

Laboratório 03

Eric Guimarães Caldas Jardim, Lucas Picinin Campos Lutti, Matheus Fontes

Almeida Moreira Silva, Nando Augusto Veloso Tupinambá

¹Instituto de Ciências Exatas e Informática
Pontifícia Universidade de Minas Gerais (PUC Minas)
Belo Horizonte – MG – Brasil

{1349194, 1254051, 1322257, 1337014}@sga.pucminas.br

Resumo. *O trabalho desenvolvido consiste em construir um dataset contendo a análise de pull request de repositórios mais populares no GitHub, sendo que estes devem possuir pelo menos 100 PRs (Merged ou Closed). Nesse sentido, a terceira tarefa do Laboratório de Medição e Experimentação de Software tem como objetivo analisar a atividade de code review no meio dos principais projetos presentes no Github, trazendo um olhar crítico de todos os aspectos que possam influenciar no merge de um pull request.*

1. Introdução

É fato que a atividade de code review é extremamente importante durante a confecção de um projeto e tem se tornado cada vez mais presente nos dias atuais, tendo um aumento exponencial em seu uso de forma concomitante com o crescimento dos processos de desenvolvimentos ágeis. Para melhor exemplificar, essa atividade é a inspeção do código de um possível contribuidor por parte do desenvolvedor do projeto, onde o mesmo analisa a solicitação recém-criada e define se a mesma pode ser anexada ao programa principal. Nesse sentido, podem ser observados diversos sistemas famosos que são open source, ou seja, permitem que usuários externos façam suas contribuições e solicitem um pull request, como Linux, Mozilla Firefox, Python, entre outros.

Deste modo, foi proposta uma atividade para que, baseado em algumas métricas pré-definidas, o grupo pudesse analisar as tendências de PRs nos principais repositórios do GitHub, podendo ter um conhecimento maior do que faz um código ser anexado ao programa principal (merged) ou rejeitado (closed).

Para análise dos repositórios extraídos, perguntas norteadoras foram definidas para que houvesse um parâmetro de pesquisa, tal como a métrica utilizada para resposta de cada uma:

- RQ 01. Qual a relação entre o tamanho dos PRs e o feedback final das revisões?

Métrica - Tamanho do pull request (Linhas de código e número de arquivos)

Hipótese - Pull requests maiores tendem a ser mais sólidos e contribuem de maneira mais direta para um programa principal.

- RQ 02. Qual a relação entre o tempo de análise dos PRs e o feedback final das revisões?

Métrica - Intervalo entre a criação do PR e o merged ou closed

Hipótese - Em sua maioria, pull request que possuem um maior tempo de análise tendem a ter um melhor feedback final

- RQ 03. Qual a relação entre a descrição dos PRs e o feedback final das revisões?

Métrica - Número de caracteres do corpo de descrição do PR

Hipótese - Quanto maior a descrição do pull request, maior será o entendimento do responsável por aprovar ou não o mesmo e, deste modo, o merged pode ocorrer de forma majoritária.

- RQ 04. Qual a relação entre as interações nos PRs e o feedback final das revisões?

Métrica - Número de participantes e números de comentários

Hipótese 1 - Quanto mais participantes, maior a chance do pull request ficar desconexo e sem serventia para o programa principal.

Hipótese 2 - As linhas de comentários não interferem na qualidade do PR.

- RQ 05. Qual a relação entre o tamanho dos PRs e o número de revisões realizadas?

Métrica - Tamanho do pull request (Linhas de código e número de arquivos)

Hipótese - Quanto maior o PR, maior o número de revisões

- RQ 06. Qual a relação entre o tempo de análise dos PRs e o número de revisões realizadas?

Métrica - Intervalo entre a criação do PR e o merged ou closed

Hipótese - Quanto maior o tempo de análise, maior o número de revisões realizadas

- RQ 07. Qual a relação entre a descrição dos PRs e o número de revisões realizadas?

Métrica - Número de caracteres do corpo de descrição do PR

Hipótese - Quanto maior a descrição, menor o número de revisões, uma vez que o entendimento é facilitado

- RQ 08. Qual a relação entre as interações nos PRs e o número de revisões realizadas?

Métrica - Número de participantes e números de comentários

Hipótese - Quanto maior o número de participantes, maior o número de revisões

2. Metodologia (neste tópico deve ficar claro COMO foi realizado o seu trabalho)

3. Resultados

4. Conclusões e trabalhos futuros

Referências