UNIVERSIDADE DO MINDELO
DEPARTAMENTO DE ENGENHARIA E CIÊNCIAS DO MAR

CURSO DE LICENCIATURA EM ENGENHARIA INFORMÁTICA E SISTEMA COMPUTACIONAIS

Relatório

Tema: Relatório Data Visualization

Autor: Natanael Duarte, Nº5073

Professor: Estanislau Lima

Mindelo, 2024

UNIVERSIDADE DO MINDELO
DEPARTAMENTO DE ENGENHARIA E CIÊNCIAS DO MAR

CURSO DE LICENCIATURA EM
ENGENHARIA INFORMÁTICA E SISTEMAS COMPUTACIONAIS

RELATÓRIO

Tema: Relatório de Data Visualization

Autor: Natanael Duarte, Nº 5073

Professor: Estanislau Lima

Mindelo, 2024

Índice

INTRODUÇÃO

Este relatório técnico aborda a análise de uma base de dados de tráfego web, coletada através do AWS CloudWatch, com o objetivo de detectar atividades suspeitas e tentativas de ataque. Os dados foram gerados monitorando o tráfego para um servidor web de produção, utilizando várias regras de detecção para identificar padrões anômalos. O objetivo desta análise é aprimorar as técnicas de detecção de ameaças em ambientes de nuvem.

DESCRIÇÃO DO CONJUNTO DE DADOS

O conjunto de dados contém registos de tráfego web que incluem informações sobre o tempo de acesso, endereços IP, URLs acessadas, códigos de resposta HTTP, entre outros. Estas informações são cruciais para identificar atividades suspeitas, como tentativas de exploração de vulnerabilidades ou acessos não autorizados.

OPERAÇÕES REALIZADAS

1. Leitura e Exploração Inicial dos Dados:

- Importação dos dados utilizando pandas.

- Visualização inicial das primeiras linhas do dataset para entender a estrutura dos dados.

2. Limpeza de Dados:

- Tratamento de valores ausentes.

- Conversão de tipos de dados para facilitar a análise.

3. Análise Exploratória de Dados (EDA):

-Geração de estatísticas descritivas.

-Criação de gráficos para visualizar distribuições e identificar padrões:

-Histogramas para distribuição de códigos de resposta HTTP.

-Gráficos de barras para os endereços IP mais frequentes.

-Gráficos de linha para monitorar o tráfego ao longo do tempo.

4. Deteção de Anomalias:

-Utilização de técnicas de deteção de anomalias para identificar padrões suspeitos.

-Aplicação de modelos de classificação para diferenciar entre tráfego normal e suspeito.

5. Visualização dos Resultados:

- Foi usado o google colab para implementar códigos em python para gerar gráficos usando o dataset.

CONCLUSÕES

A análise revelou várias informações importantes sobre o tráfego web:

Padrões de Tráfego:

- A maioria do tráfego foi classificada como normal, mas foram identificadas algumas anomalias que podem indicar atividades suspeitas.

- Certos endereços IP apareceram com frequência significativamente maior, sugerindo possíveis tentativas de exploração.

Códigos de Resposta HTTP:

- A distribuição dos códigos de resposta HTTP ajudou a identificar comportamentos anômalos, como um número elevado de respostas 404 (não encontrado) e 500 (erro do servidor).

Tendências Temporais:

- A análise temporal do tráfego revelou períodos específicos com aumento de atividade suspeita, o que pode estar relacionado a tentativas de ataque coordenadas.

RECOMENDAÇÕES

Com base na análise realizada, foram feitas as seguintes recomendações para aprimorar a segurança cibernética:

1. Monitoramento Contínuo:

- Implementar um monitoramento contínuo do tráfego web para identificar e responder rapidamente a atividades suspeitas.

2. Regras de Deteção de Anomalias:

- Refinar as regras de detecção de anomalias para melhorar a precisão na identificação de tráfego suspeito.

3. Análise de Endereços IP:

- Investigar endereços IP que aparecem com frequência elevada para determinar se são maliciosos e tomar medidas apropriadas, como bloqueio ou mitigação.

4. Automatização de Respostas:

- Desenvolver e implementar scripts automatizados para responder a incidentes de segurança, reduzindo o tempo de resposta e mitigando potenciais danos.

ARQUIVOS E GRÁFICOS

Os gráficos gerados foram salvos em alta resolução para melhor visualização e interpretação. O notebook contendo todas as operações realizadas e os gráficos está no ficheiro: DataAnalysis.pynb