UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Um mapeamento sistemático sobre ferramentas para detecção de Flaky Tests

Pedro Anderson Costa Martins

Orientadora: Dra. Carla Ilane Moreira Bezerra

Trabalho de Conclusão de Curso I
Dezembro de 2022

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Motivação

- Manifestação de Flaky Tests são comuns e podem trazer muitos problemas;

- Escassez de estudos comparativos mais abrangentes entre ferramentas de detecção de flaky tests

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Objetivos

Gerais:

- Identificar ferramentas adotadas para detecção de flaky tests e realizar uma análise qualitativa entre as soluções encontradas a fim de gerar um catálogo das ferramentas.

Específicos:

- Realizar um mapeamento sistemático para coletar ferramentas de detecção de Flaky Tests existentes.
- Realizar uma análise qualitativa das ferramentas selecionadas.
- Elaborar um catálogo com ferramentas para detecção de flaky tests.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Fundamentação Teórica

Testes de Software

- O teste de software é uma cadeia de processos que busca certificar que o código desenvolvido realmente exerça o que foi planejado, ao mesmo tempo que não apresente imprevistos para o usuário final (MYERS et al., 2011)

- Corresponde a uma parte considerável do processo de software
  - Pode consumir muito tempo e recursos

- Automação

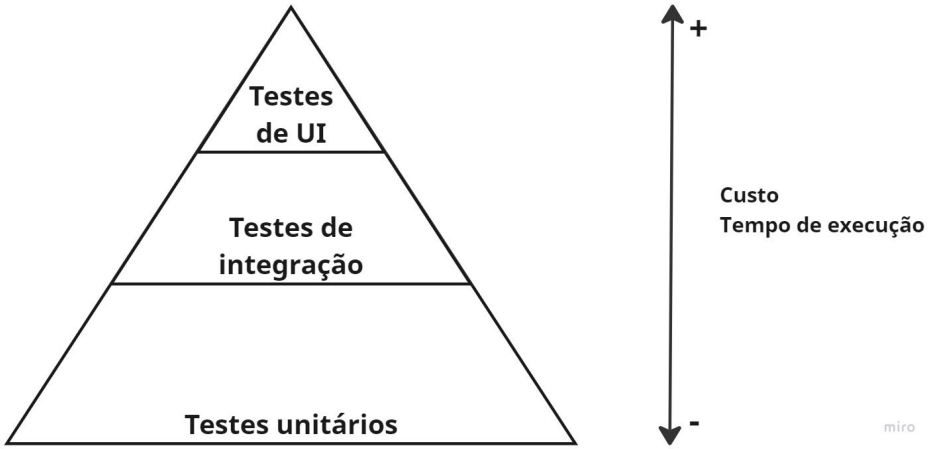
UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Fundamentação Teórica

Testes de Software Automatizados

- Scripts codificados com a utilização de frameworks e são executados cada vez que o software em desenvolvimento é testado (SOMMERVILLE, 2011).
  - Testes de regressão

- Pirâmide de Testes
  - Testes unitários
  - Testes de integração
  - Testes de UI



Fonte: Elaborado pelo autor.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Fundamentação Teórica

Flaky Tests

- Flaky Tests são testes que se comportam de forma não determinística, portanto, podem gerar resultados de aprovação e falha quando executados repetidamente no mesmo código em teste (GRUBER; FRASER, 2022).

- Causas
- Correção
- Estratégias de detecção
- Ferramentas de detecção

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Fundamentação Teórica

Flaky Tests - Causas e correções

- No trabalho de Luo et al. (2014), 201 commits relacionados a flaky tests foram inspecionados :

Commits inspecionados	201
Espera Assíncrona	74
Simultaneidade	32
Dependência da Ordem de Teste	19
Vazamento de recursos	11
Rede	10
Tempo	5
E/S	4
Aleatoriedade	4
Operações com Ponto Flutuante	3
Coleções Não Ordenadas	1

Categorias	Tipo de Correção
Espera Assíncrona	Adicionar/modificar *waitFor* Adicionar/modificar *sleep* Reordenar execução
Simultaneidade	Bloquear operação atômica Tornar determinístico Alterar condição Alterar asserção
Dependência da Ordem de Teste	Configurar/alterar estado Remover dependência Mesclar testes

Fonte: Adaptado de Luo et al. (2014)

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Fundamentação Teórica

Flaky Tests - Estratégias de detecção

- Detectar a origem de uma falha é o primeiro passo para corrigi-la.
- Detecção automática de falhas e que potencialmente podem ser flaky tests

- Técnicas de detecção
  - Reexecução de testes
  - Cobertura de código

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Fundamentação Teórica

Flaky Tests - Ferramentas de detecção

- Algumas ferramentas foram citadas nos trabalhos de Lam et al. (2020), Parry et al. (2021) entre outros.

Ferramenta	Descrição	Fonte
DTDetector	Detecta testes dependentes de ordem revertendo a ordem do conjunto de testes, embaralhando-a, executando cada k-permutação isoladamente ou apenas executando permutações que provavelmente irão expor uma dependência de ordem de teste, com base na análise conservadora de acesso de campo estático entre testes e monitoramento do uso de arquivos.	Zhang *et al.* (2014)
OrcalePolish	Pode identificar indiretamente testes dependentes de ordem ou testes que têm o potencial de se tornarem dependente da ordem.	Huo e Clause (2014)
Shaker	Introduz estresse de CPU e memória durante execuções repetidas do conjunto de testes em uma tentativa de aumentar a probabilidade de manifestar *flaky tests* nas categorias de espera assíncronas e simultaneidade.	Silva *et al.* (2020)
FlakeFlagger	Uma técnica para detectar *flaky tests* sem exigir reexecuções do conjunto de testes, usando uma máquina modelo de aprendizagem. Requer uma combinação de dados de teste dinâmicos, como cobertura de linha e dados estáticos, como recursos do código-fonte do teste.	Alshammari *et al.* (2021)

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Trabalhos Relacionados

Trabalhos	Realiza um mapeamento sistemático	Apresenta ferramentas de detecção de flaky tests	Realiza uma análise qualitativa	Cataloga as ferramentas de detecção
Aljedaani et al. (2021)	X			X
Habchi et al. (2022)			X	
Zolfaghari et al. (2021)	X	X		
Trabalho Proposto	X	X	X	X

Fonte: Elaborado pelo autor.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Metodologia

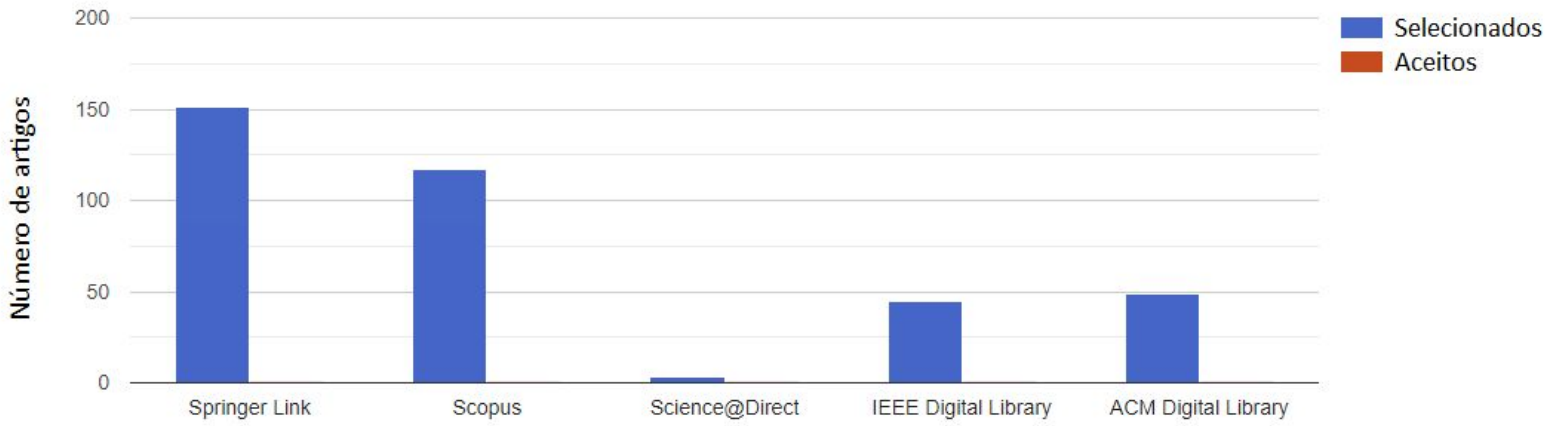
Fonte: Elaborado pelo autor.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Resultados preliminares

365 artigos encontrados na primeira etapa de filtragem do mapeamento

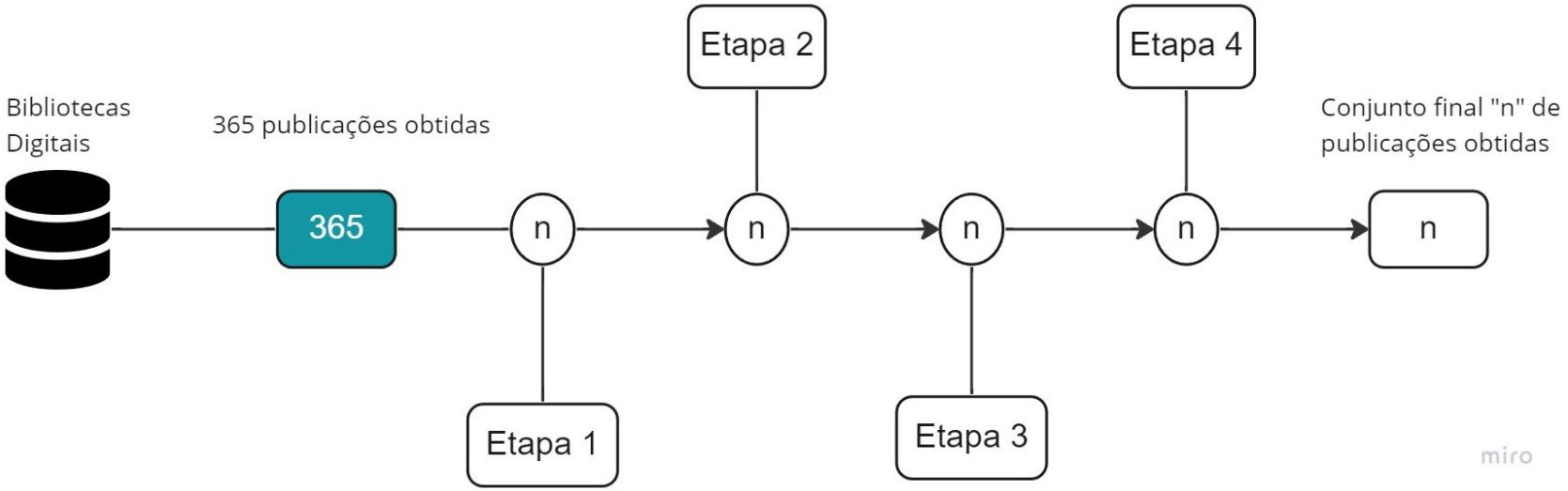


Fonte: Elaborado pelo autor.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Resultados preliminares



Fonte: Elaborado pelo autor.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Cronograma

Atividades	2022/2023							
	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul
Defesa do TCC 1	x							
Mapeamento Sistemático: Planejamento	x	x						
Mapeamento Sistemático: Execução		x	x					
Mapeamento Sistemático: Análise			x	x	x			
Catalogação					x	x		
Análise Qualitativa						x	x	
Escrita do Trabalho Final	x	x	x	x	x	x	x	
Revisão final da Monografia						x	x	
Defesa do Trabalho Final								x

Fonte: Elaborado pelo autor.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Referências

MYERS, G.; SANDLER, C.; BADGETT, T. The Art of Software Testing. Wiley, 2011. (ITPro collection). ISBN 9781118133156. Disponível em: https://books.google.com.br/books?id= GjyEFPkMCwcC.

POLO, M.; REALES, P.; PIATTINI, M.; EBERT, C. Test automation. IEEE Software, v. 30, n. 1, p. 84–89, 2013.

SOMMERVILLE, I. Engenharia de software. Pearson Prentice Hall, 2011. ISBN 9788579361081. Disponível em: https://books.google.com.br/books?id=H4u5ygAACAAJ.

COHN, M. Succeeding with Agile: Software Development Using Scrum. AddisonWesley, 2010. (A Mike Cohen signature book). ISBN 9780321579362. Disponível em: https://books.google.com.br/books?id=IdT6AgAAQBAJ.

GRUBER, M.; LUKASCZYK, S.; KROIß, F.; FRASER, G. An empirical study of flaky tests in python. In: 2021 14th IEEE Conference on Software Testing, Verification and Validation (ICST). [S. l.: s. n.], 2021. p. 148–158.

LUO, Q.; HARIRI, F.; ELOUSSI, L.; MARINOV, D. An empirical analysis of flaky tests. In: Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering. New York, NY, USA: Association for Computing Machinery, 2014. (FSE 2014), p. 643–653. ISBN 9781450330565. Disponível em: https://doi.org/10.1145/2635868.2635920.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ

Referências

PARRY, O.; KAPFHAMMER, G. M.; HILTON, M.; MCMINN, P. A survey of flaky tests. ACM Trans. Softw. Eng. Methodol., Association for Computing Machinery, New York, NY, USA, v. 31, n. 1, oct 2021. ISSN 1049-331X. Disponível em: https://doi.org/10.1145/3476105.

ZOLFAGHARI, B.; PARIZI, R. M.; SRIVASTAVA, G.; HAILEMARIAM, Y. Root causing, detecting, and fixing flaky tests: state of the art and future roadmap. Software: Practice and Experience, Wiley Online Library, v. 51, n. 5, p. 851–867, 2021.

LAM, W.; WINTER, S.; WEI, A.; XIE, T.; MARINOV, D.; BELL, J. A large-scale longitudinal study of flaky tests. Proc. ACM Program. Lang., Association for Computing Machinery, New York, NY, USA, v. 4, n. OOPSLA, nov 2020. Disponível em: https://doi.org/10.1145/3428270.

ALJEDAANI, W.; PERUMA, A.; ALJOHANI, A.; ALOTAIBI, M.; MKAOUER, M. W.; OUNI, A.; NEWMAN, C. D.; GHALLAB, A.; LUDI, S. Test smell detection tools: A systematic mapping study. In: Evaluation and Assessment in Software Engineering. New York, NY, USA: Association for Computing Machinery, 2021. (EASE 2021), p. 170–180. ISBN 9781450390538. Disponível em: https://doi.org/10.1145/3463274.3463335.

HABCHI, S.; HABEN, G.; PAPADAKIS, M.; CORDY, M.; TRAON, Y. L. A qualitative study on the sources, impacts, and mitigation strategies of flaky tests. In: 2022 IEEE Conference on Software Testing, Verification and Validation (ICST). [S. l.: s. n.], 2022. p. 244–255.

UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ

Obrigado pela atenção!

Dúvidas?