Detecção de URLs Maliciosas

Mineração de Dados Aplicada

Julio Cesar da Silva Rodrigues[1]

[1]Universidade Federal de São João del-Rei
Curso de Ciência da Computação
julio.csr.271@aluno.ufsj.edu.br

4 de Maio de 2023

Conteúdo

Conteúdo

URLs Maliciosas

- Via rápida e direta para aplicar crimes cibernéticos;

URLs Maliciosas

- Via rápida e direta para aplicar crimes cibernéticos;
- Potencial de infecção exponencial;

URLs Maliciosas

- Via rápida e direta para aplicar crimes cibernéticos;
- Potencial de infecção exponencial;
- Brasil no Top 15 com maior número de vítimas [Abranet(2019)];

URLs Maliciosas

- Via rápida e direta para aplicar crimes cibernéticos;
- Potencial de infecção exponencial;
- Brasil no Top 15 com maior número de vítimas [Abranet(2019)];
- Evolução nas técnicas de camuflagem e detecção.

Conteúdo

Perguntas e Objetivos

- Quais são as principais características que definem a natureza de uma URL?

Perguntas e Objetivos

- Quais são as principais características que definem a natureza de uma URL?
- Existe vantagem na aplicação de *instance selection* em balanceamento?

Perguntas e Objetivos

- Quais são as principais características que definem a natureza de uma URL?
- Existe vantagem na aplicação de *instance selection* em balanceamento?
- É possível obter resultados equiparáveis com mais classes?

Perguntas e Objetivos

- Quais são as principais características que definem a natureza de uma URL?
- Existe vantagem na aplicação de *instance selection* em balanceamento?
- É possível obter resultados equiparáveis com mais classes?
- Influência de algoritmos em resultados com a base final.

Conteúdo

Base de Artigos Selecionados

1. Detecting Malicious URLs Using Machine Learning Techniques: Review and Research Directions (IEEE *Xplore®*);

Base de Artigos Selecionados

1. Detecting Malicious URLs Using Machine Learning Techniques: Review and Research Directions (IEEE *Xplore®*);
2. A Survey of Intelligent Detection Designs of HTML URL Phishing Attacks (IEEE *Xplore®*);

Base de Artigos Selecionados

1. Detecting Malicious URLs Using Machine Learning Techniques: Review and Research Directions (IEEE *Xplore®*);
2. A Survey of Intelligent Detection Designs of HTML URL Phishing Attacks (IEEE *Xplore®*);
3. A Comparative Survey of Instance Selection Methods applied to NonNeural and Transformer-Based Text Classification (ACM Digital Library);

Base de Artigos Selecionados

1. Detecting Malicious URLs Using Machine Learning Techniques: Review and Research Directions (IEEE *Xplore®*);
2. A Survey of Intelligent Detection Designs of HTML URL Phishing Attacks (IEEE *Xplore®*);
3. A Comparative Survey of Instance Selection Methods applied to NonNeural and Transformer-Based Text Classification (ACM Digital Library);
4. An Assessment of Lexical, Network, and Content-Based Features for Detecting Malicious URLs Using Machine Learning and Deep Learning Models (Hindawi).

Etapas

- Parcial I - 25/05/2023:

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.
- Parcial II - 13/06/2023:

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.
- Parcial II - 13/06/2023:
  1. Análise e seleção de atributos;

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.
- Parcial II - 13/06/2023:
  1. Análise e seleção de atributos;
  2. Balanceamento da base de dados.

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.
- Parcial II - 13/06/2023:
  1. Análise e seleção de atributos;
  2. Balanceamento da base de dados.
- Final - 27/06/2023:

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.
- Parcial II - 13/06/2023:
  1. Análise e seleção de atributos;
  2. Balanceamento da base de dados.
- Final - 27/06/2023:
  1. Construção dos modelos de *machine learning*;

Etapas

- Parcial I - 25/05/2023:
  1. Análise exploratória aprofundada da base;
  2. Construção de novos atributos.
- Parcial II - 13/06/2023:
  1. Análise e seleção de atributos;
  2. Balanceamento da base de dados.
- Final - 27/06/2023:
  1. Construção dos modelos de *machine learning*;
  2. Validação dos resultados e finalização da escrita do artigo.

Referências I

Abranet.
Relatório aponta que cada url maliciosa no brasil afeta 18 usuários.
https://www.abranet.org.br/Noticias/Relatorio-aponta-que-cada-URL-maliciosa-no-Brasil-afeta-18-usuarios-2585.html?UserActiveTemplate=site, 2019.

Malak Aljabri, Hanan S. Altamimi, Shahd A. Albelali, Maimunah Al-Harbi, Haya T. Alhuraib, Najd K. Alotaibi, Amal A. Alahmadi, Fahd Alhaidari, Rami Mustafa A. Mohammad, and Khaled Salah.
Detecting malicious urls using machine learning techniques: Review and research directions.
IEEE Access, 10:121395–121417, 2022.
doi: 10.1109/ACCESS.2022.3222307.

Sultan Asiri, Yang Xiao, Saleh Alzahrani, Shuhui Li, and Tieshan Li.
A survey of intelligent detection designs of html url phishing attacks.
IEEE Access, 11:6421–6443, 2023.
doi: 10.1109/ACCESS.2023.3237798.

Referências II

Washington Cunha, Felipe Viegas, Celso França, Thierson Rosa, Leonardo Rocha, and Marcos André Gonçalves.
A comparative survey of instance selection methods applied to nonneural and transformer-based text classification.
ACM Comput. Surv., jan 2023.
ISSN 0360-0300.
doi: 10.1145/3582000.
URL https://doi.org/10.1145/3582000.
Just Accepted.

Rami Mustafa A. Mohammad undefined Samiha Mirza Dina H. Alhamed Hanan S. Altamimi Sara Mhd. Bachar Chrouf Malak Aljabri, Fahd Alhaidari.
An assessment of lexical, network, and content-based features for detecting malicious urls using machine learning and deep learning models.
Hindawi, 10:14, 2022.
doi: https://doi.org/10.1155/2022/3241216.