Automatizado	Informações Consideradas	Domínio	Subjetivo	Local
Sim	Índices de Confiança	GSD/Equip. Virtuais	Sim	Sim

3.4 Projeto e Implementação

Devido à facilidade para a obtenção de dados de projetos por meio da própria *Application Programming Interface* (API) do GitHub¹ e de projetos, como o GHTorrent², que fornecem uma base de dados com os dados dos projetos, nós projetamos e implementamos uma instância *framework* ARSENAL-GSD para trabalhar com o sistema de versionamento GitHub.

A Figura 3.5 apresenta o diagrama de casos de uso para o *framework* ARSENAL-GSD contendo 12 casos de uso, que são descritos a seguir:

Atualizar Grafo de Confiança: Este caso de uso é responsável por atualizar os valores do grafo de confiança.

Pré-condições: o grafo de relacionamentos do projeto alvo deve existir.

Pós-condições: atualiza-se o grafo de relacionamentos e gera-se um novo grafo de confiança.

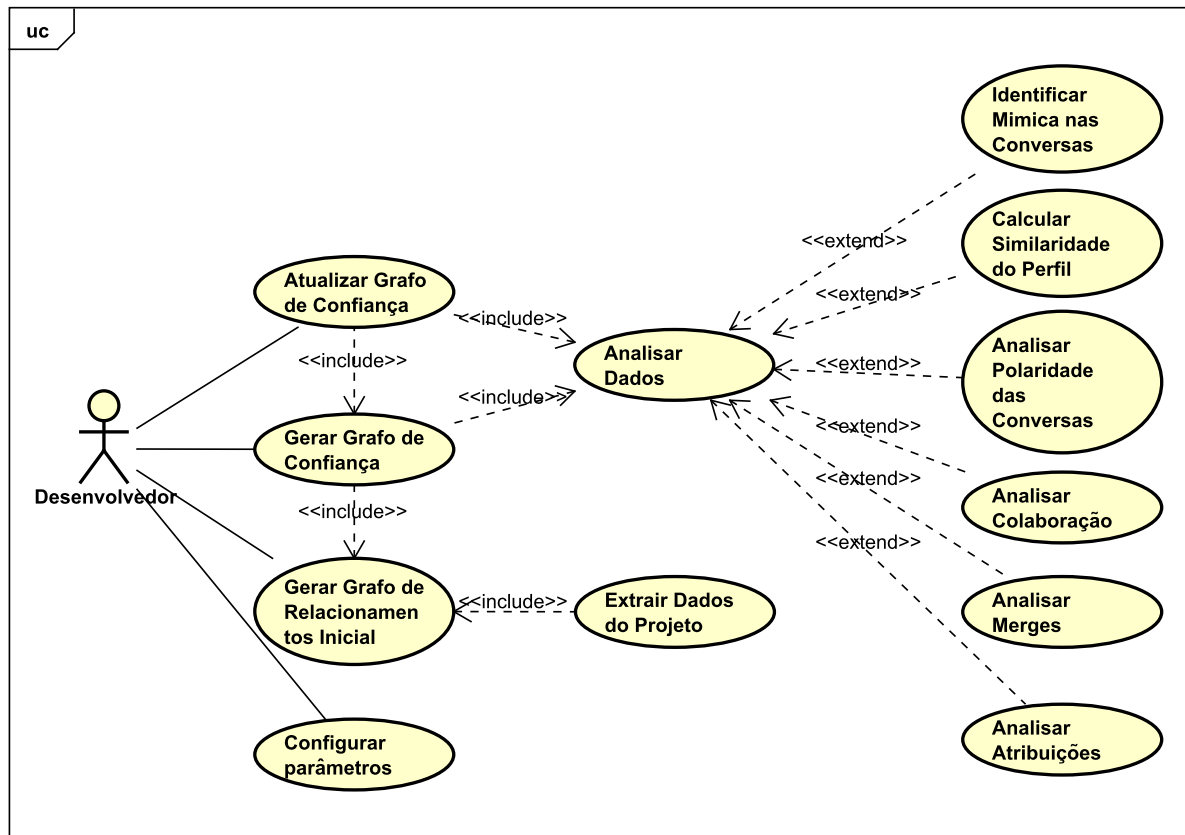
1. Remover as arestas antigas.
2. <<include>> Gerar Grafo de Confiança

Gerar Grafo de Confiança: Este caso de uso é responsável pela análise dos índices de confiança a fim de gerar, a partir dos dados de projeto de um sistema de versionamento, o grafo de confiança.

¹<https://developer.github.com/v3/>

²<http://gtorrent.org/>

Figura 3.5: Diagrama de casos de uso do *framework* ARSENAL-GSD.



Pré-condições: *framework* configurado e dados do projeto.

Pós-condições: o grafo de confiança é gerado.

1. <<include>> Analisar Indício (Podem ser incluídos uma ou mais extensões deste caso de uso).
2. Criar um novo grafo (grafo de confiança).
3. Copiar os nós do grafo de relacionamentos para novo grafo.
4. Criar uma aresta do vértice A para o vértice B e uma de B para A se existir pelo menos uma aresta entre A e B com rótulo $PR:\{prId\}$ no grafo de relacionamentos.
5. Definir o valor das arestas como a média ponderada dos valores das arestas geradas no caso de uso analisar indício e seu respectivo peso.
6. Retorna o grafo.

Gerar Grafo de Relacionamentos Inicial: Caso de uso responsável por gerar o grafo de relacionamentos para um projeto.

Pré-condições: dados do projeto.

Pós-condições: grafo de relacionamentos.

1. <<include>> Extrair Dados do Projeto.
2. Criar um grafo com um nó para cada participante do projeto.
3. Para cada par de participantes de uma *pull request* adicionar uma aresta com o rótulo PR:{prId} ao grafo.
4. Retornar o grafo resultante.

Configurar parâmetros: Caso de uso que realiza a configuração dos parâmetros para uso do *framework*. A configuração envolve a definição das técnicas de extração de indícios (representadas no diagrama de casos de uso pelo caso de uso **Analisar Indício** e suas extensões) que serão utilizados pelo *framework*, o peso dessas técnicas, o intervalo de tempo que será considerado para extração dos dados e o projeto alvo. As técnicas consideradas dependem da disponibilidade dos dados utilizados por eles no sistema de versionamento. A escolha dos pesos deve ser baseada em dados históricos.

Pré-condições: o *framework* encontra-se desconfigurado.

Pós-condições: o *framework* encontra-se configurado e pronto para uso.

1. Informar o projeto alvo.
2. Informar as técnicas de extração de indícios que serão utilizadas.
3. Definir o peso de cada técnica adicionada.
4. Definir o intervalo de tempo para extração dos dados.

Analisar Dados: Este caso de uso é responsável por analisar os dados do projeto em busca de um indício de confiança. Cada caso de uso que estende esse representa a análise de um dado diferente, uma técnica de extração de indícios diferente. Cada dado analisado pode levar a pelo menos um indício.

Pré-condições: o grafo de relacionamentos inicial e os dados do projeto alvo.

Pós-condições: arestas parciais referentes a análise dos dados são adicionadas ao grafo de relacionamentos inicial.

1. Analisar o grafo de relacionamento e os dados do projeto.
2. Gerar um valor para os dados analisados.
3. Adicionar uma aresta final com o valor gerado ao grafo de relacionamentos inicial.

Extrair Dados do Projeto: Este caso de uso é responsável por extrair do sistema de versionamento os dados de um projeto.

Pré-condições: o projeto alvo deve existir no sistema de versionamento e o *framework* deve estar configurado.

Pós-condições: os dados do projeto dentro do intervalo de tempo são extraídos.

1. Extrair as *pull request* que foram atualizadas dentro do intervalo considerado.
2. Extrair os comentários de cada *pull request* realizados dentro do intervalo considerado.
3. Extrair os participantes envolvidos (inclui os projetos assistidos, usuários seguidos e localização) nas *pull requests* dentro do intervalo considerado.
4. Extrair a última atribuição das *pull requests*, se esta estiver dentro do intervalo considerado.

Identificar Mímica nas Conversas: Este caso de uso é responsável por gerar arestas finais com a média de similaridade entre as conversas.

Pré-condições: o grafo de relacionamentos inicial e os dados do projeto alvo devem existir.

Pós-condições: o grafo de relacionamento ganha arestas finais com a média similaridade entre as conversas.

1. Analisar a similaridade de cada comentário com os comentários anteriores utilizando a similaridade do cosseno gerando uma aresta parcial com o valor da similaridade.
2. Criar uma aresta final entre cada par de participantes com a média do valor das arestas parciais.

Calcular Similaridade do Perfil: Este caso de uso é responsável por gerar as arestas finais com a similaridade do perfil dos membros do projeto.

Pré-condições: o grafo de relacionamentos inicial e os dados do projeto alvo devem existir.

Pós-condições: o grafo de relacionamento ganha arestas finais com a similaridade do perfil dos membros.

1. Para cada par de membros que tenha pelo menos uma aresta com rótulo PR: {prId} entre eles no grafo de relacionamentos, calcular:
 - Similaridade entre os usuários seguidos (similaridade de Jaccard).
 - Similaridade entre os projetos assistidos (similaridade de Jaccard).
 - Similaridade entre as culturas com base nos índices de Geert Hofstede.
2. Criar uma aresta final entre o par de participantes com a média das três similaridades.

Analisar Polaridade das Conversas: Este caso de uso é responsável por gerar as arestas parciais e finais com a proporção de conversas positivas entre os membros do projeto.

Pré-condições: o grafo de relacionamentos inicial e os dados do projeto alvo devem existir.

Pós-condições: o grafo de relacionamento ganha arestas finais com a proporção de conversas positivas entre os participantes.

1. Extrair o alvo dos comentários (criador da *pull request* e @{usuário}).
2. Analisar a polaridade dos comentários.
3. Criar uma aresta parcial do criador do comentário para cada alvo com peso 1 se o comentário for positivo, caso contrário com peso 0.
4. Analisar as arestas parciais de um participante para outro geradas no passo anterior e gerar uma aresta final com a porcentagem de comentários positivos em relação ao total de comentários entre os dois participantes.

Analisar Colaboração: Este caso de uso é responsável por gerar arestas finais com a porcentagem de comunicação com um vértice em relação ao total.

Pré-condições: O grafo de relacionamentos do projeto alvo deve existir.

Pós-condições: O grafo de relacionamento ganha arestas finais que representam a porcentagem de colaboração entre os participantes.

1. Contar a quantidade de arestas com rótulo PR: {prId} que saem de um nó para outro.
2. Criar uma aresta final tendo como peso o valor obtido no passo anterior dividido pelo total de arestas com rótulos PR: {prId} diferentes que saem do nó.

Analisar *Merges*: Este caso de uso é responsável por gerar arestas finais com a porcentagem de *pull requests* de um usuário que tiveram *merge* realizado e foram visualizadas por outro usuário.

Pré-condições: o grafo de relacionamentos inicial do projeto alvo e os dados do projeto devem existir.

Pós-condições: o grafo de relacionamento ganha arestas finais com a porcentagem de *pull requests* que tiveram *merge* realizado.

1. Para cada *pull request* que teve *merge* realizado adicionar uma aresta parcial entre os participantes da mesma e o seu criador.
2. Criar uma aresta final com a porcentagem das *pull requests* que tiveram *merge* realizado.

Analisar *Atribuições*: Este caso de uso é responsável por gerar uma aresta final caso um usuário atribua uma *pull request* a outro usuário.

Pré-condições: O grafo de relacionamentos do projeto alvo deve existir.

Pós-condições: O grafo de relacionamento ganha arestas finais de atribuição de *pull request*.

1. Criar uma aresta parcial do participante A para o participante B se A atribuiu uma *pull request* a B.
2. Criar uma aresta final do participante A para o participante B com valor 1 se existe pelo menos uma aresta parcial de A para B.

A Figura 3.6 apresenta o diagrama de componentes do *framework* ARSENAL-GSD, composto de 4 componentes, conforme descrito a seguir:

Graph Esse componente fornece ao *framework* os grafos que serão utilizados como grafo de relacionamentos inicial, grafo de relacionamentos e grafo de confiança. Alterando este componente, é possível alterar como o *framework* mantém seus dados. A implementação padrão utiliza os grafos fornecidos pela API JGraphT (Naveh e Contributors, 2015).

VS Data Extractor Esse componente extrai os dados do sistema de versionamento. Ele extrai as informações do perfil de um usuário, informações e conversas das *pull requests*. Em nossa implementação para o GitHub, utilizamos a GitHub Java API³, uma interface Java para a GitHub API (GitHub, 2015). Além de extrair os dados esse componente também gera o grafo de relacionamentos inicial.

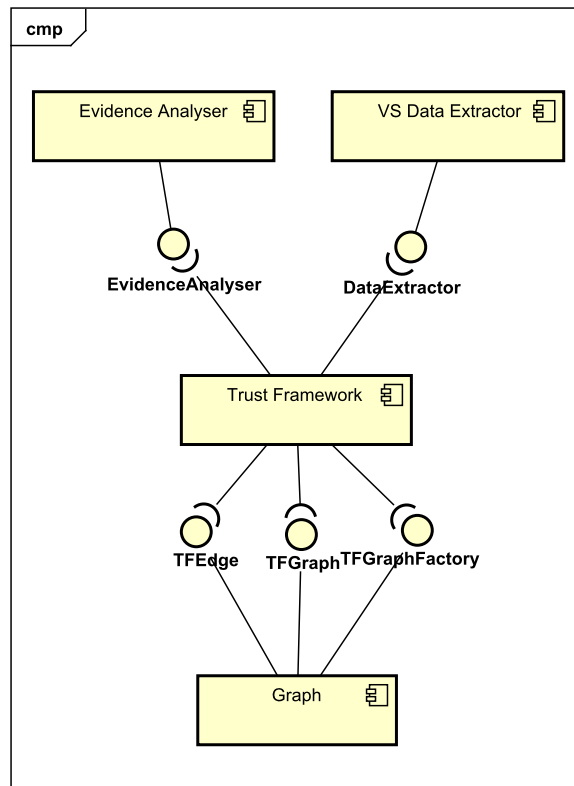
Evidence Analyser Esse componente fornece classes que implementam a interface *EvidenceAnalyser*, representando uma técnica de extração de indício. Cada uma dessas classes vai analisar os dados extraídos do sistema de versionamento e gerar um valor que é armazenado no grafo de relacionamentos e são utilizados para estimar a existência de confiança. Nós fornecemos seis técnicas para extração de indícios: mímica, atribuições, comunalidade, polaridade, *merges* e colaboração. Essas técnicas de extração de indícios serão detalhadas na Seção 3.4.1. Para adicionar mais indícios ao *framework* é necessário somente uma nova implementação da interface *EvidenceAnalyser* para uma nova técnica de extração de indícios.

Trust Framework Esse é o componente principal. Ele fornece meios de configurar e utilizar o *framework*. Esse componente é responsável por pegar os dados do **VS Data Extractor**, transmitir para o **Evidence Analyser** junto do grafo de relacionamentos, para que esse possa ser atualizado. A partir do grafo de relacionamentos esse componente gera o grafo de confiança utilizando a fórmula de confiança descrita no decorrer dessa seção.

Note que providenciando novas implementações para **VS Data Extractor** é possível estender o ARSENAL-GSD para outros sistemas de versionamento. Contudo, como cada sistema de versionamento pode conter dados diferentes, o componente **Evidence Analyser** está vinculado ao componente **VS Data Extractor**. Se quisermos oferecer suporte a outro sistema de versionamento, pode ser necessário substituir o componente **Evidence Analyser** por um que dê suporte ao novo sistema de versionamento. Também é possível estender o conjunto de indícios considerando implementando outras classes que estendam a interface *EvidenceAnalyser*.

³<https://github.com/eclipse/egit-github/tree/master/org.eclipse.egit.github.core>

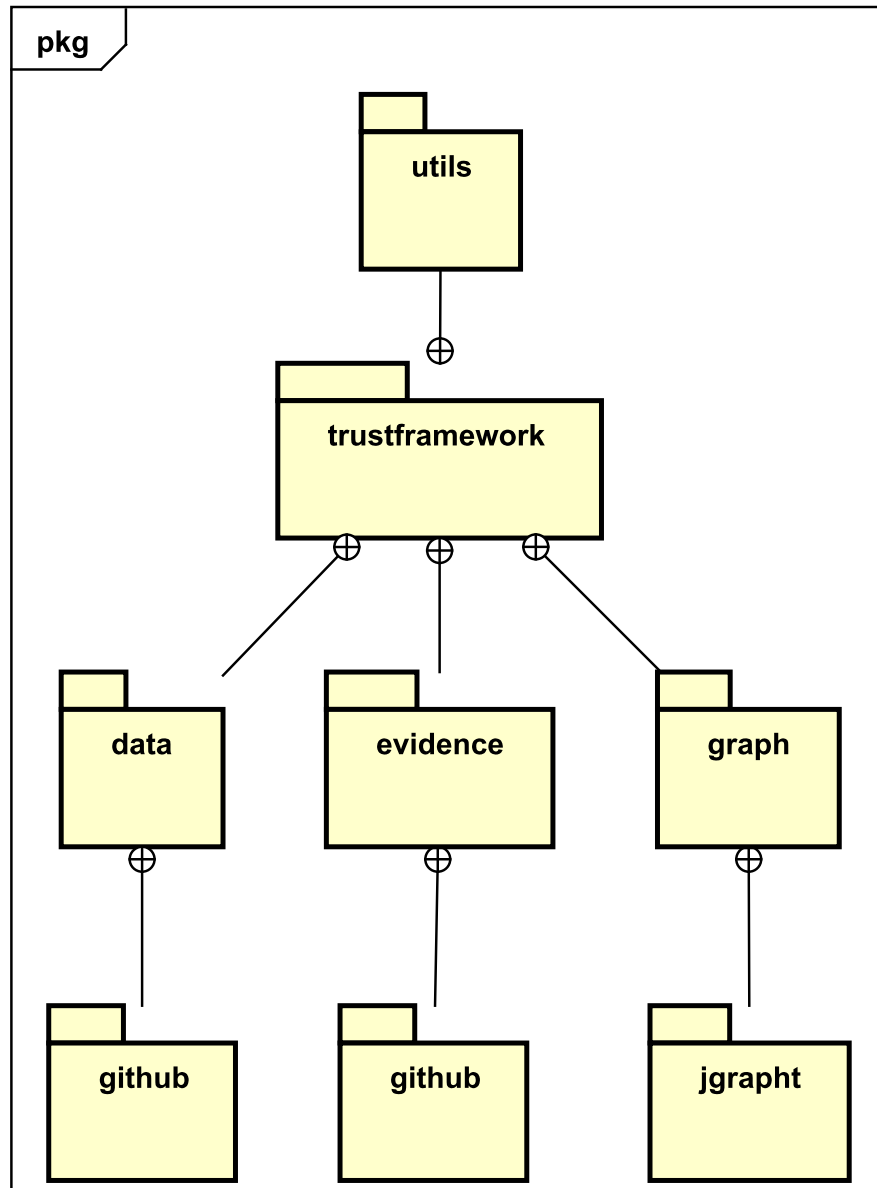
Figura 3.6: Diagrama de componentes para o *framework* ARSENAL-GSD.



Os componentes do *framework* ARSENAL-GSD podem ser mapeados em oito pacotes (Figura 3.7):

- br.uem.trustframework** pacote principal, que contém todas as classes do *framework*.
- br.uem.trustframework.graph** pacote com as classes/interfaces relacionadas com a estrutura de grafos utilizada pelo *framework*.
- br.uem.trustframework.graph.jgrapht** contém uma implementação dos grafos utilizando a API JGraphT.
- br.uem.trustframework.data** pacote com as classes/interfaces para extração dos dados de um sistema de versionamento e geração do grafo de relacionamentos.
- br.uem.trustframework.data.github** contém implementação das classes de extração para o GitHub.
- br.uem.trustframework.evidence** pacote com as classes/interfaces que representam as técnicas de extração de indícios para o *framework*.

Figura 3.7: Diagrama de pacotes do *framework* ARSENAL-GSD.



br.uem.trustframework.evidence.github contém a implementação das técnicas de extração de indícios discutidas neste trabalho para o GitHub.

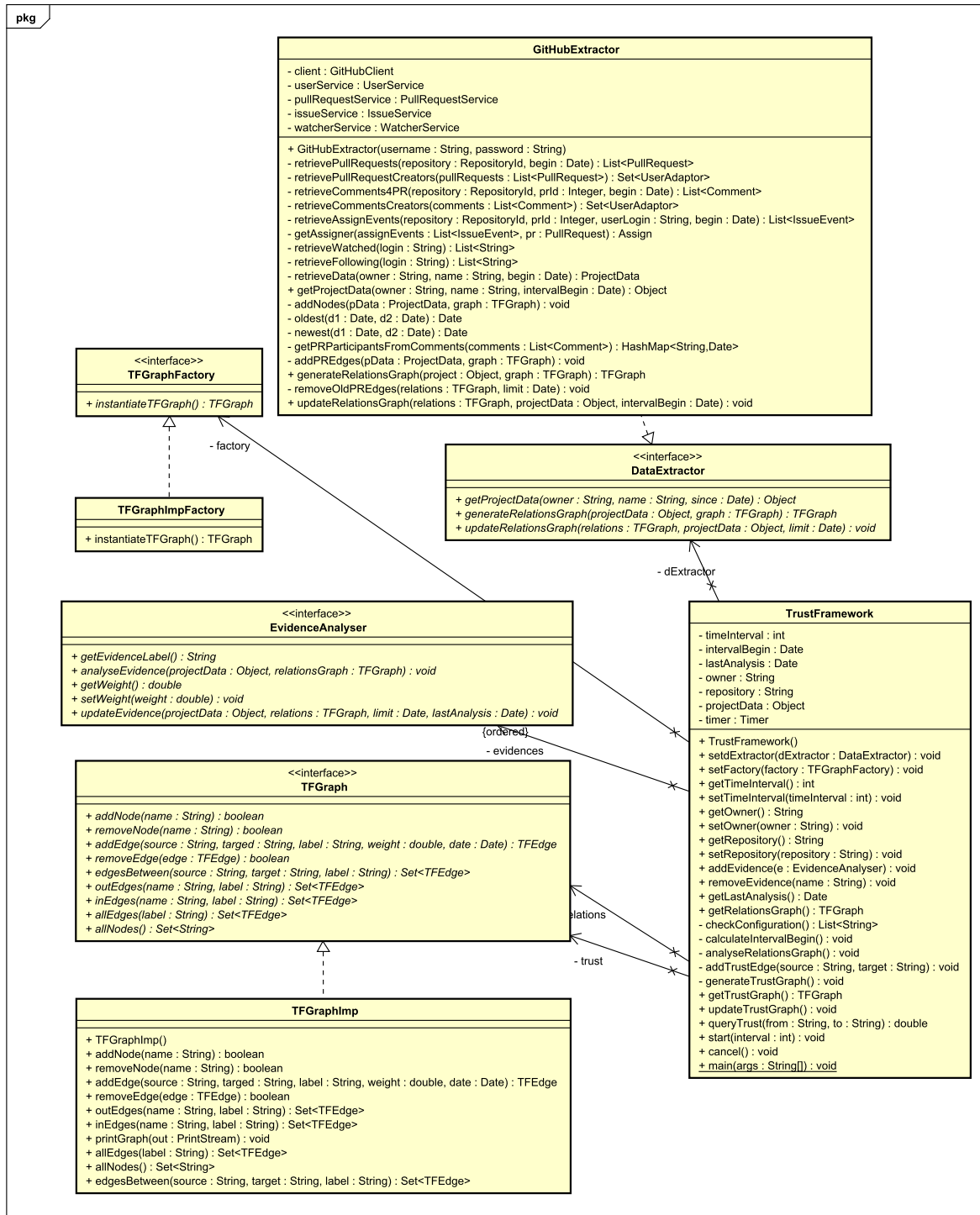
br.uem.trustframework.utils pacote com classes auxiliares, utilizadas pelas classes dos pacotes acima, como por exemplo, a classe para o cálculo das similaridades (jaccard e cosseno).

Esses pacotes contém as classes do *framework*. Como um diagrama de classes com todas as classes do *framework* não ficaria legível, o diagrama da Figura 3.8 mostra

apenas as classes principais do *framework* ARSENAL-GSD. Nesse diagrama, temos classe principal do *framework*, a `TrustFramework`, que mantém duas instâncias de `TFGraph` (o grafo de relacionamentos e o grafo de confiança), uma instância de `TFGraphFactory` para instanciar os grafos quando necessário e um conjunto de instâncias da interface `EvidenceAnalyser`. Cada instância de `EvidenceAnalyser` representa uma técnica de extração de indícios utilizada para se obter um ou mais indícios de confiança. As demais classes são implementações de suas respectivas interfaces.

Além dessas classes, temos as implementações disponíveis para o GitHub da interface `EvidenceAnalyser`: `Colaboration`, `TaskDelegation`, `MessageSentiment`, `Profile-Comunality`, `KnowledgeAcceptance` e `ConversationMimicy`, que analisam a colaboração, a atribuição de *pull requests*, a polaridade dos comentários, a similaridade do perfil, a quantidade de *merges* e a mímica de vocabulário, respectivamente. `TFEdge`, representa a aresta utilizada pelo *framework*, em nosso estudo de caso sua implementação, `Labeled-WeightedEdge`, foi por meio da API `JGraphT`. Temos também a classe `ProjectData`, que armazena os dados do projeto do GitHub e outras classes auxiliares que são utilizadas no *framework*.

Figura 3.8: Diagrama de classes do *framework* ARSENAL-GSD.



Os padrões *Abstract Factory*, *Facade* e *Adapter* foram empregados na implementação. *Abstract Factory* foi empregado por meio da interface *TFGraphFactory* utilizada para criar grafos dentro do *framework*. Por sua vez, o *Facade* foi utilizado com todas as interfaces, facilitando a utilização das classes que as implementam, agrupando os métodos para extrair dados, executar as técnicas de extração de indícios e gerar os resultados em um único método. O padrão *Adapter* foi utilizado na classe *UserAdaptor* a fim de adaptar a classe *User* fornecida pela GitHub Java API para uso no *framework*.

Além dos diagramas apresentados, foram feitos os diagramas de sequência para os casos de uso da Figura 3.5. Todos os diagramas de sequência, foram criados, contudo apresentamos, para efeito de ilustração, somente os diagramas para os casos de uso **Extrair Dados do Projeto** (Figura 3.9) e **Gerar Grafo de Confiança** (Figura 3.10). A seguir são feitas a descrição textual de todos os diagramas de sequência gerados:

Analisar Atribuições: A análise das atribuições começa com a geração das arestas parciais. O conjunto de atribuições do projeto é recuperado e para cada atribuição no conjunto adicionamos uma aresta parcial. Por fim geramos as arestas finais, recuperando todas as arestas parciais, e para cada par de participantes se existir pelo menos uma aresta parcial entre eles, adicionamos uma aresta final.

Analisar Colaboração: Para analisar as colaborações não há necessidade de gerar arestas parciais, uma vez que as arestas de interação são utilizadas. Para começar recuperamos todos os nós do grafo de relacionamentos e para cada nó mapeamos a quantidade de interações com cada participante. Calculado o número de interações criamos uma aresta final dividindo o resultado pelo total de interações.

Analisar Merges: A análise dos *merges* começa com a geração das arestas parciais, recuperando um mapeamento, de comentários para cada *pull request*. Para cada valor no mapeamento recuperamos a *pull request* e caso a mesma tenha realizado *merge*, adicionamos uma aresta parcial para cada participante que comentou na *pull request*, quem realizou o *merge* e quem atribuiu alguém à *pull request*. Por fim, geramos as arestas finais. Para tanto, recuperamos as arestas parciais e para cada par fonte-alvo distinto e geramos uma aresta final com proporção entre as arestas parciais e as arestas de interação entre eles.

Analisar Polaridade das Conversas: Para analisar a polaridade das conversas, iniciamos com a geração das arestas parciais recuperando os participantes envolvidos no projeto e o mapeamento de comentários por *pull request*. Para cada valor no mapeamento, recuperamos a *pull request*. Para cada comentário, extraímos os

alvos do comentário, incluímos o criador da *pull request* aos alvos e para cada alvo criamos uma aresta parcial partindo do criador do comentário, tendo peso 1, se a polaridade for positiva, e 0, caso contrário. Para gerar as arestas finais, recuperamos o conjunto de participantes envolvidos e, para cada participante, recuperamos as arestas parciais que partem dele. Para cada alvo distinto dessas arestas criamos uma aresta final com a proporção de arestas parciais para o alvo com valor 1 em relação ao total de arestas parciais para o alvo.

Atualizar Grafo de Confiança: A atualização do grafo de confiança começa recalculando o começo do intervalo de tempo considerado, então os dados do projeto dentro deste intervalo são extraídos e o grafo de relacionamentos é atualizado. Para cada indício executamos a atualização das arestas parciais e finais (todas as arestas finais e as parciais antigas são removidas) e por fim geramos o grafo de confiança novamente.

Calcular Similaridade do Perfil: Assim como a análise das colaborações, o cálculo da similaridade de perfil não utiliza arestas parciais. São geradas somente arestas finais. Para isso recuperamos todas as arestas de interação e para cada aresta criamos a aresta final entre a fonte e alvo. O cálculo da similaridade utiliza a similaridade de Jaccard para calcular a similaridade de usuários seguidos e de projetos assistidos. Além dessas similaridades, também utilizamos a similaridade entre a cultura dos participantes por meio do índices de Geert Hofstede. As arestas finais são criadas calculando-se a média simples com os valores obtidos das três similaridades.

Configurar: Para configurar o *framework* basta informar o proprietário do repositório, o nome do repositório, o intervalo de tempo considerado, o extrator de dados do projeto, uma instancia de `TFGraphFactory`, e as técnicas de extração de indícios. No diagrama foram adicionadas 6 técnicas de extração: colaboração, mímica nas conversas, *merges*, polaridade das conversas, similaridade do perfil e atribuições.

Extrair Dados do Projeto: A extração de dados do projeto consiste em recuperar as *pull requests* (1.1.1) do projeto, os comentários para cada *pull request* (1.1.4), os participantes do projeto (1.1.3 e 1.1.5), as atribuições (1.1.6) e, para cada participante, o conjunto de usuários que ele segue (1.1.9) e os projetos que ele acompanha (1.1.11).

Gerar Grafo de Confiança: Para gerar o grafo de confiança devemos obter o grafo de relacionamentos (1.1.1) e, para cada indício gerar as arestas finais (1.1.2). Tendo as arestas finais, podemos gerar o grafo de confiança. Para cada par de nós do grafo

de relacionamentos que possuam pelo menos uma aresta de interação, adicionamos o mesmo no grafo de confiança (1.1.2 e 1.2.3). Para cada indício buscamos as arestas finais entre esses nós (1.2.4.1 e 1.2.5.1) e adicionamos uma aresta no grafo de confiança cujo peso é a média ponderada do valor das arestas finais com o peso de cada indício (1.2.4.2 e 1.2.5.2).

Gerar Grafo de Relacionamentos Inicial: O grafo de relacionamentos é gerado adicionando um nó para cada participante envolvido no projeto e criando as arestas de interação. As arestas de interação são criadas analisando cada *pull request* e criando uma aresta entre cada par de participantes da mesma. Além dos usuários que comentaram a *pull request*, os que realizaram o *merge* ou atribuíram uma *pull request* a um usuário também são considerados participantes.

Identificar Mímica nas Conversas: Para identificar mímica de vocabulário nas conversas, primeiramente geramos as arestas parciais. Recuperamos o mapeamento de comentários por *pull request*. Para cada valor no mapeamento, recuperamos a *pull request* e para cada comentário calculamos a similaridade entre o mesmo e a descrição da *pull request* criando uma aresta parcial. Comparamos também cada comentário com os comentários anteriores criando mais arestas parciais. Criadas as arestas parciais, passamos para as finais. Recuperamos os participantes do projeto e para cada participante recuperamos as arestas parciais. As arestas finais são criadas tendo como peso a média simples dos pesos das arestas parciais para cada alvo.

Figura 3.9: Diagrama de sequência - Extrair Dados do Projeto

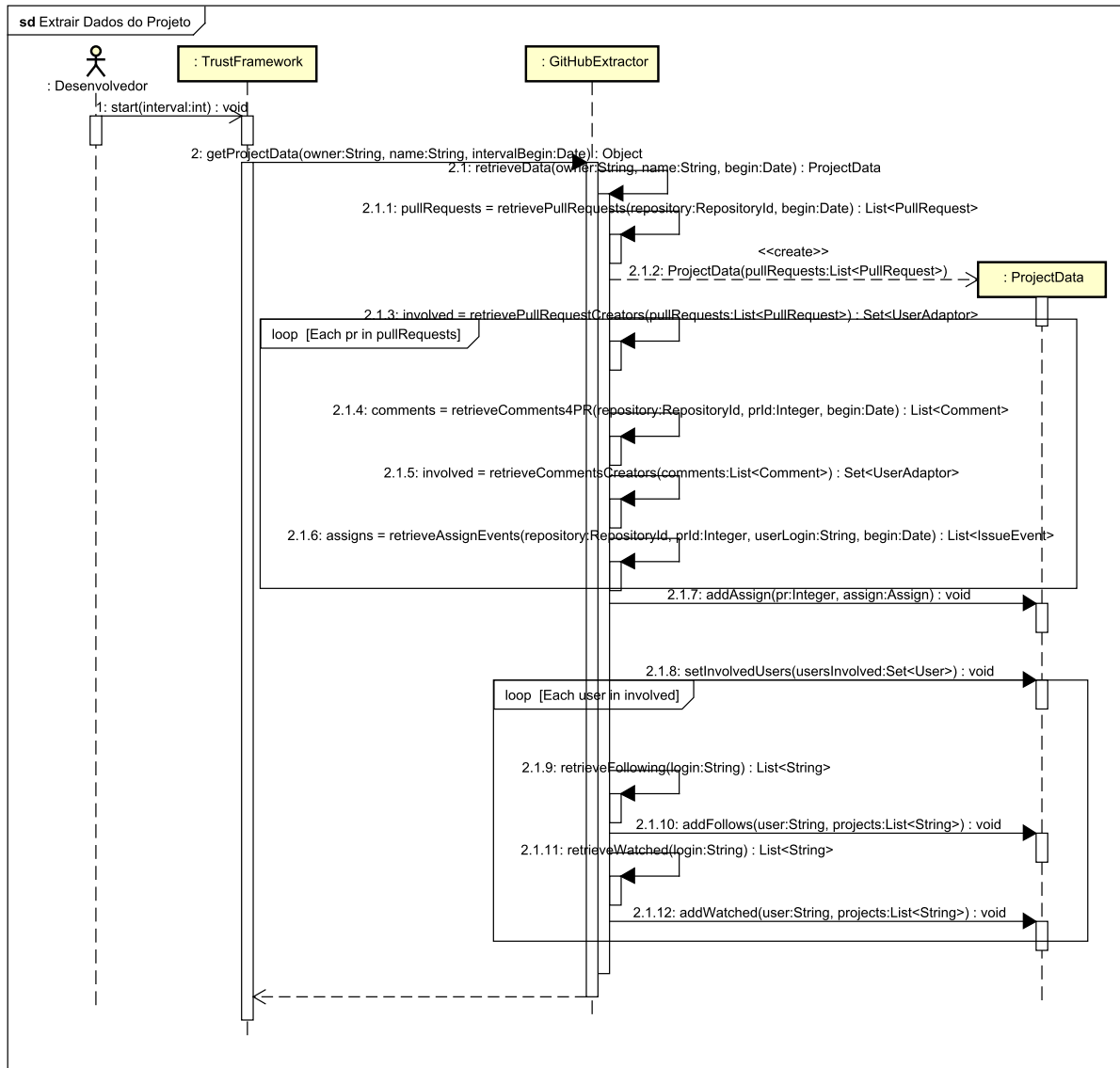
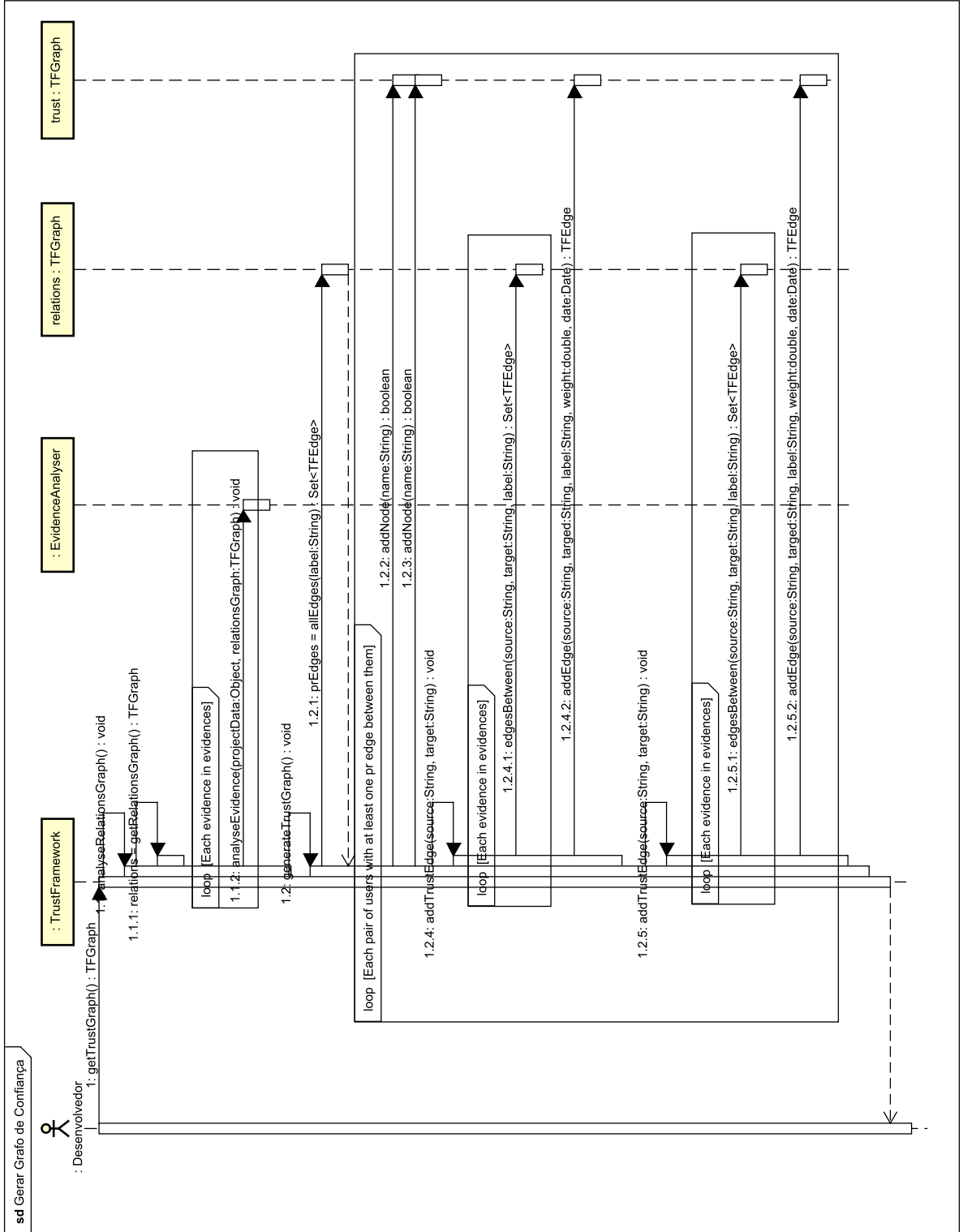


Figura 3.10: Diagrama de sequência - Gerar Grafo de Confiança



3.4.1 Técnicas de extração de indícios

A fim de extrair os valores dos indícios de confiança, utilizamos um conjunto de fórmulas, as quais chamamos de técnicas de extração de indícios na Figura 3.2. Para extrair a mímica de vocabulário, calculamos quão similar é o vocabulário utilizado nos comentários para uma *pull request*. Nós calculamos quão similar é o vocabulário de dois comentários utilizando a similaridade de cosseno sobre a frequência de palavras. Na Equação 3.1, $SC(c_1, c_2)$ é a similaridade de cosseno entre dois comentários, FV é o vetor com a frequência de palavras de cada comentário.

$$SC(c_1, c_2) = \frac{FV_1 \cdot FV_2}{\|FV_1\| \|FV_2\|} \quad (3.1)$$

O valor de mímica do membro m_1 para o membro m_2 , $MM(m_1, m_2)$ é a média das similaridades entre os comentários de m_1 e os comentários de m_2 que os precedem, em relação ao momento da escrita, na mesma *pull request*. Na Equação 3.2 PR_{12} é o conjunto de *pull requests* nas quais m_1 e m_2 interagem, C_1 é o conjunto de comentários de m_1 na *pull request* pr_i e C_{2j} é o conjunto de comentários de m_2 que precedem o comentário c_j .

$$MM(m_1, m_2) = \frac{\sum_{pr_i=1}^{|PR_{12}|} \sum_{c_j=1}^{|C_1|} \sum_{c_k=1}^{|C_{2j}|} SC(c_j, c_k)}{\sum_{pr_i=1}^{|PR_{12}|} \sum_{c_j=1}^{|C_1|} \sum_{c_k=1}^{|C_{2j}|} 1} \quad (3.2)$$

A Comunalidade é determinada calculando-se a média entre a similaridade de três atributos do perfil dos usuários do GitHub: (I) usuários seguidos, (II) projetos assistidos e (III) localização. (I) e (II) são calculados utilizando a similaridade de Jaccard, (Equação 3.3) na qual C_1 e C_2 são os conjuntos avaliados. Nesse caso, o conjunto de usuários seguidos e o conjunto de projetos assistidos por ambos os membros. Por sua vez, (III) utiliza uma variante da distância euclidiana (Dodd et al., 2013) (Equação 3.4) sobre os valores dos índices de Geert Hofstede (Hofstede et al., 2010). Os índices de Geert Hofstede caracterizam a cultura de uma região utilizando seis índices: distância do poder, individualismo, masculinidade, aversão à incerteza, orientação de longo prazo e indulgência. Na Equação 3.4 L_i é a localização, K é a quantidade de índices, I_{ik} é o valor de um índice k para uma localização L_i e V_k é a variância do índice k

$$JS(C_1, C_2) = \frac{C_1 \cap C_2}{C_1 \cup C_2} \quad (3.3)$$

$$ED(L_1, L_2) = \sqrt{\sum_{k=1}^K \frac{(I_{1k} - I_{2k})^2}{V_k}} \quad (3.4)$$

Portanto, a comunalidade $CM(m_1, m_2)$ entre os membros m_1 e m_2 é dada pela Equação 3.5, na qual F_i , W_i e L_i são, respectivamente, o conjunto de usuários seguidos, o conjunto de projetos assistidos e a localização de m_i .

$$CM(m_1, m_2) = \frac{JS(F_1, F_2) + JS(W_1, W_2) + \frac{1}{1+ED(L_1, L_2)}}{3} \quad (3.5)$$

Nós utilizamos a polaridade dos comentários para estimar benevolência, habilidade e tom positivo. O valor de polaridade entre dois membros é dado pela Equação 3.6, na qual $P(m_1, m_2)$ é o valor de polaridade do membro m_1 para o membro m_2 , C_{+12} é a quantidade de comentários com polaridade positiva de m_1 para m_2 e C_{12} é o total de comentários de m_1 para m_2 . Comentários de m_1 para m_2 são comentários que m_1 escreveu na *pull request* criada por m_2 ou comentários nos quais m_1 mencionou m_2 . Note que podem existir *pull requests* criadas por m_1 em que m_2 não interage e, nesse caso, as *pull requests* não são consideradas.

$$P(m_1, m_2) = \frac{C_{+12}}{C_{12}} \quad (3.6)$$

Dependabilidade e delegação podem ser observadas por meio de atribuições de um membro por outro em uma *pull request*. O valor de atribuição de um membro m_1 para um membro m_2 é 1 se m_1 atribuiu pelo menos uma *pull request* para m_2 ou 0 caso contrário. A Equação 3.7 calcula o valor de atribuição de m_1 para m_2 , $A(m_1, m_2)$ baseada na quantidade de *pull requests* atribuídas para m_2 por m_1 , PR_{s12} .

$$A(m_1, m_2) = \begin{cases} 1, & \text{if } PR_{s12} \geq 1 \\ 0, & \text{otherwise} \end{cases} \quad (3.7)$$

Nós inferimos os valores de aceitação de conhecimento e habilidade de um membro da quantidade de *pull requests* que tiveram *merge*. Na Equação 3.8 $M(m_1, m_2)$ é a proporção de *pull requests* criadas por m_2 em que m_1 e m_2 interagiram, que teve *merge*, PR_{a12} é a quantidade de *pull request* criadas por m_2 em que m_2 e m_1 interagiram, que teve *merge*, PR_{12} é a quantidade de *pull request* criadas por m_2 em que m_1 e m_2 interagiram.

$$M(m_1, m_2) = \frac{PR_{a12}}{PR_{12}} \quad (3.8)$$

Nós estimamos colaboração como a proporção de interações dada pela Equação 3.9. Nessa equação, $C(m_1, m_2)$, é a proporção de interações entre m_1 e m_2 em relação ao total de interações de m_1 . I_{12} é o número de interações entre m_1 e m_2 , e I_1 é a quantidade de interações de m_1 com todos os outros membros.

$$C(m_1, m_2) = \frac{I_{12}}{I_1} \quad (3.9)$$

Finalmente, estimamos os valores de confiança entre os membros utilizando a Equação 3.10, que é a média ponderada das fórmulas listadas anteriormente. A melhor maneira de definir os pesos α_i seria por meio do uso de dados históricos obtidos de projetos anteriores para aprender os pesos. Contudo, não temos conhecimento de uma base de dados com informação de confiança entre os membros em sistemas de versionamento. Assim nós definimos os pesos baseados no número de indícios calculados por meio do uso de cada fórmula apresentada acima.

$$T(m_1, m_2) = \frac{\alpha_1 MM(m_1, m_2) + \alpha_2 CM(m_1, m_2) + \alpha_3 P(m_1, m_2)}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6} + \frac{\alpha_4 A(m_1, m_2) + \alpha_5 M(m_1, m_2) + \alpha_6 C(m_1, m_2)}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6} \quad (3.10)$$

Primeiro, consideramos que cada indício tem o mesmo peso. Escolhemos o peso 2, porque nós iremos propagá-lo para os antecedentes de confiabilidade e técnicas de extração de indícios e se o valor fosse menor, iríamos obter muitas casas decimais.

Consideramos que cada antecedente de confiabilidade tem o mesmo peso para calcular a confiabilidade, então, propagando o peso 2 da confiabilidade para seus antecedentes, obtemos um peso 0,5 para cada antecedente.

Propagando os pesos dos antecedente e todos os indícios (exceto confiabilidade, que é propagada por meio de seus antecedentes) obtemos os valores $\alpha_1 = 2$, $\alpha_2 = 0.5$, $\alpha_3 = 2.75$, $\alpha_4 = 2.5$, $\alpha_5 = 2.25$ e $\alpha_6 = 2$, conforme mostrado na Figura 3.2.

Como exemplo, iremos propagar os pesos para $MM(m_1, m_2)$ e $A(m_1, m_2)$. $MM(m_1, m_2)$ infere somente mímica, que tem peso 2. Assim quando propagamos o peso para a fórmula, essa também ganha peso 2. $A(m_1, m_2)$ por sua vez, infere delegação com peso 2 e dependabilidade com peso 0,5. Assim, propagando esses pesos para a fórmula, essa ganha peso 2,5. Realizando esse procedimento para cada fórmula, obtemos todos os valores α .

Na Figura 3.2, os números entre parênteses representam os pesos de cada indício e antecedente da confiabilidade, exceto os que aparecem nas técnica de extração de indícios, que são os valores α obtidos propagando os pesos dos indícios e antecedentes da confiabilidade.

Como mencionado na descrição do componente **Evidence Analyser**, cada técnica de extração de indício está codificada em uma classe implementando a interface *EvidenceAnalyser* no componente *EvidenceAnalyser*. A fim de implementar essas classes, utilizamos as seguintes APIs:

- Lucene (Apache, 2015a): Utilizada para gerar o vetor de frequência de palavras necessário para o cálculo da mímica de vocabulário.
- Commons Math (Apache, 2015b): Fornece os métodos necessários para calcular a similaridade de cosseno.
- Sentistrength (Thelwall et al., 2010a): Utilizado como API de análise de sentimentos. A partir da API obtemos os valores de polaridade para cada comentário e calculamos o valor de polaridade ($P(m_1, m_2)$).
- AlchemyApi (IBM, 2015): Fornece várias funcionalidades referentes a análises linguísticas, incluindo análise de sentimentos. Contudo, nós a utilizamos a fim de extrair o alvo dos comentários (pessoas as quais o comentário foi direcionado) que foram utilizados para calcular o valor de polaridade.
- Google Maps Geocoding API (Google, 2015): O GitHub permite que todo usuário forneça a sua localização, então alguns usuários fornecem mais informações outros fornecem menos. Utilizamos a API Maps para descobrir de qual país é o endereço fornecido por cada usuário.

3.4.2 Funcionamento

Para começar a utilizar o *framework* ARSENAL-GSD é necessário configurá-lo. Para isso, precisamos informar o projeto alvo (proprietário/repositório), as técnicas de extração de indícios que serão utilizadas e seus pesos, o tamanho da janela temporal, a frequência de utilização e o mecanismo para gerar grafos.

Uma vez que o *framework* começa ser executado, ele extrai os dados do projeto do GitHub. Os dados de entrada podem ser classificados em três categorias:

- Perfil: que inclui projetos assistidos, usuários seguidos e localização;
- Conversas: que inclui os comentários das *pull requests*;
- *Pull Requests*: que inclui o estado da *pull request* e o usuário atribuído a ela.

Por sua vez, os dados do GitHub são processados para extrair os indícios apresentados na Seção 2.4.3. Os dados obtidos do GitHub são entregues a cada instancia da interface *EvidenceAnalyser*. Essas instancias irão adicionar arestas parciais e finais ao grafo de relacionamentos.

Os dados de perfil são utilizados para extrair o valor de comunalidade utilizando a fórmula $CM(m_1, m_2)$.

Utilizando as fórmulas $MM(m_1, m_2)$, $P(m_1, m_2)$ e $C(m_1, m_2)$ extraímos a mímica, a polaridade e a colaboração, respectivamente, a partir das conversas. O valor de mímica é extraído calculando a similaridade entre dois comentários. Polaridade é obtida usando uma ferramenta de análise de sentimentos nos comentários. O valor de polaridade por sua vez é utilizado para estimar tom positivo, benevolência e habilidade. Analisando a participação dos membros nas conversas calculamos os valores de colaboração.

As fórmulas $A(m_1, m_2)$ e $M(m_1, m_2)$ extraem os valores de atribuição e *merge* respectivamente a partir dos dados das *pull requests*. Utilizamos as atribuições para inferir delegação e dependabilidade. O valor de *merge*, por sua vez, infere aceitação de conhecimento e habilidade, uma vez que o *merge* é a aceitação da *pull request* de alguém, e a *pull request* é o resultado de sua habilidade.

Dependabilidade, habilidade, comunalidade e benevolência não são indícios de confiança. Contudo, a partir deles podemos estimar a confiabilidade que, por sua vez, é um indício de confiança. A confiabilidade combinada com mímica, delegação, tom positivo, aceitação de conhecimento e colaboração nos permite estimar a existência de confiança entre membros usando a Equação 3.10. Essa estimativa é mostrada por meio do grafo de confiança.

Com grafo de confiança gerado, podemos, por exemplo, utilizar os valores de confiança para sugerir membros para uma equipe que tenha uma chance maior de ter um alto nível de confiança. Além disso, como o *framework* processa os últimos comentários e automaticamente atualiza os valores de confiança, ele nos permite monitorar a confiança entre os membros, de forma que o gerente pode tomar providencias quando mudanças negativas forem percebidas na confiança da equipe.

3.5 Considerações Finais

Esta seção descreveu o *framework* ARSENAL-GSD para estimar a existência de confiança entre membros de equipes de DGS. Apresentamos: (i) as suas características comparadas as características dos trabalhos relacionados, (ii) o projeto e implementação para uma instancia do *framework* utilizando o sistema de versionamento GitHub, onde descrevemos

os casos de uso, componentes, pacotes, classes e técnicas de extração de indícios utilizadas na implementação, e (iii) o funcionamento do *framework*. A próxima seção apresenta uma *survey* realizada para investigar a importância de cada indício de confiança, dos antecedentes da confiabilidade, a validade das técnicas de extração de indícios e a validade da fórmula para estimar a existência de confiança.

Validação

4.1 Considerações Iniciais

Na Seção 3 descrevemos o *framework* ARSENAL-GSD para estimar a existência de confiança entre membros de equipes DGS que utilizam sistemas de versionamento. No decorrer da seção apresentamos os indícios considerados e quais as técnicas de extração para estes indícios. Devido à falta de uma base de dados com anotações de confiança entre membros em um sistema de versionamento consideramos que todos os indícios têm o mesmo valor para calcular a confiança, embora essa não seja a melhor maneira de definir os pesos.

O ideal para definir os pesos e avaliar a eficiência do *framework* seria a execução de algoritmos de aprendizado sobre uma base anotada, a partir da qual poderíamos aprender a melhor configuração de pesos para as técnicas de extração de indícios. Após determinados os pesos, a execução de experimentos utilizando o *framework* na base de dados anotada e comparando o resultado obtido com as anotações nos permitiriam avaliar a eficiência do mesmo. Como não foi possível realizar essas tarefas, optamos pela realização de um estudo empírico.

Estudos empíricos são necessários para nos ajudar a avaliar e validar os resultados obtidos durante uma pesquisa, permitindo dizer cientificamente quando alguma coisa é melhor do que outra. São dois os principais tipos de pesquisa: a qualitativa e a quantitativa. A pesquisa qualitativa tem como foco o estudo dos objetos em seu cenário natural, tentando entender um fenômeno baseado nas explicações que as pessoas oferecem. A pesquisa quantitativa tem como foco quantificar um relacionamento ou comparar dois

ou mais grupos, tentando estabelecer uma relação de causa e efeito. Os estudos empíricos por sua vez estão divididos em quatro tipos principais (Wohlin et al., 2006):

Experimento controlado interessado em um pequeno escopo, é realizado geralmente em laboratório. Costumam ser altamente controlados, podendo ser chamados de experimentos controlados.

Estudo de caso normalmente estudam projetos reais monitorando e coletando dados com um objetivo específico. Dependendo dos dados coletados, análises estatísticas podem ser realizadas.

Survey pode atingir um grande número de pessoas, uma vez que é possível enviar questionários por meio eletrônico. Geralmente é realizada em retrospecto, abordando algo que já é utilizado há algum tempo. Os principais meios de se obter informações em uma *survey* são entrevistas e questionários.

Análise *Post-mortem* Assim como as *surveys*, aborda algo no passado. Tem como alvo um projeto, ou atividade que já foi concluída. Geralmente é realizada observando a documentação ou entrevistando pessoas envolvidas.

A próxima subseção 4.2 apresenta uma *survey* aplicada à pessoas com conhecimento a respeito de desenvolvimento global de software para avaliar quais indícios de confiança são mais importantes e se as fórmulas utilizadas pelo *framework* são adequadas. Por fim, a subseção 4.3 apresenta as considerações finais dessa seção.

4.2 A Survey

Segundo Kasunic (2005, tradução nossa), uma *survey* é “uma abordagem para levantamento e análise de dados na qual os participantes respondem à questões ou afirmações desenvolvidas antecipadamente”, cujo desenvolvimento é constituído de sete etapas. Quando conduzida apropriadamente, é possível generalizar os resultados para uma grande quantidade de pessoas a partir de um subconjunto das mesmas. As sete etapas para a realização de uma *survey* estão detalhadas na 4.1.

As subseções que seguem apresentam os sete passos realizados em a nossa *survey*.

4.2.1 Identificar os objetivos de pesquisa

Esta *survey* teve como objetivo responder duas questões:

Tabela 4.1: Sete etapas para realização de uma *survey* (Kasunic, 2005).

1. Identificar os objetivos de pesquisa
O que você quer alcançar com a <i>survey</i> ? Que informação já existe sobre o problema que você está perguntando? A <i>survey</i> deve começar com uma afirmação do problema e como a <i>survey</i> irá responder as questões sobre o problema.
2. Identificar e caracterizar o público alvo
Quem, especificamente, irá responder a <i>survey</i> ? Quais suposições você pode fazer sobre conhecimento deles a respeito das questões que você tem em mente, da terminologia que eles entendem, da vontade deles de participar da <i>survey</i> , e assim por diante?
3. Projetar o plano de amostragem
Quão grande é a população do público alvo? O público alvo pode ser enumerado? Como você irá garantir que aqueles que responderem a <i>survey</i> serão representantes do público alvo?
4. Projetar e escrever o questionário
Os objetivos e questões internas da <i>survey</i> devem ser traduzidas em um questionário, cuidadosamente elaborado, para facilitar a análise e interpretação.
5. Realização do teste piloto
O questionário deve ser “testado” com membros do público alvo para remover problemas e melhorar o questionário.
6. Distribuir o questionário
O questionário deve ser distribuído para os membros selecionados do público alvo como definido pelo plano de amostragem.
7. Analisar os resultados e escrever o relatório
Os resultados devem ser coletados e traduzidos em gráficos apropriados que facilitem o entendimento. Os gráficos podem ser compilados em um relatório e interpretações, inferências, generalizações e ressalvas podem ser feitas baseadas nas evidências fornecidas pelos resultados.

- **P1:** Considerando que os pesos de cada uma das técnicas de extração de indícios não foram adequadamente definidos. Dentre os indícios apresentados na Figura 3.2, quão importante cada um deles é para o cálculo da confiança?
- **P2:** Várias fórmulas foram apresentadas para extrair os indícios (técnicas de extração de indícios). Essas fórmulas são adequadas para a captura dos indícios?

4.2.2 Identificar e caracterizar o público alvo

Nosso público alvo são pessoas da área de engenharia de software com graduação em um curso da área, preferencialmente aqueles com conhecimento em DGS e sistemas de versionamento que façam parte da indústria. Como a confiança é algo que todas as pessoas têm algum conhecimento, pois faz parte das relações pessoais e de trabalho, consideramos

também pessoas com conhecimento na área de processamento de linguagem natural, uma vez que o trabalho envolve análise de sentimentos.

4.2.3 Projetar o plano de amostragem

A seleção de participantes foi por conveniência. O questionário foi distribuído para pessoas conhecidas que combinavam com a descrição de público alvo para esta *survey*. O questionário foi distribuído para 22 pessoas, contudo obtivemos apenas 13 respostas.

4.2.4 Projetar e escrever o questionário

O questionário foi desenvolvido utilizando a plataforma Google Forms¹ e seguindo às diretrizes de Kasunic (2005). Ele foi dividido em quatro seções, sendo elas: (i) descrição da *survey*, assim como o termo de participação; (ii) caracterização do participante com seis questões a respeito do nível de estudo, profissão e conhecimento a respeito dos assuntos abordados; (iii) validação dos indícios, com 10 questões a respeito da importância dos indícios e antecedentes de confiabilidade e (iv) validação das fórmulas, com 11 questões a respeito das fórmulas utilizadas para a extração dos indícios. Além dessas questões, as seções (ii) e (iii) continham uma questão aberta para sugestões/críticas. Para que os participantes pudessem responder às questões, um documento com a descrição do *framework*, dos indícios considerados e das fórmulas foi disponibilizado. Esse documento, assim como o formulário, pode ser visto nos anexos A e B, respectivamente.

Para referência futura, as questões do questionário foram:

1. Caracterização do Participante

- (a) Nível de Estudo
- (b) Qual é a sua instituição/empresa atual?
- (c) Qual seu cargo atual?
- (d) Qual o seu conhecimento e experiência com desenvolvimento global de software (DGS)?
- (e) Qual o seu conhecimento e experiência com o GitHub?
- (f) Qual o seu conhecimento e experiência com o Análise de Sentimentos/Mi-neração de Opinião?

2. Validação dos Indícios

¹<https://www.google.com/forms/about/>

- (a) Para cada indício quão importante ele é para o cálculo da probabilidade de confiança (0-sem importância a 9-muito importante)? (Questão em tabela - 6 questões ao todo)
- (b) Para cada antecedente da confiabilidade, quão importante ele é ao determinar a confiabilidade de alguém? (Questão em tabela - 4 questões ao todo)

3. Validação das Fórmulas

- (a) Para cada indício, a fórmula utilizada para calculá-lo é válida e adequada? (Questão em tabela - 6 questões ao todo)
- (b) Para cada antecedente da confiabilidade, a fórmula utilizada para calculá-lo é válida e adequada? (Questão em tabela - 4 questões ao todo)
- (c) A fórmula para o cálculo da confiança é válida e adequada?

4.2.5 Realização do teste piloto

O teste piloto foi realizado com dois doutores, um da área de desenvolvimento global de software e um da área de processamento de linguagem natural. Os problemas levantados foram corrigidos.

4.2.6 Analisar os resultados e escrever o relatório

A Figura 4.1 apresenta os gráficos com as frequências para as respostas das perguntas de caracterização dos participantes. Dos participantes que responderam ao questionário, 5 são graduados, 5 são mestres e 3 são doutores, dos quais a maioria (6) são professores/pesquisadores em instituições de ensino superior. Quanto aos conhecimentos sobre o domínio do trabalho, exceto pela análise de sentimentos, em que a grande maioria respondeu nenhum/básico, os participantes estão bem divididos entre médio/alto e nenhum/básico com poucas ocorrências das respostas nenhum.

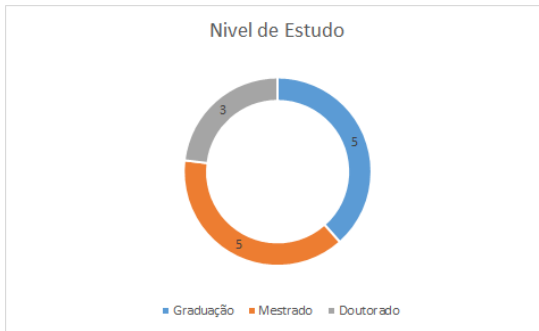
Os gráficos das figuras 4.2 e 4.3 apresentam as frequências das respostas a respeito da importância dos indícios para o cálculo da confiança e dos antecedentes para o cálculo da confiabilidade, respectivamente. Para essas perguntas obtivemos respostas variadas. Alguns valores tiveram uma frequência maior, mas nenhum gráfico apresentou uma resposta com 75% de frequência ou mais.

Por fim, as figuras 4.4, 4.5 e 4.6 apresentam as respostas a respeito da validade das fórmulas para extrair os indícios, antecedentes de confiabilidade e da própria fórmula

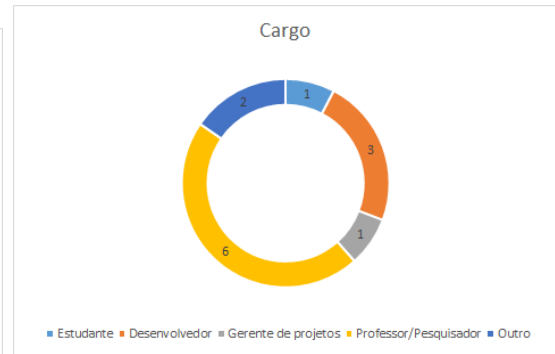
da confiança, respectivamente. Observamos que para todas as fórmulas, 10 ou mais participantes escolheram **Sim**, **boa** o **suficiente** ou **Sim**, **perfeita**.

A 4.2 apresenta as estatísticas descritivas para as respostas das figuras 4.2 e 4.3. A partir dessa tabela podemos responder à questão de pesquisa **P1**, utilizando-se das médias obtidas como peso dos indícios e antecedentes. Esses pesos, representam a importância de cada indício para calcular a probabilidade de confiança e de cada antecedente para calcular a confiabilidade de alguém. Desta forma, aplicando o mesmo procedimento para propagação dos pesos dos indícios de confiança e antecedentes de confiabilidade para as técnicas de extração de indícios utilizado para na Seção 3.4, obtemos os novos valores para α : $\alpha_1 = 5.15$, $\alpha_2 = 7.82$, $\alpha_3 = 1.78$, $\alpha_4 = 9.05$, $\alpha_5 = 8.27$ e $\alpha_6 = 7.15$. A Figura 4.7 apresenta o *framework* ARSENAL-GSD com os novos valores de α_{1-6} .

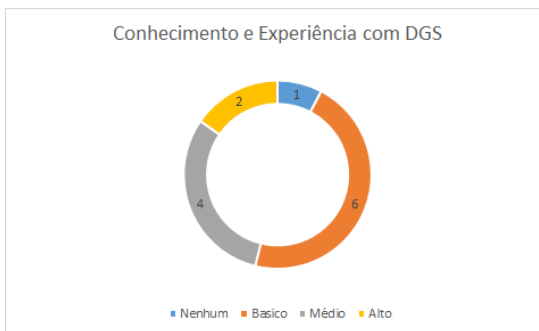
Figura 4.1: Frequência para as respostas de caracterização dos participantes.



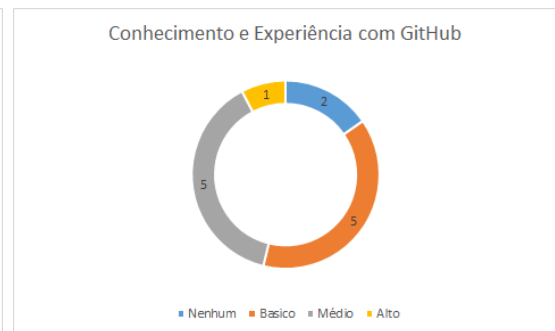
(a) Nível de estudo dos participantes



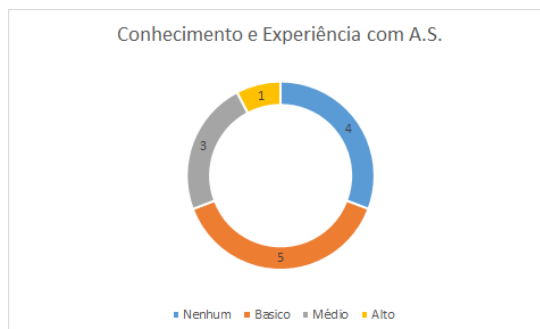
(b) Cargo dos participantes



(c) Experiência com DGS



(d) Experiência com GitHub



(e) Experiência com análise de sentimentos

Figura 4.2: Frequência para as respostas de validação dos indícios.

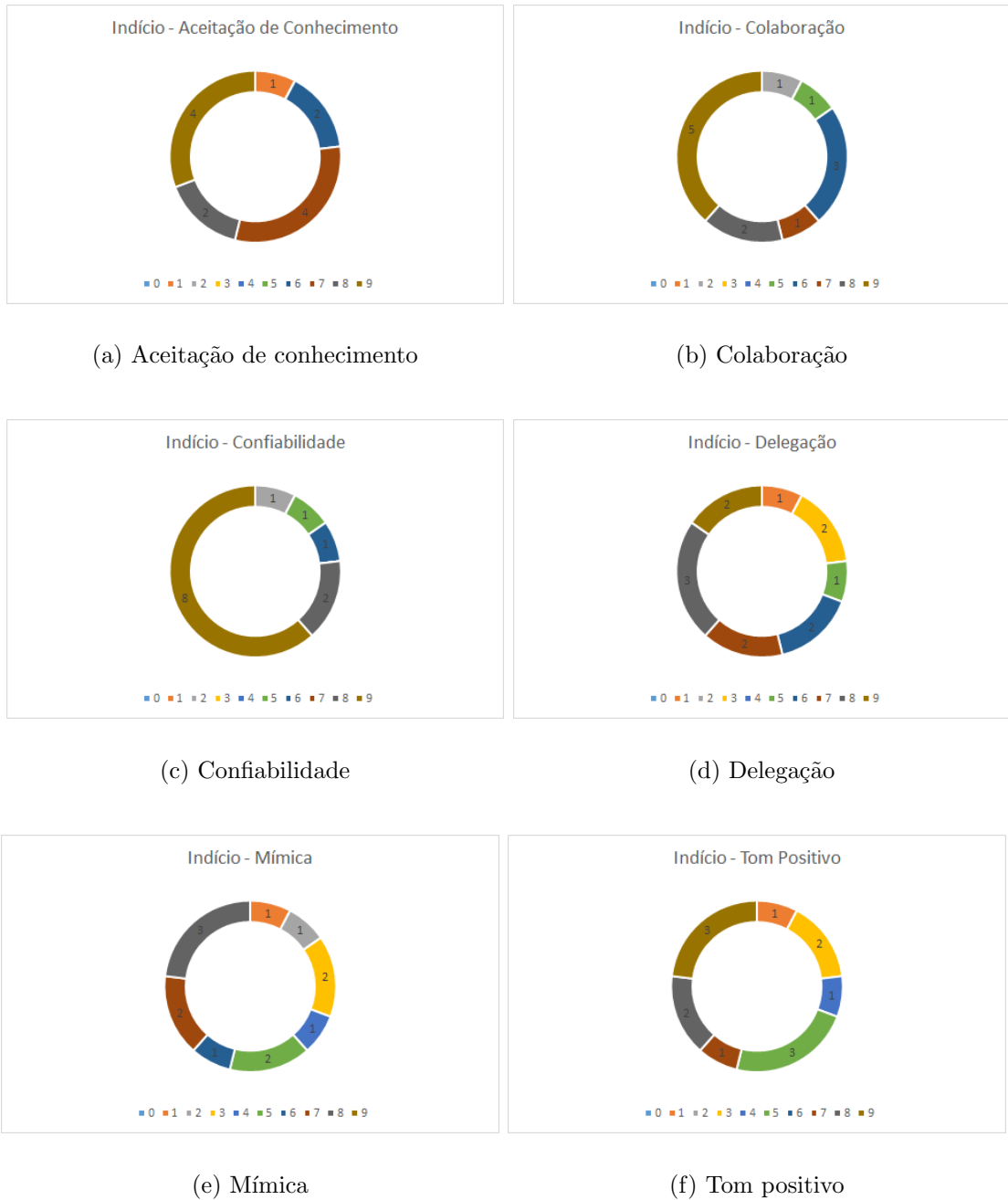


Figura 4.3: Frequência para as respostas de validação dos antecedentes.

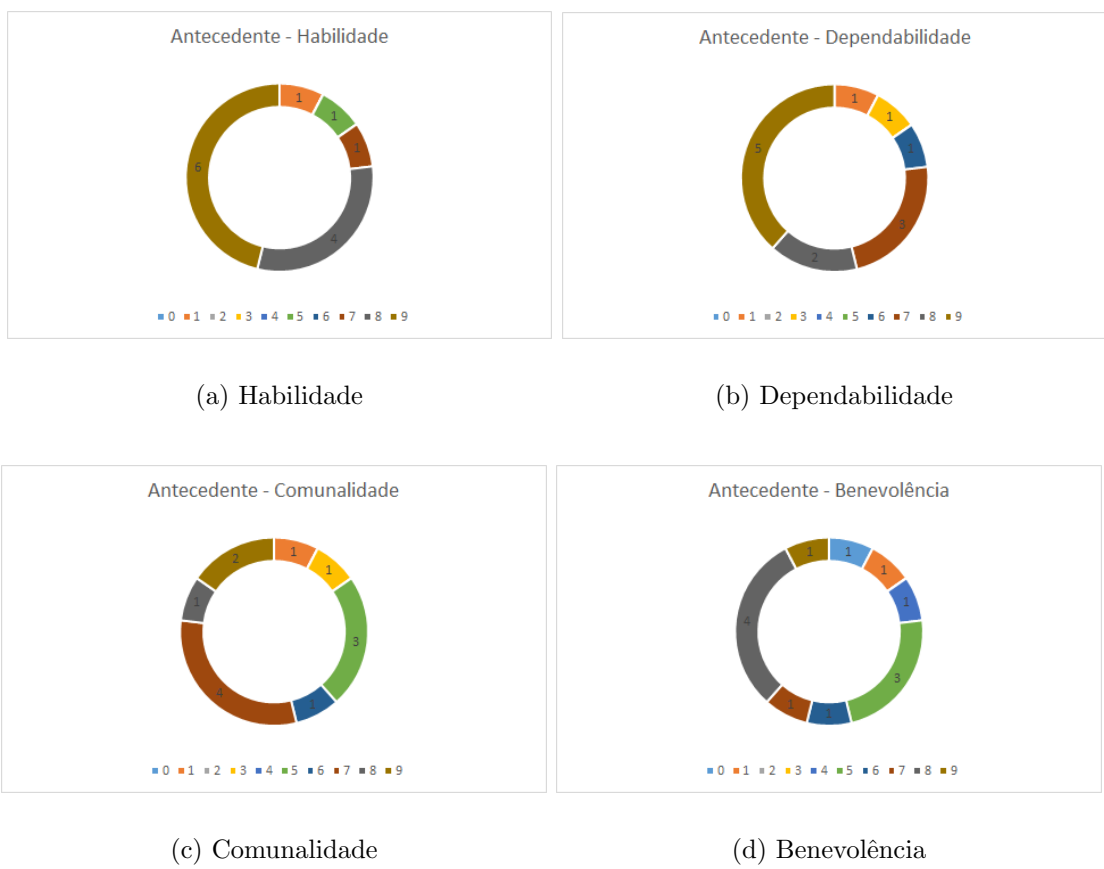


Figura 4.4: Frequência para as respostas de validação das fórmulas para os indícios.



(a) Aceitação de conhecimento

(b) Colaboração



(c) Confiabilidade

(d) Delegação



(e) Mímica

(f) Tom positivo

Figura 4.5: Frequência para as respostas de validação das fórmulas para os antecedentes.



(a) Habilidade

(b) Dependabilidade



(c) Comunalidade

(d) Benevolência

Figura 4.6: Frequência para as respostas de validação da fórmula de confiança.

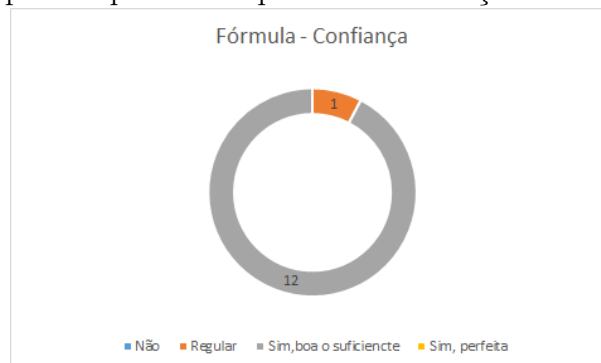


Figura 4.7: Framework ARSENAL-GSD com os pesos ajustados.

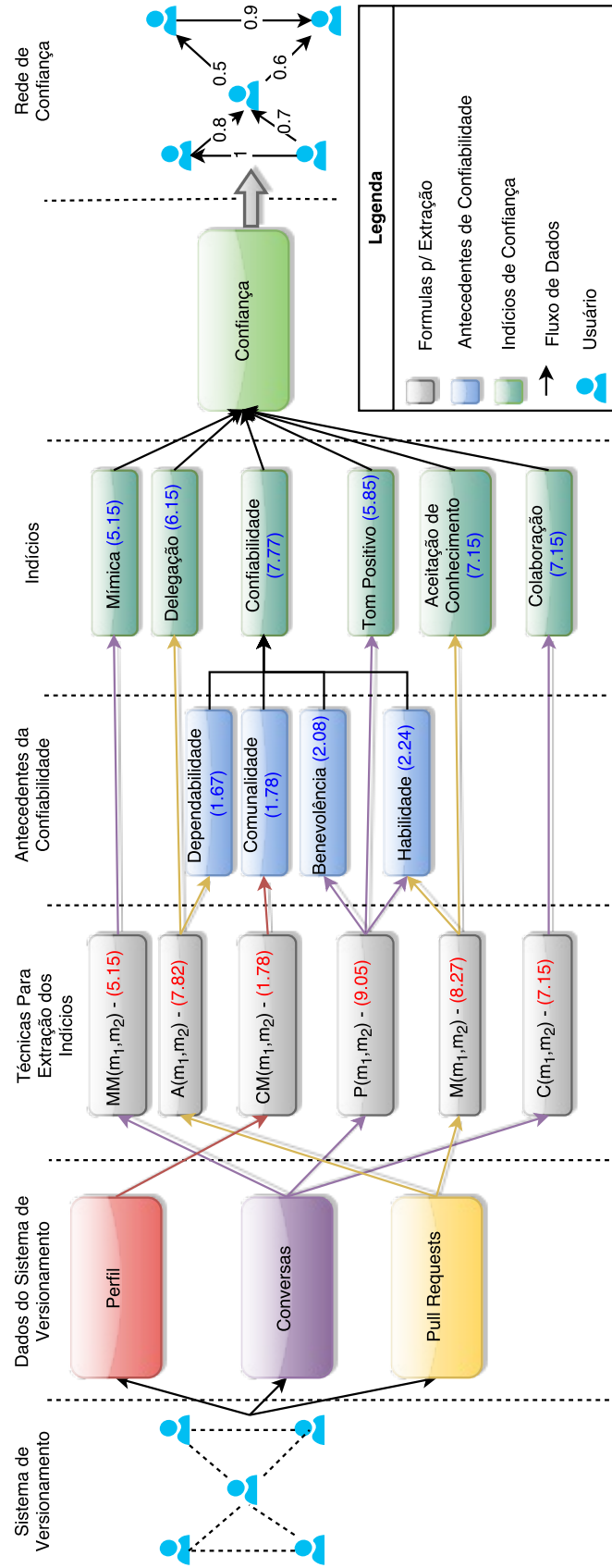


Tabela 4.2: Estatística descritiva indícios e antecedentes da confiabilidade.

Indícios/Antecedentes	Média	Variância	Desvio Padrão
Aceitação de Conhecimento	7,15	4,64	2,15
Colaboração	7,15	4,47	2,12
Confiabilidade	7,77	4,69	2,17
Delegação	6,15	6,31	2,51
Mímica	5,15	5,81	2,41
Tom Positivo	5,85	7,14	2,67
Habilidade	7,62	5,26	2,29
Dependabilidade	5,69	7,73	2,78
Comunalidade	6,08	5,24	2,29
Benevolência	7,08	6,24	2,50

Com os dados da *survey* decidimos analisar se existe uma diferença estatisticamente relevante entre os valores de importância dos indícios de confiança e antecedentes da confiabilidade fornecidos por participantes com nenhum conhecimento ou conhecimento teórico a respeito de DGS e participantes com conhecimento médio ou alto. Para isso utilizamos o teste estatísticos T-test, se os dados seguirem uma distribuição normal, ou o teste de Mann–Whitney, se os dados não seguirem uma distribuição normal. Para testar se os dados seguem uma distribuição normal ou não utilizamos o teste de normalidade de Shapiro-Wilk que possui duas hipóteses:

Hipótese nula (H0): Os dados seguem uma distribuição normal.

Hipótese alternativa (H1): Os dados não seguem uma distribuição normal.

Na 4.3 separamos os participantes em dois conjuntos de acordo com o conhecimento e experiência em DGS: Nenhum/Teórico e Médio/Alto. Para ambos os conjuntos e para cada indicio, aplicamos o teste de Shapiro-Wilk para determinar se os dados seguem uma distribuição normal. O *p-value* resultado do teste pode ser visto na segunda e terceira coluna, quando o valor é menor do que 0.05, rejeitamos a hipótese nula (**H0**) do teste, aceitando a hipótese alternativa (**H1**) e, portanto, o conjunto não segue uma distribuição normal (valores em vermelho), caso contrário não rejeitamos a hipótese nula e os dados seguem uma distribuição normal (valores em azul). Quando para um indicio os dois conjuntos seguem uma distribuição normal podemos utilizar um teste paramétrico, esta análise é apresentada na última coluna.

Utilizando os resultados da 4.3, optamos por utilizar o *T-test* ou o teste de Mann-Whitney. Para o *T-test* as hipóteses são:

Hipótese nula (H0): Não há diferença entre as médias.

Tabela 4.3: Resultado teste Shapiro-Wilk para normalidade.

Indícios/Antecedentes	Nenhum/Teórico	Médio/Alto	Paramétrico?
Aceitação de Conhecimento	0,03626	0,42070	Não
Colaboração	0,18710	0,05872	Sim
Confiabilidade	0,00106	0,03294	Não
Delegação	0,60060	0,21170	Sim
Mímica	0,36500	0,19120	Sim
Tom Positivo	0,65220	0,32040	Sim
Habilidade	0,01095	0,09114	Não
Dependabilidade	0,00176	0,31730	Não
Comunalidade	0,13860	0,50500	Sim
Benevolência	0,39110	0,18130	Sim

Hipótese alternativa (H1): A diferença entre as médias é diferente de 0.

já para o teste de Mann-Whitney as hipóteses são:

Hipótese nula (H0): O desvio de localização é igual a 0.

Hipótese alternativa (H1): O desvio de localização é diferente de 0.

A 4.4 apresenta para cada indício de confiança e antecedente da confiabilidade, qual foi o teste utilizado e o resultado do mesmo. Observe que todos os resultados foram superiores a 0.05, desta forma não podemos rejeitar a hipótese nula (**H0**). Portanto podemos concluir, com 95% de certeza que os indícios de confiança e antecedentes de confiabilidade tem a mesma importância independente do conhecimento em DGS. E com base na 4.2 em que a menor média foi 5.15 observamos que todos esses indícios são importantes para a confiança. Em uma pergunta aberta sobre possíveis indícios de confiança sugeriram a utilização de quantidade de defeitos em *commits* e retrabalho. Quanto maior o número de defeitos gerados pelos *commits* de uma pessoa menor a confiabilidade nela, de forma similar quanto mais os membros tiverem de refazer o trabalho de outro membro, menor a confiabilidade nele. Ambas as sugestões de indícios não são relevantes uma vez que utilizamos *pull requests*, uma vez que, elas tendem a ser aceitas somente quando estão corretas (não seria necessário alguém refazer) e não causariam problemas ao projeto (não possuem defeitos). Sendo assim estas sugestões estão de certa formas cobertas pelo indício aceitação de conhecimento extraído com a análise dos *merges*

Para tabular as respostas sobre a validade das fórmulas consideramos que 0=Não, 1=Regular, 2=Sim, boa o suficiente e 3=Sim, perfeita. Na 4.5 apresentamos as estatísticas descritivas para as respostas sobre a validade das fórmulas. Como utilizamos uma escala Likert nessas respostas não há sentido em apresentar a média, variância

Tabela 4.4: Resultado testes.

Indícios/Antecedentes	Teste	Resultado
Aceitação de Conhecimento	Mann–Whitney	0,41760
Colaboração	T-test	0,43100
Confiabilidade	Mann–Whitney	0,87040
Delegação	T-test	0,11020
Mímica	T-test	0,05239
Tom Positivo	T-test	0,14850
Habilidade	Mann–Whitney	0,64710
Dependabilidade	Mann–Whitney	0,41550
Comunalidade	T-test	0,05786
Benevolência	T-test	0,73570

e desvio padrão uma vez que não significarão nada. Portanto, a tabela apresenta a mediana, moda e amplitude (diferença entre o maior e o menor valor). Como podemos observar todas as fórmulas obtiveram mediana (valor que divide os dados ao meio) e moda (valor que mais se repete) igual a 2, podemos então considerar que as fórmulas foram consideradas válidas, respondendo assim à questão P2. Contudo nenhuma das fórmulas foi considerada perfeita, logo ainda cabem ajustes. Colaboração, delegação e dependabilidade foram as fórmulas que tem amplitude 3 pois obtiveram uma resposta em cada extremo, mas a maioria das respostas ficaram em **Sim**, **boa** o **suficiente** como todas as outras fórmulas, o que para nós é um bom resultado.

Tabela 4.5: Estatística descritiva validade das fórmulas para indícios e antecedentes.

Técnica	Mediana	Moda	Amplitude
Aceitação de Conhecimento	2	2	1
Colaboração	2	2	3
Confiabilidade	2	2	1
Delegação	2	2	3
Mímica	2	2	2
Tom Positivo	2	2	1
Habilidade	2	2	1
Dependabilidade	2	2	3
Comunalidade	2	2	2
Benevolência	2	2	1
Confiança	2	2	1

De forma similar ao que fizemos com a importância dos indícios e antecedentes, analisamos se existe uma dependência entre o nível de conhecimento/experiência em DGS e aceitação e rejeição das fórmulas. Para isso consideramos as respostas **Não** e

Regular como rejeição, já as respostas `Sim`, `boa` o `suficiente` e `Sim`, `perfeita` foram consideradas aceitação. A fim analisar esta dependência utilizamos o teste exato de Fisher que tem as seguintes hipóteses:

Hipótese nula (H0): As duas variáveis (nível de conhecimento/experiência e aceitação) são independentes.

Hipótese alternativa (H1): As duas variáveis são dependentes.

A 4.6 apresenta a distribuição das respostas agrupadas em Aceito/Rejeito de acordo com o conhecimento/experiência em DGS assim como o *p-value* obtido no teste exato de Fisher. Como nenhuma das fórmulas *p-value* < 0.05 a hipótese nula (H0) não pode ser rejeitada e, portanto, consideramos que não existe dependência entre o conhecimento/experiência em DGS e o aceite/rejeite das fórmulas.

Tabela 4.6: Distribuição das respostas agrupando por conhecimento em DGS e Aceito/-Rejeitado juntamente com o resultado do teste exato de Fisher.

Técnica	Nenhum/Teórico		Médio/Alto		Fisher
	Aceito	Rejeito	Aceito	Rejeito	
Aceitação de Conhecimento	6	1	5	1	1
Colaboração	6	1	5	1	1
Confiabilidade	5	2	5	1	1
Delegação	5	2	6	0	0,4615
Mímica	7	0	5	1	0,4615
Tom Positivo	5	2	5	1	1
Habilidade	6	1	5	1	1
Dependabilidade	6	1	6	0	1
Comunalidade	6	1	6	0	1
Benevolência	6	1	5	1	1
Confiança	6	1	6	0	1

Ao final da *survey* havia uma questão aberta para que os participantes fornecessem sugestões ou críticas sobre o trabalho. Abaixo listamos as sugestões/críticas fornecidas:

- *Aceitação de conhecimento teria relação com a dependabilidade:* alguns dos indícios de confiança e antecedentes da confiabilidade tem relação entre si, mas estas relações não foram abordadas em nosso trabalho. Fica esta análise como uma sugestão para trabalhos futuros.
- *É estranho a delegação ser binária:* optamos na delegação por analisar se existe confiança para delegar uma tarefa, uma *pull request*. Alguém delegar mais *pull*

requests para uma pessoa do que para outra significa que ela deu mais trabalho para uma do que para outra, seria o indício de quantidade de trabalho. Contudo não é possível calcular esse indício uma vez que não sabemos quanto é uma grande quantidade de trabalho no contexto do projeto com as informações que temos.

- *A mímica não é muito expressiva para a confiança uma vez que quanto mais tempo os membros trabalham juntos eles acabam “falando a mesma língua”*: do nosso ponto de vista esses membros passam a “falar a mesma língua” justamente porque a confiança se desenvolve entre os membros com o passar do tempo.
- *Utilização da estatística TF-IDF para calcular a mímica*: segundo a sugestão de um participante essa seria mais robusta para calcular a similaridade do vocabulário, sendo assim fica como trabalho futuro analisar esta sugestão.
- *Estudar interferência de mais de dois membros nas fórmulas*: o nosso trabalho foca na confiança entre dois membros, não há a necessidade de analisar a interferência dos outros membros, queremos justamente o que um membro acha do outro.
- *Simplificar a fórmula de confiança, para levar em conta computações já realizadas*: o *framework* leva em conta as computações já realizadas, mas faz isso por meio do grafo de relacionamentos. Os dados que podem ser reutilizados são mantidos no grafo de relacionamentos para a próxima estimativa da confiança.

4.2.7 Ameaças a Validade

Abaixo listamos as ameaças à validade deste trabalho:

Ameaças à Validade de Conclusão: a maior ameaça nesse sentido foi a quantidade de participantes que responderam a *survey*, dos 22 convidados somente 13 se dispuseram a responder. Contudo a grande maioria dos participantes tem conhecimento em DGS e desses 50% tem conhecimento prático. Como focamos na confiança, algo que todos têm conhecimento, as contribuições dos participantes foram significativas.

Ameaças à Validade de *Constructo*: O teste piloto foi realizado a fim de garantir que o questionário e o documento necessário para responde-lo fossem adequados. As sugestões obtidas no teste piloto foram analisadas e os problemas encontrados foram corrigidos.

Ameaças à Validade Interna:

Diferença entre os participantes: Metade dos participantes não possuíam conhecimento/experiência em DGS ou este era somente teórico e a outra metade possuía conhecimento/experiência em DGS média/alta (prática). A fim de analisar o efeito desta diferença de conhecimento/experiência aplicamos testes para analisar se houve diferenças em suas respostas.

Acurácia da resposta dos participantes: Os participantes receberam um documento explicando o *framework*, os indícios considerados e fórmulas utilizadas sobre as quais se baseavam as perguntas do questionário.

Efeito de fadiga: Consideramos que um participante levaria em média 30 minutos para ler e responder o questionário, mas para garantir que não houvesse problema com fadiga o questionário ficou disponível *online* durante um período de 30 dias.

Outros fatores importantes: é provável que não houve influência entre os participantes uma vez que o questionário estava disponível *online* e cada participante realizou em um horário diferente, muito provavelmente em um local diferente, uma vez que vários participantes trabalham em locais diferentes, alguns em estados diferentes.

Ameaças à Validade Externa: Preferencialmente tentamos conseguir participantes que trabalhem ou trabalharam com DGS, mas poucos participantes responderam ao convite, então estendemos para pessoas que tem conhecimento mesmo que apenas teórico sobre DGS. Como o principal tema abordado no questionário é confiança acreditamos que esta dificuldade foi amenizada. Outra ameaça foi o método de seleção dos participantes que se deu por conveniência.

4.3 Considerações Finais

Nesta seção descrevemos a *survey* realizada para validar os indícios e fórmulas utilizadas no *framework* ARSENAL-GSD e analisamos os resultados obtidos por meio dela. Com os resultados pudemos definir os pesos dos indícios de confiança, antecedentes de confiabilidade e conseqüentemente das técnicas de extração de indícios. As técnicas de extração de indícios por sua vez foram validadas, contudo melhorias ainda podem ser realizadas. A próxima seção apresenta como utilizar o *framework* por meio de uma aplicação simples.

Utilizando o ARSENAL-GSD

5.1 Considerações Iniciais

As seções anteriores apresentaram o *framework* ARSENAL-GSD, como ele foi concebido, suas características, o projeto e implementação de uma instância para o GitHub. Também validamos os indícios e técnicas de extração de indícios utilizados por meio de uma *survey*. Nesta seção mostramos como utilizar o ARSENAL-GSD de forma prática, por meio de uma aplicação simples que gera os grafos de relacionamentos inicial e o de confiança para um repositório do GitHub.

5.2 Instanciando o ARSENAL-GSD

Quadro 5.1: Instanciando o *framework* ARSENAL-GSD.

```
1 tf = new TrustFramework();
2 ...
3 tf.setFactory(new TFGraphImpFactory());
4 tf.setdExtractor(new GitHubExtractor(username, password));
5 tf.setOwner(owner);
6 tf.setRepository(repo);
7 tf.setTimeInterval(timeInteval);
8 tf.addEvidence(new ConversationMimicry(1));
9 ...
10 TFGraph relGraph = tf.getRelationsGraph();
11 TFGraph trustGraph = tf.getTrustGraph();
```

O código do Quadro 5.1 mostra como instanciar o *framework* ARSENAL-GSD. Primeiramente instanciamos a classe `TrustFramework` (linha 1). Em seguida, configuramos essa instância informando o *Factory* para os grafos, o extrator de dados do projeto (no exemplo utilizamos o GitHub), o proprietário do repositório, o nome do repositório e o intervalo de tempo a ser considerado (linhas 3-7). Também faz parte da configuração informar as técnicas de extração de indícios e seus pesos. A linha 8 adiciona a técnica de extração de mímica com peso 1. Configurada a instância, basta chamar o método `getRelationsGraph` (linha 10) para obter o grafo de relacionamentos ou `getTrustGraph` (linha 11) para obter o grafo de confiança.

5.3 Um Exemplo de Uso

Para exemplificar a utilização do *framework* desenvolvemos uma aplicação simples que configura e gera os grafos para um repositório do GitHub. A Figura 5.1 apresenta a tela inicial da aplicação. Nessa tela informamos o nome e senha de um usuário do GitHub, os dados do projeto alvo (proprietário e nome do repositório), a quantidade de dias a ser considerada para extração e análise dos dados e, por fim, quais as técnicas de extração de indícios que serão utilizadas na análise.

Figura 5.1: Tela da aplicação de exemplo.

Ao clicar no botão *Initial Relations Graph*, geramos o grafo de relacionamentos da Figura 5.2. O grafo gerado considera um período de 120 dias e contém 4 usuários. Observe que o usuário `loyalt` não interagiu com ninguém e por isso não tem nenhuma aresta com os outros usuários. Clicando no botão *Trust Graph*, geramos o grafo de confiança apresentado na Figura 5.3. Como `loyalt` não interagiu com nenhum usuário, ele não

aparece no grafo de confiança, uma vez que o valor da estimativa de confiança com os outros seria igual a zero.

Figura 5.2: Grafo de relacionamentos inicial.

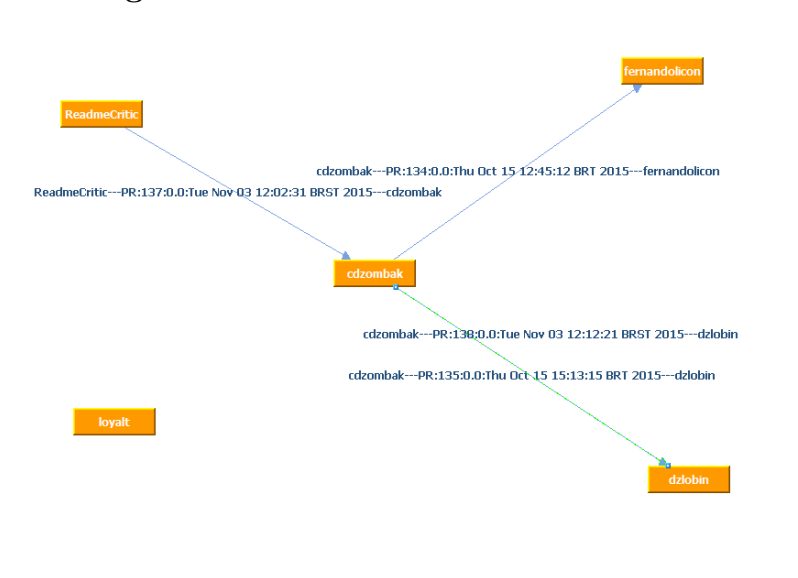
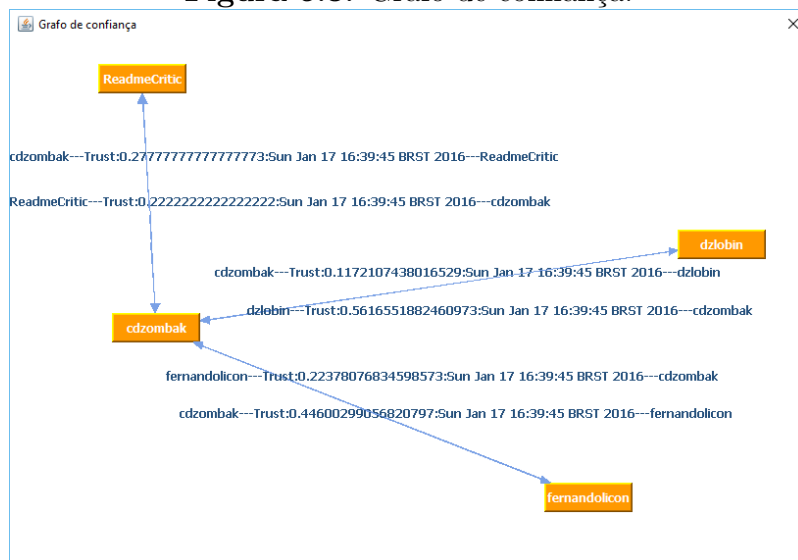


Figura 5.3: Grafo de confiança.



5.3.1 Arestas Parciais e Finais

A fim de mostrar como utilizamos as técnicas de extração de indícios apresentadas na Seção 3, vamos demonstrar como são criadas as arestas parciais e finais considerando as

pull requests 135 e 138 (figuras 5.4 e 5.5), nas quais existe uma interação entre **cdzombak** e **dzlobin**. A Figura 5.6 apresenta as arestas parciais e finais entre **cdzombak** e **dzlobin** adicionadas ao gráfico de relacionamentos durante o cálculo da estimativa de confiança. A seguir explicamos como foram calculados os valores de cada uma dessas arestas:

Sentiment Para cada comentário de **dzlobin** na *pull request* 135 e 138 adicionamos uma aresta **PSentiment** com o peso 1, uma vez que a polaridade encontrada nesses comentários é positiva. A aresta final **Sentiment** de **dzlobin** para **cdzombak** tem peso 1, pois é a proporção de arestas **PSentiment** com peso 1.

Similarity O valor da aresta **Similarity** representa a similaridade entre **cdzombak** e **dzlobin**, que foi calculada com base na localização, projetos assistidos e usuários seguidos.

Collaboration O valor das arestas **Collaboration** são calculados com base no grafo de relacionamentos inicial. Observando a Figura 5.2, notamos que **dzlobin** interage nas *pull requests* 135 e 138, em ambos os casos com **cdzombak**, o que leva ao valor 1 (100% das interações de **dzlobin** são com **cdzombak**). Já **cdzombak** interage nas *pull requests* 134, 135, 137 e 138, e dessas quatro somente duas são com **dzlobin**, o que faz com que a aresta **Collaboration** de **cdzombak** para **dzlobin** tenha valor 0,5.

KnowledgeAcceptance As *pull requests* 135 e 138 foram criadas por **cdzombak** e somente a 138 teve *merge*, o que cria a aresta **PKnowledgeAcceptance** com valor 1. A aresta **KnowledgeAcceptance** de **cdzombak** para **dzlobin** tem valor 0,5, pois metade das *pull requests* de **cdzombak** em que **dzlobin** interagiu teve *merge*.

Mimicry Os comentários de **dzlobin** nas *pull requests* 135 e 138 não tem nenhuma palavra igual às descrições fornecidas por **cdzombak**, por isso, as duas arestas **PMimicry** de **cdzombak** para **dzlobin** tem valor 0. Como a arestas **Mimicry** é a média simples das arestas **PMimicry**, seu valor também é 0.

Assign Como neste exemplo não existe nenhuma atribuição, as arestas **PAssign** e **Assign** não foram adicionadas ao grafo.

5.3.2 Utilizando as Saídas

O *framework* ARSENAL-GSD fornece duas saídas: o grafo de relacionamentos e o grafo de confiança. Com o grafo de relacionamentos o gerente pode analisar quais membros da

Figura 5.4: Pull request 135.

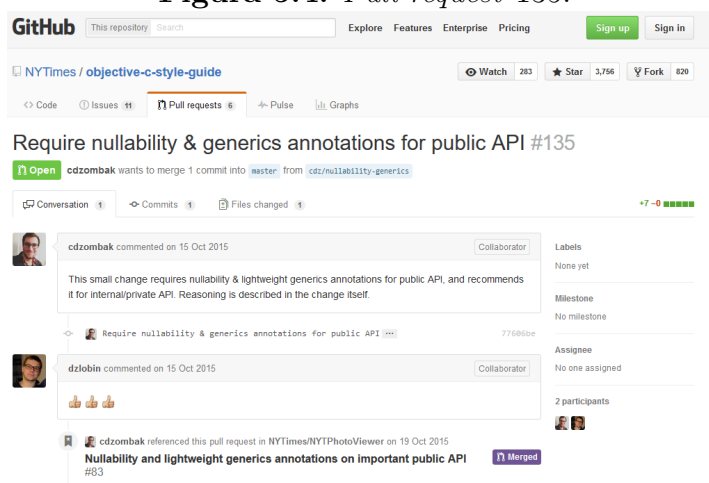
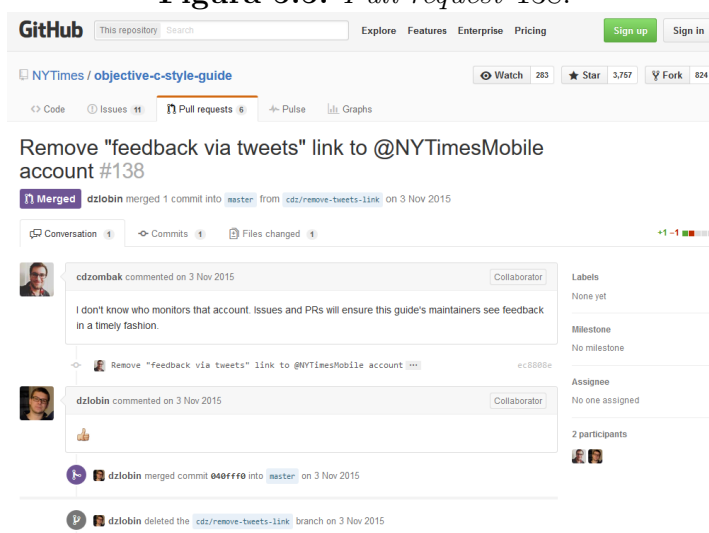


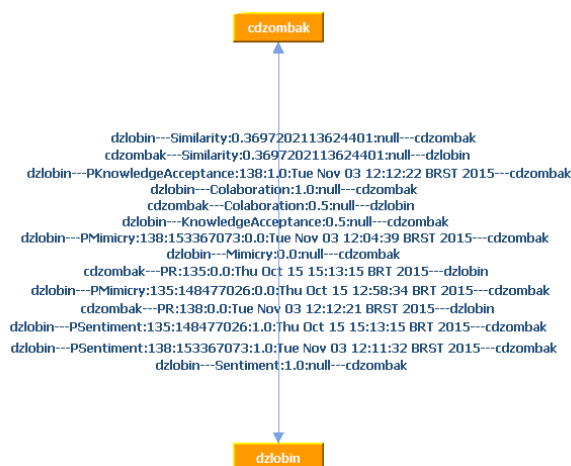
Figura 5.5: Pull request 138.



equipe interagem mais dentro da equipe, os que estão mais envolvidos no projeto. Esse tipo de membro tem grande chance de ter um conhecimento mais amplo do projeto, sendo ideal para o treinamento de novos membros, e para coordenar a equipe. No caso do grafo de relacionamentos da Figura 5.2, observamos que **cdzombak** seria esse tipo de membro, uma vez que ele está envolvido em todas as interações, e é um membro bastante ativo comparado aos outros.

Já o grafo de confiança é a principal saída do *framework*. Imagine a seguinte situação: o gerente de projetos deve alocar duas pessoas para realizar uma nova tarefa. Quais dois membros ele deve escolher para que a atividade seja realizada o mais rápido possível, mas com qualidade? O ARSENAL-GSD pode auxiliar o gerente a tomar essa decisão. Durante

Figura 5.6: Arestas entre **cdzombak** e **dzlobin** no grafo de relacionamentos.



este trabalho, mostramos diversos benefícios apresentados por equipes com alto nível de confiança, inclusive, vários desses benefícios foram considerados indícios da existência de confiança, como a qualidade das saídas, coesão, motivação. Essas características são essenciais para o DGS e garantem maior eficiência para a equipe, portanto, a melhor opção seria alocar os membros que possuem um maior nível de confiança entre eles. Considerando o grafo da Figura 5.3 e levando em conta que todos os membros possuem as habilidades necessárias para realizar a tarefa em questão, a melhor opção seria alocar **cdzombak** e **dzlobin** ($0,18 + 0,56 = 0,74$) para a equipe.

Os dois grafos também podem ser utilizados em conjunto para monitorar o nível de confiança entre os membros da equipe. Suponhamos que em um novo grafo de confiança, gerado 30 dias após o grafo de confiança da Figura 5.3, o valor da aresta de **dzlobin** para **cdzombak** caísse para 0,42, uma queda de 25%. O gerente deve descobrir o que gerou essa queda na confiança e avaliar quais medidas ele deve tomar para restabelecer a confiança. O grafo de relacionamentos pode ser utilizado em conjunto a fim de observar qual componente da estimativa sofreu queda.

5.4 Considerações Finais

Esta seção demonstrou como utilizar *framework* ARSENAL-GSD. Por meio de um exemplo sua utilização em um repositório real existente no GitHub. Apresentamos os parâmetros de entrada necessários para instanciar e configurar o *framework*, como são geradas as arestas parciais e finais do grafo de relacionamentos, os grafos gerados como

saída para o *framework* e como esses grafos podem ser utilizados em benefício de uma equipe DGS. A próxima seção apresenta as conclusões para este trabalho.

Conclusão

A eficiência de uma equipe de DGS está diretamente ligada à confiança entre os membros da equipe. A existência de confiança diminui os custos de monitoramento e o risco de risco moral, aumenta a comunicação, facilita a cooperação, coordenação e compartilhamento de conhecimento e informações que melhoram a qualidade dos produtos gerados.

Motivado pela importância da confiança nessas equipes, este trabalho teve como objetivo desenvolver um meio automático de estimar a confiança entre os membros de uma equipe de DGS a partir de indícios de confiança observados nas interações em sistemas de versionamento com o auxílio da análise de sentimentos. Realizamos um levantamento dos indícios de confiança, uma vez que não encontramos meios de calcular a confiança diretamente, e modelos de confiança. Com esse levantamento, constatamos que os modelos apresentados necessitavam de avaliação dos membros após cada interação, que é um problema se houverem muitas interações, ou ignoravam a subjetividade quando dispensavam essas avaliações (automáticos) e utilizavam métricas pré-definidas para medir o sucesso de uma interação.

Para resolver esses problemas, desenvolvemos o *framework* ARSENAL-GSD, a partir de um subconjunto dos indícios de confiança encontrados que podem ser extraídos de forma automática a partir dos dados das interações realizadas em um sistema de versionamento. Optamos por utilizar sistemas de versionamento, uma ferramenta colaborativa amplamente utilizada no desenvolvimento de software, como fonte de dados, uma vez que o foco do *framework* é DGS. Com a estimativa desses indícios de confiança, dispensamos a necessidade de avaliação dos membros a cada interação, como ocorre com alguns dos modelos apresentados na literatura. Para garantir a subjetividade ao *framework*, empregamos a análise de sentimento, uma das principais características do *framework*,

para extrair alguns dos indícios de confiança, por exemplo, o tom positivo nas conversas. Os indícios obtidos por meio da análise de sentimentos, juntamente com o indício mímica garantem ao *framework* um certo nível de subjetividade após a remoção da necessidade de avaliações.

6.1 Limitações

Algumas limitações podem ser observadas nesse trabalho. A primeira é a possibilidade de faltarem termos relevantes nas *strings* de busca de ambos os mapeamentos sistemáticos da literatura, o que pode ter levado a falta de algum trabalho relevante para o nosso trabalho. Outra limitação relacionada ao mapeamento sistemático é relacionada a interpretação dos trabalhos encontrados. Alguns trabalhos podem ter sido interpretados erroneamente ocasionando em sua exclusão.

Temos também uma limitação ligada a escolha dos indícios considerados. Proatividade, comunicação orientada à tarefas e motivação são indícios que podem ser observados nas conversas, mas não é possível identifica-los diretamente a partir da análise de sentimentos. Já, gerência de tempo, delegação de autoridade e divisão de papeis são exemplos de indícios que provavelmente ocorrem, mas as informações disponíveis nos sistemas de versionamento não são suficientes para extraí-los.

Por fim temos uma limitação referente a *survey*, devido a baixa quantidade de participantes, desses somente metade possuía conhecimento pratico em DGS. Além disso temos as outras ameaças a validade do experimento.

6.2 Contribuições

A principal contribuição do trabalho se encontra no mapeamento dos indícios de confiança e elementos dos modelos de confiança que podem ser capturados utilizando análise de sentimentos. Nós também contribuimos com uma implementação de uma instância do *framework* que funciona com o GitHub. Esperamos que o *framework* forneça uma estimativa melhor da existência de confiança em comparação aos modelos automáticos encontrados na literatura, uma vez que ele utiliza análise de sentimentos e um rico conjunto de indícios. Empregando análise de sentimentos, podemos extrair algo único de cada pessoa, adicionando assim subjetividade para nossas estimativas, que é uma característica importante da confiança. Essa subjetividade não pode ser capturada com o uso de métricas, que são geralmente utilizadas nos modelos automáticos.

Gerentes de equipes DGS podem se beneficiar do *framework* ARSENAL-GSD tanto por meio do grafo de relacionamentos quanto do grafo de confiança. O grafo de relacionamentos permite ao gerente observar quais membros são mais ativos e interagem mais dentro da equipe. Esses membros podem ser uma boa opção para coordenar equipes ou inteirar novos membros sobre o andamento do projeto. O grafo de confiança, a principal saída do *framework*, permite ao gerente criar uma equipe com altos níveis de confiança ou então monitorar variações no nível de confiança em equipes existentes, para que ações possam ser tomadas quando houver uma queda no nível de confiança. Como vimos ao longo deste trabalho, um alto nível de confiança entre os membros reduz os custos de monitoramento, o risco moral, permite o desenvolvimento de comportamentos benéficos para a equipe (vários indícios de confiança são exemplos desses comportamentos) e melhora a qualidade das saídas. No caso de haver uma queda no nível de confiança da equipe, o grafo de confiança pode ser analisado para descobrir em qual técnica de extração de indícios ocorreu a queda, auxiliando assim o gerente na adoção de medidas corretivas.

Como nós não temos um conjunto de dados do GitHub anotados com informações de confiança, inicialmente consideramos que todos os indícios e antecedentes da confiabilidade possuem o mesmo peso. Contudo, após realizarmos uma *survey* para avaliar a importância dos indícios, antecedente e as técnicas de extração de indícios observamos que cada indício de confiança e antecedente da confiabilidade tem uma importância diferente. Com os resultados da *survey*, recomendamos novos pesos para os indícios.

Todas as técnicas de extração de indícios foram consideradas boas o suficiente, mas não perfeitas e, portanto, ainda podem ser melhoradas. Entretanto, somente um participante sugeriu como melhorar a técnica de extração da mímica. As sugestões e críticas fornecidas pelos participantes foram analisadas e as que se aplicam serão consideradas em novas versões e trabalhos futuros. Como todos os indícios foram considerados relevantes e as técnicas de extração desses indícios foram consideradas adequadas, apesar de não termos avaliado o *framework* em si, essas evidências nos levam a crer que o *framework* é adequado para estimar a existência de confiança entre membros de equipes DGS.

Este trabalho obteve uma publicação na conferência *International Conference on Enterprise Information Systems* (ICEIS) de 2016, intitulada *Estimating Trust in Virtual Teams: A Framework based on Sentiment Analysis*.

6.3 Trabalhos Futuros

Como trabalhos futuros, nós pretendemos: (i) adicionar novos indícios de confiança ao ARSENAL-GSD, (ii) melhorar as técnicas de extração de indícios que não estão

perfeitas como constatado com a *survey*, (iii) replicar a *survey* a fim de comprovar os resultados obtidos, (iv) monitorar um projeto real, permitindo que colemos informações de confiança sobre os membros da equipe a fim de comparar com os resultados fornecidos pelo ARSENAL-GSD, (v) implementar aplicações que utilizem o ARSENAL-GSD, como um recomendador de equipes e um monitorador de confiança da equipe.

REFERÊNCIAS

ADALI, S.; ESCRIVA, R.; GOLDBERG, M.; HAYVANOVYCH, M.; MAGDON-ISMAIL, M.; SZYMANSKI, B.; WALLACE, W.; WILLIAMS, G. Measuring behavioral trust in social networks. In: *IEEE International Conference on Intelligence and Security Informatics*, 2010, p. 150–152.

AL-ANI, B.; BIETZ, M. J.; WANG, Y.; TRAINER, E.; KOEHNE, B.; MARCZAK, S.; REDMILES, D.; PRIKLADNICKI, R. Globally distributed system developers: their trust expectations and processes. In: *Conference on Computer Supported Cooperative Work*, ACM, 2013, p. 563–574.

Disponível em <http://dl.acm.org/citation.cfm?id=2441840>

AL-ANI, B.; MARCZAK, S.; REDMILES, D.; PRIKLADNICKI, R. Facilitating contagion trust through tools in global systems engineering teams. *Information and Software Technology*, v. 56, n. 3, p. 309–320, 2014.

Disponível em <http://linkinghub.elsevier.com/retrieve/pii/S0950584913001936>

AL-ANI, B.; WILENSKY, H.; REDMILES, D.; SIMMONS, E. An understanding of the role of trust in knowledge seeking and acceptance practices in distributed development teams. In: *International Conference on Global Software Engineering (ICGSE)*, 2011, p. 25–34.

APACHE Apache lucene - welcome to apache lucene. <https://lucene.apache.org/>, accessed: 2015-11-05, 2015a.

APACHE Math – commons math: The apache commons mathematics library. <http://commons.apache.org/proper/commons-math/index.html>, accessed: 2015-11-05, 2015b.

AWAL, G. K.; BHARADWAJ, K. K. Team formation in social networks based on collective intelligence - an evolutionary approach. *Applied Intelligence*, v. 41, n. 2, p. 627–648, 2014.

Disponível em <http://link.springer.com/article/10.1007/s10489-014-0528-y>

BERRY, G. R. A cross-disciplinary literature review: Examining trust on virtual teams. *Performance Improvement Quarterly*, v. 24, n. 3, p. 9–28, 2011.

Disponível em <http://doi.wiley.com/10.1002/piq.20116>

BIRD, S.; KLEIN, E.; LOPER, E. *Natural language processing with python*. "O'Reilly Media, Inc.", 2009.

BORBORA, Z. H.; AHMAD, M. A.; OH, J.; HAIGH, K. Z.; SRIVASTAVA, J.; WEN, Z. Robust features of trust in social networks. *Social Network Analysis and Mining*, v. 3, n. 4, p. 981–999, 2013.

Disponível em <http://link.springer.com/article/10.1007/s13278-013-0136-6>

CHENG, C.-L.; XU, X.-L.; GAO, B.-Z. METrust: A mutual evaluation-based trust model for p2p networks. *International Journal of Automation and Computing*, v. 9, n. 1, p. 63–71, 2012.

Disponível em <http://link.springer.com/article/10.1007/s11633-012-0617-5>

DODD, O.; FRIJNS, B.; GILBERT, A. On the role of cultural distance in the decision to cross-list. *European Financial Management*, p. n/a–n/a, 2013.

Disponível em <http://dx.doi.org/10.1111/j.1468-036X.2013.12038.x>

FAN, Z.-P.; SUO, W.-L.; FENG, B.; LIU, Y. Trust estimation in a virtual team: A decision support method. *Expert Systems with Applications*, v. 38, n. 8, p. 10240–10251, 2011.

Disponível em <http://www.sciencedirect.com/science/article/pii/S0957417411002661>

FILSTAD, C.; GOTTSCHALK, P. Knowing in mobile organisations - trust and knowledge sharing in virtual teams. *International Journal of Mobile Learning and Organisation*, v. 4, n. 3, p. 269–280, 2010.

Disponível em <http://dx.doi.org/10.1504/IJML0.2010.033555>

GITHUB Github api v3. <https://developer.github.com/v3/>, accessed: 2015-11-12, 2015.

GOMES, V.; MARCZAK, S. Problems? we all know we have them. do we have solutions too? a literature review on problems and their solutions in global software development. In: *International Conference on Global Software Engineering (ICGSE)*, 2012, p. 154–158.

Disponível em <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6337353>

GOOGLE The google maps geocoding api — google maps geocoding api — google developers. <https://developers.google.com/maps/documentation/geocoding/intro>, accessed: 2015-11-05, 2015.

GUZMAN, E. Visualizing emotions in software development projects. In: *First IEEE Working Conference on Software Visualization (VISSOFT)*, 2013, p. 1–4.

HERBSLEB, J. D.; MOITRA, D. Global software development. *IEEE Software*, v. 18, n. 2, p. 16–20, 2001.

Disponível em http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=914732

HOFSTEDE, G.; HOFSTEDE, G. J.; MINKOV, M. *Cultures and organizations: Software of the mind*. 3 ed. New York: McGraw-Hill USA, 2010.

HUANG, D.; ARASAN, V. Email-based social network trust. In: *IEEE Second International Conference on Social Computing (SocialCom)*, 2010, p. 363–370.

HUYNH, T. D.; JENNINGS, N. R.; SHADBOLT, N. R. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, v. 13, n. 2, p. 119–154, 2006.

Disponível em <http://link.springer.com/article/10.1007/s10458-005-6825-4>

IACONO, C.; WEISBAND, S. Developing trust in virtual teams. In: *Hawaii International Conference on System Sciences*, 1997, p. 412–420 vol.2.

IBM Alchemy api — powering the new ai economy. <http://www.alchemyapi.com/>, accessed: 2015-11-05, 2015.

IZQUIERDO, J.; COSENTINO, V.; ROLANDI, B.; BERGEL, A.; CABOT, J. Gila: Github label analyzer. In: *International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 2015, p. 479–483.

JARVENPAA, S. L.; KNOLL, K.; LEIDNER, D. E. Is anybody out there? antecedents of trust in global virtual teams. *Journal of Management Information Systems*, v. 14, n. 4, p. pp. 29–64, 1998.

Disponível em <http://www.jstor.org/stable/40398291>

JARVENPAA, S. L.; LEIDNER, D. E. Communication and trust in global virtual teams. *Journal of Computer-Mediated Communication*, v. 3, n. 4, p. 0–0, 1998.

Disponível em <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1998.tb00080.x/full>

JARVENPAA, S. L.; SHAW, T. R.; STAPLES, D. S. Toward contextualized theories of trust: The role of trust in global virtual teams. *Information Systems Research*, v. 15, n. 3, p. 250–267, 2004.

Disponível em <http://pubsonline.informs.org/doi/abs/10.1287/isre.1040.0028>

JIMÉNEZ, M.; PIATTINI, M. Problems and solutions in distributed software development: A systematic review. In: BERKLING, K.; JOSEPH, M.; MEYER, B.; NORDIO, M., eds. *Software Engineering Approaches for Offshore and Outsourced Development*, n. 16 in Lecture Notes in Business Information Processing, Springer Berlin Heidelberg, p. 107–125, 2009.

Disponível em http://link.springer.com/chapter/10.1007/978-3-642-01856-5_8

JONGELING, R.; DATTA, S.; SEREBRENİK, A. Choosing your weapons: On sentiment analysis tools for software engineering research. In: *Proc. ICSME*, IEEE, 2015, p. 531–535.

KASUNIC, M. *Designing an effective survey*. Relatório Técnico CMU/SEI-2005-HB-004, Software Engineering Institute, Carnegie Mellon University, 2005.

KHAN, M. S. Role of trust and relationships in geographically distributed teams: exploratory study on development sector. *International Journal of Networking and Virtual Organisations*, v. 10, n. 1, p. 40–58, 2012.

Disponível em <http://inderscience.metapress.com/index/N669244035618636.pdf>

KUO, F.-Y.; YU, C.-P. An exploratory study of trust dynamics in work-oriented virtual teams. *Journal of Computer-Mediated Communication*, v. 14, n. 4, p. 823–854, 2009.

Disponível em <http://doi.wiley.com/10.1111/j.1083-6101.2009.01472.x>

LAVRAC, N.; LJUBIC, P.; URBANCIC, T.; PAPA, G.; JERMOL, M.; BOLLHALTER, S. Trust modeling for networked organizations using reputation and collaboration estimates. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, v. 37, n. 3, p. 429–439, 2007.

LI, J.; ZHENG, X.; WU, Y.; CHEN, D. A computational trust model in c2c e-commerce environment. In: *IEEE International Conference on E-Business Engineering, ICEBE 2010*, Shanghai, China, 2010a, p. 244 – 249.

Disponível em <http://dx.doi.org/10.1109/ICEBE.2010.27>

LI, L.; OUYANG, C.; XIE, H.; SHI, X. The trust model of virtual team: Organizing method bases on the agent relation network. In: *2010 International Conference on Biomedical Engineering and Computer Science (ICBECS)*, 2010b, p. 1–5.

LI, M.; ALESSIO, B.; ZHOU, W. Ost: A transaction based online social trust model for social network and file sharing security. In: *IEEE/IFIP 8th International Conference on Embedded and Ubiquitous Computing (EUC)*, 2010c, p. 826–832.

LIU, B. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, v. 5, n. 1, p. 1–167, 2012.

Disponível em <http://dx.doi.org/10.2200/S00416ED1V01Y201204HLT016>

MAYER, R. C.; DAVIS, J. H.; SCHOORMAN, F. D. An integrative model of organizational trust. *The Academy of Management Review*, v. 20, n. 3, p. pp. 709–734, 1995.

Disponível em <http://www.jstor.org/stable/258792>

MCKNIGHT, D. H.; CHERVANY, N. L. *The meanings of trust*. Relatório Técnico WP9604, University of Minnesota, Management Information Systems Research Center, 1996.

MCNAB, A. L.; BASOGLU, K. A.; SARKER, S.; YU, Y. Evolution of cognitive trust in distributed software development teams: a punctuated equilibrium model. *Electronic Markets*, v. 22, n. 1, p. 21–36, 2012.

Disponível em <http://link.springer.com/10.1007/s12525-011-0081-z>

MITCHELL, A.; ZIGURS, I. Trust in virtual teams: Solved or still a mystery? *ACM SIGMIS Database*, v. 40, n. 3, p. 61–83, 2009.

Disponível em <http://doi.acm.org/10.1145/1592401.1592407>

MUMBI, C. K. An investigation of the role of trust in virtual project management success. 2007.

Disponível em <http://researchrepository.murdoch.edu.au/215/>

NAVEH, B.; CONTRIBUTORS Welcome to jgrapht - a free java graph library. <http://jgrapht.org/>, accessed: 2015-11-12, 2015.

NEPAL, S.; PARIS, C.; POUR, P.; BISTA, S.; FREYNE, J. A social trust based friend recommender for online communities “invited paper”;. In: *Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on*, 2013, p. 419–428.

NUNAMAKER, J. F.; REINIG, B. A.; BRIGGS, R. O. Principles for effective virtual teamwork. *Communications of the ACM*, v. 52, n. 4, p. 113–117, 2009.

Disponível em <http://doi.acm.org/10.1145/1498765.1498797>

O'CONCHUIR, E.; HOLMSTROM, H.; AGERFALK, P.; FITZGERALD, B. Exploring the assumed benefits of global software development. In: *International Conference on Global Software Engineering, 2006. ICGSE '06*, 2006, p. 159–168.

O'DONOVAN, J.; SMYTH, B.; EVRIM, V.; MCLEOD, D. Extracting and visualizing trust relationships from online auction feedback comments. In: *20th International Joint Conference on Artificial Intelligence, IJCAI'07*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007, p. 2826–2831 (*IJCAI'07*,).

Disponível em <http://dl.acm.org/citation.cfm?id=1625275.1625730>

PANG, B.; LEE, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, v. 2, n. 1-2, p. 1–135, 2008.

Disponível em <http://dx.doi.org/10.1561/15000000011>

PANGIL, F.; CHAN, J. The mediating effect of knowledge sharing on the relationship between trust and virtual team effectiveness. *Journal of Knowledge Management*, v. 18, n. 1, p. 92–106, 2014.

PAUL, S.; HE, F. Time pressure, cultural diversity, psychological factors, and information sharing in short duration virtual teams. In: *2012 45th Hawaii International Conference on System Science (HICSS)*, 2012, p. 149–158.

PEÑARROJA, V.; ORENGO, V.; ZORNOZA, A.; HERNÁNDEZ, A. The effects of virtuality level on task-related collaborative behaviors: The mediating role of team trust. *Computers in Human Behavior*, v. 29, n. 3, p. 967–974, 2013.

Disponível em <http://www.sciencedirect.com/science/article/pii/S0747563212003603>

RICO, R.; ALCOVER, C.-M.; SANCHEZ-MANZANARES, M.; GIL, F. The joint relationships of communication behaviors and task interdependence on trust building and change in virtual project teams. *Social Science Information*, v. 48, n. 2, p. 229–255, 2009.

Disponível em <http://ssi.sagepub.com/cgi/doi/10.1177/0539018409102410>

ROBBES, R.; LANZA, M. Versioning systems for evolution research. In: *Eighth International Workshop on Principles of Software Evolution*, 2005, p. 155–164.

RUSMAN, E.; VAN BRUGGEN, J.; SLOEP, P.; KOPER, R. Fostering trust in virtual project teams: Towards a design framework grounded in a TrustWorthiness ANtecedents (TWAN) schema. *International Journal of Human-Computer Studies*, v. 68, n. 11, p. 834–850, 2010.

Disponível em <http://www.sciencedirect.com/science/article/pii/S1071581910000881>

RUSMAN, E.; VAN BRUGGEN, J.; SLOEP, P.; VALCKE, M.; KOPER, R. The mind's eye on personal profiles: A cognitive perspective on profile elements that inform initial trustworthiness assessments and social awareness in virtual project teams. *Computer Supported Cooperative Work (CSCW)*, v. 22, n. 2-3, p. 159–179, 2013.

Disponível em <http://link.springer.com/10.1007/s10606-012-9171-5>

SCISSORS, L. E.; GILL, A. J.; GERGLE, D. Linguistic mimicry and trust in text-based cmc. In: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, CSCW '08*, New York, NY, USA: ACM, 2008, p. 277–280 (*CSCW '08*,).

Disponível em <http://doi.acm.org/10.1145/1460563.1460608>

SENGUPTA, B.; CHANDRA, S.; SINHA, V. A research agenda for distributed software development. In: *International Conference on Software Engineering*, ACM, 2006, p. 731–740.

Disponível em <http://dl.acm.org/citation.cfm?id=1134402>

SHERCHAN, W.; NEPAL, S.; PARIS, C. A survey of trust in social networks. *ACM Computing Surveys (CSUR)*, v. 45, n. 4, p. 47:1–47:33, 2013.

Disponível em <http://doi.acm.org/10.1145/2501654.2501661>

SIAKAS, K. V.; SIAKAS, E. The need for trust relationships to enable successful virtual team collaboration in software outsourcing. *International Journal of Technology, Policy and Management*, v. 8, n. 1, p. 59–75, 2008.

Disponível em <http://inderscience.metapress.com/index/v113677052t2j84k.pdf>

SKOPIK, F.; TRUONG, H. L.; DUSTDAR, S. Viete - enabling trust emergence in service-oriented collaborative environments. In: *WEBIST 2009 - Fifth International Conference on Web Information Systems and Technologies*, 2009, p. 471–478.

SOCHER, R.; PERELYGIN, A.; WU, J. Y.; CHUANG, J.; MANNING, C. D.; NG, A. Y.; POTTS, C. Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, Citeseer, 2013, p. 1642.

STRIUKOVA, L.; RAYNA, T. The role of social capital in virtual teams and organisations: corporate value creation. *International Journal of Networking and Virtual Organisations*, v. 5, n. 1, p. 103–119, 2008.

Disponível em <http://inderscience.metapress.com/index/D0441XN2L6554MQW.pdf>

THELWALL, M.; BUCKLEY, K.; PALTOGLOU, G.; CAI, D.; KAPPAS, A. Sentiment in short strength detection informal text. *Journal of the American Society for Information Science and Technology*, v. 61, n. 12, p. 2544–2558, 2010a.

Disponível em <http://dx.doi.org/10.1002/asi.v61:12>

THELWALL, M.; BUCKLEY, K.; PALTOGLOU, G.; CAI, D.; KAPPAS, A. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, v. 61, n. 12, p. 2544–2558, 2010b.

Disponível em <http://dx.doi.org/10.1002/asi.21416>

TUREK, P.; WIERZBICKI, A.; NIELEK, R.; HUPA, A.; DATTA, A. Learning about the quality of teamwork from wikiteams. In: *International Conference on Social Computing (SocialCom)*, 2010, p. 17–24.

WOHLIN, C.; HÖST, M.; HENNINGSSON, K. Empirical research methods in web and software engineering. In: MENDES, E.; MOSLEY, N., eds. *Web Engineering*, Springer Berlin Heidelberg, p. 409–430, 2006.

Disponível em http://dx.doi.org/10.1007/3-540-28218-1_13

YIN, G.; JIANG, F.; CHENG, S.; LI, X.; HE, X. Autrust: A practical trust measurement for adjacent users in social networks. In: *Cloud and Green Computing (CGC), 2012 Second International Conference on*, 2012, p. 360–367.

YU, Y.; WANG, H.; YIN, G.; LING, C. X. Reviewer recommender of pull-requests in github. In: *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, IEEE, 2014, p. 609–612.

YUSOF, S.; ZAKARIA, N. Exploring the state of discipline on the formation of swift trust within global virtual teams. In: *2012 45th Hawaii International Conference on System Science (HICSS)*, 2012, p. 475–482.

ZHANG, Y.; CHEN, H.-J.; JIANG, X.-H.; SHENG, H.; WU, Z.-H. RCTrust: A combined trust model for electronic community. *Journal of Computer Science and Technology*, v. 24, n. 5, p. 883–892, 2009.

Disponível em <http://link.springer.com/article/10.1007/s11390-009-9279-3>

A

Documento para Realização da Survey

A.1 Visão Geral

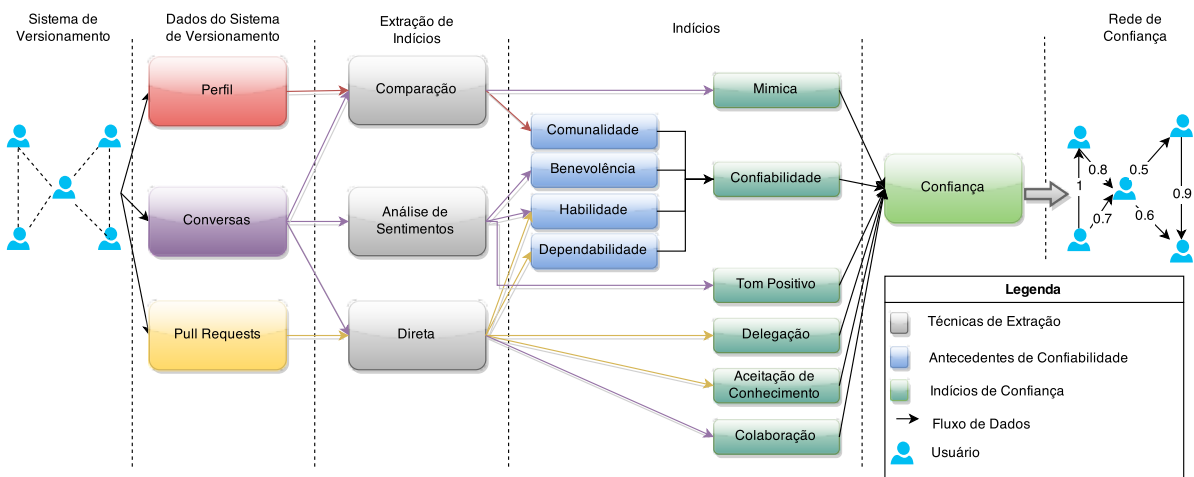


Figura 1.1: *Framework* proposto para estimar a confiança.

O *framework* (Figura 1.1) tem como entrada três tipos de dados:

- Perfil: projetos em que um usuário participou, endereço, usuários seguidos;
- Conversas: os comentários nas *pull requests*;
- *Pull Requests*: estado da *pull request*, usuários atribuídos a ela.

Esses dados por sua vez são processados para extração de alguns dos indícios:

- **Aceitação de conhecimento:** As pessoas tendem a aceitar conhecimento de quem elas confiam (Al-Ani et al., 2011). Em (Turek et al., 2010) a movimentação do texto de um autor por outro em um artigo da Wikipédia foi considerada um sinal de confiança, pois expressa a tendência de um autor acreditar na credibilidade do outro.
- **Confiabilidade:** confiança e confiabilidade estão altamente correlacionadas, quando a confiança em alguém é alta a confiabilidade também é. A confiabilidade é avaliada com base em seus antecedentes, que são divididos em cinco categorias (Rusman et al., 2010): (i) Comunalidade, que diz respeito a quão similares as pessoas são, (ii) Capacidade, que se refere às capacidades intelectuais da pessoa, (iii) Benevolência, que se refere ao nível de cortesia e positividade que o confiado exhibe em relação ao confiante, (iv) Normas internalizadas, se referem às normas morais do confiado e (v) Dependabilidade, que se refere a quão responsável é uma pessoa. Destes cinco antecedentes, comunalidade, benevolência, habilidade e dependabilidade podem ser capturados nos sistemas de versionamento.
- **Tom positivo:** em equipes de alta confiança os membros mostram entusiasmo nas conversas, elogiam e encorajam os outros. (Jarvenpaa et al., 1998) observou que as discussões eram resolvidas tão gentilmente que quase não eram percebidas (Jarvenpaa et al., 1998).
- **Colaboração:** equipes de alta confiança exibem grande quantidade de interações, comunicação e fornecem *feedback* sobre o trabalho dos companheiros (Jarvenpaa et al., 1998; Kuo e Yu, 2009; Mitchell e Zigurs, 2009; Rusman et al., 2010).
- **Divisão e delegação de tarefas** ocorrem com a existência de confiança (Rusman et al., 2010).
- **Mímica:** uso de um vocabulário comum em comunicação mediada por computador acontecem quando existe confiança entre as pessoas (Scissors et al., 2008).

Os dados podem ser comparados para encontrar coisas em comum no perfil de dois usuários (comunalidade), bem como a repetição de palavras/frases/abreviações nas conversas (mímica). Uma das características deste *framework* é a utilização da análise de sentimentos nas conversas. A polaridade encontrada nos comentários pode indicar benevolência, tom positivo e, nesse caso específico, habilidade, uma vez que os comentários

são a respeito dos *commits*, que por sua vez são resultado das habilidades do usuário. Por fim, indícios como delegação de tarefas, aceitação de conhecimento, colaboração, dependabilidade e habilidade podem ser extraídos a partir do estado da *pull request*, das atribuições e da participação nas conversas.

Dependabilidade, habilidade, comunalidade e benevolência não são indícios da confiança, contudo, a partir deles podemos chegar a uma estimativa da confiabilidade, que é indício da confiança. A confiabilidade juntamente com outros indícios fornece uma estimativa da confiança entre os usuários. Essa estimativa é fornecida por meio de um grafo, no qual os nós representam os usuários e as arestas representam a existência de confiança. O peso das arestas, variando de 0 a 1, mostram a intensidade dessa confiança.

De posse do grafo com os valores de confiança é possível utilizá-lo para sugerir uma equipe que venha a ter um alto nível de confiança. Além disso, o fato do *framework* ser dinâmico (processamento dos comentários mais recentes atualiza automaticamente os valores de confiança) possibilita o monitoramento da confiança entre os membros, para que medidas sejam tomadas caso se perceba variações negativas na confiança da equipe.

A.2 Fórmulas

Partindo da Figura 1.1, tendo como caso de uso o GitHub, existem seis indícios de confiança, sendo que o indício confiabilidade é derivado dos quatro antecedentes da confiabilidade. Esses indícios são extraídos a partir de três tipos de dados. A Figura 1.2 substitui as técnicas de extração da Figura 1.1 pelas fórmulas utilizadas e que serão apresentadas nessa seção.

Para extrair um valor de mímica das conversas, calculamos quão similar é o vocabulário utilizado nas conversas. A similaridade do vocabulário de dois comentários foi calculada utilizando a similaridade do cosseno sobre a frequência das palavras. Na Equação A.1, $SC(c_1, c_2)$ é a similaridade de cosseno entre dois comentários, FV é o vetor de frequência de palavras dos comentários.

$$SC(c_1, c_2) = \frac{FV_1 \cdot FV_2}{\|FV_1\| \|FV_2\|} \quad (\text{A.1})$$

O valor de mímica do participante p_1 para o participante p_2 , $MM(p_1, p_2)$ é a média das similaridades entre os comentários de p_1 e os comentários de p_2 que o antecedem na mesma *pull request*. Na Equação A.2 PR_{12} é o conjunto de *pull requests* em que p_1 e p_2 interagem, C_1 é o conjunto de comentários de p_1 na *pull request pr* e C_{21} é o conjunto de comentários de p_2 que antecedem o comentário c_1 .

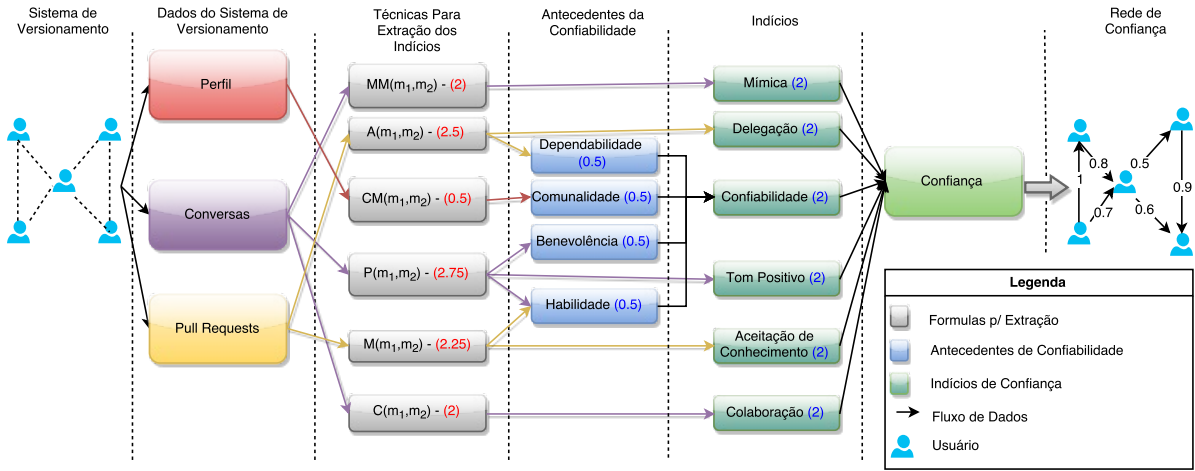


Figura 1.2: Fórmulas utilizadas para extração dos indícios no *framework*

$$MM(p_1, p_2) = \frac{\sum_{pr}^{PR_{12}} \sum_{c_1}^{C_1} \sum_{c_2}^{C_{21}} SC(c_1, c_2)}{\sum_{pr}^{PR_{12}} \sum_{c_1}^{C_1} \sum_{c_2}^{C_{21}} 1} \quad (\text{A.2})$$

A comunalidade é calculada pela média da similaridade de três atributos do perfil dos usuários do GitHub: (I) os usuários seguidos, (II) os projetos assistidos (projetos que o usuário acompanha) e (III) a localização. (I) e (II) são calculados utilizando a similaridade de jaccard (Equação A.3) onde C_1 e C_2 são os conjuntos avaliados. (III) por sua vez utiliza uma variante da distância euclidiana (Dodd et al., 2013, Equação A.4) sobre os valores dos índices de Geert Hofsted (Hofstede et al., 2010). Os índices de Geert Hofsted caracterizam a cultura de uma região, ao todo são seis índices: distância do poder, individualismo, masculinidade, aversão à incerteza, orientação de longo prazo e indulgência. Nessa equação L_i é o local, K é a quantidade de índices, I_{ik} é o valor do índice k para um local L_i e V_k é a variância do índice k . Caso os índices não existam para um determinado país, será considerada a maior distância possível.

$$SJ(C_1, C_2) = \frac{C_1 \cap C_2}{C_1 \cup C_2} \quad (\text{A.3})$$

$$DE(L_1, L_2) = \sqrt{\sum_{k=1}^K \frac{(I_{1k} - I_{2k})^2}{V_k}} \quad (\text{A.4})$$

Portanto, a comunalidade entre os participante p_1 e p_2 , $CM(p_1, p_2)$ é dada pela Equação A.5, onde F_{p_i} , W_{p_i} e L_i são respectivamente o conjunto dos usuários seguidos, conjunto dos projetos assistidos e a localização do participante p_i .

$$CM(p_1, p_2) = \frac{SJ(F_{p_1}, F_{p_2}) + SJ(W_{p_1}, W_{p_2}) + \frac{1}{1+DE(L_1, L_2)}}{3} \quad (\text{A.5})$$

Benevolência, habilidade e tom positivo podem ser extraídos a partir da polaridade dos comentários. O valor de polaridade entre dois participantes é dado pela Equação A.6, onde $P(p_1, p_2)$ é o valor da polaridade do participante p_1 para o participante p_2 , C_{+12} é a quantidade de comentários com polaridade positiva de p_1 para p_2 e C_{12} é o total de comentários de p_1 para p_2 .

$$P(p_1, p_2) = \frac{C_{+12}}{C_{12}} \quad (\text{A.6})$$

Dependabilidade e delegação podem ser observados por meio das atribuições de um participante por outro em uma *pull request*. O valor para as atribuições de um participante p_1 para um participante p_2 é 1 se p_1 atribuiu pelo menos uma *pull request* a p_2 ou 0 caso contrário. A Equação A.7 calcula o valor de atribuição de p_1 para p_2 , $A(p_1, p_2)$, com base na quantidade de *pull requests* atribuídas a p_2 por p_1 , PR_{s12} .

$$A(p_1, p_2) = \begin{cases} 1, & \text{se } PR_{s12} \geq 1 \\ 0, & \text{caso contrario} \end{cases} \quad (\text{A.7})$$

Aceitação de conhecimento e também a habilidade de um participante podem ser observados a partir da quantidade de *pull requests* que sofreram *merge*. Na Equação A.8, $M(p_1, p_2)$ é a proporção de *pull requests* criadas por p_2 que sofreram *merge* e que tiveram a interação de p_1 , PR_{a12} é a quantidade de *pull request* criadas por p_2 que sofreram *merge* e que tiveram a interação de p_1 , PR_{12} é a quantidade de *pull request* criadas por p_2 e que tiveram a interação de p_1 .

$$M(p_1, p_2) = \frac{PR_{a12}}{PR_{12}} \quad (\text{A.8})$$

O último indício, colaboração, é dado pela Equação A.9. Nessa equação $C(p_1, p_2)$ é proporção de interações de p_1 com p_2 em relação ao total de interações de p_1 . I_{12} é a quantidade de interações de p_1 com p_2 e I_1 é a quantidade de interações de p_1 .

$$C(p_1, p_2) = \frac{I_{12}}{I_1} \quad (\text{A.9})$$

Por fim, para calcular a probabilidade de existir confiança entre os participantes utilizamos a Equação A.10, a média ponderada das fórmulas apresentadas anteriormente. A melhor forma para definir os pesos α_i seria utilizar dados de projetos anteriores para aprender os pesos. Contudo, não temos conhecimento de nenhuma base com informações de confiança entre membros em sistemas de versionamento, então definimos os pesos com base na utilização das fórmulas pelos indícios. Por exemplo, considerando que cada indício da Figura 1.1 tem o mesmo peso, no caso 2, os antecedentes de confiabilidade terão peso 0.5 (derivado do peso 2 do indício confiabilidade, uma vez que cada antecedente também tem o mesmo peso para o cálculo da confiabilidade). A antecedente habilidade pode ser obtida da análise dos *merges* e da polaridade das conversas, cada um sendo responsável por metade do valor da habilidade (0.25). Ao somar todas as contribuições da polaridade das conversas temos 0.25 vindos da habilidade, 0.5 da benevolência e 2 do indício tom positivo que resulta em um peso 2.75. Portanto definimos que $\alpha_1 = 2$, $\alpha_2 = 0.5$, $\alpha_3 = 2.75$, $\alpha_4 = 2.5$, $\alpha_5 = 2.25$ e $\alpha_6 = 2$.

Na Figura 1.2 os números entre parênteses de cor azul são os pesos dos indícios e antecedentes de confiabilidade. Já os de cor vermelha são os valores α obtidos através da soma dos pesos dos indícios que os utilizam diretamente.

$$T(p_1, p_2) = \frac{\alpha_1 MM(p_1, p_2) + \alpha_2 CM(p_1, p_2) + \alpha_3 P(p_1, p_2)}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6} + \frac{\alpha_4 A(p_1, p_2) + \alpha_5 M(p_1, p_2) + \alpha_6 C(p_1, p_2)}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6} \quad (\text{A.10})$$

B

Questionário - Survey

Validação de indícios e fórmulas do Trust Framework

Esta survey tem como objetivo avaliar os indícios de confiança e as formulas utilizadas pelo framework de confiança desenvolvido pelo aluno de mestrado em ciência da computação da Universidade Estadual de Maringá – UEM, Guilherme Augusto Maldonado da Cruz, orientado pelas professoras Dra. Valéria Delisandra Feltrim e Dra. Elisa Hatsue Moriya Huzita.

O framework de confiança foi desenvolvido para identificar a presença de confiança entre membros de equipes de desenvolvimento distribuído de software utilizando-se dos dados nos sistemas de versionamento utilizados por estas equipes (utilizamos como estudo e caso o GitHub). A partir das informações de confiança fornecidas pelo framework podemos criar novas equipes com membros que já confiem uns nos outros, ou até mesmo, monitorar a confiança entre os membros de equipes existentes uma vez que o estabelecimento da confiança nessas equipes é um desafio.

Este questionário é constituído de três seções a partir desta: a seção de caracterização do participante (7 questões), a seção de validação dos indícios (3 questões) e a seção de validação das fórmulas (4 questões). Para responder as questões o participante deve ler o documento disponibilizado abaixo com a descrição do framework, indícios e fórmulas utilizadas. Todas as informações coletadas por este questionário serão confidenciais e serão utilizadas somente para validação do framework. Os resultados obtidos serão apresentados na dissertação sem qualquer informação que ligue as repostas aos participantes.

O questionário tem uma duração aproximada de 25 minutos e deve ser respondido até o dia 15 de novembro de 2015. Em caso de duvidas envie um e-mail para guilherme.maldonado.cruz@gmail.com.

Ao clicar em CONTINUAR declaro estar ciente de que os resultados coletados a meu respeito serão confidenciais e de que não receberei nenhum tipo de ônus pela participação, com exceção ao aprendizado de novas técnicas/tecnologias.

DESCRIÇÃO DO FRAMEWORK: <https://drive.google.com/file/d/0B1hL5LhDbGnCNXpzT1BQZWpqMTg/view?usp=sharing>

PRÓXIMA

Nunca envie senhas pelo Formulários Google.

Figura 2.1: Questionário - Parte 1

Validação de indícios e fórmulas do Trust Framework

***Obrigatório**

Caracterização do Participante

Esta seção contém questões a respeito do participante, como nível de estudo e conhecimento nas áreas utilizadas pelo framework

Nome *

Sua resposta _____

Nível de Estudo *

Graduação

Mestrado

Doutorado

Qual é a sua instituição/empresa atual? *

Caso não esteja ligado a alguma instituição/empresa preencha o campo com NENHUMA

Sua resposta _____

Qual seu cargo atual? *

Se escolher a opção outros informe o cargo no campo desta opção

Estudante

Desenvolvedor

Gerente de Projetos

Outro: _____

Figura 2.2: Questionário - Parte 2

Qual o seu conhecimento e experiência com desenvolvimento global de software (DGS)? *

- Nenhum
- Básico (somente conhecimento teórico)
- Médio (já trabalhei com DGS em não mais do que três projetos diferentes)
- Alto (eu já trabalhei com DGS em mais de três projetos diferentes)

Qual o seu conhecimento e experiência com o GitHub? *

- Nenhum
- Básico (somente conhecimento teórico)
- Médio (já utilizei o GitHub em não mais do que três projetos diferentes)
- Alto (eu utilizei o GitHub em mais de três projetos diferentes)

Qual o seu conhecimento e experiência com o Análise de Sentimentos/Mineração de Opinião? *

- Nenhum
- Básico (somente conhecimento teórico)
- Médio (já utilizei em não mais do que três projetos diferentes)
- Alto (eu utilizei em mais de três projetos diferentes)

VOLTAR

PRÓXIMA

Nunca envie senhas pelo Formulários Google.

Figura 2.3: Questionário - Parte 3

Validação de indícios e fórmulas do Trust Framework

*Obrigatório

Validação dos Indícios

Esta seção contém questões a respeito dos indícios de confiança utilizados pelo framework.

Antes de responder as perguntas desta seção leia o arquivo com a descrição do framework (<https://drive.google.com/file/d/0B1hL5LhDbGnCNXpzT1BQZWPqMTg/view?usp=sharing>).

Framework de Confiança



Importancia dos Indícios *

Para cada indício (itens verdes na figura acima) quão importante ele é para o cálculo da probabilidade de confiança?

	0 - Nenhuma	1	2	3	4	5
Mímica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Confiabilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tom Positivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Delegação	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aceitação de Conhecimento	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Colaboração	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 2.4: Questionário - Parte 4

Importancia dos Antecedentes de Confiabilidade *

Para cada antecedente da confiabilidade (itens azuis na figura acima), quão importante ele é ao determinar a confiabilidade de alguém?

	0 - Nenhuma	1	2	3	4	5
Habilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Benevolência	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comunalidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dependabilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

< >

Além dos indícios levantados, você consideraria importante mais algum outro indicio que possa ser extraído de um sistema de versionamento? Qual (quais)? Por que?

Sua resposta

VOLTAR

PRÓXIMA

Nunca envie senhas pelo Formulários Google.

Figura 2.5: Questionário - Parte 5

Validação de indícios e fórmulas do Trust Framework

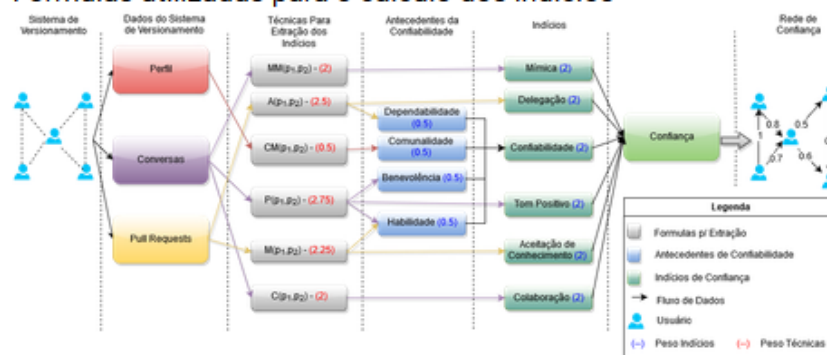
*Obrigatório

Validação das Fórmulas

Esta seção contém questões a respeito das fórmulas utilizados pelo framework para analisar os dados em busca dos indícios de confiança.

Antes de responder as perguntas desta seção leia o arquivo com a descrição do framework (<https://drive.google.com/file/d/0B1hL5LhDbGnCNXpzT1BQZWpqMTq/view?usp=sharing>).

Fórmulas utilizadas para o calculo dos indícios



Validade das formulas para o cálculo dos indícios *

Para cada indício, a formula utilizada para calcula-lo é válida e adequada?

	Não	Regular	Sim, boa o suficiente	Sim, perfeita
Mímica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Confiabilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tom Positivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Delegação	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aceitação de Conhecimento	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Colaboração	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 2.6: Questionário - Parte 6

Validade das formulas para o cálculo dos antecedentes de confiabilidade *

Para cada antecedente da confiabilidade, a formula utilizada para calcula-lo é válida e adequada?

	Não	Regular	Sim, o suficiente	Sim, perfeita
Habilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Benevolência	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comunalidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dependabilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fórmula p/ cálculo da confiança

$$T(p_1, p_2) = \frac{\alpha_1 M(p_1, p_2) + \alpha_2 CM(p_1, p_2) + \alpha_3 P(p_1, p_2)}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6} + \frac{\alpha_4 A(p_1, p_2) + \alpha_5 M(p_1, p_2) + \alpha_6 C(p_1, p_2)}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6}$$

A formula para o cálculo da confiança é valida e adequada? *

- Não
- Regular
- Sim, o suficiente
- Sim, perfeita

Suções/Criticas

Você tem alguma sugestão de adequação/critica a respeito das formulas utilizadas?

Sua resposta

VOLTAR

ENVIAR

Figura 2.7: Questionário - Parte 7