

Protocolo da Pesquisa

Mauricio Aniche (6892132)

1. Projeto do Estudo

O objetivo desta pesquisa é descobrir qual a relação entre maus cheiros de código de teste e possíveis problemas no código de produção. Para alcançar o objetivo, este estudo propõe a comparação de métricas de qualidade do código de produção com métricas de qualidade de códigos de teste.

Para cada métrica de qualidade de código de teste, o código de produção será separado em dois grupos: o grupo dos códigos de produção cujos testes automatizados **apresentam** determinado mau cheiro, e o grupo dos códigos de produção cujos testes automatizados **não apresentam** aquele mau cheiro.

Em seguida, por meio de testes estatísticos, este estudo discute se a diferença entre os valores das métricas de ambos os grupos é significativa.

2. Hipóteses

Este estudo propõe um conjunto extenso de hipóteses. Para simplificar a leitura deste esboço, as hipóteses serão representadas da seguinte forma:

H_n = Códigos de produção cujos testes de unidade apresentam o mau cheiro "X" apresentam um desempenho pior de acordo com a métrica de qualidade "Y".

Como representado acima, X representa o conjunto de maus cheiros de teste, enquanto Y representa o conjunto de métricas de qualidade sobre o código de teste.

O conjunto X contém as seguintes métricas: tamanho do cenário dos testes, quantidade de asserções no teste, quantidade de objetos dublês, quantidade de métodos de teste para um método de produção, quantidade de métodos de teste para uma classe de produção.

Já o conjunto Y contém as seguintes métricas: complexidade ciclomática, coesão (LCOM), acoplamento aferente, número de linhas por método, número de métodos por classe.

Existe uma relação entre ambos os conjuntos, que é mostrado na tabela abaixo, e que serão testados nas hipóteses acima:

Mau cheiro de teste	Métrica de qualidade (Hn)
Tamanho do cenário dos testes	LCOM (H1), Número de linhas por método (H2), Acoplamento Aferente (H3).
Quantidade de asserções no teste	Número de linhas por método (H4), Complexidade Ciclométrica (H5).
Quantidade de objetos dublês	Acoplamento Aferente (H6).
Quantidade de métodos de teste para um método de produção	LCOM (H7), Número de linhas por método (H8), Complexidade Ciclométrica (H9).
Quantidade de métodos de teste para uma classe de produção	LCOM (H10), Número de métodos por classe (H11).

3. Grupo de Dados

Cada mau cheiro no código possui uma heurística para detecção. Essa heurística devolve um número, pertencente a uma escala ordinal, onde, quanto maior o número, maior a presença daquele mau cheiro.

Para separar os métodos de produção testados por esses métodos de teste, serão ignorados aqueles cuja "quantidade de mau cheiro" esteja na média.

Para isso, as métricas estão separadas em 3 grupos:

Primeiro quartil	Segundo quartil	Terceiro quartil
Usado na análise como "códigos de produção cujo teste não apresenta o mau cheiro".	Será ignorado.	Usado na análise como "códigos de produção cujo teste apresenta o mau cheiro".

4. Teste Estatístico

Este estudo usa o teste de **Wilcoxon** para comparar a diferença de qualidade entre o primeiro e o terceiro quartil. O teste de Wilcoxon é um teste não-paramétrico que verifica se existe diferença na média dos valores dos grupos.

O teste poderá mudar, de acordo com a distribuição dos dados encontrados.