

The page features a decorative graphic consisting of three blue circles of varying sizes, each composed of concentric rings of different shades of blue. These circles are arranged in a vertical line, with the largest at the top, a medium one in the middle, and the largest at the bottom. Two thin, light blue lines intersect at the top left and extend diagonally across the page, framing the circles and the text.

EvolTrack-SocialNetwork: Planejamento da Avaliação Estudo Exploratório

**Andréa Magalhães
14/11/2011**

SUMÁRIO

1. PLANEJAMENTO DO ESTUDO EXPLORATÓRIO	3
1.1. Definição do Estudo	3
1.1.1. <i>Objetivo Global</i>	3
1.1.2. <i>Objetivo do Estudo</i>	3
1.1.3. <i>Questões</i>	3
1.2. Planejamento do Estudo	4
1.2.1. <i>Definição de Hipóteses</i>	4
1.2.2. <i>Contexto</i>	5
1.2.3. <i>Variáveis</i>	7
1.2.4. <i>Instrumentação e Preparação</i>	8
1.2.5. <i>Validade do Estudo</i>	9
1.3. Validade do Planejamento	11
REFERÊNCIAS BIBLIOGRÁFICAS.....	12
APÊNDICE I – TERMO DE CONSENTIMENTO LIVRE ESCLARECIDO	14
APÊNDICE II – FORMULÁRIO DE CARACTERIZAÇÃO DO PARTICIPANTE .	16
APÊNDICE III – FORMULÁRIO PARA REALIZAÇÃO DO ESTUDO	18
APÊNDICE IV – QUESTIONÁRIO DE AVALIAÇÃO DO ESTUDO	21
APÊNDICE V – ORÁCULO	24

1. Planejamento do Estudo Exploratório

Esta seção descreve o planejamento da avaliação da EvolTrack-SocialNetwork através de um estudo exploratório inicial. Para a elaboração deste planejamento foram utilizados como exemplo trabalhos do Grupo de Reutilização de Software¹ (OLIVEIRA, 2011, VASCONCELOS, 2007) e dos outros grupos de pesquisa de Engenharia de Software da COPPE/UFRJ (BARRETO e ROCHA, 2009, BARRETO, 2011, TRAVASSOS *et al.*, 2002).

1.1. Definição do Estudo

A definição do estudo deve declarar claramente o que se pretende avaliar e qual é o escopo da avaliação.

1.1.1. Objetivo Global

O propósito desse estudo é avaliar a ferramenta de visualização e análise de redes sociais EvolTrack-SocialNetwork em relação ao apoio oferecido à percepção da colaboração durante o acompanhamento de projetos de desenvolvimento de software.

1.1.2. Objetivo do Estudo

O objetivo do estudo é apresentado de acordo com o paradigma GQM (*Goal Question Metric*) (BASILI *et al.*, 1994), como descrito a seguir.

Analisar a ferramenta de visualização e análise de redes sociais EvolTrack-SocialNetwork

Com o propósito de caracterizar

Com respeito à precisão e cobertura

Do ponto de vista do pesquisador da área de colaboração

No contexto de projetos de desenvolvimento de software

1.1.3. Questões

A fim de avaliar os resultados, o paradigma GQM (BASILI *et al.*, 1994) é adotado. Assim, esta seção apresenta as questões e métricas definidas.

Q1: Os participantes são capazes de perceber a colaboração no projeto de desenvolvimento de software?

Esta percepção da colaboração será medida através das respostas dadas pelos participantes às tarefas do estudo. Logo, as seguintes métricas foram definidas:

M1: Precisão

¹ Site: <http://reuse.cos.ufrj.br>

A precisão, que é um indicador de exatidão, se refere ao número de respostas corretas dadas em relação ao total de respostas dadas. O seu cálculo é feito através da seguinte fórmula:

$$\text{Precisão} = \frac{\text{número de respostas corretas dadas}}{\text{total de respostas dadas}}$$

M2: Cobertura

A cobertura, que é um indicador de eficácia e completude, se refere ao número de respostas corretas dadas em relação ao total de respostas corretas esperadas. O seu cálculo é feito através da seguinte fórmula:

$$\text{Cobertura} = \frac{\text{número de respostas corretas dadas}}{\text{total de respostas corretas esperadas}}$$

M3: Média Harmônica

A fim de verificar se os valores de precisão e cobertura obtidos são razoáveis, foi utilizada uma média harmônica², que estabelece uma relação de compromisso entre os valores de precisão e cobertura. Uma vez que essas duas medidas são conflitantes entre si, é recomendado o uso da média harmônica para obter uma média balanceada entre precisão e cobertura. Quanto mais o valor da média harmônica se aproxima de 1, melhor é essa relação de compromisso. A média harmônica é calculada da seguinte forma:

$$\text{Média Harmônica} = \frac{2 * \text{Precisão} * \text{Cobertura}}{\text{Precisão} + \text{Cobertura}}$$

1.2. Planejamento do Estudo

O planejamento apresenta o projeto do experimento, ou seja, a maneira como o experimento será conduzido. Neste momento, também, é definida a instrumentação a ser utilizada e são analisadas as ameaças à validade do estudo.

1.2.1. Definição de Hipóteses

Um experimento geralmente é formulado através de hipóteses. A hipótese principal se chama hipótese nula e declara que não há nenhum relacionamento estatisticamente significativo entre a causa e o efeito. O objetivo principal do experimento é então rejeitar a hipótese nula a favor de uma ou algumas hipóteses alternativas. A decisão sobre a rejeição de uma hipótese

² A média harmônica costuma ser particularmente recomendada para análise de valores que sejam inversamente proporcionais.

nula pode ser tomada com base nos resultados de sua verificação utilizando um teste estatístico (TRAVASSOS *et al.*, 2002).

Neste estudo, a seguinte hipótese foi definida:

- **Hipótese Nula (H0):** Não é possível compreender como a colaboração acontece em um projeto de desenvolvimento de software, utilizando a ferramenta EvolTrack-SocialNetwork.

Hipótese Alternativa (H1): É possível compreender como a colaboração acontece em um projeto de desenvolvimento de software, utilizando a ferramenta EvolTrack-SocialNetwork.

1.2.2. Contexto

O contexto do experimento descreve as condições em que o experimento está sendo executado (TRAVASSOS *et al.*, 2002).

a) Participantes

Os participantes do estudo, selecionados por conveniência, são alunos e ex-alunos de pós-graduação da Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Como trata-se de um estudo exploratório inicial, foram selecionados como participantes pesquisadores da área de colaboração para capturar a visão de especialistas da área.

As sessões do estudo serão individuais. Todos os participantes irão utilizar a ferramenta. Não haverá nenhum tipo de compensação para os participantes.

b) Tarefas

Neste estudo foi definido um conjunto de tarefas a serem realizadas, visando caracterizar o apoio à percepção da colaboração. As tarefas propostas exploram tanto a visualização quanto a análise de redes sociais. A visualização de redes sociais adota a representação visual das informações, através de diferentes formas, para diminuir a sobrecarga cognitiva do usuário e facilitar a compreensão e exploração dos dados através de imagens. A análise de redes sociais é uma forma de compreender a interação e a organização social de um grupo, através das propriedades (como centralidade, densidade e etc.) das redes sociais.

Estas tarefas foram definidas considerando os diferentes tipos de percepção (ARAUJO, 2000, DOURISH e BELLOTTI, 1992) que devem ser oferecidos a um participante.

As tarefas são divididas em três grupos, com diferentes níveis de complexidade, baseados no trabalho de Oliveira (2011):

- **Tarefas de Filtragem (TF):** Este grupo consiste de tarefas mais simples, de forma que, se algum participante não estiver apto a realizá-las, o mesmo deve ser desconsiderado na análise dos resultados (por indicar possíveis problemas no entendimento da ferramenta ou das tarefas do experimento). Exemplos destas tarefas são:
 - Visualização
 - Quais atores compõem a rede social?
 - Análise
 - Qual é a densidade da rede sócio-técnica?

- **Tarefas Básicas (TB):** Este grupo consiste de tarefas básicas, que podem ser realizadas por meio de fatos extraídos da visualização. Neste grupo de tarefas, o aspecto da percepção é tratado. Exemplos destas tarefas são:
 - Visualização
 - O que mudou na rede social da versão 1 para a versão 2?
 - Com quem João pode estar em conflito?
 - Análise
 - Qual é a distância (em número de arestas) entre a classe X e a classe Y na rede técnica?
 - Quem tem a maior centralidade de grau na rede sócio-técnica?

- **Tarefas de Assimilação (TA):** Este grupo consiste de tarefas mais difíceis e complexas, exigindo que o participante utilize o conhecimento prévio e o raciocínio de forma a obter o entendimento e interpretar as informações contidas nos instrumentos da avaliação. Exemplos destas tarefas são:
 - Visualização
 - Qual artefato tem indícios de que precisa ser refatorado?
 - Para quem você passaria a tarefa de trabalhar no artefato Z? Por que?
 - Com quem João precisa se comunicar para coordenar o trabalho?
 - Análise
 - De acordo com as características abaixo, como você classificaria a coordenação do projeto?

Os participantes não serão comunicados da existência e do objetivo desta divisão de tarefas em grupos, de forma a não influenciar a execução das atividades.

O oráculo será gerado com base nas respostas de especialistas no domínio e validado por meio das respostas dadas pelos participantes. O oráculo é apresentado no Apêndice V.

c) Dados Utilizados

Este estudo exploratório irá utilizar dados de projetos de desenvolvimento de software livre. A escolha do ambiente de projetos de software livre para a realização dos estudos iniciais se deve ao fato de que estes projetos publicam os seus processos de desenvolvimento, códigos, artefatos e dados sobre os produtos livremente na Internet. Portanto, fornecem uma oportunidade única para as experimentações em Engenharia de Software, devido à diversidade, complexidade e representatividade de dados disponíveis em uma escala global.

Este estudo foca em redes sociais obtidas a partir da mineração das relações existentes implicitamente entre os desenvolvedores de software por meio das atualizações de código fonte realizadas pelos desenvolvedores de software. Assume-se que estas relações podem ser uma fonte de informação adequada para entender a coordenação do grupo (SANTOS *et al.*, 2011). A partir do código-fonte é possível verificar quais atividades são ou foram realizadas pelos membros do grupo de trabalho e assim verificar o andamento do projeto.

Em particular, este estudo irá utilizar o projeto de software livre Torrent Episode Downloader³ extraído do portal SourceForge. O projeto foi selecionado levando-se em conta o fato de ter sido desenvolvido em Java para satisfazer às restrições tecnológicas da ferramenta. Também foi selecionado um projeto com pelo menos dois desenvolvedores. Projetos com somente um desenvolvedor foram descartados, pois não seria possível realizar a análise das propriedades de redes sociais visto que nenhuma relação seria estabelecida entre outros membros. Os dados minerados desse projeto serão usados para apresentar as redes técnica, sócio-técnica e social.

1.2.3. Variáveis

Há dois tipos de variáveis do experimento: independentes e dependentes. As variáveis independentes referem-se à entrada do processo de experimentação. Essas variáveis apresentam a causa que afeta o resultado do processo de experimentação. O objetivo das variáveis independentes é identificar, no contexto de execução do estudo, os pontos que exercem ou podem exercer alguma interferência sobre os resultados obtidos.

³ <http://sourceforge.net/projects/ted/>

As variáveis dependentes referem-se à saída do processo de experimentação (TRAVASSOS *et al.*, 2002), ou seja, são aquelas que realmente queremos avaliar no estudo de caso. Estas variáveis são definidas de acordo com os objetivos e as questões que se pretende responder com o estudo.

A variável independente no estudo é o nível de conhecimento em relação à colaboração e redes sociais. Já as variáveis dependentes são: precisão e cobertura.

1.2.4. Instrumentação e Preparação

Esta seção define quais instrumentos devem ser aplicados durante a avaliação e como a avaliação será preparada.

Para este estudo, foram projetados cinco instrumentos principais que serão entregues impressos aos participantes:

- Termo de consentimento livre e esclarecido (ver Apêndice I);
- Formulário de caracterização do participante (ver Apêndice II);
- Formulário para realização do estudo (ver Apêndice III);
- Questionário para avaliação do estudo (ver Apêndice IV);
- Material de treinamento.

O primeiro instrumento é o termo de consentimento livre e esclarecido, que declara ao participante o objetivo do estudo, os limites da sua participação e suas responsabilidades durante a avaliação. O termo também informa que os dados da avaliação não estão sujeitos a serem utilizados para classificar o desempenho dos participantes e garante a confidencialidade dos dados. Este termo deve ser assinado pelo participante antes do início do estudo.

Em seguida, o participante será solicitado a preencher um formulário de caracterização que permite ao pesquisador determinar o seu perfil e auxilia na análise posterior dos dados obtidos por meio do estudo.

Ainda antes de iniciar o estudo, será feito um breve treinamento com os participantes visando explicar os principais conceitos de redes sociais e explicar a utilização da ferramenta.

De forma a facilitar a imersão do participante no contexto do estudo, é criada uma situação fictícia, na qual o participante é colocado na posição de gerente de um projeto de desenvolvimento de software cujo objetivo é criar um produto inovador. Assim, o formulário para a realização do estudo, apresenta ao participante esse contexto de trabalho e as instruções para a realização das tarefas. Este formulário vem acompanhado

da apresentação dos dados na ferramenta. Este conjunto de informações deve ser utilizado pelo participante ao longo de toda a execução do estudo. Por fim, este formulário também é o instrumento para a coleta dos resultados da tarefa.

Finalmente, após a execução da tarefa, é entregue ao participante um questionário para avaliação do estudo, que pretende obter informações qualitativas acerca do estudo, do ponto de vista do participante. Este questionário inclui sugestões para a ferramenta EvolTrack-SocialNetwork e considerações sobre o procedimento de avaliação.

1.2.5. Validade do Estudo

É comum que haja questões que possam impactar ou limitar a validade dos resultados dos estudos. Estas questões são denominadas ameaças à validade. Há quatro tipos de validade dos resultados do experimento: validade de conclusão, validade interna, validade de construção e validade externa (TRAVASSOS *et al.*, 2002, WOHLIN *et al.*, 1999).

A *validade interna* define se o relacionamento observado entre o tratamento e o resultado é causal, e não é o resultado da influência de outro fator que não é controlado ou mesmo que não foi medido. Durante a avaliação da validade interna, uma maior atenção deve ser prestada aos participantes, ou seja, à seleção da população, à maneira da divisão nas classes, ao modo da aplicação dos tratamentos e aos aspectos sociais.

A *validade externa* define as condições que limitam a habilidade de generalizar os resultados do estudo para outros contextos fora do ambiente avaliado. Durante a avaliação da validade externa, a interação do tratamento com as pessoas, o lugar e o tempo devem ser considerados.

A *validade de construção* considera os relacionamentos entre a teoria e a observação, ou seja, se o tratamento reflete bem a causa e o resultado reflete bem o efeito. Durante a avaliação da validade de construção, deve-se tomar cuidado com o comportamento incorreto do lado dos participantes ou do experimentador.

Por fim, a *validade de conclusão* é relacionada à habilidade de chegar a uma conclusão correta a respeito dos relacionamentos entre o tratamento e o resultado do experimento. Durante a avaliação da validade de conclusão é necessário considerar os conceitos como a escolha do teste estatístico, a escolha do tamanho do conjunto dos participantes e a confiabilidade das medidas.

As ameaças a validade identificadas para o presente estudo são:

Validade interna

- Caso o estudo não seja executado em um único dia por todos os participantes, um ponto que poderia influenciar o resultado do estudo seria a troca de informações

entre os participantes que já haviam realizado o estudo e os que ainda não o haviam realizado. Para evitar este problema, será requisitado explicitamente que os participantes não troquem informações a respeito do estudo;

- Os participantes do estudo são colegas de trabalho e membros do mesmo grupo de pesquisa do pesquisador. Dessa forma, pode haver uma tentativa de agradar o pesquisador. Para minimizar esse risco, tanto nos instrumentos utilizados no estudo como no treinamento realizado no início do estudo, buscou-se deixar claro que os participantes devem ser imparciais na sua análise;
- A própria ferramenta de apoio pode influenciar os resultados, caso os participantes enfrentem dificuldades não esperadas na realização do estudo (lentidão, erros no servidor, etc.). A interação com a ferramenta também pode influenciar na maneira como os participantes realizam a atividade. Assim, será ministrado um treinamento sobre a utilização da ferramenta e durante o estudo será observada a interação com a ferramenta para avaliar se algum fator não planejado em relação à ferramenta influencia na execução da atividade;
- O entendimento dos participantes sobre as questões dos formulários é diretamente influenciado pela forma como as questões foram elaboradas; se a questão tiver sido mal formulada, o estudo pode ser afetado negativamente (WOHLIN *et al.*, 1999). A análise dos instrumentos utilizados pelos outros pesquisadores envolvidos no estudo visa justamente reduzir esta interferência.

Validade externa

- Normalmente, ambientes acadêmicos não simulam totalmente as condições existentes em um ambiente de desenvolvimento de software. Em contrapartida, Carver *et al.* (2003) apontam que a realização de estudos experimentais tendo alunos como participantes permite obter uma série de benefícios para os pesquisadores (e.g., obtenção de evidências preliminares para confirmar ou rejeitar hipóteses e treinamento de jovens investigadores no campo da pesquisa empírica) e os estudantes (e.g., atividades práticas e a educação em tópicos que são alvo de pesquisas recentes);
- Uma ameaça à validade externa é o fato de o experimento considerar um único projeto de software livre. Assim, experimentos com uma quantidade maior ou com outros tipos de projetos devem ser executados;

- Não é possível representar todas as situações possíveis de um projeto de software. Apesar das tarefas do estudo terem sido elaboradas com base em situações comuns no contexto em que estão inseridas, é necessário verificar se os objetivos da ferramenta EvolTrack-SocialNetwork são atingidos em outras circunstâncias.

Validade de construção

- Uma possibilidade é que as medidas selecionadas não sejam bons indicadores da viabilidade da proposta do trabalho. Para diminuir esse risco, foram escolhidas medidas que visam capturar as informações necessárias para responder às questões do estudo;
- Há uma hierarquia entre os participantes do estudo (funcionários) e o responsável pela execução do estudo (chefe). A fim de minimizar os efeitos desta ameaça, tanto nos instrumentos utilizados no estudo como no treinamento realizado no início do estudo, buscou-se deixar claro que os resultados do estudo não iriam influenciar o resultado do desempenho dos participantes nas suas atividades. No entanto, não é possível garantir se os participantes executaram o estudo da mesma forma como o teriam executado em outro cenário;
- O agrupamento das tarefas por tipo auxilia a análise dos dados. No entanto, embora algumas destas tarefas possam ter grau de dificuldade maior do que o de outras tarefas, o mesmo peso foi atribuído a todas as tarefas. Isto pode influenciar os resultados. Devido à subjetividade na avaliação do grau de dificuldade (o que introduziria viés na análise dos dados), optou-se por manter esta configuração.

Validade de conclusão

- A principal ameaça à validade de conclusão neste estudo é o tamanho da amostra, com um número pequeno de participantes, não sendo o ideal do ponto de vista estatístico. Por isso, este estudo apresenta uma limitação nos resultados que serão considerados apenas como indícios. Além disso, será realizada apenas uma análise qualitativa dos resultados.

1.3. Validade do Planejamento

Conforme sugerido por Mafra e Travassos (2006), o plano do estudo e os instrumentos devem ser validados por outros pesquisadores antes de dar prosseguimento à execução do estudo. De

preferência, devem ser selecionados pesquisadores que não possuam interesse direto nos resultados do estudo, para minimizar a presença de viés.

Neste trabalho, este planejamento do estudo exploratório não será validado, pois trata-se de um estudo inicial apenas. Além disso, o presente estudo é que contribui para a validação do planejamento e dos instrumentos que serão aplicados posteriormente em um estudo de viabilidade mais completo. Este estudo exploratório também não inclui um piloto.

Referências Bibliográficas

- ARAUJO, R., 2000, *Ampliando a Cultura de Processos de Software: Um Enfoque Baseado em Groupware e Workflow*. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- BARRETO, A. O. S.; ROCHA, A. R. C. DA, 2009, "Definição e Monitoração de Objetivos de Software Alinhados ao Planejamento Estratégico". *Simpósio Brasileiro de Qualidade de Software (SBQS)*, p. 114-128, Ouro Preto, MG, Brasil.
- BARRETO, A. S., 2011, *Uma Abordagem para Definição de Processos baseada em Reutilização Visando à Alta Maturidade em Processos*. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- BASIL, V. R.; CALDIERA, G.; ROMBACH, H. D., 1994, "The goal question metric approach", *Encyclopedia of software engineering*, v. 1, p. 528-532.
- CARVER, J.; JACCHERI, L.; MORASCA, S. et al., 2003, "Issues in using students in empirical studies in software engineering education". *International Software Metrics Symposium*, p. 239- 249, Sydney, Australia.
- DOURISH, P.; BELLOTTI, V., 1992, "Awareness and coordination in shared workspaces". In: *Conference on Computer Supported Cooperative Work (CSCW)*, p. 107-114, Toronto, Ontario, Canada.
- MAFRA, S. N.; TRAVASSOS, G. H., 2006, *Estudos Primários e Estudos Secundários apoiando a Busca por Evidência em Engenharia de Software, Relatório Técnico ES-687/06, PESC-COPPE*.
- OLIVEIRA, M. S. DE, 2011, *PREViA: Uma Abordagem para a Visualização da Evolução de Modelos de Software*. Dissertação de Mestrado, COPPE/UFRJ, Rio de Janeiro.
- SANTOS, T. A. L.; ARAUJO, R. M.; MAGDALENO, A. M., 2011, "Bringing Out Collaboration in Software Development Social Networks". *International Conference on Product Focused Software Development and Process Improvement - Short Papers*, p. 18-21, Torre Canne, Italy.
- TRAVASSOS, G. H.; GUROV, D.; AMARAL, E. A. G., 2002, *Introdução à Engenharia de Software Experimental, Relatório Técnico ES-590/02, PESC-COPPE*.
- VASCONCELOS, A. P. V. DE, 2007, *Uma Abordagem de Apoio à Criação de Arquiteturas de Referência de Domínio baseada na Análise de Sistemas Legados*. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

WOHLIN, C.; RUNESON, P.; HÖST, M., 1999, *Experimentation in Software Engineering: An Introduction*. 1 ed. Springer.