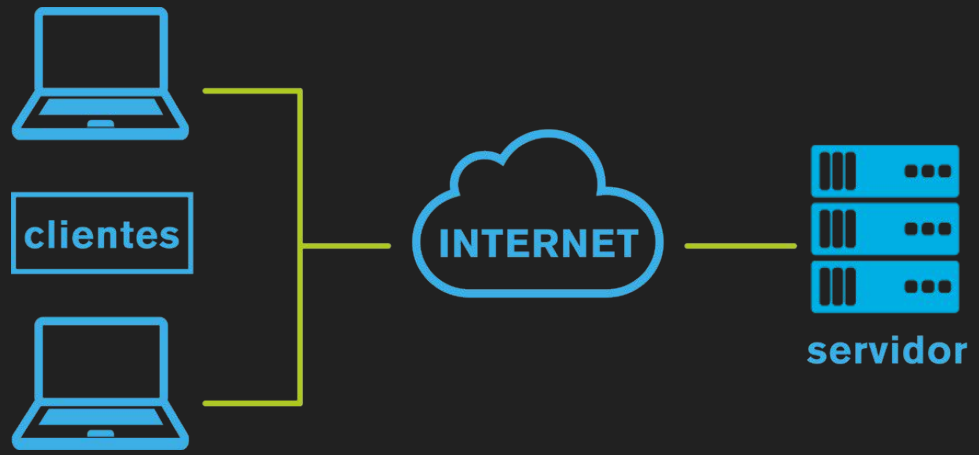
Detección de anomalías mediante aprendizaje automático en tráfico de servidores web

Mitsiu Alejandro Carreño Sarabia

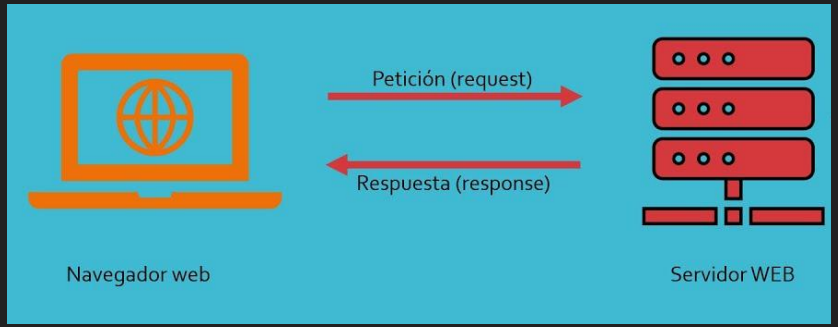
Antecedentes y Contexto

Arquitectura Cliente-Servidor

- Se solicitan recursos
- Sistema centralizado optimiza recursos
- Concepto fundamental en la construcción del internet



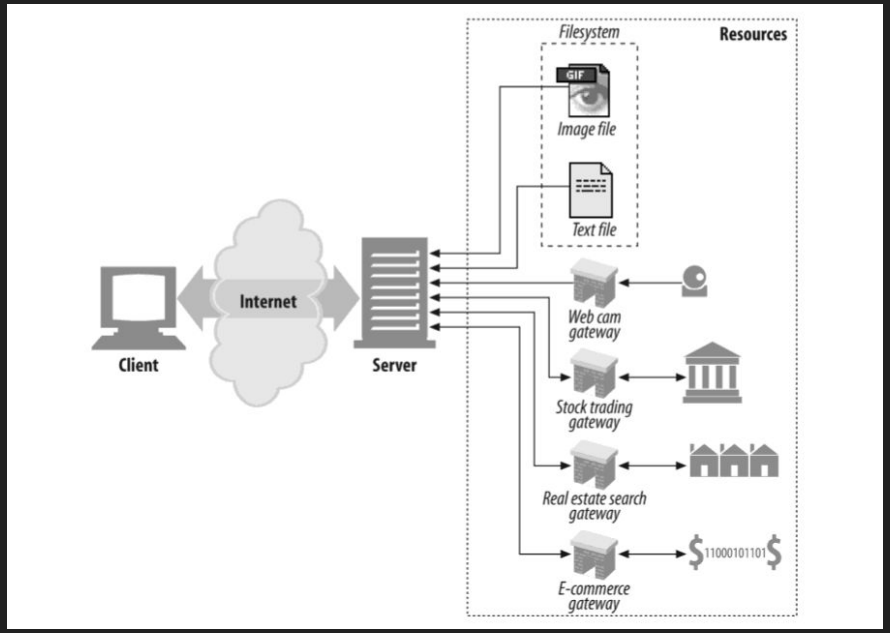
Antecedentes y Contexto



Comunicación Solicitud/Respuesta

- Modalidad de un mensaje único por nodo
- La comunicación es iniciada por el cliente
- El servidor solo puede contestar, no solicitar
- El cliente debe esperar la respuesta

Antecedentes y Contexto



Protocolo HTTP

- Divide los mensajes en pequeños bloques
- Permite transferir cualquier tipo de información
- Es un protocolo de transmisión “segura”

Antecedentes y Contexto

Protocolo HTTP - Métodos

- Indican una acción
- Siempre debe estar en la solicitud

Método	Acción
Get	Obtener
Post	Crear
Put	Actualizar
Delete	Borrar
…	…

Antecedentes y Contexto

Protocolo HTTP - Código de estatus

- Indican un resultado
- Siempre debe estar en la respuesta

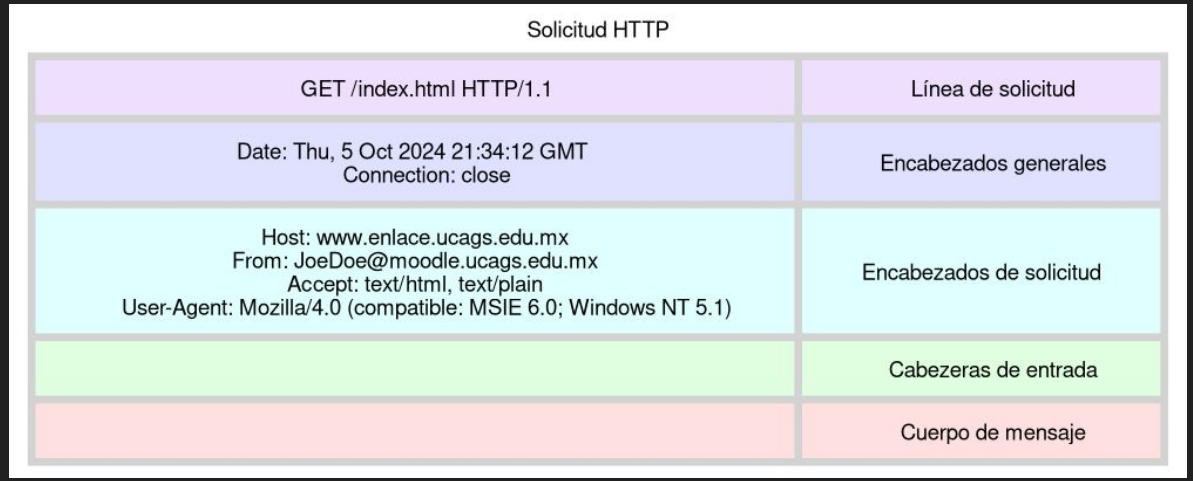
Clase	Rango
Respuestas informativas	100-199
Respuestas satisfactorias	200-299
Redirecciones	300-399
Errores de cliente	400-499
Errores del servidor	500-599

Antecedentes y Contexto

Protocolo HTTP

- Cada bloque de se divide en capas

- Contiene meta información

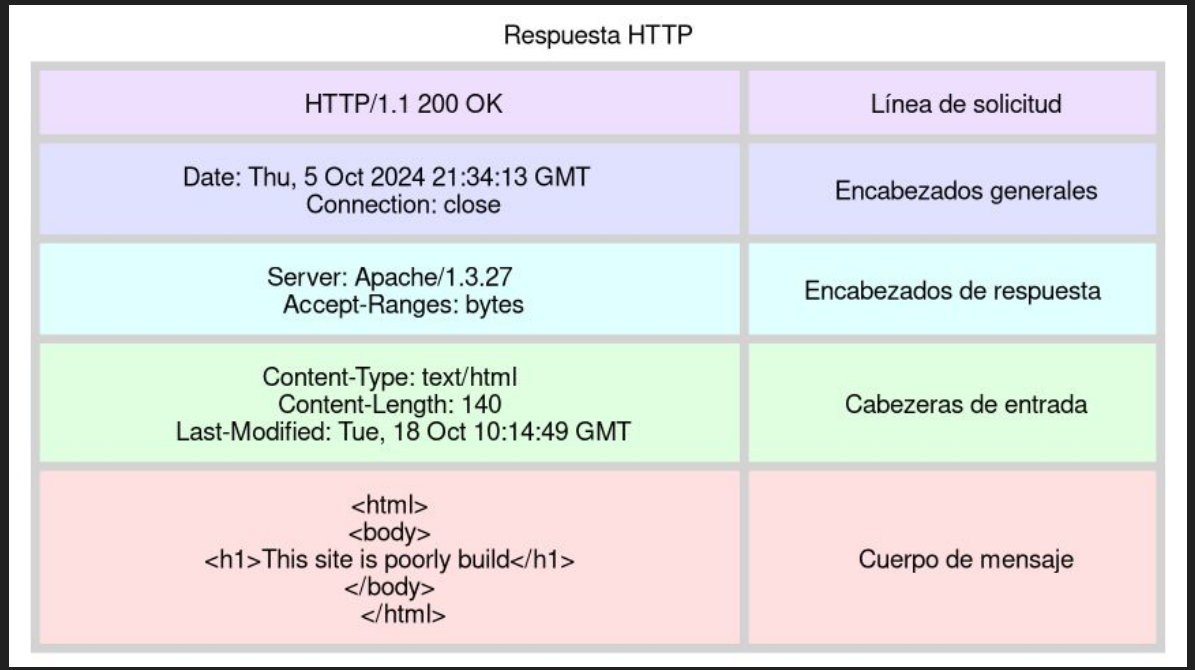


Antecedentes y Contexto

Protocolo HTTP

- Cada bloque de se divide en capas

- El cuerpo del mensaje también se conoce como payload



Variables

45.166.93.223 - - [23/Aug/2024:00:00:20 +0000] "POST /monitoring?o=1141627&p=4505121011335168&r=us HTTP/1.1" 308 57 "https://mastermanual.mx/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/128.0.0.0 Safari/537.36"

- Dirección IP
- Usuario
- Fecha
- Método HTTP
- Recurso solicitado
- Parámetros
- Versión HTTP
- Código de estatus HTTP
- Bytes intercambiados (payload)
- HTTP Referer
- User-Agent

Problemática

- La adopción generalizada del internet en la vida cotidiana, genera la necesidad de:
    - Proteger información
    - Gestionar infraestructura
    - Entender los patrones de uso

- La cantidad de información generada es grande, un análisis manual no es viable.

Hipótesis

- Se plantea la hipótesis de que el tráfico recabado presente patrones claros y definidos de uso, lo cuál permitirá comparar el tráfico actual con registros históricos, y calcular su grado de diferencia.

Objetivos de investigación

- Analizar grandes cantidad es de datos de manera automática
- Permitir la actualización de patrones, ajustando el comportamiento normal y detectando anomalías

Justificación

- Al ser datos de bitácoras de registro, proveen una mirada confiable al uso del servidor

- Entender los usos típicos y diferenciarlos de los atípicos es una herramienta poderosa para múltiples ámbitos:
    - Desarrollo de producto
    - Ajustes de infraestructura / minimizar costos
    - Detección de tendencias
    - Detección de ciberataques

Marco teórico

Diversos estudios se han realizado respecto al análisis de bitácoras de registro.

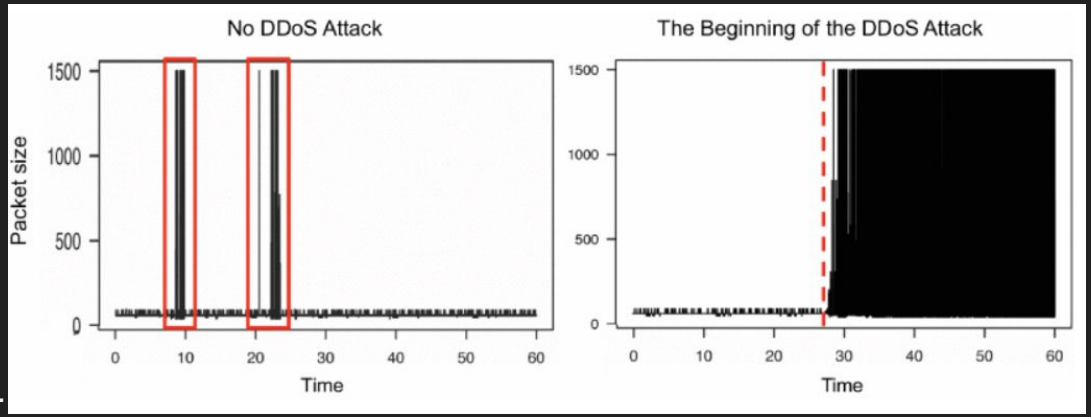
A pesar de contar con diferencias metodológicas resultan relevantes:

- A Distributed Architecture for DDoS Prediction and Bot Detection - Bruno Martins et al (2022)
- An Empirical Analysis of Anomaly Detection Methods for Multivariate Time Series - Dongwen Li et al (2023)
- Practical Anomaly Detection over Multivariate Monitoring Metrics for Online Services (PER) - Jinyang Liu et al (2023)

Marco teórico

A Distributed Architecture for DDoS Prediction and Bot Detection - Bruno Martins et al (2022)

- Su estudio está centrado en la detección de ataques DDoS
- Ante un ataque **sobre provisionar y delegar tráfico** son medidas comunes
- Estas medidas son más efectivas con una **detección más temprana**.

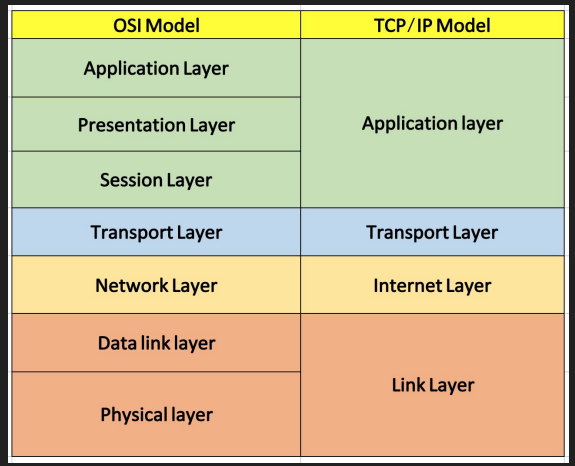


Marco teórico

Features	Description
Protocol Quantity	Quantity of protocols each node uses
Average TTL	Average *Time to live* of the sent packets
TCP Window Size	Size of the windows in the TCP protocol
Percent TCP	Percentage of packets that use TCP protocol
Percent UDP	Percentage of packets that use UDP protocol
Percent DNS	Percentage of packets that use DNS protocol
Percent ICMP	Percentage of packets that use ICMP protocol
Percent Others	Percentage of packets that use other protocols
Source Privileged Ports	Quantity of ports on the source that are privileged (<1024)
Source Not Privileged Ports	Quantity of ports on the source that are not privileged (>1024)
Destination Port Quantity	Quantity of destination ports contacted by the node
Frame Length	Packet's frame length (Layer 2 on the OSI model)
In Degree	Graph feature described in the text
Out Degree	Graph feature described in the text
In Degree Weight	Graph feature described in the text
Out Degree Weight	Graph feature described in the text
Node betweenness centrality	Graph feature described in the text
Local clustering coefficient	Graph feature described in the text
Eigenvector centrality	Graph feature described in the text

A Distributed Architecture for DDoS Prediction and Bot Detection - Bruno Martins et al (2022)

- Una diferencia crucial es que sus variables están en una capa más abajo del modelo OSI (network)

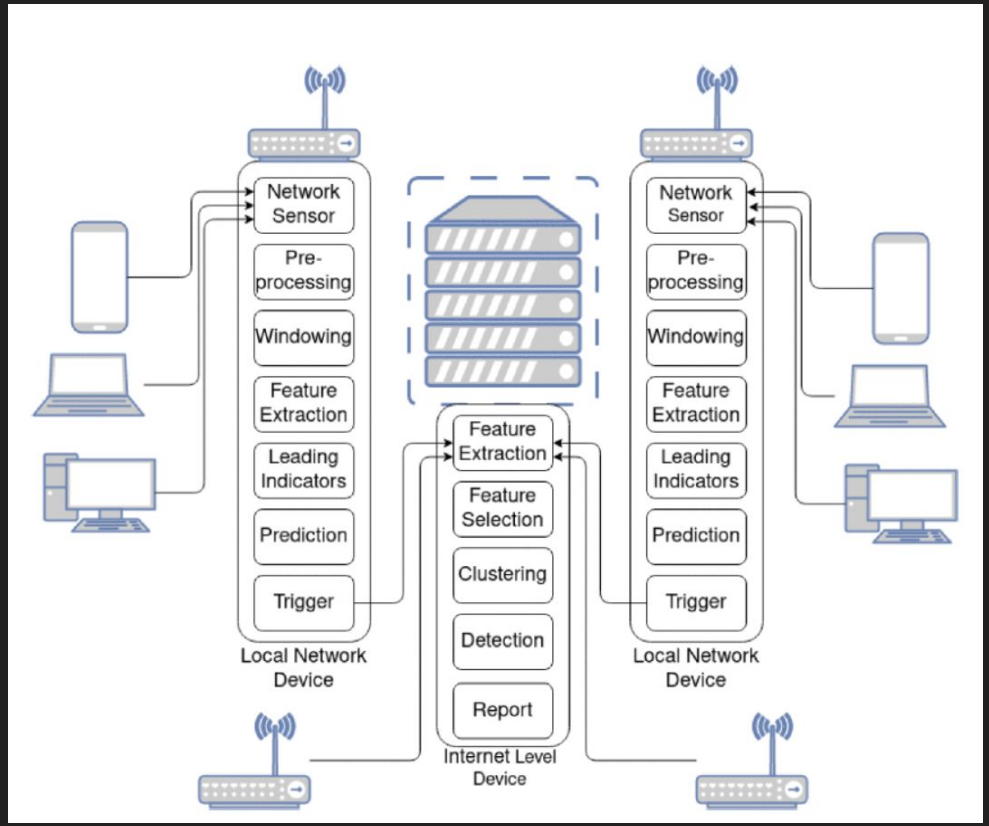


Marco teórico

A Distributed Architecture for DDoS Prediction and Bot Detection - Bruno Martins et al (2022)

- Basado en la **teoría de la metaestabilidad, la cual ofrece aprendizaje estadístico no supervisado** e identifica la inminencia de un ataque DDoS.

- La arquitectura propuesta cuenta con dos niveles: local e internet



Marco teórico

A Distributed Architecture for DDoS Prediction and Bot Detection - Bruno Martins et al (2022)

- Evaluaron tres datasets de ataques reales y simulados

- Máquinas de soporte vectorial, Redes neuronales profundas, árboles de decisión, k-vecinos

Features	CTU-13/S04	CTU-13/S11	CAIDA
Protocol Quantity			
Average TTL	✓	✓	✓
TCP Window Size			
Percent TCP			✓
Percent UDP	✓	✓	
Percent DNS	✓	✓	
Percent ICMP	✓	✓	✓
Percent Others			✓
Source Privileged Ports	✓	✓	
Source Not Privileged Ports			✓
Destination Port Quantity	✓	✓	✓
Frame Length	✓	✓	
In Degree			✓
Out Degree	✓	✓	
In Degree Weight			✓
Out Degree Weight	✓	✓	
Node betweenness centrality			
Local clustering coefficient		✓	
Eigenvector centrality	✓	✓	

Marco teórico

An Empirical Analysis of Anomaly Detection Methods for Multivariate Time Series - Dongwen Li et al (2023)

- Evaluaron múltiples datasets con distintos intervalos de tiempo, contextos y fuentes

TABLE I: Detailed information of the experimental datasets. (The symbol '#' denotes the amount of data, while the symbol '%' denotes the percentage of anomalies.)

Dataset	Source	Scenarios	#Entities	#Metrics	Time Interval	#Train	#Test	Anomalies (%)
D1	A global content service provider	Web services.	26	49	30 sec	14400	23040	0.05
D2	An Internet service provider	Network operation service.	107	22	15 min	672	672	0.02
SMD	An Internet company	/	28	38	1 min	28479	28479	0.04
ASD	An Internet company	/	12	19	5 min	8640	4320	0.05
SMAP	NASA	Global measurements of soil moisture and its freeze-thaw status.	54	25	1 min	2818	7331	0.13
MSL	NASA	The Mars rover Curiosity's operations.	27	55	1 min	4308	6100	0.11
SWaT	A water treatment plant	The real-world industrial water treatment plant operation status.	1	51	1 sec	496800	449919	0.12
WADI	A testbed	A single plant operation status.	1	123	1 sec	1048571	172801	0.06

Marco teórico

An Empirical Analysis of Anomaly Detection Methods for Multivariate Time Series - Dongwen Li et al (2023)

- Evaluaron múltiples modelos de detección de anomalías

TABLE II: An overview of unsupervised MTS anomaly detection models.

Model	Advantages	Data Preprocessing Method	Model Structures
DAGMM	• Based on time point. • Preserves the low-dimensional features and reconstruction error for anomaly detection.	Does standardization.	AE + Gaussian Mixture Model (GMM)
USAD	• Leverages the advantages of AE and adversarial training. • A straightforward model structure and a limited number of parameters.	Does standardization.	AE + Generative Adversarial Network (GAN)
OmniAnomaly	• Models the explicit temporal dependence. • Employs a VAE to map input observations to stochastic variables.	Uses zero to fill in missing values and does normalization.	RNN + VAE
DOMI	• Simultaneously extracts both categorical variables and low-dimensional data features. • Works better with MTS data that exhibits multiple normal patterns.	Uses zero to fill in missing values and does standardization.	1D-CNN + Gaussian Mixture Variational AE (GMVAE)
SDFVAE	• Be capable of explicitly learning the representations of time-invariant and time-varying characteristics.	Does normalization	CNN + RNN + VAE
InterFusion	• Employs a hierarchical VAE (HVAE) to learn different features independently. • Learns both low-dimensional inter-metric and temporal embeddings.	Uses zero to fill in missing values, remove extreme values and does standardization.	1D-CNN + RNN+ VAE
JumpStarter	• Clusters univariate time series in MTS. • Reconstructs MTS based on compressed sensing. • Effectively reduces initialization time.	Does normalization.	Clustering + Compressed Sensing
GDN	• Uses an attention-based GNN to learn the inter-metric dependence.	Uses mean values or zero to fill in missing values and does normalization.	Attention + GNN

Marco teórico

An Empirical Analysis of Anomaly Detection Methods for Multivariate Time Series - Dongwen Li et al (2023)

- Resultados

TABLE VIII: The performance of using the recommended algorithms.

Model/Strategy	F_1
DAGMM	0.8499
USAD	0.8269
OmniAnomely	0.7380
DOMI	0.8372
SDFVAE	0.8657
InterFusion	0.8878
JumpStarter	0.7855
GDN	0.8823
Smoothness	0.9008
Periodicity	0.8906
Metric correlation	0.8889
Anomaly types	0.8624

Gracias