

Reporte de Análisis de Dataset

Archivo analizado: CTU-IoT-Malware-Capture-20-1conn.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	2617
duration	885
orig_bytes	885
resp_bytes	885
conn_state	0
local_orig	3209
local_resp	3209
missed_bytes	0
history	0
orig_pkts	0
orig_ip_bytes	0
resp_pkts	0
resp_ip_bytes	0
tunnel_parents	3209

label	0
detailed-label	3193

Valores únicos por columnas relevantes:

id.orig_h: 3 únicos
id.orig_p: 595 únicos
id.resp_h: 40 únicos
id.resp_p: 5 únicos
proto: 2 únicos
service: 2 únicos
duration: 1193 únicos
orig_bytes: 19 únicos
resp_bytes: 19 únicos
conn_state: 5 únicos
missed_bytes: 4 únicos
history: 8 únicos
orig_pkts: 10 únicos
orig_ip_bytes: 20 únicos
resp_pkts: 10 únicos
resp_ip_bytes: 20 únicos
label: 2 únicos
detailed-label: 2 únicos

Estadísticas descriptivas (numéricas):

missed_bytes: mean=2.39, min=0.0, max=4500.0

orig_pkts: mean=6.49, min=0.0, max=8872.0

orig_ip_bytes: mean=310.30, min=0.0, max=372893.0

resp_pkts: mean=6.35, min=0.0, max=9307.0

resp_ip_bytes: mean=291.97, min=0.0, max=388975.0

Distribución de la columna 'label':

Benign	3193
Malicious	C&C 16

Distribución de la columna 'detailed-label':

-	3193
Torii	16

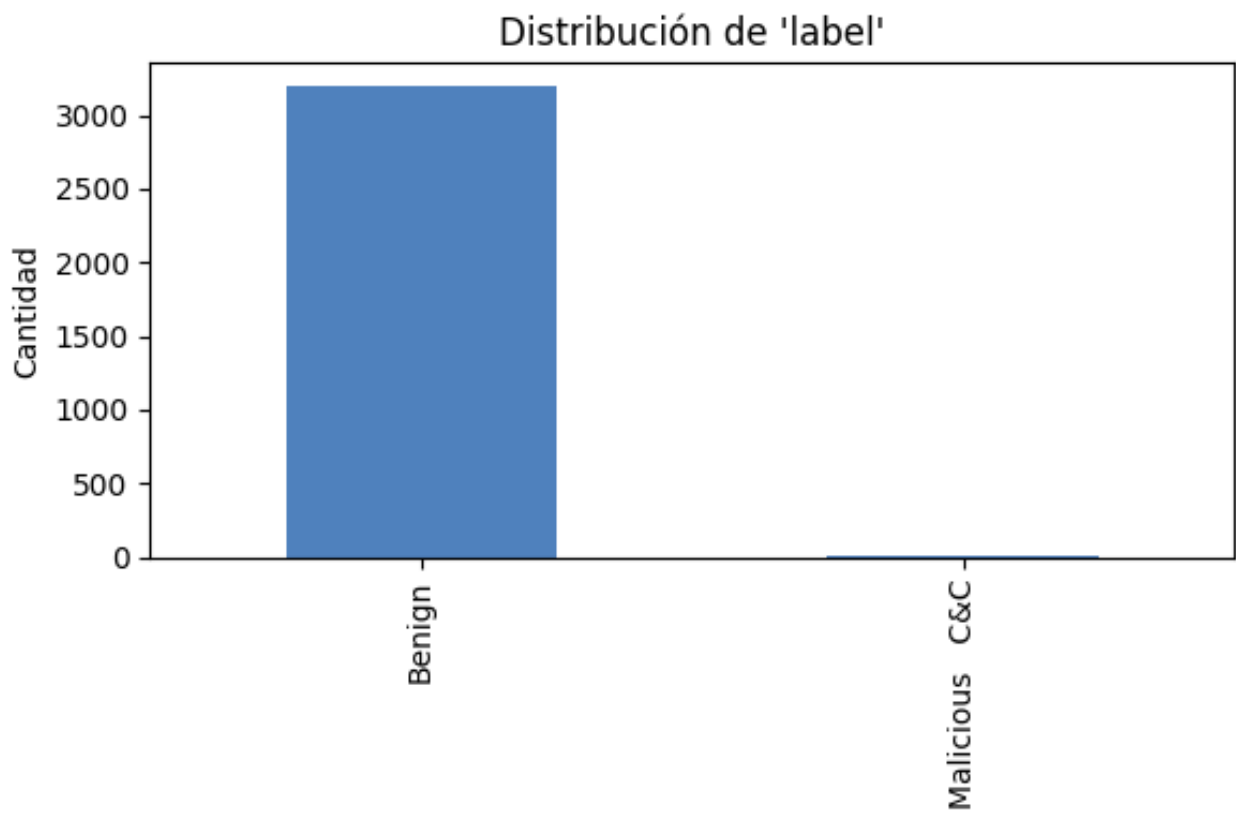
Protocolos más frecuentes (proto):

udp	3192
tcp	17

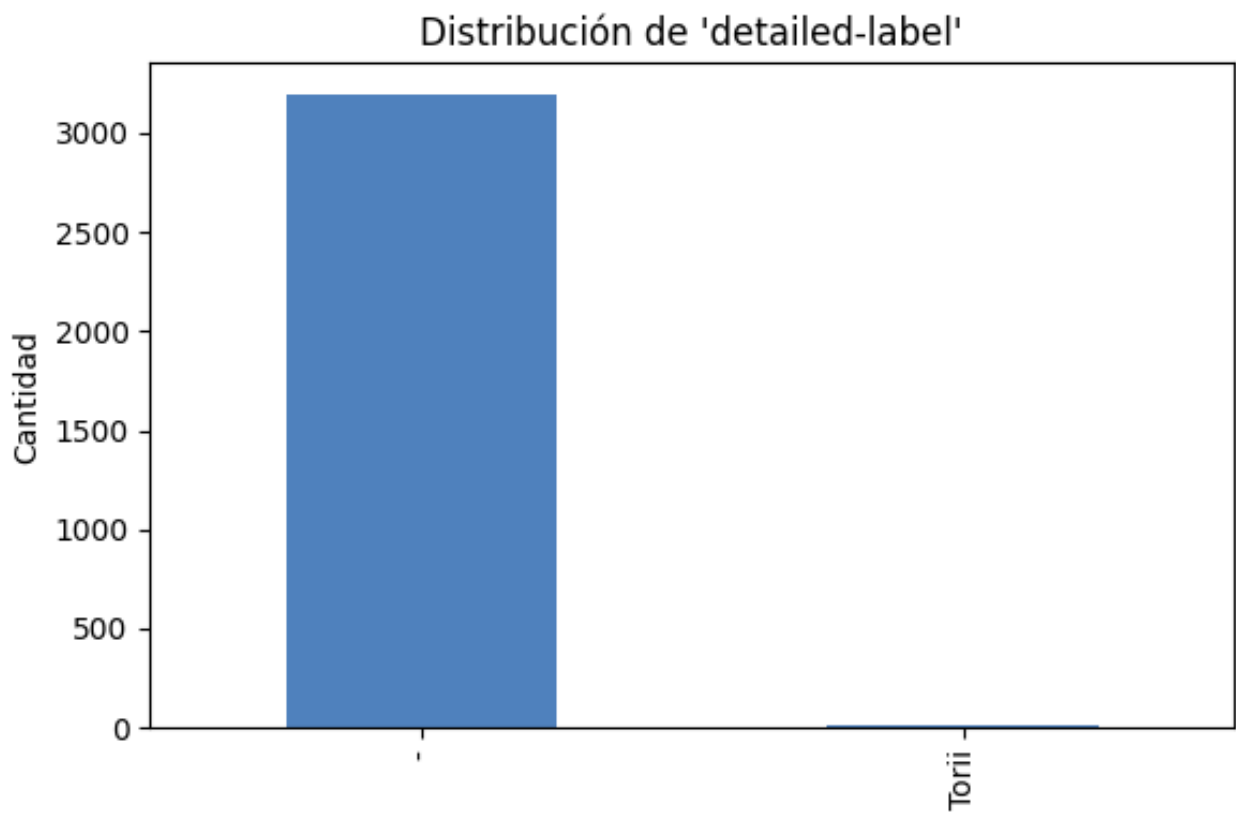
Estados de conexión más frecuentes (conn_state):

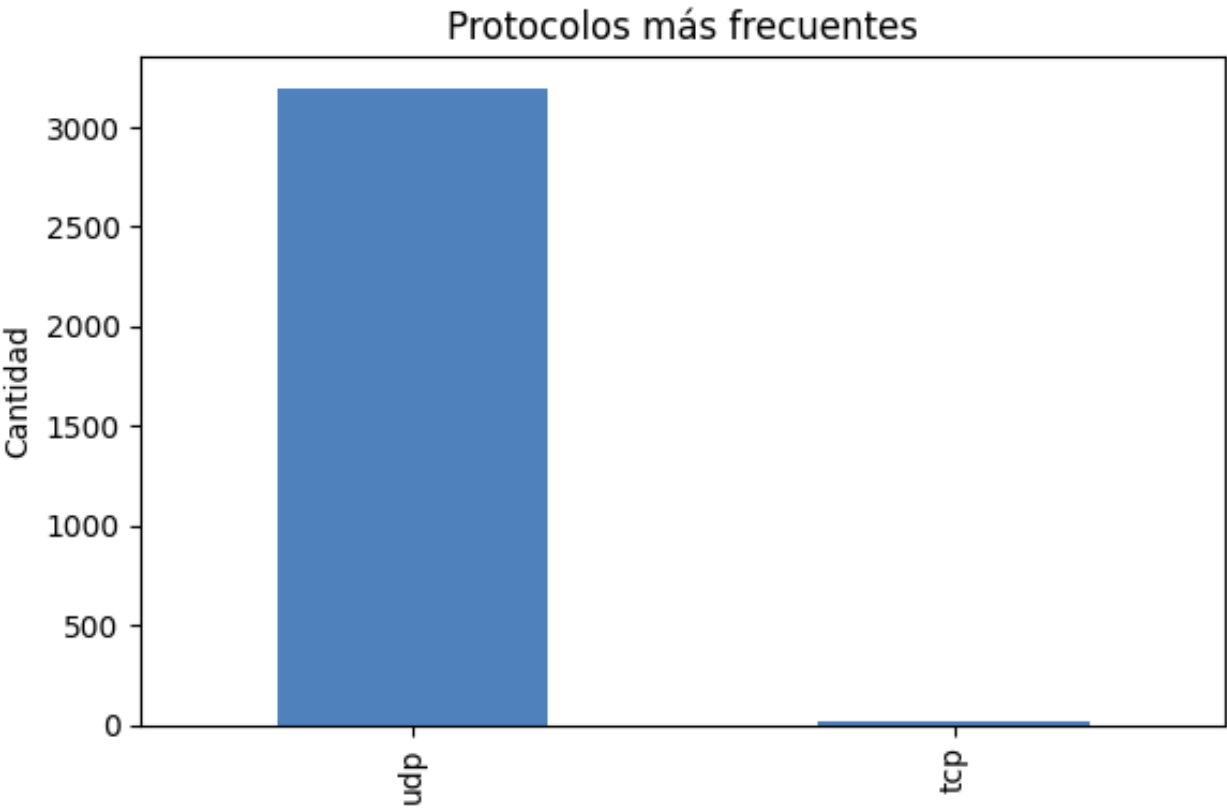
SF	2132
S0	1074
OTH	1
RSTR	1
S1	1

Label dist



Detailed label dist





Conn state dist

