

Anteprojeto de Tese de Doutorado

Sobre a Adoção de Boas Práticas de Desenvolvimento de Software em Projetos de Software Livre

Rodrigo Rocha Gomes e Souza
rodrigo@dcc.ufba.br

Dezembro de 2009

1 Objetivo

No contexto de desenvolvimento de software, boas práticas são estratégias ou atividades que, segundo mostra a experiência, contribuem para o sucesso de um projeto de software. Naturalmente, a adoção de uma boa prática pode ser dificultada pela falta de experiência dos desenvolvedores na aplicação da prática em condições específicas.

O principal objetivo desta pesquisa é entender como a adoção de boas práticas de desenvolvimento de software é afetada por variações na forma de organização da equipe de desenvolvedores e por características do projeto. Pretende-se assim identificar em quais contextos determinadas práticas são efetivamente aplicadas e entender os mecanismos que reforçam a adoção de práticas de desenvolvimento de software.

A análise será focada nas atividades de implementação, documentação e controle de qualidade. Serão investigadas, entre outras, as seguintes práticas:

- planejamento e implementação de testes antes da implementação das funcionalidades correspondentes;
- implementação de testes por pessoas diferentes daquelas que implementam funcionalidades;
- integração contínua com o repositório de código-fonte;
- correção de bugs antes da implementação de novas funcionalidades;
- revisão de código.

Crowston e Howison [1] recentemente mostraram que os padrões de interação entre desenvolvedores variam bastante de projeto para projeto. Alguns projetos são bastante centralizados em um conjunto pequeno de desenvolvedores que interagem com boa parte dos demais; outros possuem uma organização mais descentralizada, nas quais os desenvolvedores se organizam em grupos bem definidos. Não é claro, no entanto, como práticas de desenvolvimento se distribuem nessa rede social. O quanto variam as práticas entre diferentes grupos de um projeto?

Sabe-se há muito tempo que é alta a rotatividade de desenvolvedores em empresas de software, e pesquisas recentes mostram que o cenário também é comum em projetos de software livre [2]. Enquanto os processos praticados por uma empresa são mantidos graças à organização hierárquica de seus funcionários, vale a pena investigar como se dá essa manutenção em projetos de software livre, especialmente os mais descentralizados, face à saída e à entrada de desenvolvedores. Será que, na eventualidade da saída dos desenvolvedores mais produtivos de um projeto, os demais desenvolvedores perpetuam as práticas aplicadas por seus antecessores? Em que casos a saída de desenvolvedores pode causar o abandono de práticas em um projeto?

Pretende-se também estudar diferenças entre as práticas adotadas por projetos com características distintas. Características de projetos a serem estudadas incluem grau de sucesso, grau de centralização, número de desenvolvedores e tamanho do software. Esse estudo poderá levar a conclusões do tipo “a prática X não é escalável: ela raramente é usada em projetos de grande porte”.

2 Motivação

A aplicação de boas práticas de desenvolvimento está associado à melhoria na qualidade do software produzido [3]. A difusão de boas práticas dentro de um projeto é, ainda, uma forma de diminuir a discrepância entre desenvolvedores menos experientes e desenvolvedores mais experientes, o que contribui para tornar o processo de desenvolvimento de software mais previsível.

Existe, no entanto, uma lacuna entre a teoria e a prática do desenvolvimento de software [2]. Em particular, a difusão de uma boa prática depende não apenas dos benefícios preconizados por seus entusiastas, mas também de questões ligadas ao seu uso efetivo, como o custo de implantação e a escassez de boas ferramentas de apoio.

Por melhor que seja a argumentação, uma prática só é útil quando é efetivamente adotada no desenvolvimento de um software. Por isso, a compreensão dos fatores que influenciam essa adoção é uma parte importante do estudo de uma prática de desenvolvimento. Infelizmente, a maior parte do conhecimento disponível sobre tais fatores é proveniente de observações isoladas e estudos de caso [4]. Essas observações, embora de grande valia, são essencialmente tendenciosas. Uma compreensão mais completa dos fatores contribuiria para a proposta de práticas de desenvolvimento mais realistas.

3 Metodologia

Muitos projetos de software livre tornam disponíveis publicamente alguns artefatos produzidos durante o processo de desenvolvimento de software. Esses artefatos incluem código-fonte e documentação mantidos em sistemas de controle de versão, bem como relatórios de bugs e mensagens de e-mail trocadas entre desenvolvedores.

A grande disponibilidade de dados relacionados ao desenvolvimento favorece a realização de estudos em larga escala. Nesta pesquisa, práticas de desenvolvimento serão extraídas automaticamente a partir do histórico de desenvolvimento de cada projeto através de técnicas de mineração de processos.

ao longo do histórico de desenvolvimento de cada projeto sob estudo at
usadas no desenvolvimento de software

Através de técnicas de mineração de processos é possível

Técnicas de mineração de processos e estatística descritiva viabilizam a realização desta pesquisa com grandes amostras de projetos de software.

Para o estudo abrangente será selecionada uma amostra significativa de projetos de software livre hospedados em um repositório de código-fonte a ser escolhido, como o SourceForge. A amostra será estratificada de acordo com aspectos como idade do projeto, número de desenvolvedores e número de linhas de código-fonte.

Serão analisados relatórios de bugs registrados em um sistema de gerenciamento de bugs, mensagens de e-mail arquivadas em listas de discussão e arquivos mantidos sob controle de versão, como documentação e código-fonte. Ferramentas para a extração desses dados já foram usadas em pesquisas anteriores [], com bons resultados.

A análise dos dados no estudo mais abrangente será feita de forma automática através de técnicas de mineração de processos. Essa abordagem foi escolhida por tornar viável a análise de grandes volumes de dados. A mineração de processos tem como finalidade de extrair modelos de processos a partir de registros de eventos. Eventos que poderão ser extraídos dos dados coletados incluem a modificação de um arquivo fonte, o registro da resolução de um bug e o envio de uma mensagem à lista de discussão do projeto.

Inicialmente será feita uma análise exploratória sobre uma amostra reduzida de projetos de software com os métodos e ferramentas escolhidos. Essa análise terá por objetivo investigar como os dados devem ser pré-processados e identificar limitações da abordagem.

A amostra reduzida de projetos de software também será alvo de um estudo mais detalhado. Informações adicionais sobre o processo desses projetos serão coletadas através de questionários e entrevistas com desenvolvedores. Quando for possível, as informações fornecidas pelos desenvolvedores serão confrontadas com as informações extraídas automaticamente.

Atividade	2010.1	2010.2	2011.1	2011.2	2012.1	2012.2	2013.1	2013.2
??								
??								
??				•		•		
??								
??								
??						•	•	•
??								•

Tabela 1: Proposta de Cronograma de Atividades

4 Proposta de Cronograma

Referências

- [1] Kevin Crowston and James Howison. The social structure of free and open source software development. *First Monday*, 10(2), 2005.
- [2] Robert L. Glass. The relationship between theory and practice in software engineering. *Commun. ACM*, 39(11):11–13, 1996.