Aprendizaje Federado con Agrupación de modelos para la Detección de Anomalías en Dispositivos IoT Heterogéneos

Xabier Sáez de Cámara, Jose Luis Flores, Cristóbal Arellano, Aitor Urbieta y Urko Zurutuza

Mondragon Unibertsitatea, IKERLAN

2022-10-20

LA TECNOLOGÍA, NUESTRA ACTITUD

ikerlan

MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

1. Contexto.
2. Aprendizaje federado.
3. Aprendizaje federado con agrupación de modelos.
4. Banco de pruebas IoT.
5. Resultados.
6. Conclusiones.

© 2022 IKERLAN. All rights reserved

Contexto

IoT susceptible a ataques:

- Puertos/Servicios innecesarios expuestos a Internet.
- Autenticación no adecuada, uso de credenciales débiles.
- Malas prácticas de programación, uso de software vulnerable.
- Gestión inadecuada de las actualizaciones.
- …

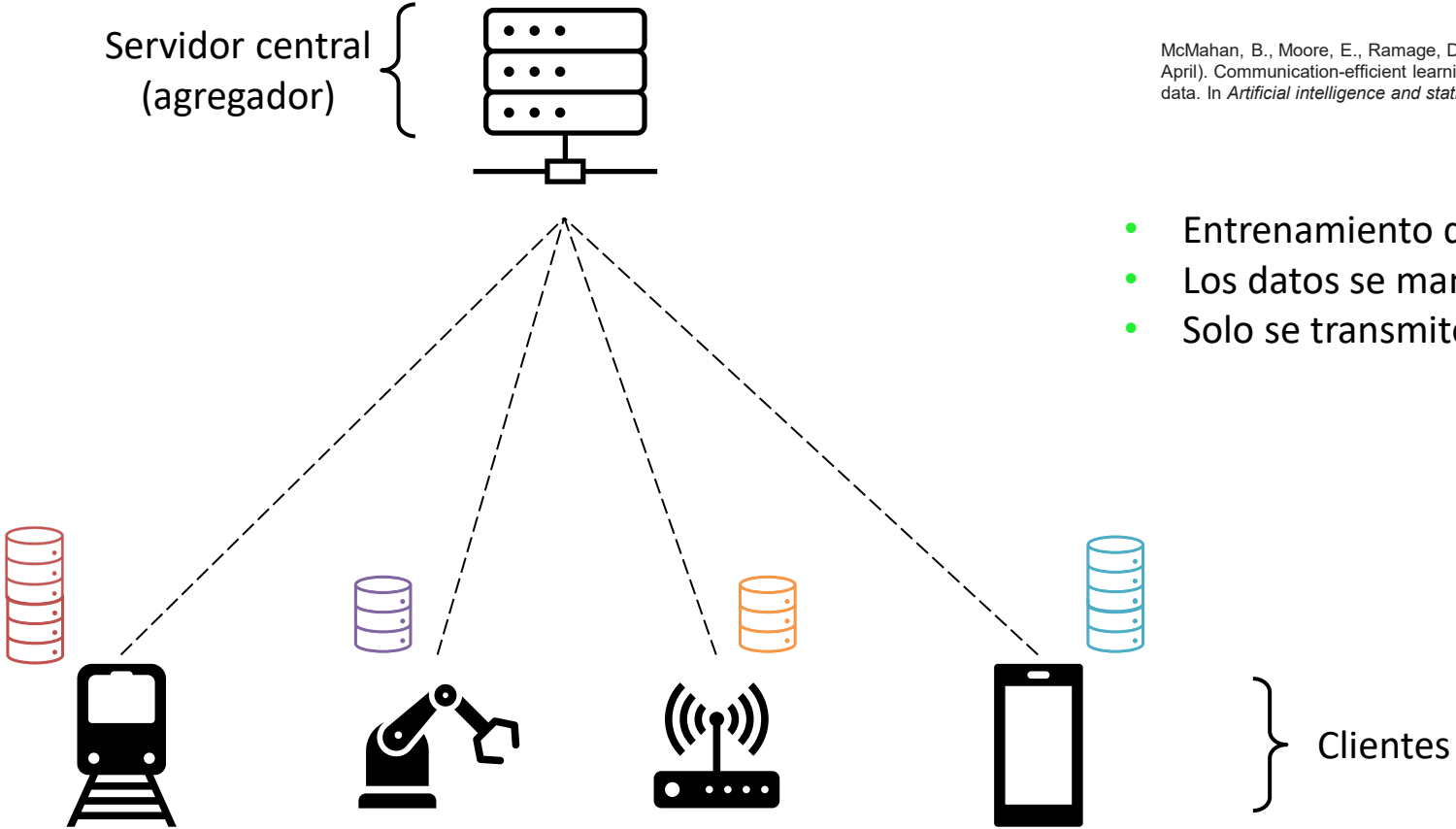
Estrategias de mitigación poco efectivas:

- Sistemas de detección de intrusiones (IDS) basados en reglas presentan dificultades en IoT:
  - Uso de técnicas de ofuscación en el malware.
  - Rápida evolución de la infraestructura del malware en IoT (del orden de 10 días).
  - Ventanas de exposición largas entre el descubrimiento del malware y la publicación de reglas para detectarlas (del orden de 100 días).

© 2022 IKERLAN. All rights reserved

Federated Learning (FL)

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.

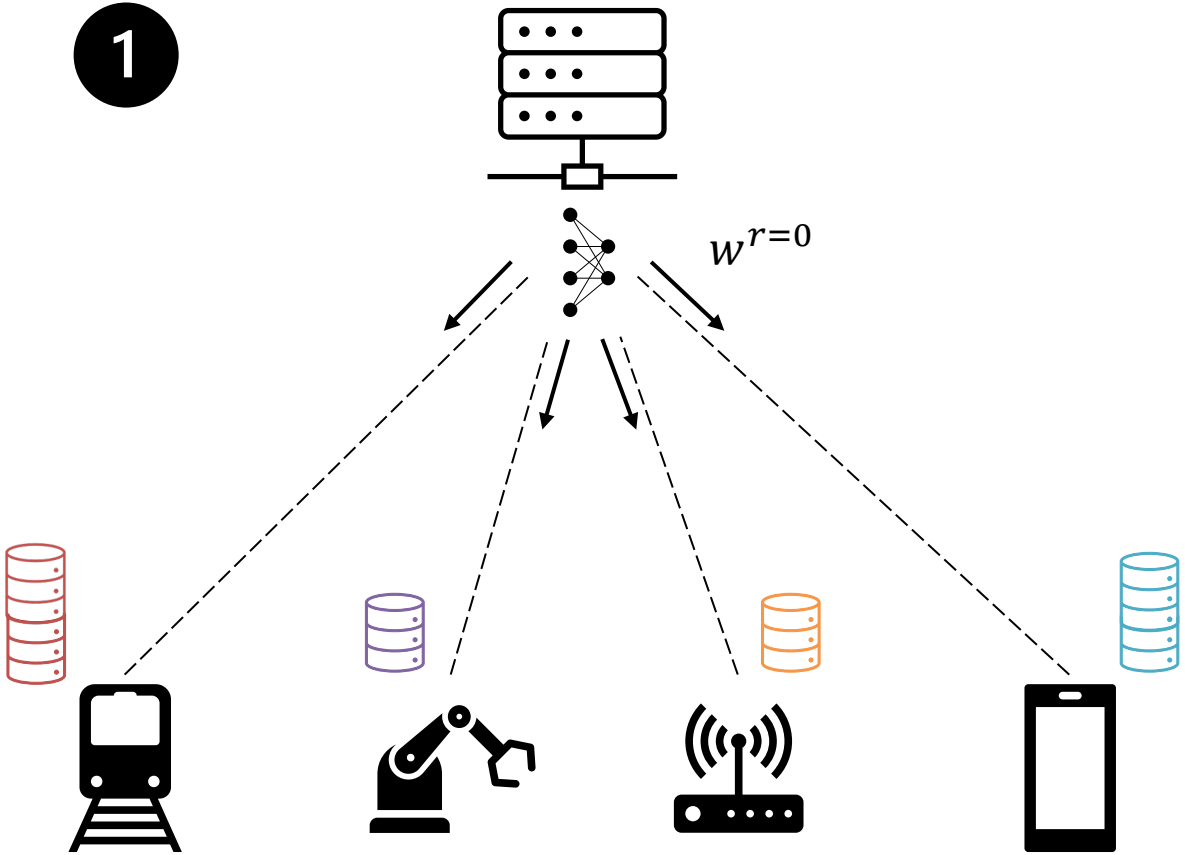


- Entrenamiento distribuido.
- Los datos se mantienen en local.
- Solo se transmiten los modelos.

© 2022 IKERLAN. All rights reserved

Federated Learning (FL)

1



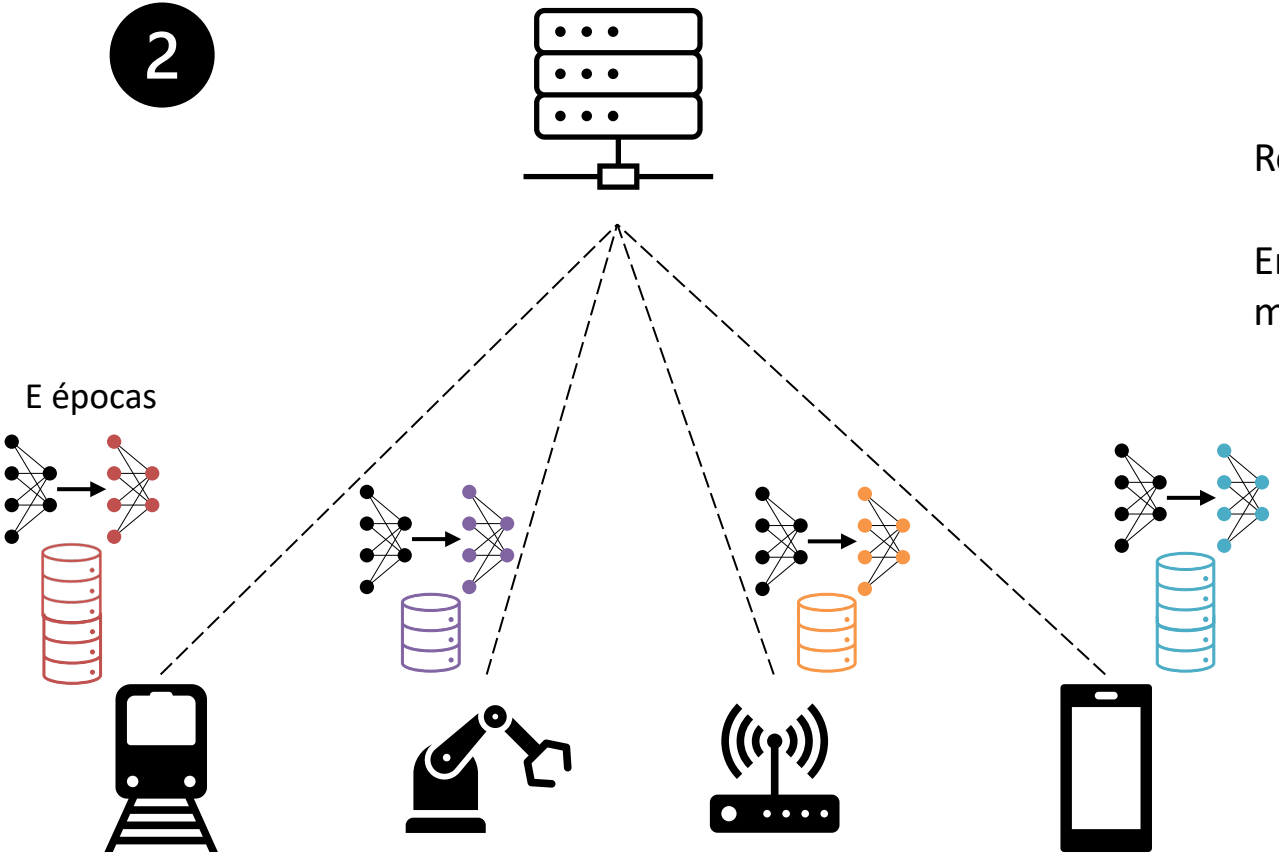
Ronda inicial $r = 0$

Inicialización y distribución del modelo.

© 2022 IKERLAN. All rights reserved

Federated Learning (FL)

2



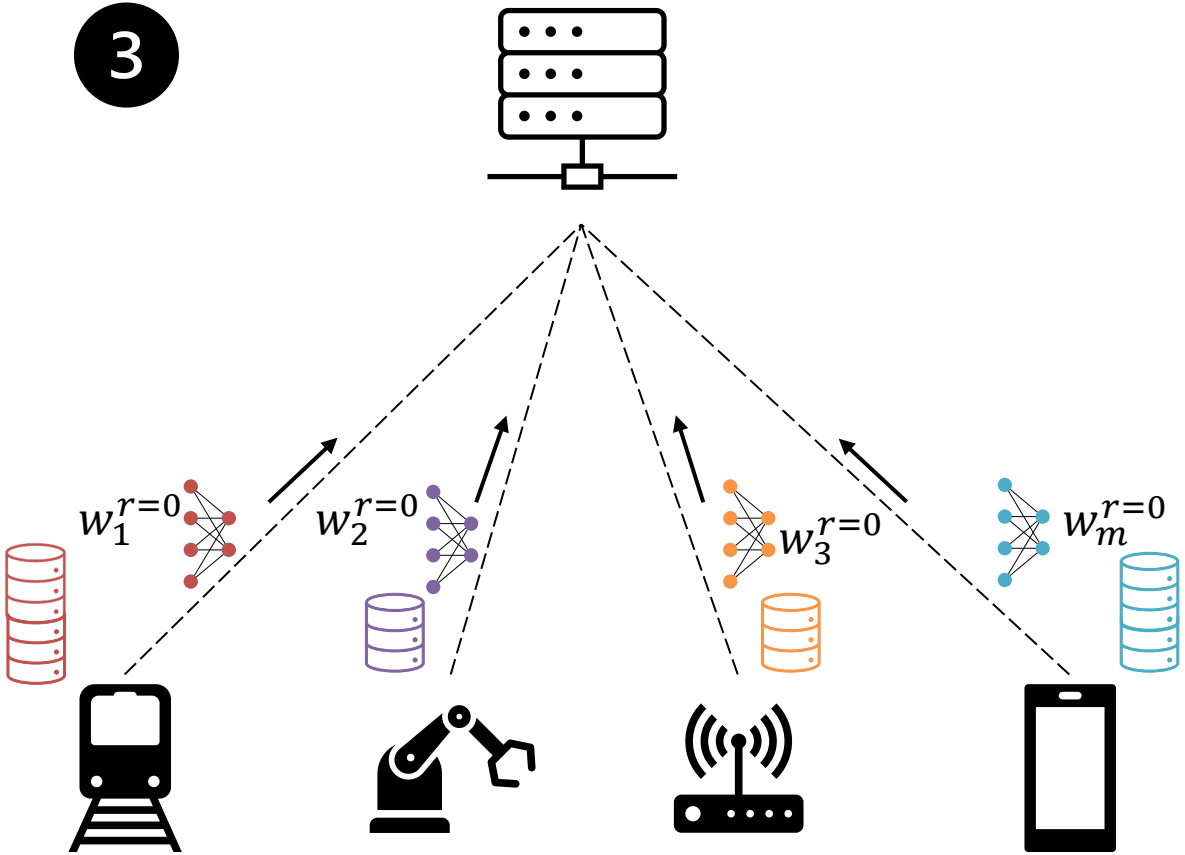
Ronda inicial $r = 0$

Entrenamiento local de los modelos (E épocas).

© 2022 IKERLAN. All rights reserved

3

Federated Learning (FL)



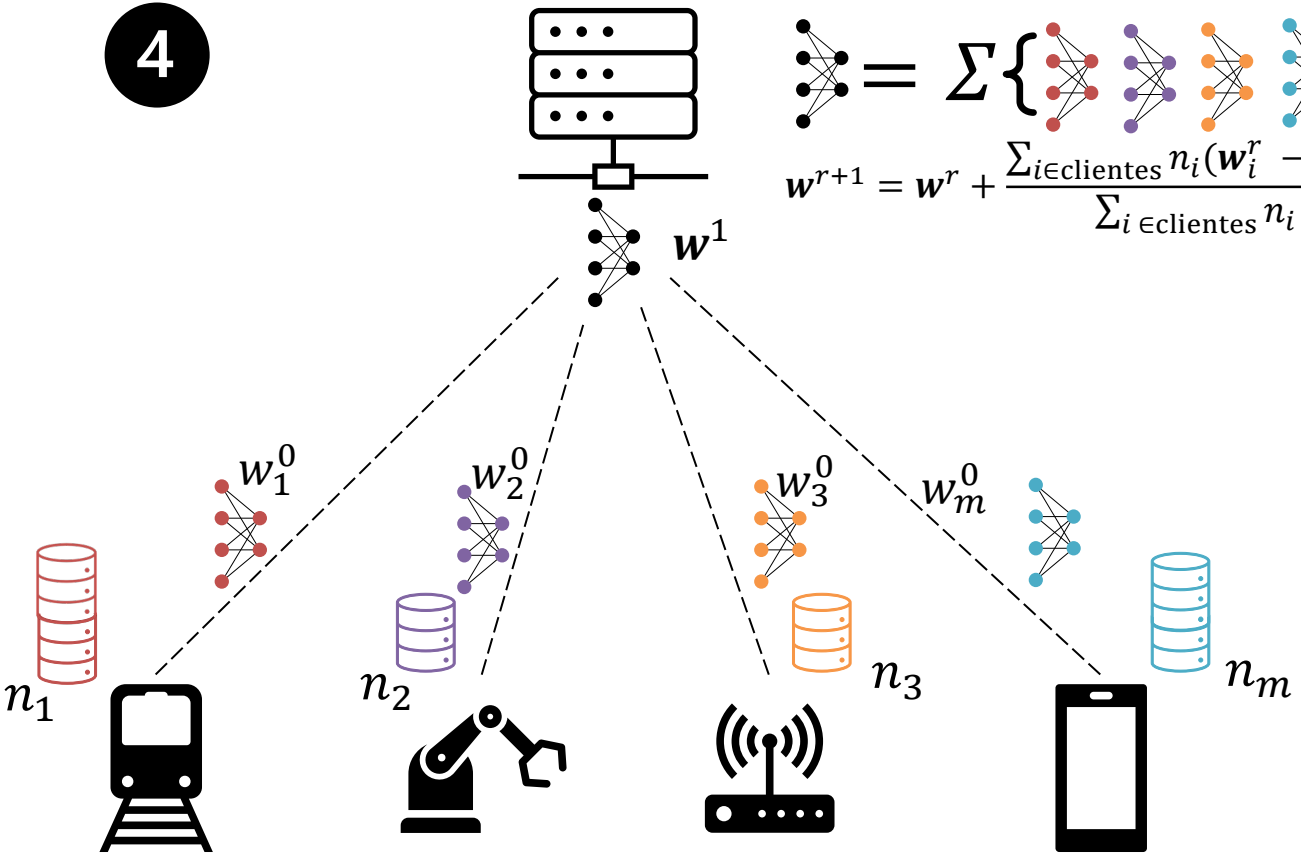
Ronda inicial $r = 0$

Transmisión de los modelos al servidor de agregación.

© 2022 IKERLAN. All rights reserved

Federated Learning (FL)

4



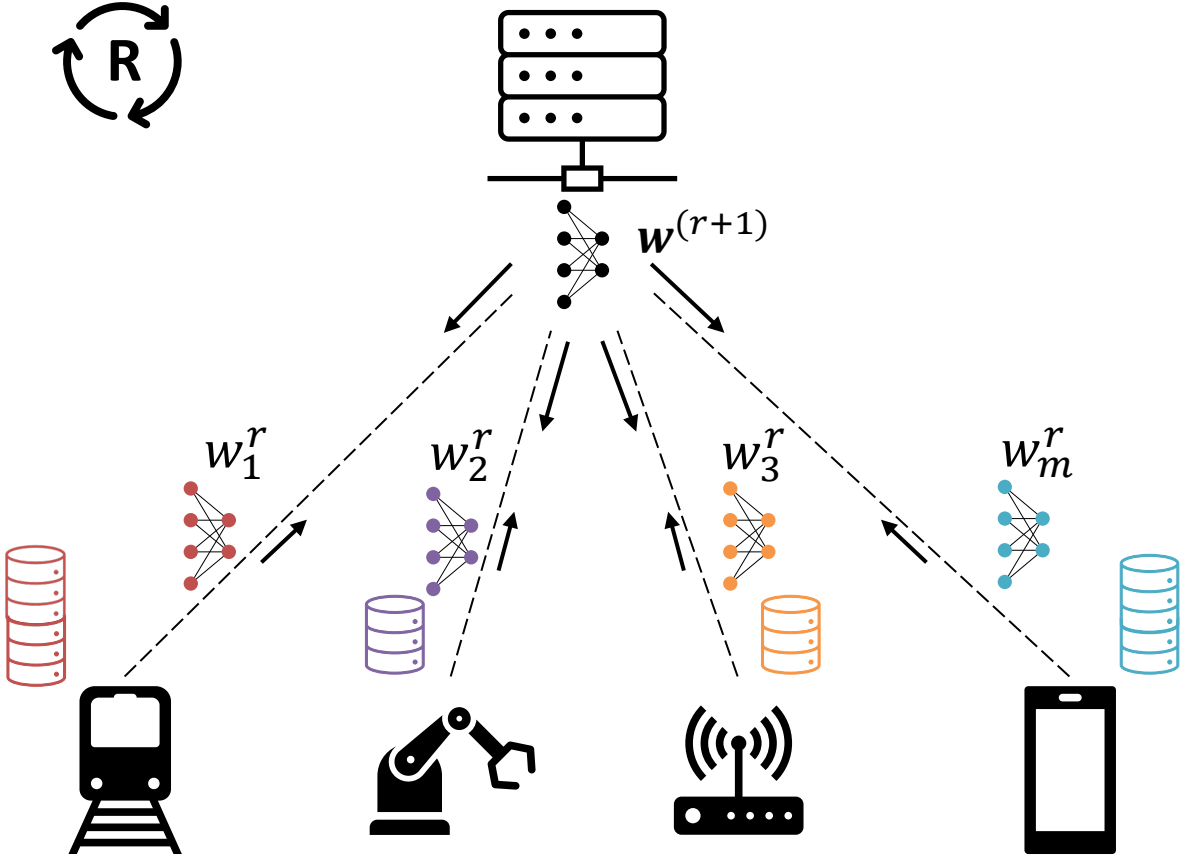
$$w^{r+1} = w^r + \frac{\sum_{i \in \text{clientes}} n_i (w_i^r - w^r)}{\sum_{i \in \text{clientes}} n_i}$$

Fin de la ronda inicial

Actualización del modelo global: FedAvg, FedAdam, FedSGDm, …

© 2022 IKERLAN. All rights reserved

Federated Learning (FL)



- ✓ Menor transmisión de datos
- ✓ Inferencia en tiempo real
- ✓ Más muestras de entrenamiento
- ✗ Coordinación compleja
- ✗ Problemas en entornos con alta heterogeneidad

© 2022 IKERLAN. All rights reserved

Aprendizaje federado con agrupación de modelos

- Agrupación basado en los parámetros de los modelos locales parcialmente entrenados.
- Integrado en la primera ronda (r=0) de FL.
- Después, se inicia un proceso de FL independiente para cada grupo.

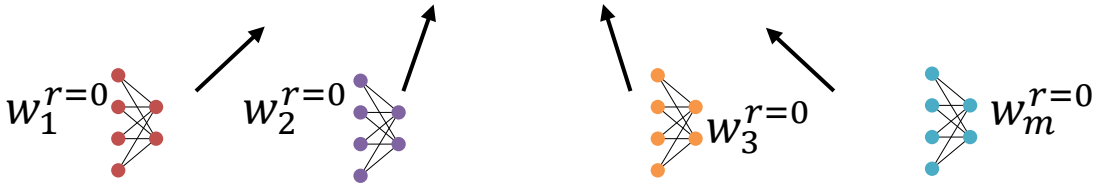
Iniciar K modelos globales

Clústering con K-means

Aplicar reducción de dimensionalidad con PCA

Colapsar cada w_i^0 a vector 1d fila a fila

$\{ w_1^0 \; w_2^0 \; w_3^0 \cdots w_m^0 \}$



© 2022 IKERLAN. All rights reserved

WHERE
TECHNOLOGY
IS AN ATTITUDE

BANCO DE PRUEBAS IOT

Banco de pruebas IoT

Emulador de red GNS3

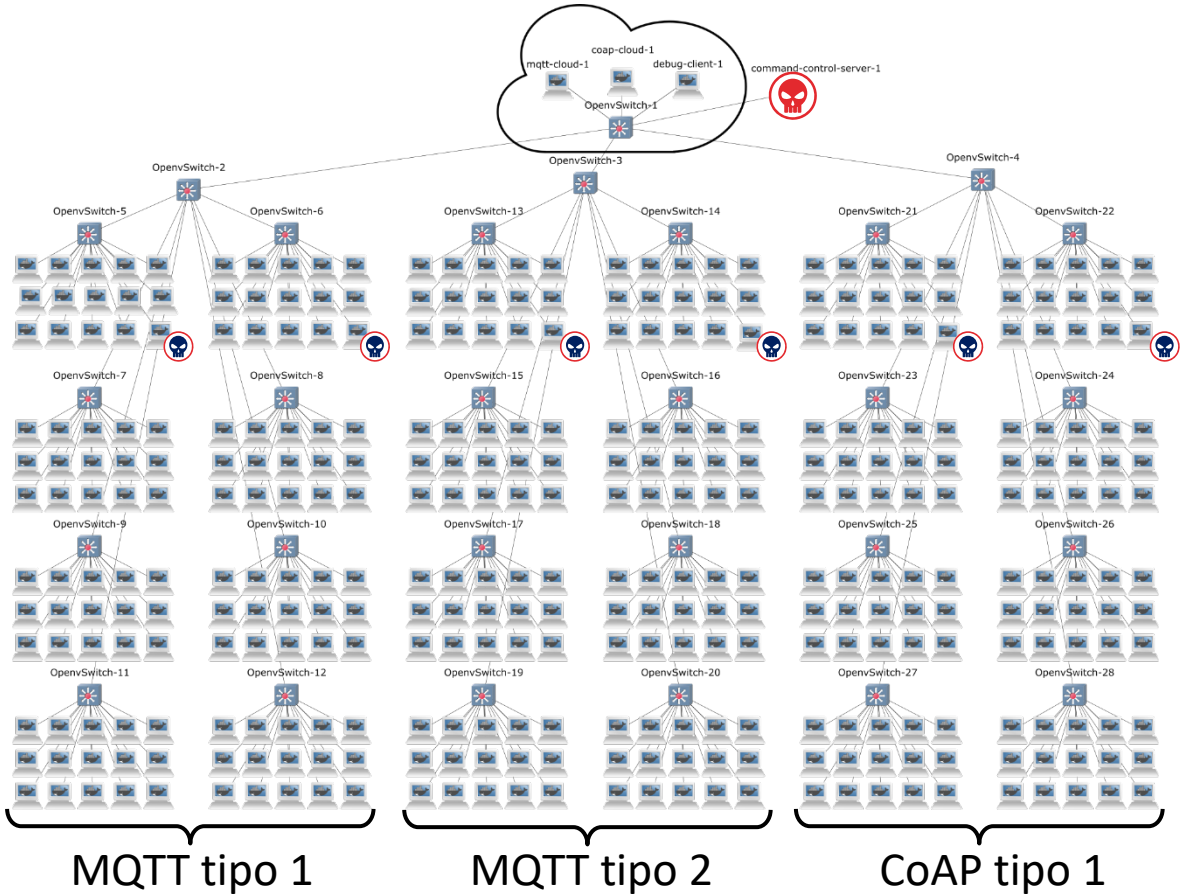
360 dispositivos IoT simulados

- MQTT tipo 1
- MQTT tipo 2
- CoAP tipo 1

Variaciones aleatorias en la periodicidad, tamaño de los mensajes transmitidos, …

Atacante

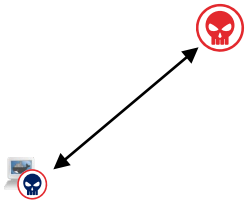
Dispositivos comprometido

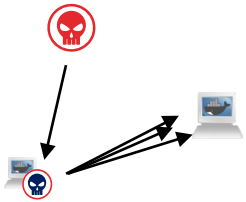


© 2022 IKERLAN. All rights reserved

Banco de pruebas IoT: ataques y comportamiento malicioso

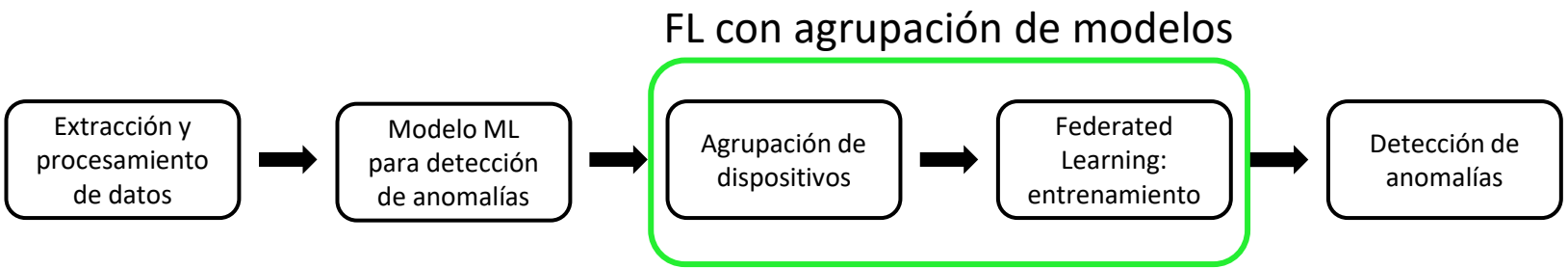
- Ataques y comportamiento malicioso:
  - Comunicación periódica con el servidor de mando y control (C&C)
    - **Merlin** (C&C y agente) https://github.com/Ne0nd0g/merlin
  - Cargar remotamente software en los dispositivos comprometidos
  - Ejecución remota de programas en la victima
  - Ataques DoS: **hping3** https://github.com/antirez/hping
    - ICMP flood
    - UDP flood
    - TCP SYN flood
    - TCP ACK flood





© 2022 IKERLAN. All rights reserved

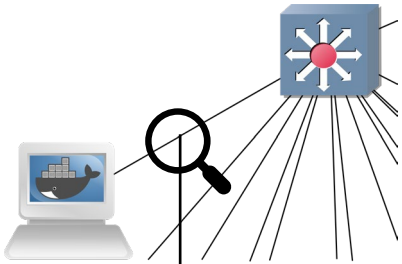
Resultados



© 2022 IKERLAN. All rights reserved

FL con agrupación de modelos

Datos y modelo ML

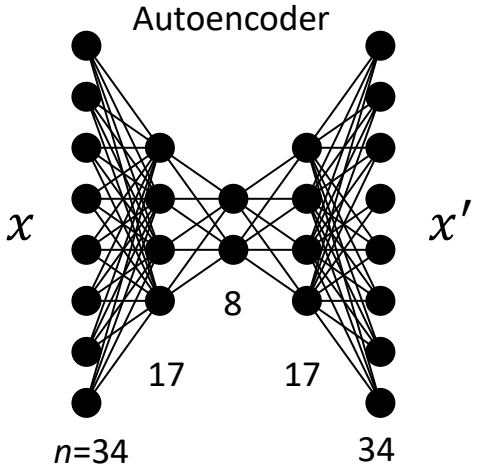


pcap

Drop IPv6
Drop ARP

Características
Tamaño del paquete
Inter arrival time
Entropía del paquete
IP type of service
IP flags
IP protocol
Src port
Dst port
TCP flags
TCP window size

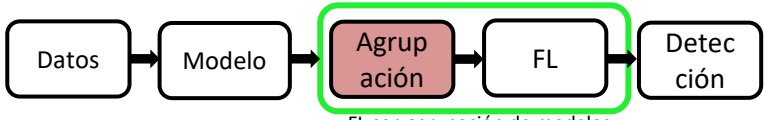
Procesamiento de características colaborativa (FL)



Entrenado en trafico normal

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - x_i')^2$$

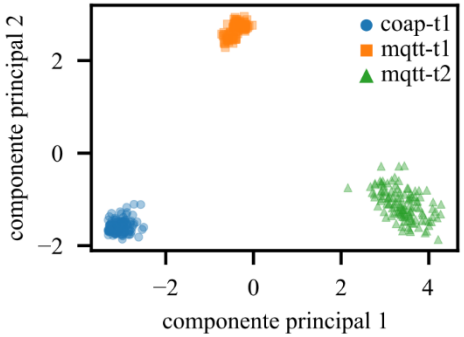
© 2022 IKERLAN. All rights reserved

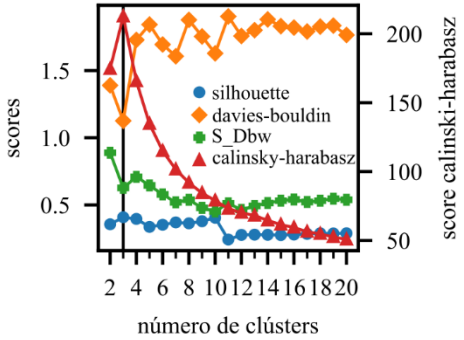


FL con agrupación de modelos

FL: agrupación

- Entrenamiento parcial durante ε épocas
- Transformar pesos del modelo a vector (1d)
- PCA: mantener 90% varianza
- Clústering k-means: k de 2 a 20



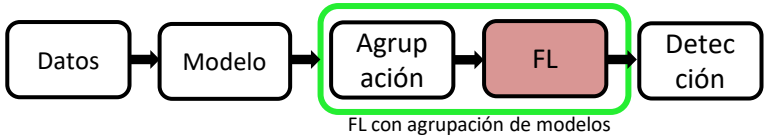


ε = 2 épocas
Adam lr = 10^{-3}

K = 3

Robusto ante distintos valores de ε.

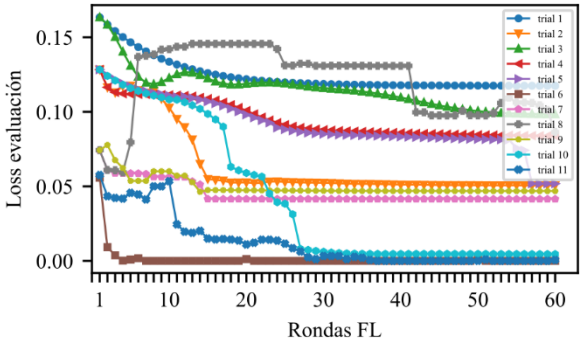
© 2022 IKERLAN. All rights reserved



FL: entrenamiento

Ajuste de hiperparámetros. → Entrenamiento final.

Optimizador cliente / servidor



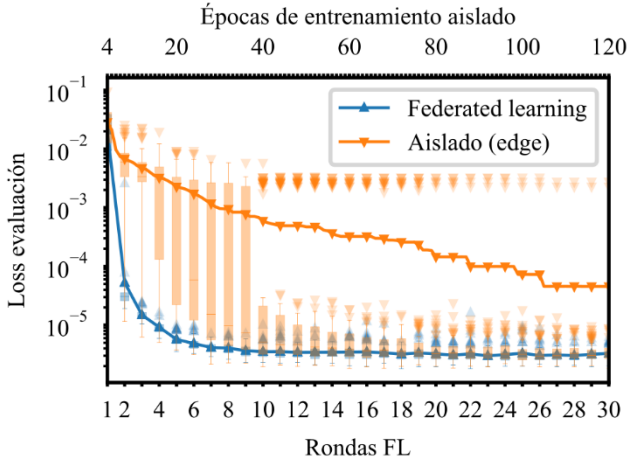
Cliente: Adam $\beta_1 = 0{,}9$ $\beta_2 = 0{,}999$

Servidor (agregación): SGD m=0

Learning rate cliente / servidor

Cliente: 0,001

Servidor (agregación): 1,0

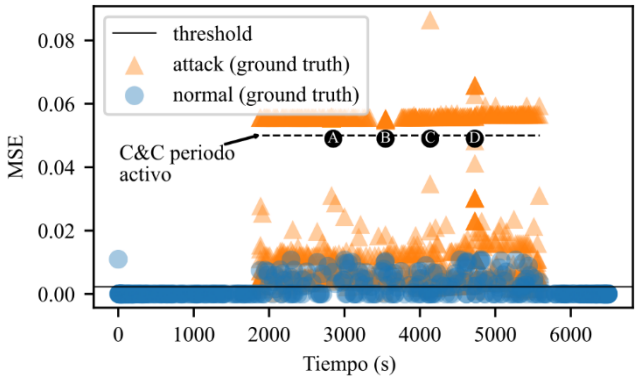


R = 30 rondas FL

E = 4 épocas locales

© 2022 IKERLAN. All rights reserved

Detección de anomalías



- **A** Cargar remotamente software hping3
- **B** ICMP flood
- **C** UDP flood
- **D** TCP SYN + ACK flood

CoAP t1 F1=0,99704

		Predicción: Ataque	Predicción: Normal
Real	Ataque	34194	0
Real	Normal	203	12405

MQTT t1 F1=0,91995

		Predicción: Ataque	Predicción: Normal
Real	Ataque	30130	5243
Real	Normal	0	13882

MQTT t2 F1=0,94711

		Predicción: Ataque	Predicción: Normal
Real	Ataque	31174	3172
Real	Normal	310	79938

© 2022 IKERLAN. All rights reserved

Conclusiones

Conclusiones:

- La fase de agrupamiento de modelos integrado en FL permite disminuir los problemas de la alta heterogeneidad sin la necesidad de usar herramientas externas, agrupamiento manual, etc.
- El agrupamiento es robusto al número de épocas locales de entrenamiento.
- FL tiene ventajas respecto al entrenamiento en la nube desde el punto de vista de la transmisión de datos.
- FL consigue una convergencia mas rápida respecto a entrenamientos no federados.

© 2022 IKERLAN. All rights reserved

MUCHAS GRACIAS

IKERLAN
P.º José María Arizmendiarrieta, 2 - 20500 Arrasate-Mondragón
T. +34 943712400 F. +34 943796944

LA TECNOLOGÍA,
NUESTRA
ACTITUD

ikerlan
MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE