Pablo Fernández Saura
pablofs@um.es

PhD student in Computer Science,
Universidad de Murcia

FL-based automated & real-time cyber-attack detection

Funded by
the European Union

Index

1. Basic research in Federated Learning
2. Applications for intrusion detection in IoT
3. Applications for fraud detection in the financial sector
4. Applications for intrusion detection in B5G
5. Future Work

Index

1 Basic research in Federated Learning

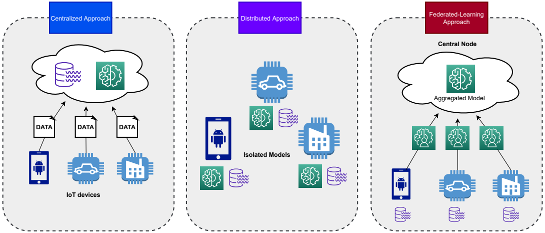
2 Applications for intrusion detection in IoT

3 Applications for fraud detection in the financial sector

4 Applications for intrusion detection in B5G

5 Future Work

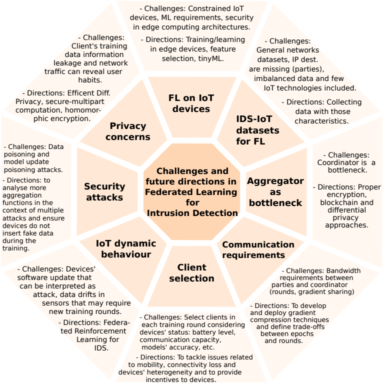
Basic research in Federated Learning



Comparison among centralized, distributed and federated learning [1]

- Training data never leaves the device
- Model training computation is decentralized
- Access to larger amounts of data
- Final models deployed closer to the users

Basic research in Federated Learning



Challenges and future directions in Federated Learning for Intrusion Detection

FL on IoT devices

- Challenges: Constrained IoT devices, ML requirements, security in edge computing architectures.

- Directions: Training/learning in edge devices, feature selection, tinyML.

IDS-IoT datasets for FL

- Challenges: General networks datasets, IP dest. are missing (parties), imbalanced data and few IoT technologies included.

- Directions: Collecting data with those characteristics.

Aggregator as bottleneck

- Challenges: Coordinator is a bottleneck.

- Directions: Proper encryption, blockchain and differential privacy approaches.

Communication requirements

- Challenges: Bandwidth requirements between parties and coordinator (rounds, gradient sharing)

- Directions: To develop and deploy gradient compression techniques and define trade-offs between epochs and rounds.

Client selection

- Challenges: Select clients in each training round considering devices' status: battery level, communication capacity, models' accuracy, etc.

- Directions: To tackle issues related to mobility, connectivity loss and devices' heterogeneity and to provide incentives to devices.

IoT dynamic behaviour

- Challenges: Devices' software update that can be interpreted as attack, data drifts in sensors that may require new training rounds.

- Directions: Federated Reinforcement Learning for IDS.

Security attacks

- Challenges: Data poisoning and model update poisoning attacks.

- Directions: to analyse more aggregation functions in the context of multiple attacks and ensure devices do not insert fake data during the training.

Privacy concerns

- Challenges: Client's training data information leakage and network traffic can reveal user habits.

- Directions: Efficent Diff. Privacy, secure-multipart computation, homomorphic encryption.

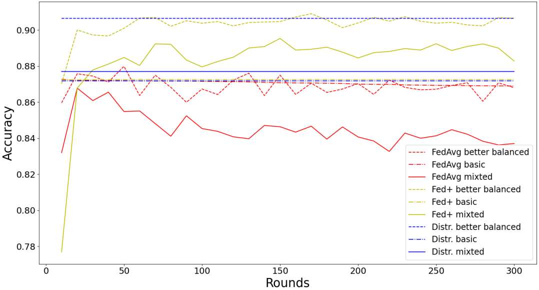
Challenges/directions in FL applied to intrusion detection in IoT [1]

- Publication in *Elsevier Computer Networks* [1]

Index

1. Basic research in Federated Learning
2. Applications for intrusion detection in IoT
3. Applications for fraud detection in the financial sector
4. Applications for intrusion detection in B5G
5. Future Work

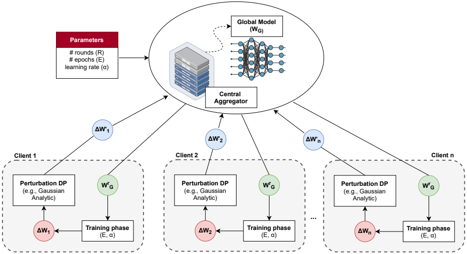
Applications for intrusion detection in IoT (review)



Comparison of avg. accuracy among scenarios [1]

- Different data distributions (basic, balanced, mixed) from ToN-IoT dataset [2]
- Different aggregation algorithms (FedAvg, Fed+)
- Multiclass Probabilistic Classification model (Logistic Regression)

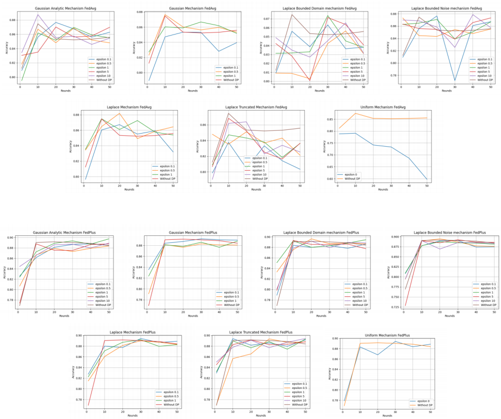
Applications for intrusion detection in IoT (Diff Privacy)



Architecture proposal for DP-based Federated Learning [3]

- Proposed workflow integrating DP/noise-adding in the FL process
- Tested and compared different noise-adding mechanisms (Gaussian, Laplace, Uniform, etc.)
- Tested and compared different privacy levels and measured the impact on accuracy

Applications for intrusion detection in IoT (Diff Privacy)



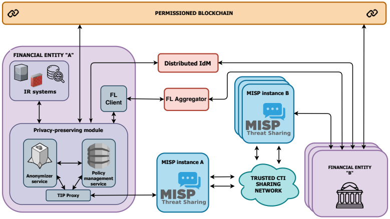
Avg. accuracy for each noise-adding mechanism (FedAvg/Fed+) [3]

- Publication in *IEEE Transactions on Industrial Informatics* [3]

Index

1. Basic research in Federated Learning
2. Applications for intrusion detection in IoT
3. Applications for fraud detection in the financial sector
4. Applications for intrusion detection in B5G
5. Future Work

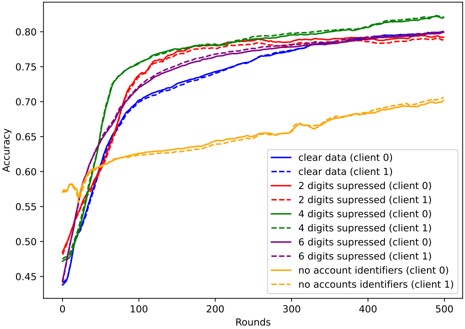
Applications for fraud detection in the financial sector



Architecture proposal for CYTILIS [4]

- Developed in the context of H2020 CyberSec4Europe project
- Evaluation using a Multi-layer Perceptron (MLP) and FL training over synthetic fraudulent transactions dataset (PaySim) [5]
- Integration with CTI platform (MISP) and DLT/Blockchain technologies

Applications for fraud detection in the financial sector

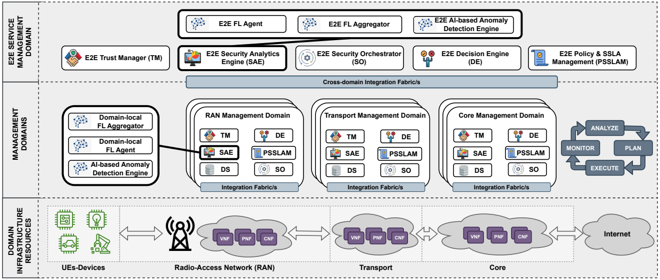


- Measured the impact on accuracy of supressing digits from transaction's origin and destination accounts
- Publication as a book chapter in *Digital Sovereignty in Cyber Security: New Challenges in Future Vision* [4]

Index

1. Basic research in Federated Learning
2. Applications for intrusion detection in IoT
3. Applications for fraud detection in the financial sector
4. Applications for intrusion detection in B5G
5. Future Work

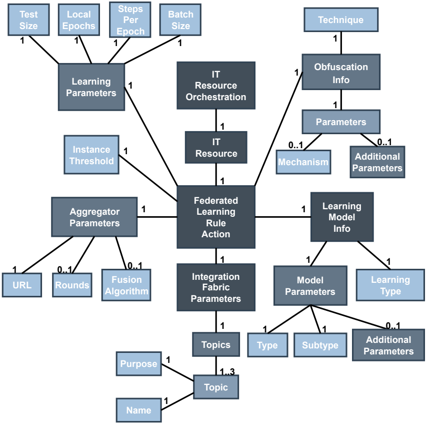
Applications for intrusion detection in B5G (FL orchestration)


E2E SERVICE MANAGEMENT DOMAIN
E2E FL Agent
E2E FL Aggregator
E2E AI-based Anomaly Detection Engine
E2E Trust Manager (TM)
E2E Security Analytics Engine (SAE)
E2E Security Orchestrator (SO)
E2E Decision Engine (DE)
E2E Policy & SSLA Management (PSSLAM)
Cross-domain Integration Fabric/s
MANAGEMENT DOMAINS
Domain-local FL Aggregator
Domain-local FL Agent
AI-based Anomaly Detection Engine
RAN Management Domain
Transport Management Domain
Core Management Domain
TM
DE
SAE
PSSLAM
DS
SO
Integration Fabric/s
ANALYZE
MONITOR
PLAN
EXECUTE
DOMAIN INFRASTRUCTURE RESOURCES
VNF
PNF
CNF
Internet
UEs-Devices
Radio-Access Network (RAN)
Transport
Core

Architecture proposal for FL orchestration in B5G

- Policy-based orchestration of FL entities (agents, aggregators)
- Crafting of a policy for deploying/configuring FL entities
- Proposed proactive/reactive workflows for intrusion detection

Applications for intrusion detection in B5G (FL orchestration)



Test Size

Local Epochs

Steps Per Epoch

Batch Size

Technique

Learning Parameters

IT Resource Orchestration

Obfuscation Info

Parameters

IT Resource

Instance Threshold

Mechanism

Additional Parameters

Aggregator Parameters

Federated Learning Rule Action

Learning Model Info

URL

Rounds

Fusion Algorithm

Integration Fabric Parameters

Model Parameters

Learning Type

Topics

Type

Subtype

Additional Parameters

Purpose

Topic

Name

Policy for orchestrating FL entities

- Publication in IEEE Future Networks World Forum 2023

Index

1. Basic research in Federated Learning
2. Applications for intrusion detection in IoT
3. Applications for fraud detection in the financial sector
4. Applications for intrusion detection in B5G
5. Future Work

Future Work

- Generation and usage of a dataset from UMU 5G testbed (user and control-plane attacks)
- Applications in Intelligent Transportation Systems (ITS) environments
- Evaluation of dynamic orchestration and integration with monitoring/mitigation mechanisms (closed loop)
- Research on Decentralized Federated Learning (DFL) frameworks and techniques
- Optimize implementation, models and data processing techniques used until now

References I

[1] Enrique Mármol Campos et al. “Evaluating Federated Learning for intrusion detection in Internet of Things: Review and challenges”. En: *Computer Networks* 203 (dic. de 2021), pág. 108661. DOI: 10.1016/j.comnet.2021.108661.

[2] Abdullah Alsaedi et al. “TON$_{I}oTTelemetryDataset : ANewGenerationDatasetofIoTandIIoTforData - DrivenIntrusionDetectionSystems$”. En: *IEEE Access* 8 (2020), págs. 165130-165150. DOI: 10.1109/ACCESS.2020.3022862.

References II

[3] Pedro Ruzafa-Alcázar et al. "Intrusion Detection Based on Privacy-Preserving Federated Learning for the Industrial IoT". En: *IEEE Transactions on Industrial Informatics* 19.2 (2021), págs. 1145-1154. DOI: 10.1109/TII.2021.3126728.

[4] Pablo Fernández Saura et al. "Privacy-Preserving Cyber Threat Information Sharing Leveraging FL-Based Intrusion Detection in the Financial Sector". En: *Digital Sovereignty in Cyber Security: New Challenges in Future Vision*. Ed. por Antonio Skarmeta et al. Cham: Springer Nature Switzerland, 2023, págs. 50-64. ISBN: 978-3-031-36096-1.

[5] Edgar Alonso Lopez-Rojas, Ahmad Elmir y Stefan Axelsson. "PAYSIM: A Financial Mobile Money simulator for Fraud Detection". En: sep. de 2016.

Pablo Fernández Saura
pablofs@um.es

PhD student in Computer Science,
Universidad de Murcia

FL-based automated & real-time cyber-attack detection

CLOUD STARS

GOBIERNO DE ESPAÑA
MINISTERIO DE ASUNTOS ECONÓMICOS Y TRANSFORMACIÓN DIGITAL

UNICO I+D

UNIVERSITAT ROVIRA i VIRGILI
BSC
UNIVERSIDAD DE MURCIA
Technische Universität München TUM

Julius-Maximilians-UNIVERSITÄT WÜRZBURG
VU
AGH
UNIVERSITÀ DI TRENTO

TU WIEN
IBM
NEARBY COMPUTING
inesc id lisboa

zh aw Zurich University
USI Università della Svizzera italiana
Imperial College London

BDV BIG DATA VALUE ASSOCIATION

Funded by the European Union