

Reporte de Análisis de Dataset

Archivo analizado: CTU-IoT-Malware-Capture-42-1conn.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	2710
duration	992
orig_bytes	992
resp_bytes	992
conn_state	0
local_orig	4426
local_resp	4426
missed_bytes	0
history	0
orig_pkts	0
orig_ip_bytes	0
resp_pkts	0
resp_ip_bytes	0
tunnel_parents	4426

label 0
detailed-label 4423

Valores únicos por columnas relevantes:

id.orig_h: 3 únicos
id.orig_p: 3058 únicos
id.resp_h: 46 únicos
id.resp_p: 8 únicos
proto: 2 únicos
service: 5 únicos
duration: 1994 únicos
orig_bytes: 42 únicos
resp_bytes: 41 únicos
conn_state: 6 únicos
missed_bytes: 4 únicos
history: 17 únicos
orig_pkts: 18 únicos
orig_ip_bytes: 79 únicos
resp_pkts: 16 únicos
resp_ip_bytes: 56 únicos
label: 3 únicos
detailed-label: 2 únicos

Estadísticas descriptivas (numéricas):

missed_bytes: mean=2.64, min=0.0, max=4380.0

orig_pkts: mean=3.04, min=1.0, max=654.0

orig_ip_bytes: mean=191.07, min=57.0, max=42791.0

resp_pkts: mean=2.10, min=0.0, max=348.0

resp_ip_bytes: mean=281.40, min=0.0, max=78523.0

Distribución de la columna 'label':

Benign	4420
Malicious FileDownload	3
Malicious C&C	3

Distribución de la columna 'detailed-label':

-	4420
FileDownload	3

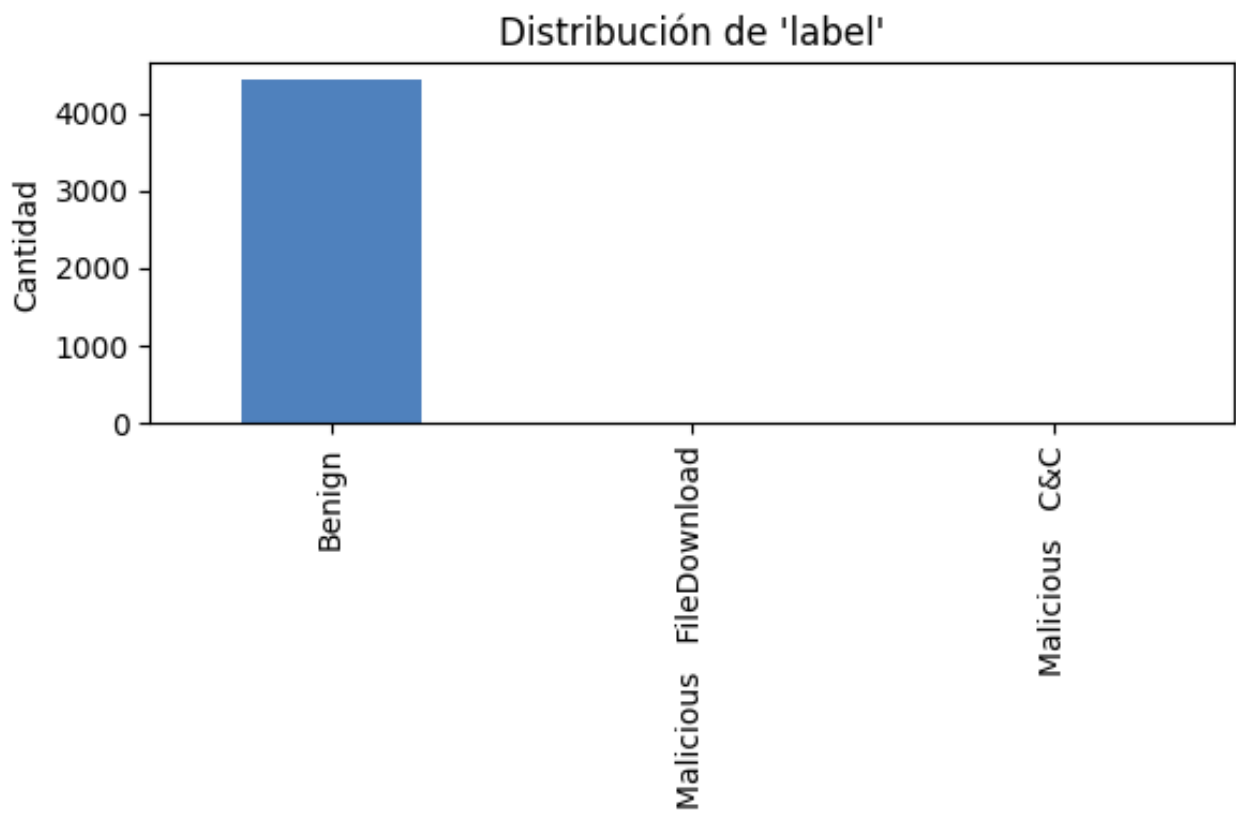
Protocolos más frecuentes (proto):

udp	2425
tcp	2001

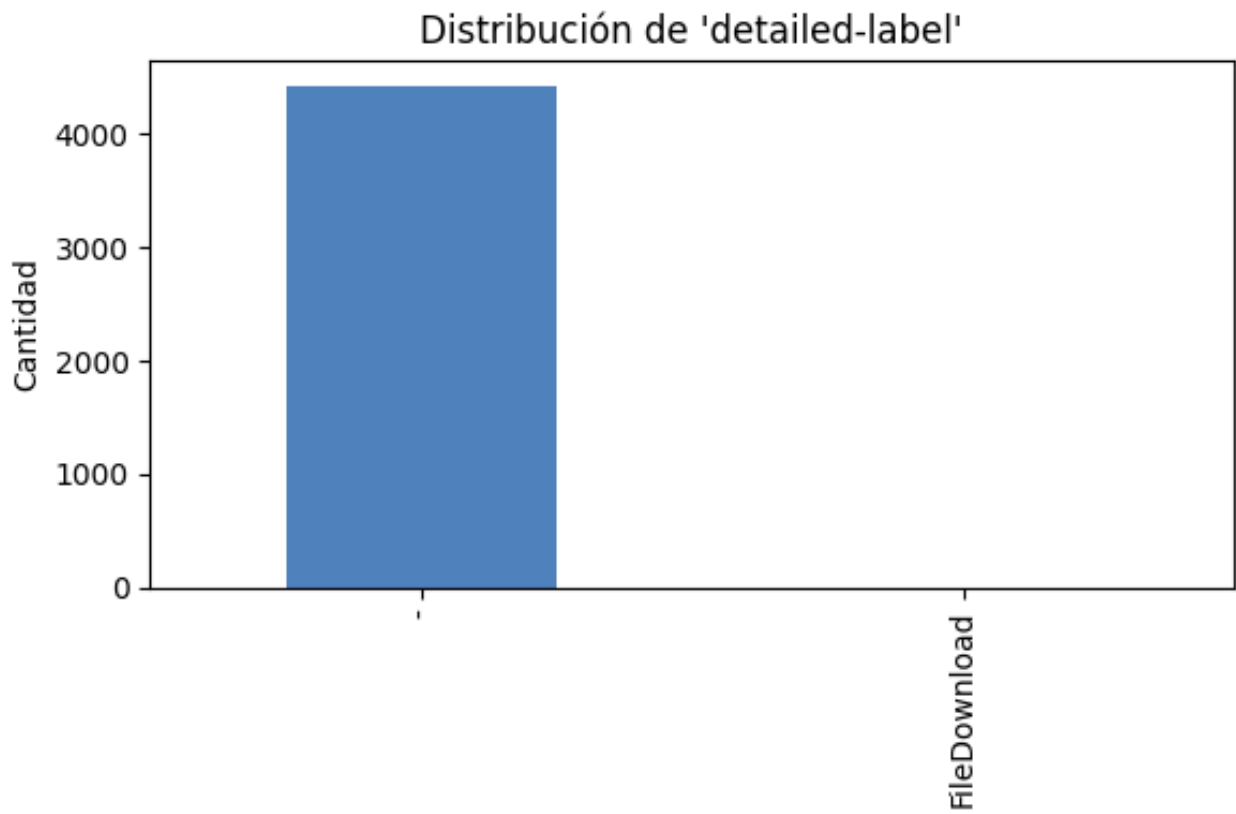
Estados de conexión más frecuentes (conn_state):

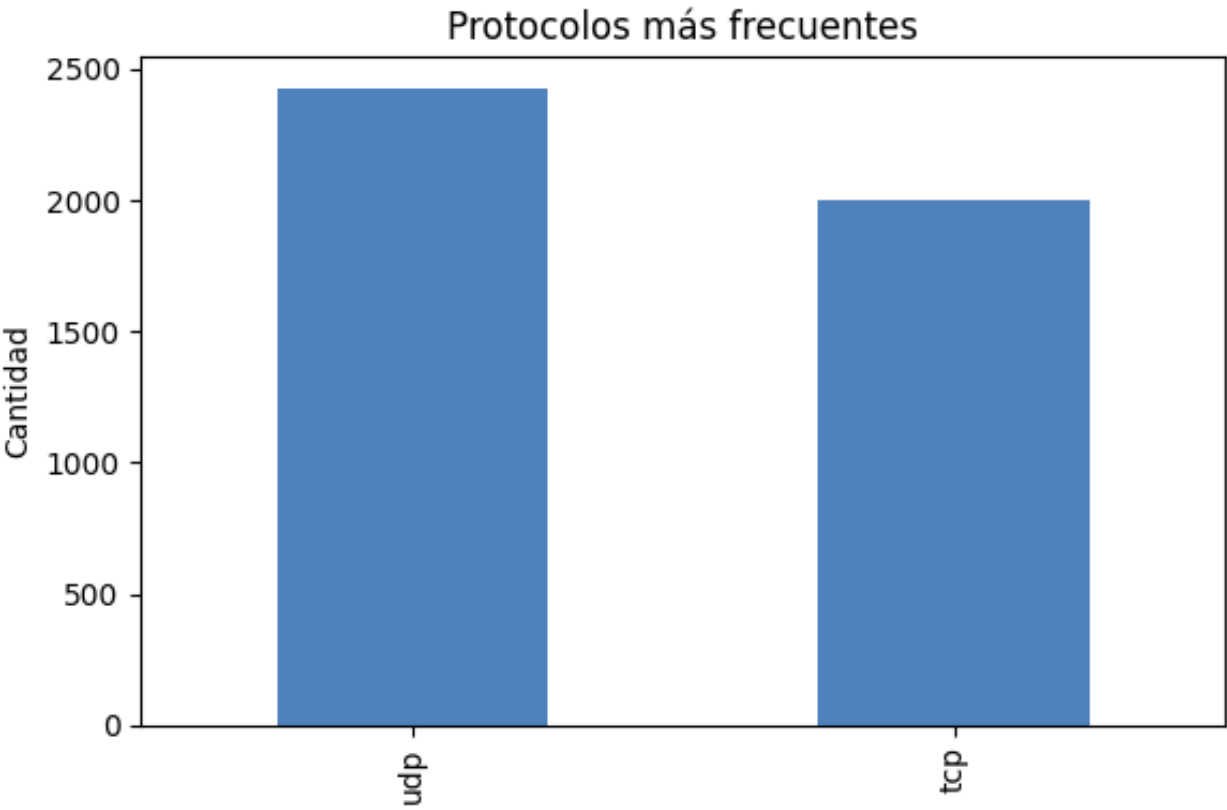
SF	1837
S0	1565
S3	873
RSTR	146
OTH	4
RSTO	1

Label dist



Detailed label dist





Conn state dist

