Métodos Probabilísticos para Engenharia Informática

**2018/2019**



**Docentes:** Prof. Carlos Bastos | Prof. António Teixeira

Implementação de um Contador Estocástico, Counting Bloom Filter e Encontrar Similares (Usando MinHash)

Tomás Batista **89296**

Rodrigo Oliveira **90514**

ÍNDICE

[INTRODUÇÃO 2](#_Toc532058028)

[Desenvolvimento 2](#_Toc532058029)

[Conclusão 4](#_Toc532058030)

[Ferramentas de Suporte 4](#_Toc532058031)

[Bibliografia 4](#_Toc532058032)

# INTRODUÇÃO

No âmbito da unidade curricular de Métodos Probabilísticos para a Engenharia Informática, foi-nos proposto a realização de um trabalho cujo objetivo é a implementação de um Contador Estocástico, um Counting Bloom Filter e Deteção de Similares usando MinHash. Cada um destes três módulos foi testado com os conteúdos dos guiões das aulas práticas. Foi também criada uma aplicação conjunta que implementa os três módulos.

# Desenvolvimento

A nossa ideia de implementação foi trabalharmos com ficheiros que contêm Passwords. Usámos um ficheiro com as 10 mil passwords mais comuns, um com as 370 passwords banidas pelo Twitter, ETC RODRIGO.

De modo a facilitar o utilizador criámos um **menu** em que o utilizador decide que módulo que experimentar usando apenas inputs de teclado.

**FOTO MENU APOS TEXTO RODRIGO**

No módulo do **Contador Estocástico** estimamos a quantidade de Passwords nos ficheiros “MostCommom10kPasswords” e “Twitter\_Passwords” (soma de ambas). O número real de passwords é 10370.

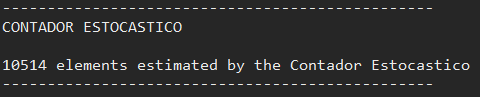
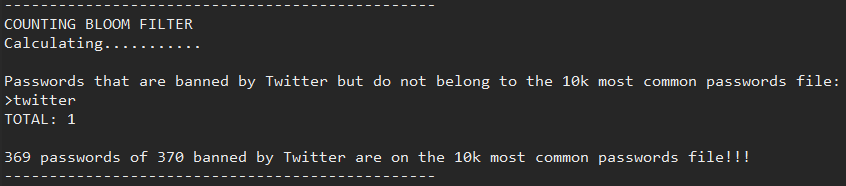


Figura 1 Teste Contador Estocástico

No módulo do **Counting Bloom Filter** testamos quais as passwords banidas pelo Twitter que não constam nas 10 mil passwords mais comuns.

Figura 2 Teste Counting Bloom Filter



No módulo do **Encontrar Similares (c/ MinHash)** RODRIGOOOOOOOOOOOOOOOOOOOOOOOOOO

**FOTO TEXTO MINHASH RODRIGO**

# Conclusão

Considerando todos os requisitos impostos no enunciado, podemos dizer que o resultado foi alcançado.

# Ferramentas de Suporte

* [GitHub/Projeto\_MPEI](https://github.com/tomas99batista/Projeto_MPEI)

# Bibliografia

* [GitHub - SecLists](https://github.com/danielmiessler/SecLists) (Ficheiros c/ passwords)