

Reporte de Análisis de Dataset

Archivo analizado: CTU-IoT-Malware-Capture-8-1conn.log.labeled.csv

Columnas y tipos de datos:

ts	float64
uid	object
id.orig_h	object
id.orig_p	int64
id.resp_h	object
id.resp_p	int64
proto	object
service	object
duration	object
orig_bytes	object
resp_bytes	object
conn_state	object
local_orig	object
local_resp	object
missed_bytes	int64
history	object
orig_pkts	int64
orig_ip_bytes	int64
resp_pkts	int64
resp_ip_bytes	int64
tunnel_parents	object
label	object
detailed-label	object

Valores nulos por columna:

ts	0
uid	0
id.orig_h	0
id.orig_p	0
id.resp_h	0
id.resp_p	0
proto	0
service	10403
duration	6185
orig_bytes	6185
resp_bytes	6185
conn_state	0
local_orig	10403
local_resp	10403
missed_bytes	0
history	0
orig_pkts	0
orig_ip_bytes	0
resp_pkts	0
resp_ip_bytes	0
tunnel_parents	10403

label	0
detailed-label	2181

Valores únicos por columnas relevantes:

id.orig_h: 2 únicos
id.orig_p: 2059 únicos
id.resp_h: 9 únicos
id.resp_p: 3 únicos
proto: 2 únicos
service: 1 únicos
duration: 2056 únicos
orig_bytes: 3 únicos
resp_bytes: 3 únicos
conn_state: 3 únicos
missed_bytes: 1 únicos
history: 4 únicos
orig_pkts: 3 únicos
orig_ip_bytes: 4 únicos
resp_pkts: 2 únicos
resp_ip_bytes: 3 únicos
label: 2 únicos
detailed-label: 2 únicos

Estadísticas descriptivas (numéricas):

missed_bytes: mean=0.00, min=0.0, max=0.0

orig_pkts: mean=1.40, min=0.0, max=3.0

orig_ip_bytes: mean=87.06, min=0.0, max=180.0

resp_pkts: mean=0.21, min=0.0, max=1.0

resp_ip_bytes: mean=15.81, min=0.0, max=88.0

Distribución de la columna 'label':

Malicious	8222
Benign	2181

Distribución de la columna 'detailed-label':

C&C	8222
-	2181

Protocolos más frecuentes (proto):

tcp	8224
udp	2179

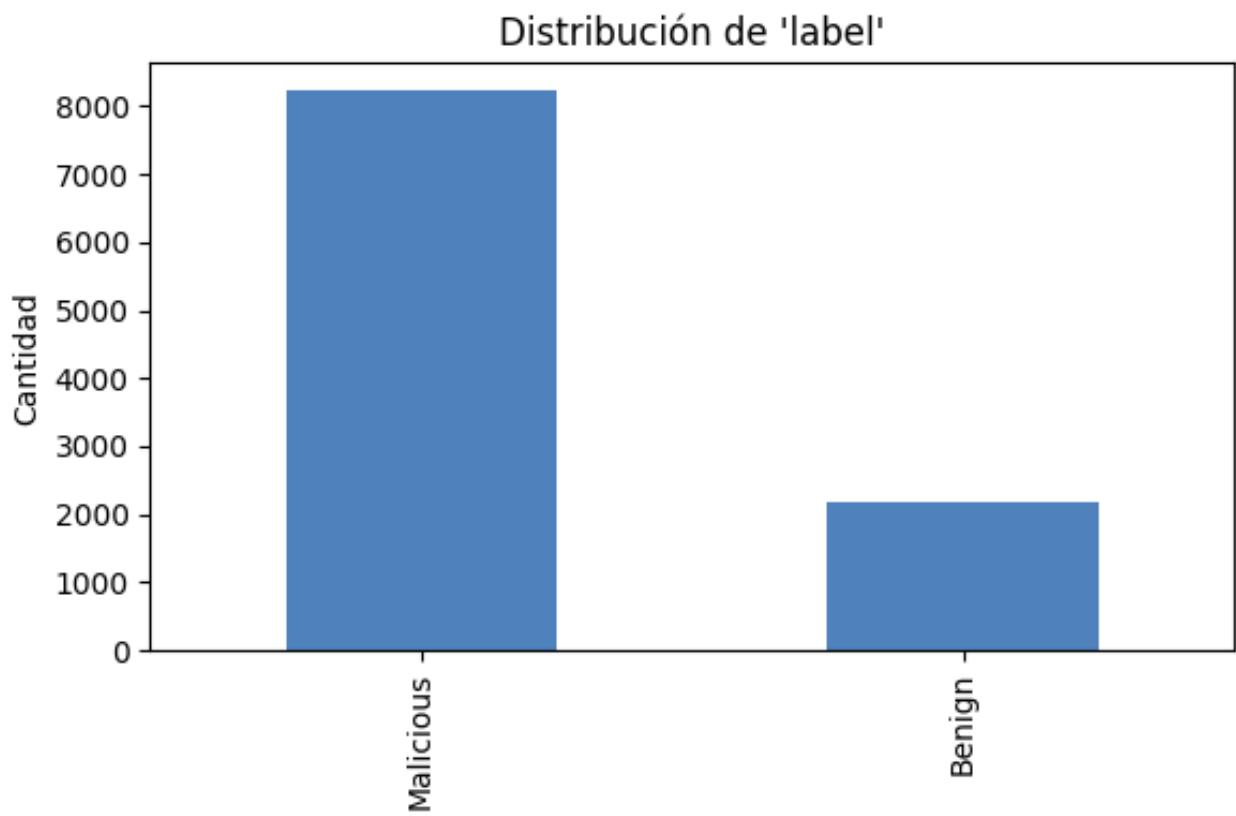
Estados de conexión más frecuentes (conn_state):

S0 8239

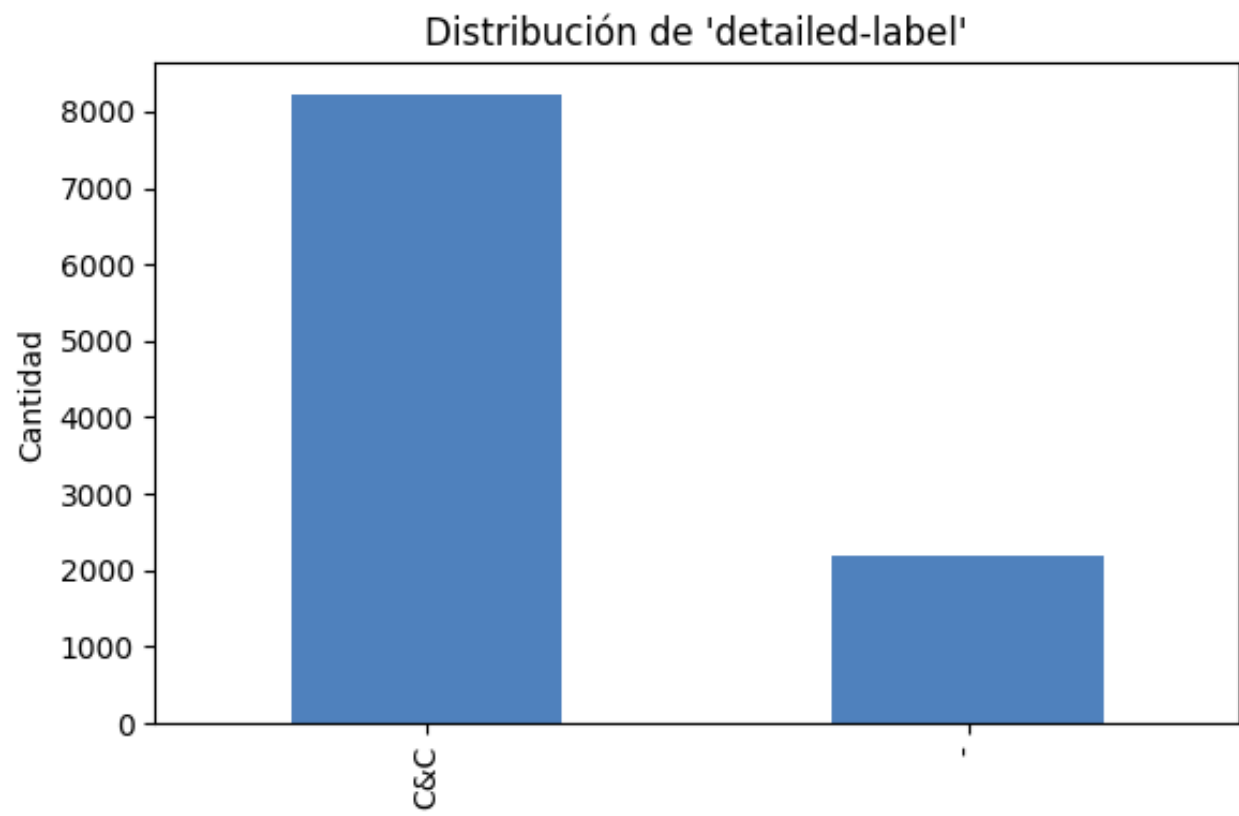
SF 2162

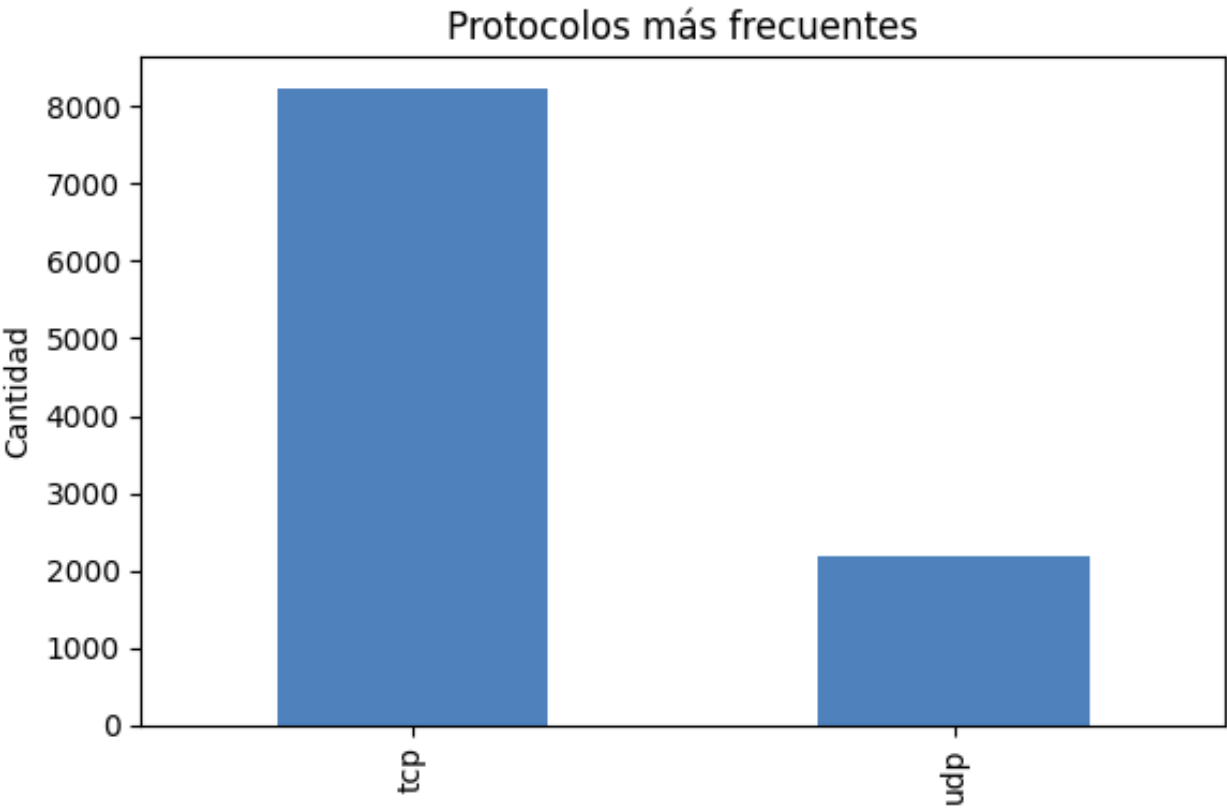
OTH 2

Label dist



Detailed label dist





Conn state dist

