Identificación de dispositivos IoT a través de huellas hardware

Autor:
Sergio Marín Sánchez

Tutores:
Gregorio Martínez Pérez
Pedro Miguel Sánchez Sánchez

UNIVERSIDAD DE MURCIA

Índice

Introducción

UNIVERSIDAD DE MURCIA

- Los dispositivos IoT traen consigo numerosos riesgos de seguridad.
- Métodos tradicionales de identificación:
  - Dirección IP.
  - Dirección MAC.
  - Certificados digitales.
- Limitaciones:
  - Direccionamiento dinámico.
  - Modificaciones del usuario.
  - Robo de credenciales.

Introducción

UNIVERSIDAD DE MURCIA

- Identificación individual de los dispositivos.
  - Mayor seguridad.
- Diferencias en la fabricación de los dispositivos.
  - Diferentes tiempos en realizar un mismo proceso.
  - Desviaciones en el reloj interno.
- Objetivos de este trabajo:
  - Diseñar una arquitectura para este sistema.
  - Recolectar marcas de tiempo de los dispositivos.
  - Generar huellas estadísticas de los dispositivos.
  - Machine Learning para evaluar el proceso de identificación.

Estado del arte

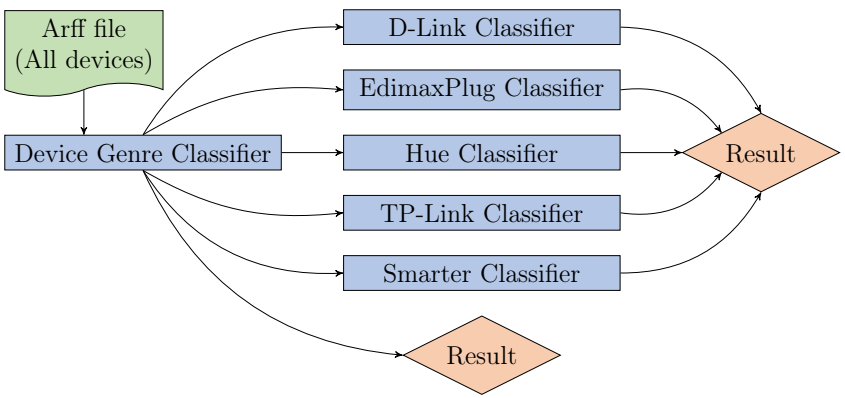
UNIVERSIDAD DE MURCIA

Referencia	Tipo de identificación	Enfoque	Tipo de aprendizaje	Resultados
Pascal Oser et al.	Tipo de dispositivo	Machine Learning	Supervisado	99.76 % de accuracy y 97.03 % de precisión
Salma Hamad et al.	Individual	Machine Learning	Supervisado	89 % de accuracy
Ahmet Aksoy et al.	Tipo y modelo del dispositivo	Machine Learning	Supervisado	Entre 42.2 % y 100 % de accuracy, con un promedio de 82 %
Hossein Jafari et al.	Individual	Machine Learning	Supervisado	96.3 % de accuracy en DNN, 94.7 % de accuracy en CNN y 76 % de accuracy en LSTM
Fabian Lanze et al.	Modelo del dispositivo	Análisis estadístico (regresión lineal)	-	Método no válido para identificar unívocamente un dispositivo.
Yair Meidan et al.	Tipo y modelo	Machine Learning	Supervisado	99.28 % de accuracy
Loh Chin Choong Desmond et al.	Individual	Machine Learning	No supervisado	Entre un 70 % y 80 % de accuracy
Este trabajo	Individual	Machine Learning	Supervisado y No supervisado	99.38 % de Accuracy, 99.39 % de Recall y 99.38 % de f-score

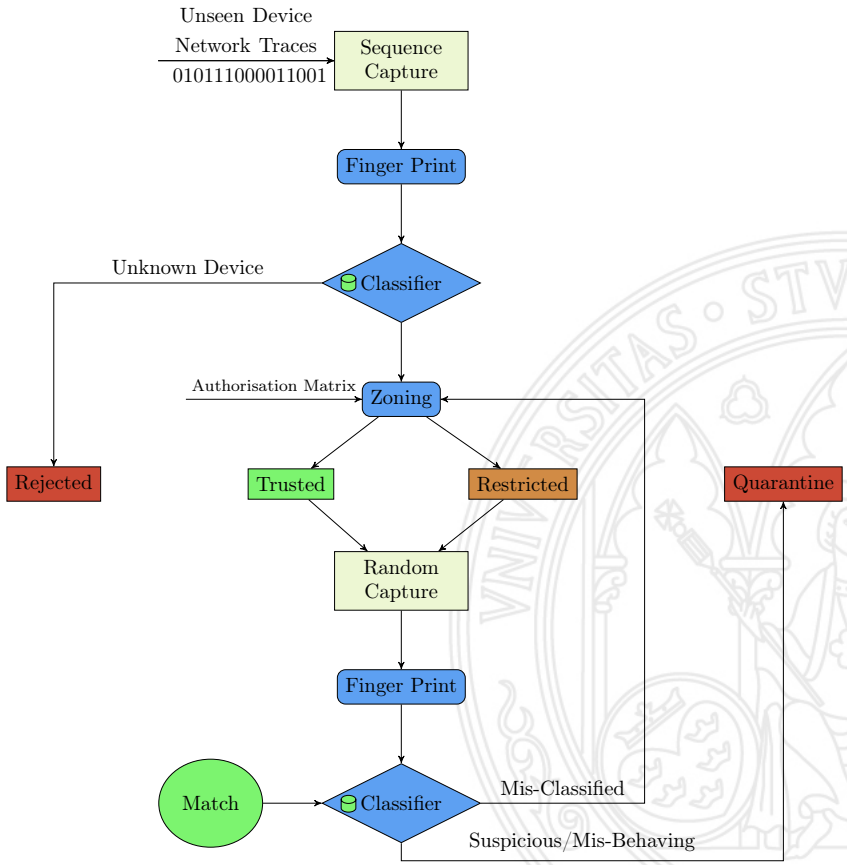
Tabla: Resultados en el estado del arte

Estado del arte

UNIVERSIDAD DE MURCIA



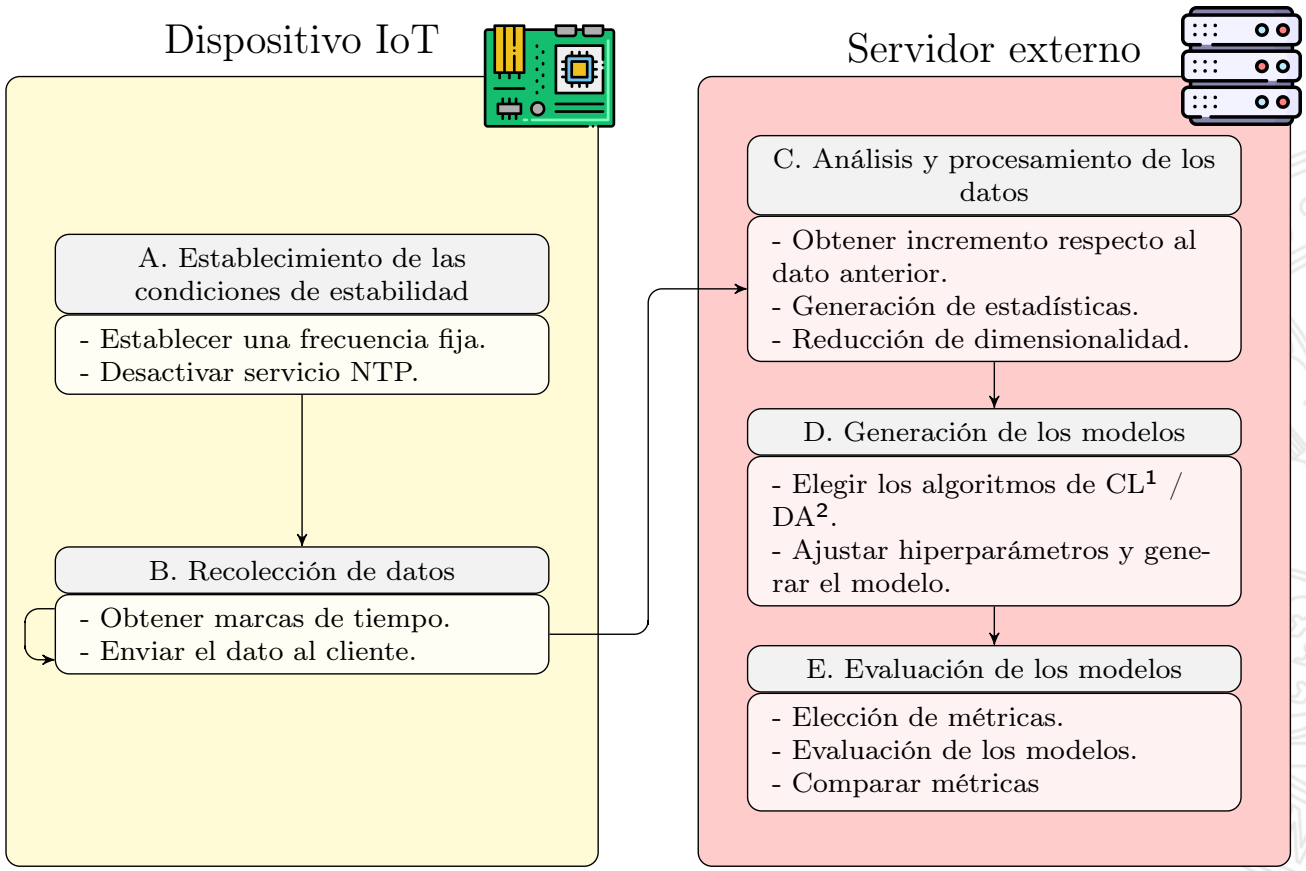
(a) Propuesta de Ahmet Aksoy et al. para identificar modelos.



(b) Propuesta de Salma Hamad et al. para identificar dispositivos.

UNIVERSIDAD DE MURCIA

- Arquitectura propuesta.



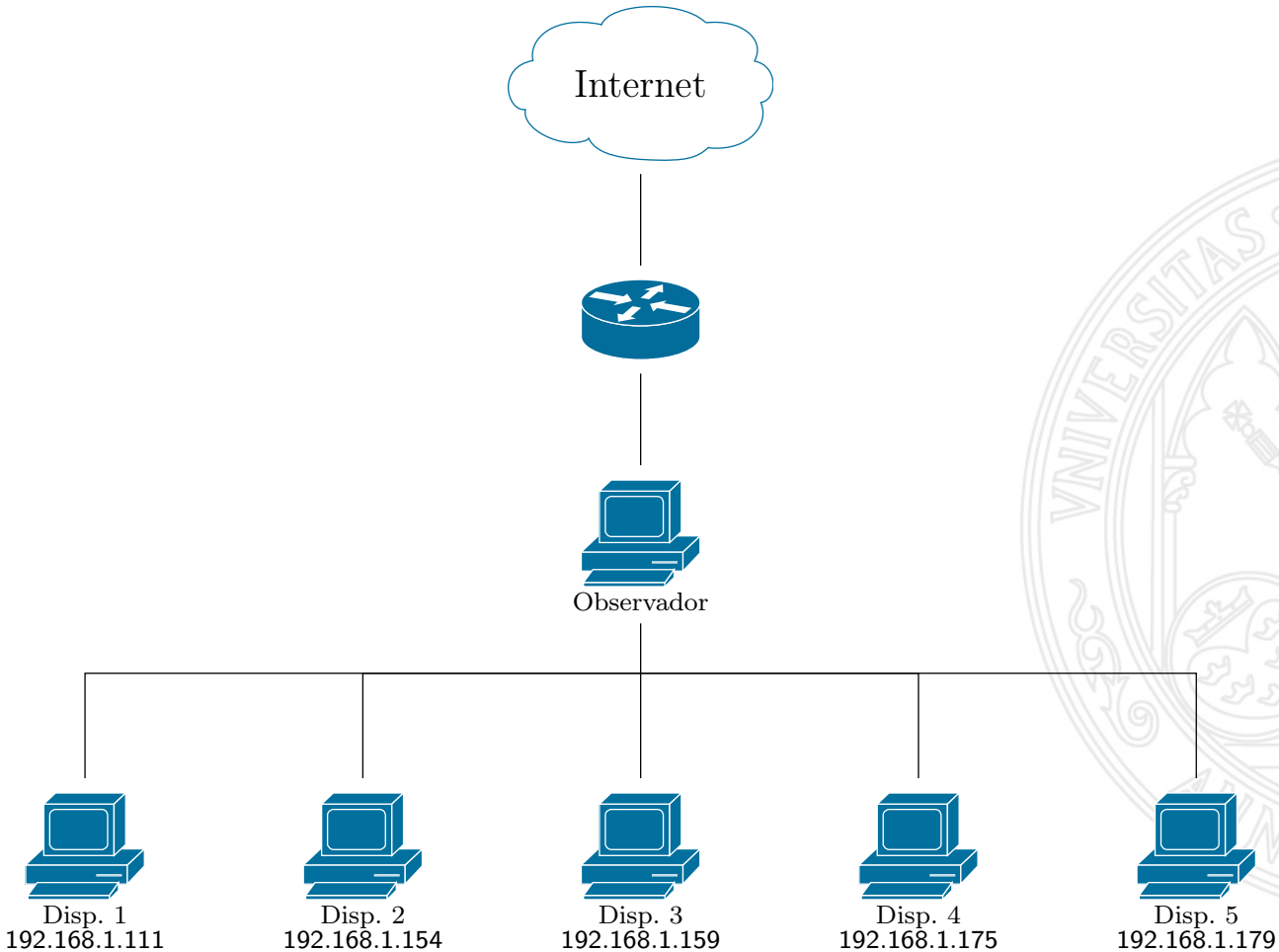
1: Clasificación
2: Detección de anomalías

Diseño

Topología

UNIVERSIDAD DE MURCIA

- Topología de la red.

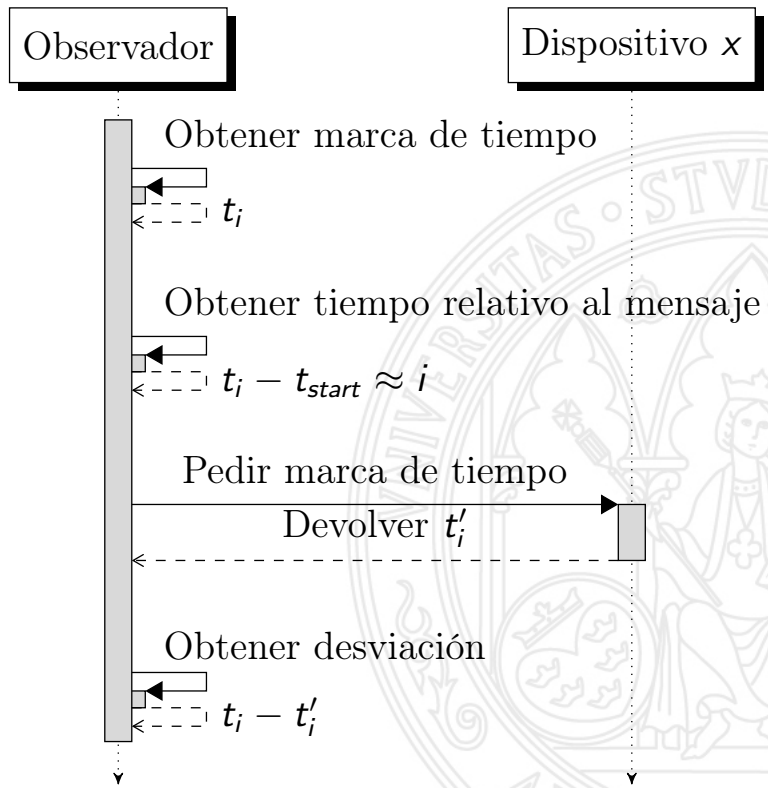


Diseño

Recolección de datos

UNIVERSIDAD DE MURCIA

- La marca de tiempo relativa a cada mensaje desde el inicio, $t_i - t_{start}$.
- La marca de tiempo absoluta del observador t_i.
- La marca de tiempo absoluta del dispositivo t'_i.
- La desviación del reloj del dispositivo respecto al del observador, $t_i - t'_i$.



UNIVERSIDAD DE MURCIA

- Evaluación de los resultados.
  - Algoritmos de ML supervisados para clasificación.
  - Algoritmos de ML no supervisados para detección de anomalías.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$f - score = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

$$Recall = \frac{TP}{TP + FN}$$

$$TNR = \frac{TN}{TN + FP}$$

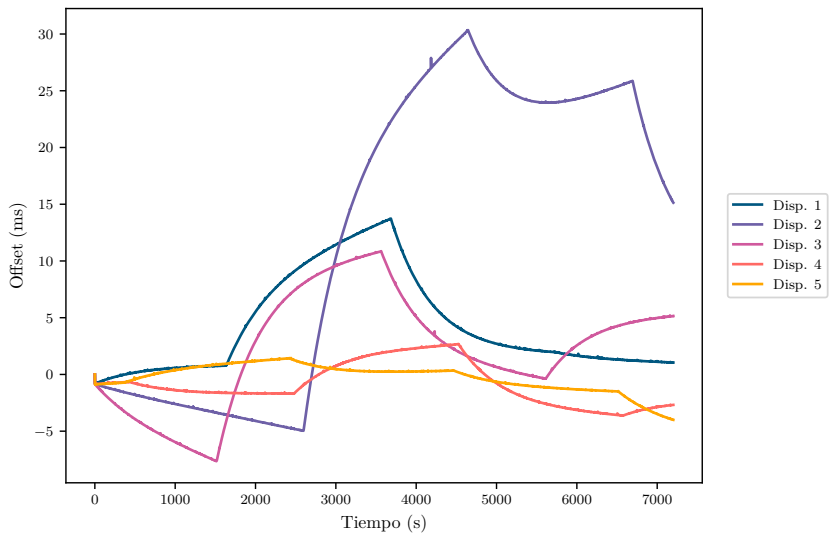
Experimentos

Comparativa de los experimentos

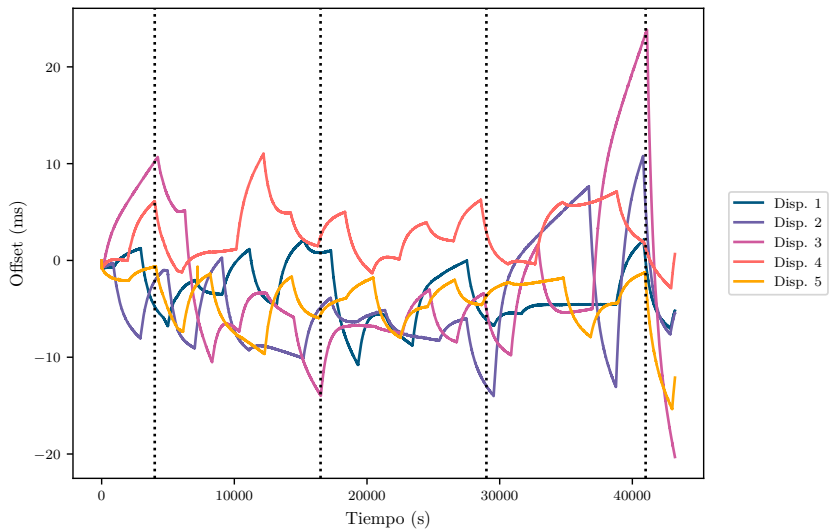
UNIVERSIDAD DE MURCIA

time	TSrock	TSrasp	offset	device
292	119238112796030	104592709716803	-14645403079227	192.168.1.111
1001191222	119239113986960	104593710167425	-14645403819535	192.168.1.111
2001485862	119240114281600	104594710453699	-14645403827901	192.168.1.111
⋮	⋮	⋮	⋮	⋮

Tabla: Ejemplo de los datos obtenidos de cada dispositivo



(a) Desviación acumulada muestra secuencial

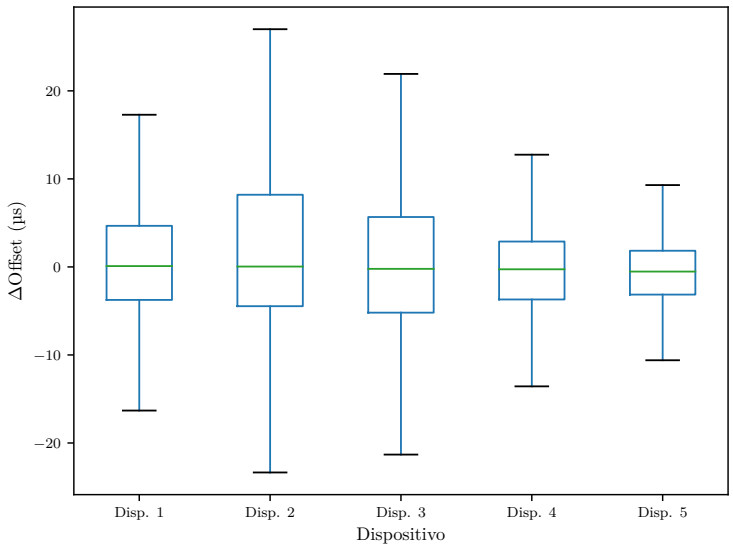


(b) Desviación acumulada muestra paralela

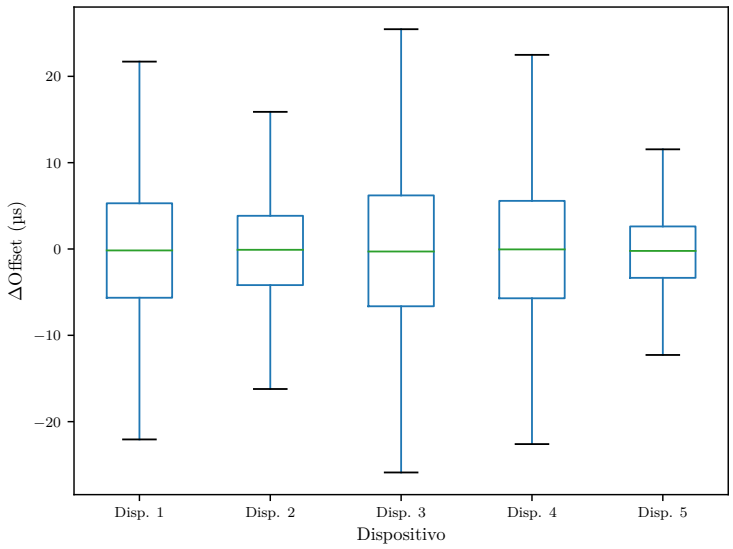
Experimentos

Comparativa de los experimentos

UNIVERSIDAD DE MURCIA



(a) Desviación muestra secuencial



(b) Desviación muestra paralela

Experimentos

Generación de estadísticas y reducción de la dimensionalidad

UNIVERSIDAD DE MURCIA

	Sum	Mean	Median	Mode	Std	IQR	Kurtosis	Skew	Max	Min	Device
1	-284.0	-4.733333333333333	-203.0	-10750.0	6531.321744049499	8739.5	-0.8026364427898236	0.266444555013173	12077.0	-10750.0	Disp. 1
2	-65895.0	-1098.25	106.5	-13344.0	3926.559099283938	2519.75	1.4605213340303709	-1.1040127142547507	7616.0	-13344.0	Disp. 2
3	96179.0	1602.9833333333333	815.0	-8136.0	5010.092595279735	6575.5	-0.39715065367509084	0.2484646585713819	12831.0	-8136.0	Disp. 3
4	109162.0	1819.3666666666666	2016.5	-10485.0	6159.084454763058	8290.5	-0.7264084617343212	-0.3858981208999922	11469.0	-10485.0	Disp. 4
5	-81317.0	-1355.2833333333333	-2127.0	-6378.0	3665.051390911538	2616.5	2.701193448053943	1.7089231615691112	10383.0	-6378.0	Disp. 5
6	19928.0	332.1333333333333	-147.0	-10750.0	6613.483928825726	10212.0	-0.8404647945245984	0.23473423365399895	12077.0	-10750.0	Disp. 1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Tabla: Datos estadísticos muestra paralela

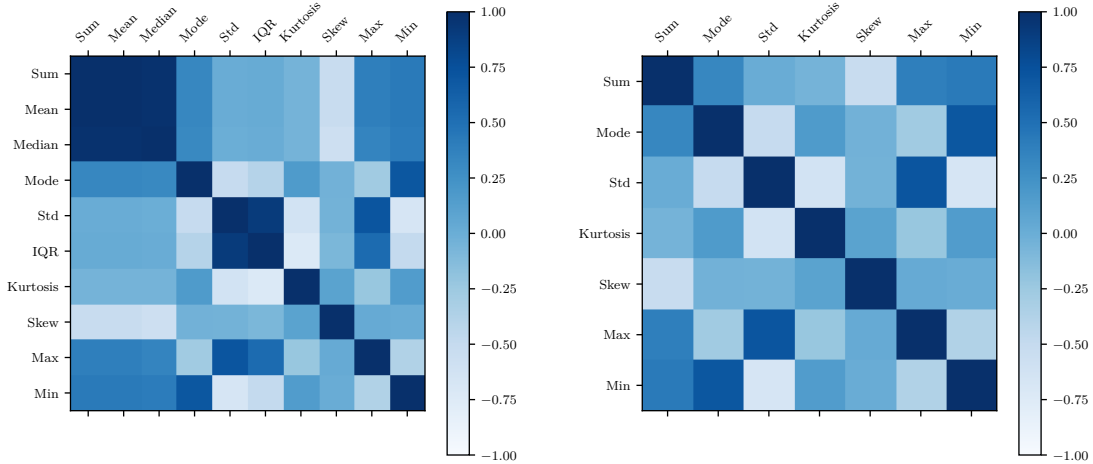
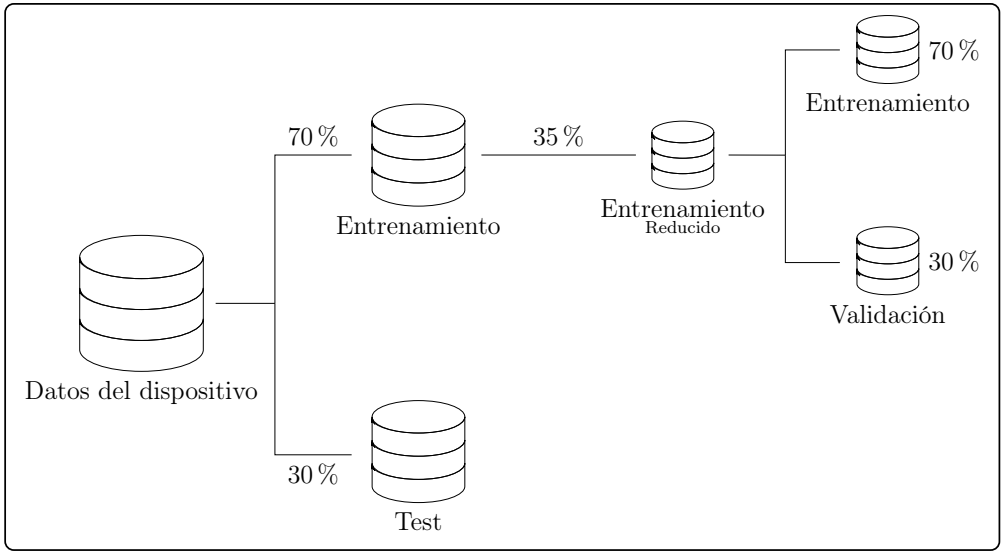


Figura: Correlación entre las variables estadísticas

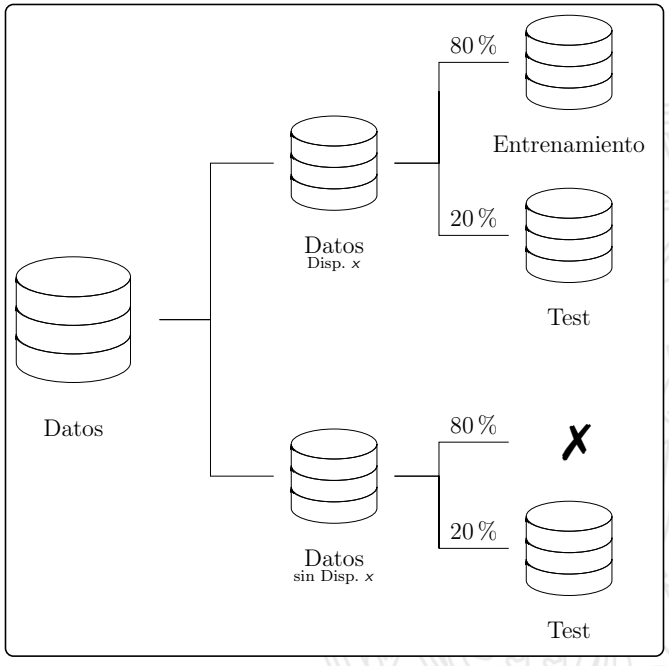
Experimentos

Particionamiento de los datos

UNIVERSIDAD DE MURCIA



(a) Particionamiento para clasificadores



(b) Particionamiento para la detección de anomalías

Resultados

Comparativa de algoritmos

UNIVERSIDAD DE MURCIA

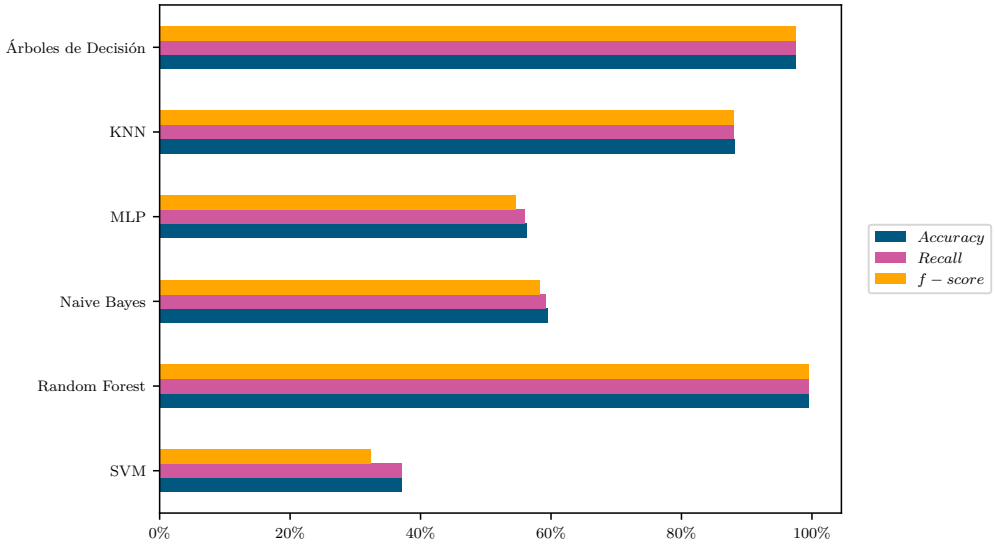


Figura: Algoritmos de clasificación

	Isolation Forest		Local Outlier Factor		OneClass-SVM	
	Recall	*TNR*	*Recall*	*TNR*	*Recall*	*TNR*
Disp. 1	95.12 %	10.36 %	98.99 %	10.68 %	49.48 %	58.24 %
Disp. 2	94.98 %	28.87 %	98.45 %	57.74 %	50.90 %	70.36 %
Disp. 3	94.80 %	13.90 %	98.73 %	13.64 %	49.84 %	52.65 %
Disp. 4	95.80 %	16.61 %	98.88 %	8.74 %	50.54 %	64.66 %
Disp. 5	94.89 %	46.99 %	98.63 %	77.98 %	49.53 %	94.59 %

Tabla: Algoritmos de detección de anomalías

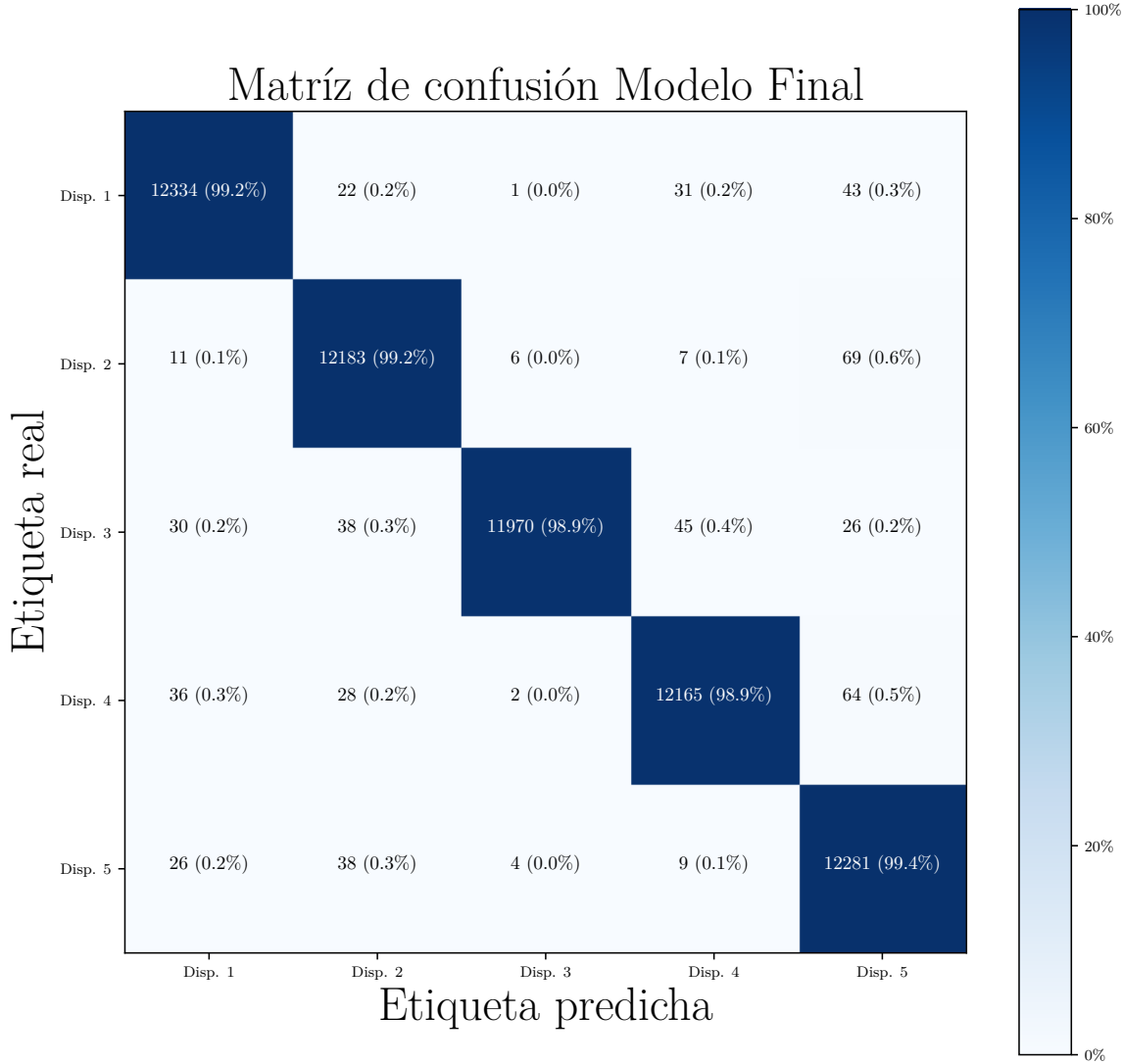
Resultados

Resultados finales

UNIVERSIDAD DE MURCIA

Resultados:

- Accuracy: 99.38 %
- f-score: 99.38 %
- Recall: 99.39 %



Conclusiones y vías futuras

UNIVERSIDAD DE MURCIA

- Conclusiones:
  - Se ha diseñado un sistema capaz de reconocer dispositivos individuales.
  - Los resultados obtenidos son similares a los del estado del arte.
  - Se ha conseguido haciendo uso únicamente de una característica hardware.
- Vías futuras:
  - Extender el sistema fuera de red local.
  - Mecanismos de identificación a tiempo real.
    - Generar huellas estadísticas de todos los dispositivos.
    - Actualización a modelos capaz de actualizarse.

Gracias por su atención.

¿Alguna pregunta?