IDENTIFICAÇÃO DE ATAQUES DE PHISHING ATRAVÉS DE MACHINE LEARNING

BIANCA DOMINGOS GUARIZI,
DALBERT MATOS MASCARENHAS

CEFET/RJ

SBSeg24 SÃO JOSÉ DOS CAMPOS
Realização: SBC Sociedade Brasileira de Computação

SOBRE MIM

Graduanda de Engenharia da Computação no CEFET/RJ;

Analista de Segurança da Informação Jr.;

Atuando profissionalmente na área de Segurança da Informação há mais de 2 anos;

Certificações ISO 27001 e Microsoft SC-900.

RESUMO

1.

CRIAÇÃO DE UMA FERRAMENTA QUE POSSA SER UTILIZADA PELO USUÁRIO

2.

DETECÇÃO DE PHISHING EM TEMPO REAL

3.

USO DE MACHINE LEARNING PARA DESENVOLVIMENTO DA SOLUÇÃO

01 MOTIVAÇÃO

02 PROBLEMA

03 SOLUÇÃO

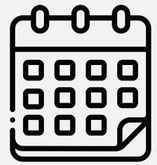
04 BASE DE DADOS E ATRIBUTOS

05 FERRAMENTA

06 RESULTADOS

07 DIFERENCIAL

08 CONCLUSÃO



AGENDA

MOTIVAÇÃO

Ataques de phishing se tornam cada vez mais sofisticados;

Treinamento e conscientização não são suficientes;

Proteção para usuários leigos.

PROBLEMA

FERRAMENTAS EXISTENTES UTILIZAM BLACKLISTS

ATAQUES ZERO-DAY

CEFET/RJ

SBSeg24 SÃO JOSÉ DOS CAMPOS

Realização: SBC Sociedade Brasileira de Computação

CEFET/RJ

SOLUÇÃO

MACHINE LEARNING

Utilização do modelo Random Forest.

EXTENSÃO PARA O NAVEGADOR

A extensão para o navegador desenvolvida realiza a coleta das URLs acessadas pelo usuário.

CÓDIGO EM PYTHON

Executa a coleta dos atributos da página Web e executa a classificação entre legítimo e phishing.

SBSeg24 SÃO JOSÉ DOS CAMPOS

Realização: SBC Sociedade Brasileira de Computação

BASE DE DADOS

PHISHING

PhishTank® Out of the Net, into the Tank.

OpenPhish

LEGÍTIMOS

MOZ

CEFET/RJ

SBSeg24 SÃO JOSÉ DOS CAMPOS

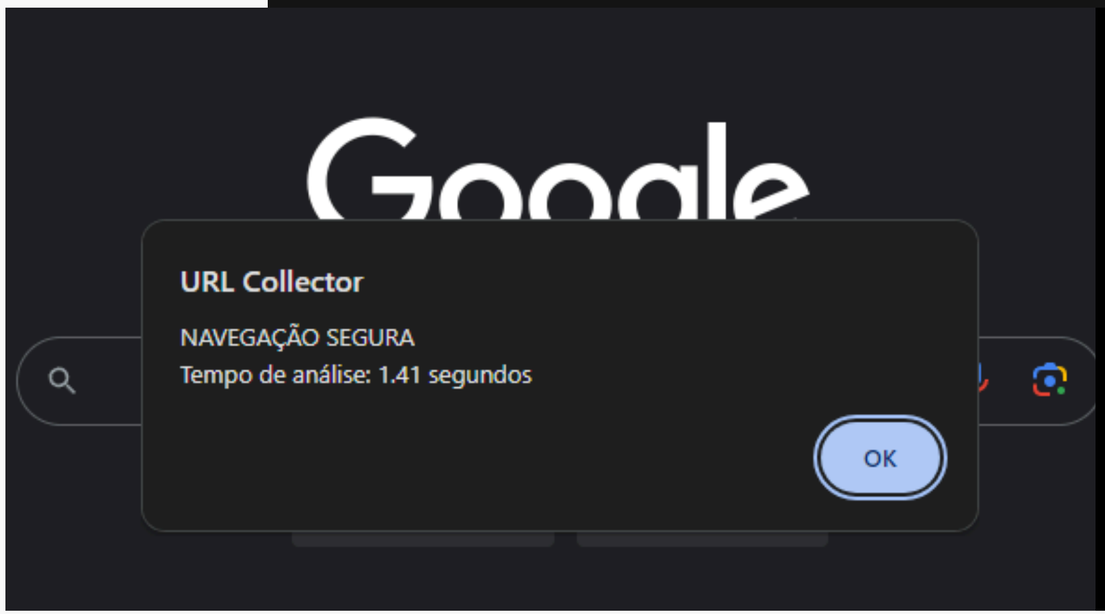
Realização: SBC Sociedade Brasileira de Computação

ATRIBUTOS

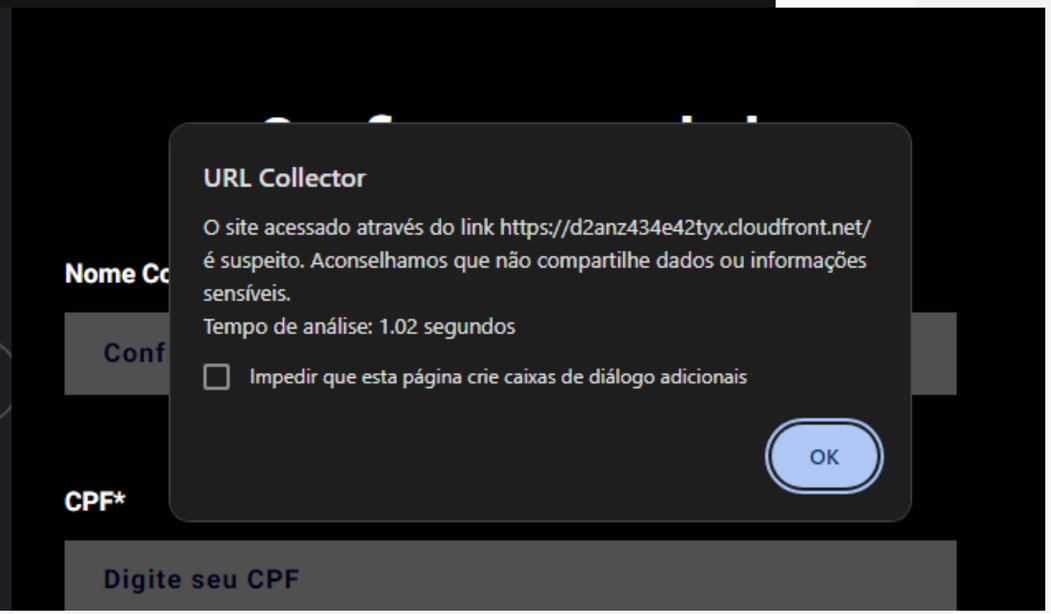
A1 - Tamanho da URL	A7 - PageRank
A2 - Utiliza HTTPS	A8 - Possui formulário HTML
A3 - Possui formato IP	A9 - Validade do Certificado SSL
A4 - Quantidade de pontos (.)	A10 - Contém caracteres '//'
A5 - Contém o caractere '@'	A11 - Possui texto 'https' na URL
A6 - Total de dias ativo	A12 - Possui caractere '-' no domínio
A13 - Possui iframe no código HTML	

FERRAMENTA

```
PS C:\Users\denis\OneDrive\Documentos\Bianca\TCC\Codigos a enviar> python3 .\code_analyse_trafic.py
Iniciando software...
 * Serving Flask app 'code_analyse_trafic'
 * Debug mode: off
INFO:werkzeug:WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
INFO:werkzeug:Press CTRL+C to quit
URL https://www.google.com/ foi identificada como navegação segura.
INFO:werkzeug:127.0.0.1 - - [14/Sep/2024 15:02:54] "POST /receive_url HTTP/1.1" 200 -
URL https://d2anz434e42tyx.cloudfront.net/ é suspeita. Aconselhamos que não compartilhe dados ou informações sensíveis.
INFO:werkzeug:127.0.0.1 - - [14/Sep/2024 15:03:48] "POST /receive_url HTTP/1.1" 200 -
```



HTTPS://GOOGLE.COM



HTTPS://D2ANZ434E42TYX.CLOUDFRONT.NET/

RESULTADOS

	Precisão	Recall	Acurácia	F1 Score
Teste	99,02%	96,04%	97,81%	97,50%
Treinamento	98,95%	96,38%	97,93%	97,65%

DIFERENCIAL

USO DE MACHINE LEARNING	ATAQUES ZERO DAY	DETECÇÃO EM TEMPO REAL
Melhor adaptação e identificação de padrões.	Identificação de ataques zero day.	Coleta de atributos e classificação em tempo real, sem se basear em blacklists.

CONCLUSÃO

CEFET/RJ

SBSeg24 SÃO JOSÉ DOS CAMPOS

Realização: SBC Sociedade Brasileira de Computação

OBRIGADA!

Use o QR Code para acessar a ferramenta no GitHub ou use o link:
https://github.com/bguarizi/findphishing