

UMA ABORDAGEM DE BIOINFORMÁTICA PARA DETECTAR PLÁGIO EM CÓDIGO-FONTE

Mestrando: Kaio Pablo Gomes

Orientadora: Simone Nasser Matos

INTRODUÇÃO

- A definição de plágio e o contexto da programação.
 - “ação de plagiar” ou “copiar o trabalho alheio” (CAMBRIDGE, 2019).
 - Códigos-fonte são modificados ou camuflados com a cópia das implementações de outros autores (DURIĆ; GAŠEVIĆ, 2012).

INTRODUÇÃO

- A preocupação associada a prática de plagiar na Educação e na Indústria de Software.
 - Segundo (SRAKA; KAUCIC, 2009), no meio acadêmico cerca de 72,5% dos alunos admitiram plagiar em atividades de programação.
 - Segundo (ULLAH et al., 2018), todo sistema de software apresenta em sua implementação até 20% de código plagiado.

INTRODUÇÃO

- Objetivo: propor uma abordagem capaz de criar um detector automático de plágio em programação.
 - Considerar a complexidade de tempo de execução.
 - Detectar os 6 níveis de (FAIDHI; ROBINSON, 1987) de modificações em códigos-fonte plagiados;

ABORDAGEM PROPOSTA

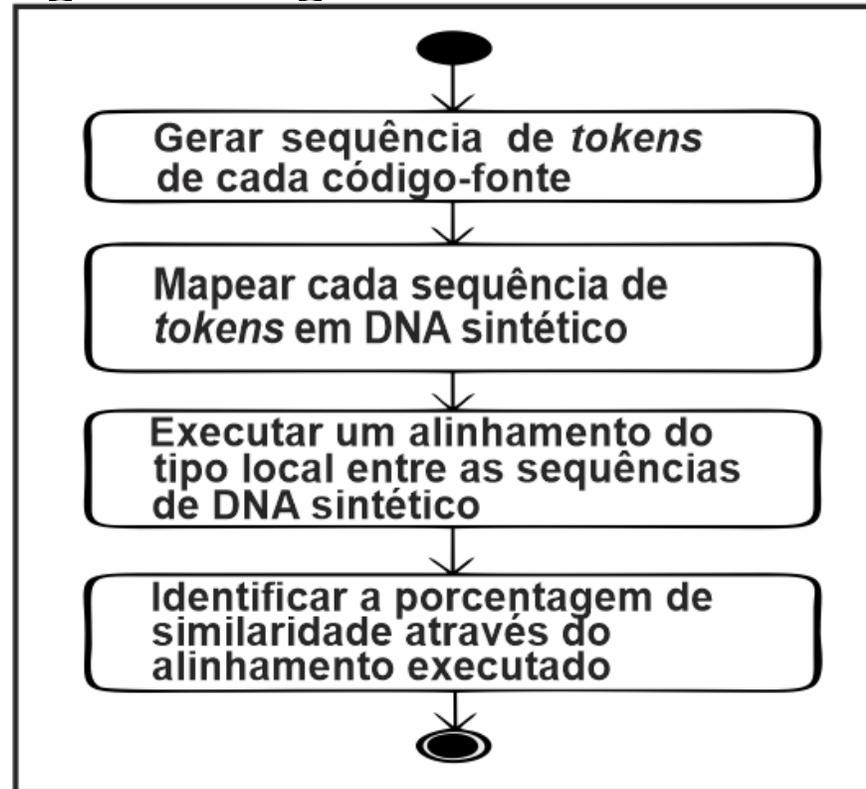
- Metodologia baseada em bioinformática de (PEDERSEN et al., 2012):
- Mapeamento sistemática realizado para entender o cenário das novas soluções.
 - Utilização de mais de uma técnica.
 - Suporte a mais de uma linguagem de programação.

FUNCIONAMENTO DA ABORDAGEM PROPOSTA

- Técnica de computação: *tokens*.
- Técnica de bioinformática: alinhamento de sequências (DNA ou Proteínas).
- Uso de ferramentas de bioinformática: BLAST.

FUNCIONAMENTO DA ABORDAGEM PROPOSTA

Figura 1 – Diagrama de atividades da abordagem



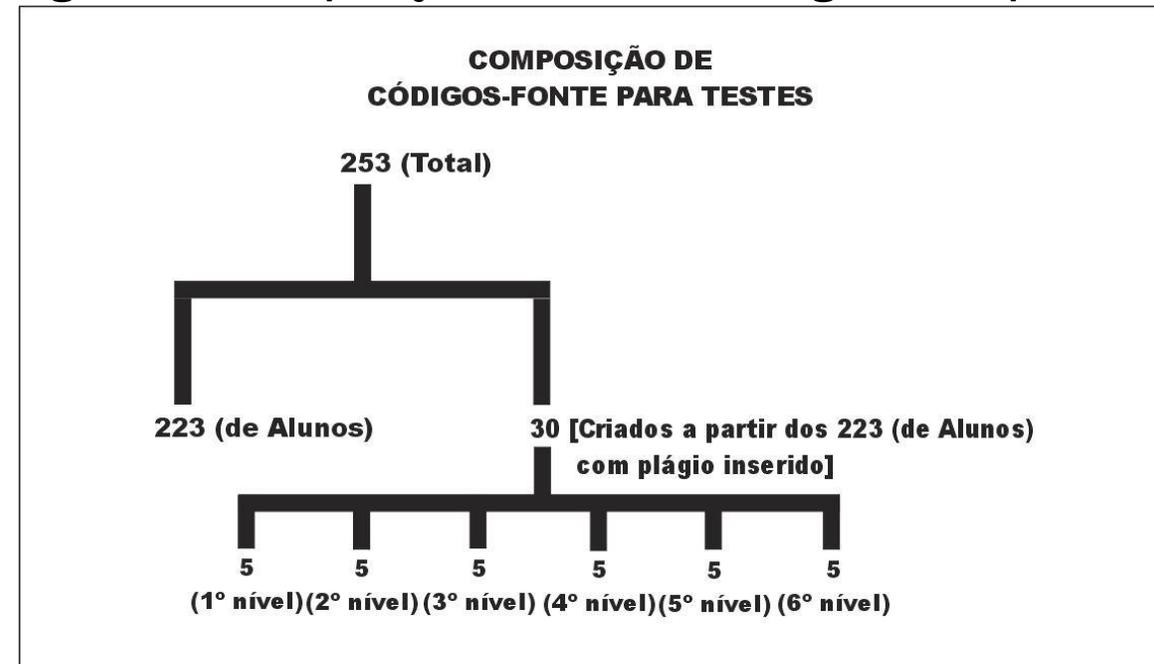
Fonte: Autoria própria

AValiação DA ABORDAGEM PROPOSTA

- Avaliar os diferentes tipos de plágios:

- 253 códigos-fonte (Linguagem C)
- 7 cenários de testes
- Parâmetros avaliativos:
-Recall e Precision
- JPLAG

Figura 2 – Composição da base de códigos-fonte para testes



Fonte: Autoria própria

AGRADECIMENTO

- CAPES;
- O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

- CAMBRIDGE, Dicionário. Plágio. 2019. <<https://dictionary.cambridge.org/pt/dicionario/portugues-ingles/plagio>>.
- DURIC, Zoran; GAŠEVIĆ, Dragan. A source code similarity system for plagiarism detection. The Computer Journal, Oxford University Press, v. 56, n. 1, p. 70–86, 2012.
- SRAKA, Dejan; KAUCIC, Branko. Source code plagiarism. In: IEEE. Information Technology Interfaces, 2009. ITI'09. Proceedings of the ITI 2009 31st International Conference on. [S.l.], 2009. p. 461–466.

REFERÊNCIAS

- ULLAH, Farhan et al. Plagiarism detection in students' programming assignments based on semantics: multimedia e-learning based smart assessment methodology. *Multimedia Tools and Applications*, Springer, p. 1–18, 2018.
- PEDERSEN, Jay et al. Blast your way through malware malware analysis assisted by bioinformatics tools. In: THE STEERING COMMITTEE OF THE WORLD CONGRESS IN COMPUTER SCIENCE, COMPUTER ENGINEERING AND APPLIED COMPUTING (WORLDCOMP). Proceedings of the International Conference on Security and Management (SAM). [S.l.], 2012. p. 1.