Federated detection and defense for IT and OT

Cyber CNI - PhD Days

Léo LAVAUR, IMT Atlantique, Cyber CNI, Rennes

2021-06-24

Advisors:

- Marc-Oliver Pahl, IMT Atlantique, Cyber CNI
- Yann Busnel, IMT Atlantique, IRISA
- Fabien Autrel, IMT Atlantique, Cyber CNI

chairecyber-cni.org/

Chaire Cyber CNI
5 industrial partners
8+ associated researchers
12 PhD students (2020/5)

References

[1] J. Kephart and D. Chess, “The vision of autonomic computing”, Computer, 2003.

[2] T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah, “Cyber threat intelligence sharing: Survey and research directions”, Computers & Security, 2019.

[3] S. Rathore, B. Wook Kwon, and J. H. Park, “BlockSecIoTNet: Blockchain-based decentralized security architecture for IoT network”, Journal of Network and Computer Applications, 2019.

[4] M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, “Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications”, IEEE Access, 2020.

[5] M.-O. Pahl, A. Kabil, E. Bourget, M. Gay, and P.-E. Brun, “A Mixed-Interaction Critical Infrastructure Honeypot,” 2020.

[6] T. D. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A.-R. Sadeghi, “DÏoT: A Federated Self-learning Anomaly Detection System for IoT,” in 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019.

References

[7] B. Li, Y. Wu, J. Song, R. Lu, T. Li, and L. Zhao, “DeepFed: Federated Deep Learning for Intrusion Detection in Industrial Cyber-Physical Systems,” IEEE Transactions on Industrial Informatics, 2020.

[8] W. Schneble and G. Thamilarasu, “Attack detection using federated learning in medical cyber-physical systems,” Aug. 2019.

[9] Y. Chen, J. Zhang, and C. K. Yeo, “Network Anomaly Detection Using Federated Deep Autoencoding Gaussian Mixture Model,” in Machine Learning for Networking, 2020.

[10] M.-O. Pahl and F. X. Aubet, “All Eyes on You: Distributed Multi-Dimensional IoT Microservice Anomaly Detection,” 14th International Conference on Network and Service Management, CNSM 2018 and Workshops, 2018.

[11] W. Zhang, T. Zhou, Q. Lu, X. Wang, C. Zhu, H. Sun, Z. Wang, S. K. Lo, and F.-Y. Wang, “Dynamic Fusion based Federated Learning for COVID-19 Detection,” arXiv, 2020.

Contents

1. Context

Collaboration to cope with large-scale attacks

Background

- Benefit from real-world use cases
- Exchange with partners* for insights
- Existing works and infrastructures in the chair (CNI testbed, datasets...)

Distributed attacks are more frequent, and also target industrial systems...

- **Mirai (2016)** ⇒ Uses TCP probing, and bruteforces logins
- **WannaCry & NotPetya (2017)** ⇒ Exploit MS17-010
- **AZORult (2018)** ⇒ Uses known C2s
- **Ryuk (2018)** ⇒ Uses Emotet / Trickbot

* Airbus Cyber, Amossy, BNP Paribas, EDF, Nokia Labs, SNCF

Background

CHAIRE
CYBER CNI
Sécurité des infrastructures critiques

- Benefit from real-world use cases
- Exchange with partners* for insights
- Existing works and infrastructures in the chair (CNI testbed, datasets...)

Distributed attacks are more frequent, and also target industrial systems...

CTI sharing

Collective situation awareness using threat intelligence

Federated detection

Information sharing to detect and infer cyberattacks

* Airbus Cyber, Amossys, BNP Paribas, EDF, Nokia Labs, SNCF, COMCYBER

Thesis objective

Four observations*:

*From 71 reviewed papers, including 15 surveys

(a) Lack of collective knowledge
There is a lack of collective knowledge in cybersecurity, and more particularly in the OT. [2]

(b) Lack of incentives
Trust and privacy are major hurdle for stakeholders to share data. [2]

(c) Insufficient resiliency
Centralized systems represent a Single Point of Failure and can induce a communication overhead. [3]

(d) Architectural isolation
The siloed architecture of detection systems is an obstacle to their effectiveness. [4]

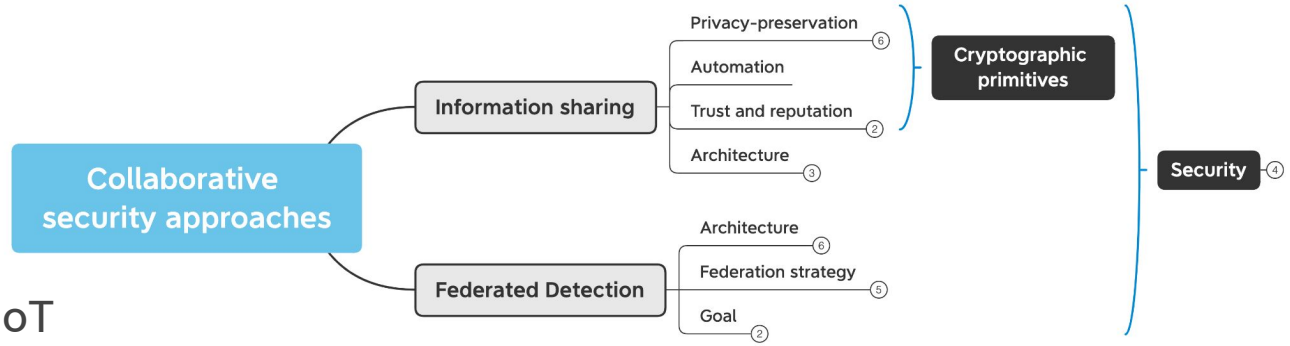
R.Q: *How to federate knowledge and defense between non-trusting parties?*

- What to collect?
- What to share?
- How to share it?

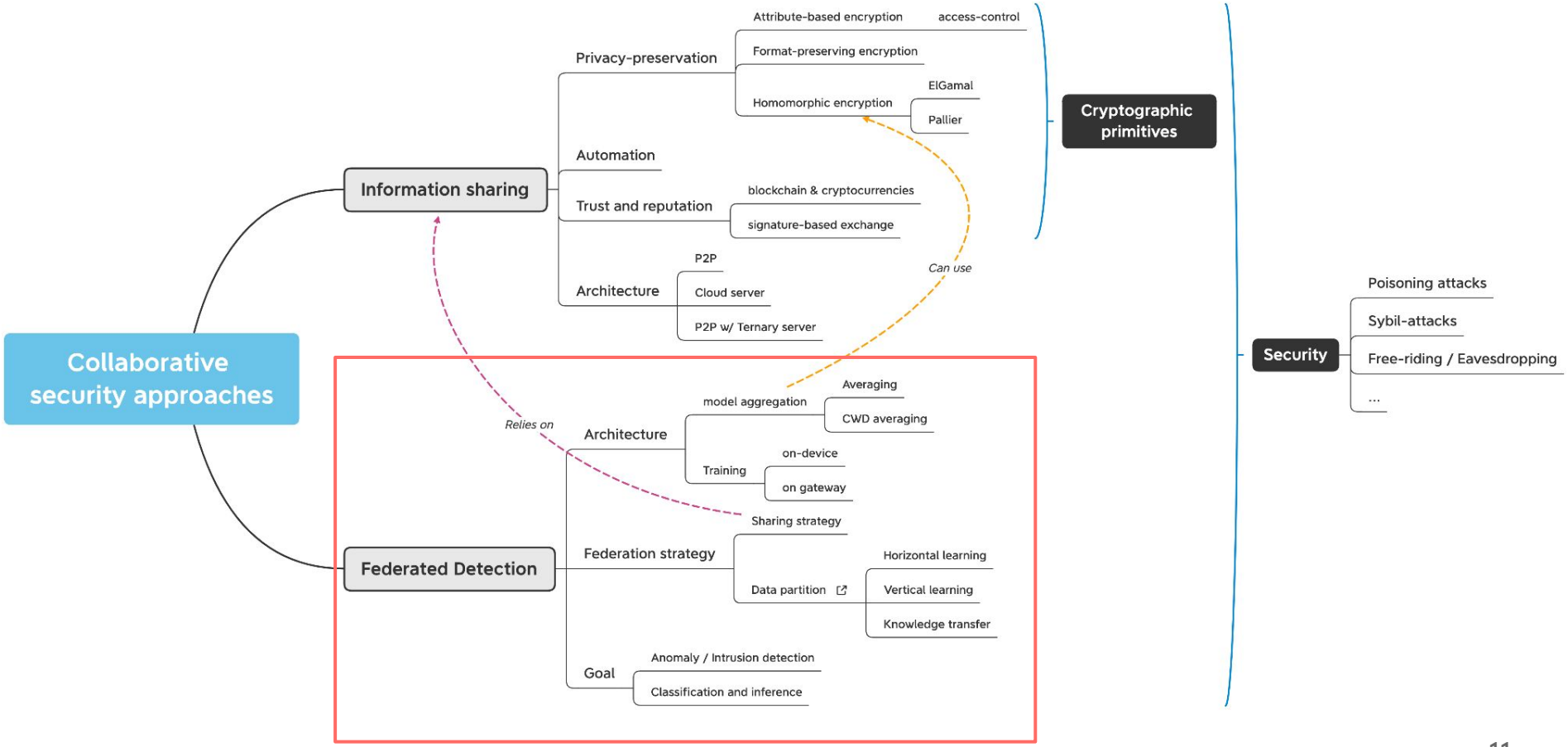
2. Current state

Writing a survey on automated collaborative security

Survey* on collaborative security for the IIoT



*From 71 reviewed papers, including 15 surveys


Collaborative security approaches
Information sharing
Privacy-preservation
Attribute-based encryption
access-control
Format-preserving encryption
Homomorphic encryption
ElGamal
Pallier
Automation
Trust and reputation
blockchain & cryptocurrencies
signature-based exchange
Architecture
P2P
Cloud server
P2P w/ Ternary server
Can use
Relies on
Federated Detection
Architecture
model aggregation
Averaging
CWD averaging
Training
on-device
on gateway
Sharing strategy
Federation strategy
Data partition
Horizontal learning
Vertical learning
Knowledge transfer
Goal
Anomaly / Intrusion detection
Classification and inference
Cryptographic primitives
Security
Poisoning attacks
Sybil-attacks
Free-riding / Eavesdropping
...

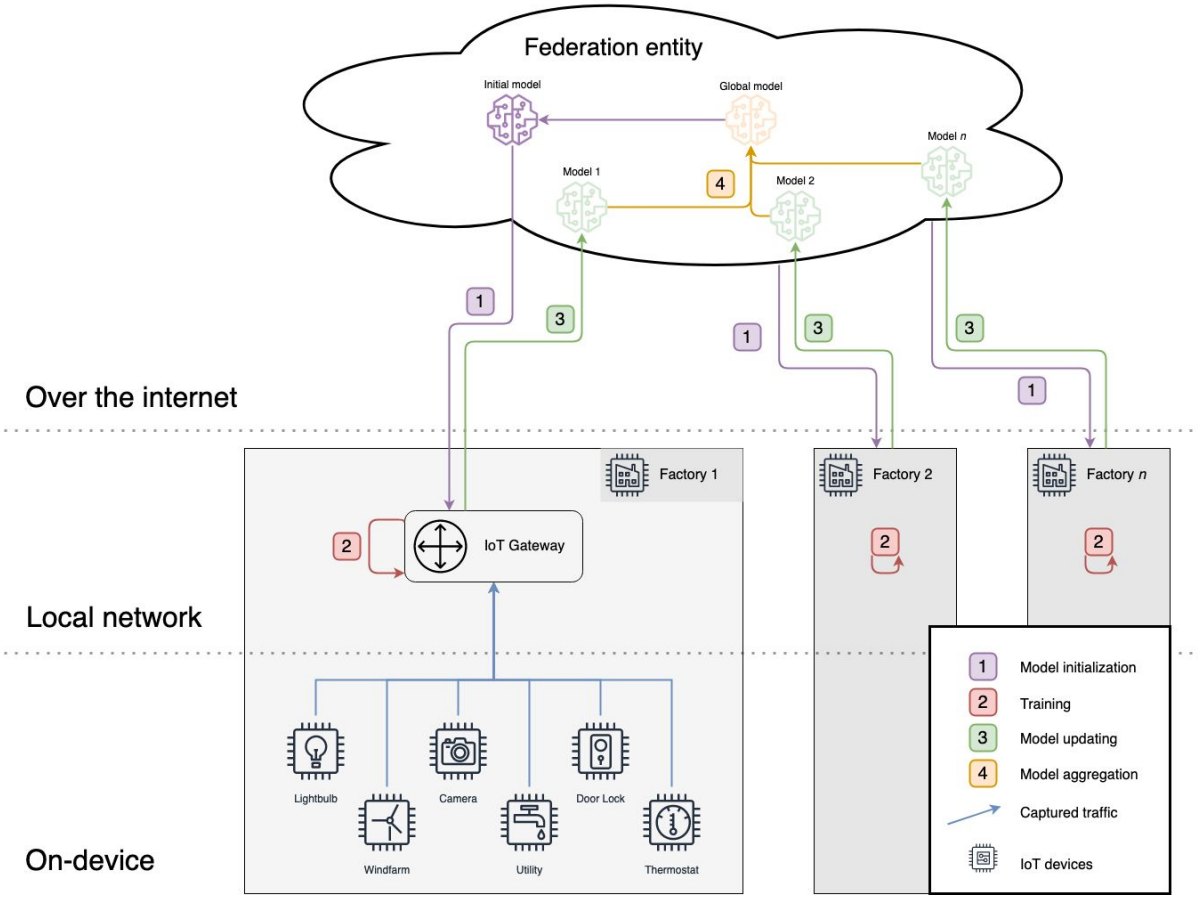
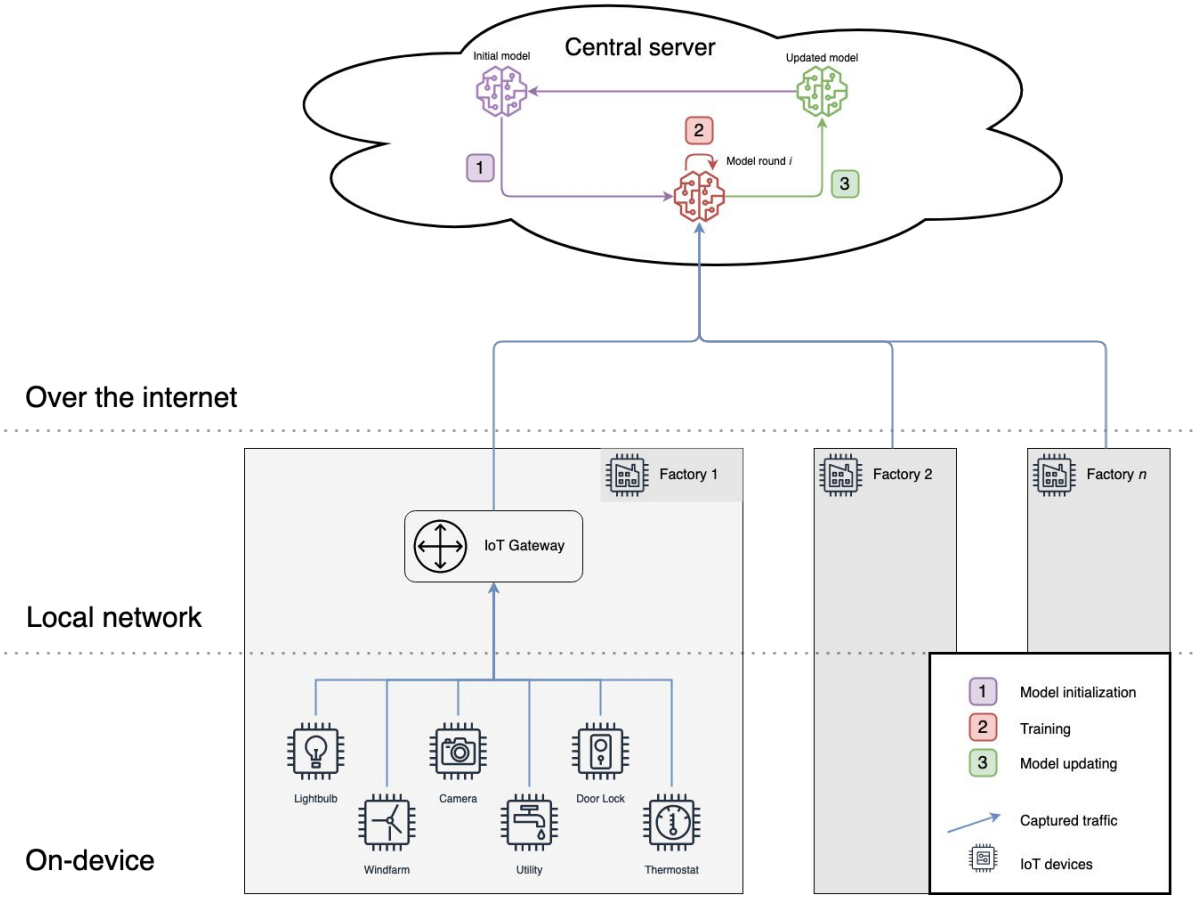


Fig. 1. Federated reference architecture

- Horizontal FL: aggregation of homogeneous models
- Note: collection of additional data could be performed using a *Honeypot Factory*



- Centralized collection and analysis of data
- Better accuracy and general performance
- Privacy issues, bandwidth overhead

Fig. 2. Centralized reference architecture

Fig. 2. Taxonomy of federated ML-based detection and defense systems (provisional)

State-of-the-Art

1. Classification of devices (either automatically, or with predefined classes)
2. Train per-class behavior models using network features / sensor values to obtain a “normal traffic” fingerprint
    - Multiple layer memory-based NN (LSTM, GRU)
    - Softmax (to classify the output in either ‘normal’ or ‘attack’, or different classes of attacks)

3. Next steps

Building experiments on the best use-cases

Reproducibility

Implement FL of the selected works:

- Nguyen *et al.* 2019 [6]
- Rathore *et al.* 2019 [3]
- Li *et al.* 2020 [7]
- Schneble *et al.* 2019 [8]
- Chen *et al.* 2020 [9]
- Pahl and Aubet 2018 [10]
- Zhang *et al.* 2020 [11]

TensorFlow + Keras

Experiment-driven research

IT networks

Detecting threats in typical IT networks with high traffic volume.

Smart factory

Detection of distributed attacks in the context of CPSs & IIoTs.

Smart building

Detecting anomalies in heterogeneous and partitioned environments.

Test beds

IT networks

Fig. 4. Airbus CyberRange

Smart building

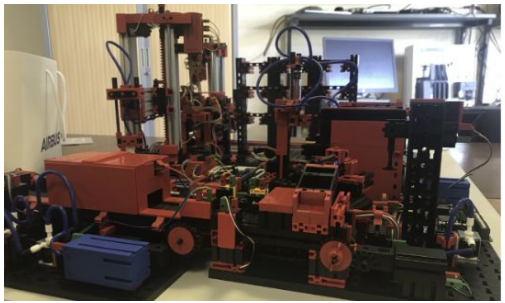


Fig. 3. Industrial testbed (Chaire CyberCNI)

Smart factory

Fig. 4. Cencyble (IMT Atlantique)

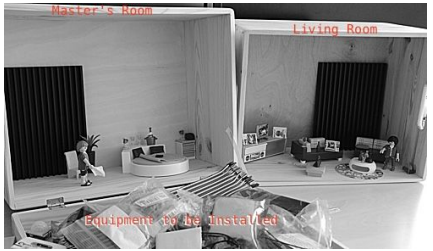


Fig. 5. S2O spaces (TUM)

Conclusion

Federated architectures for knowledge & defense between non-trusting parties

- Ongoing survey: identify the possibilities from the literature
- Next steps:
    a. reproduce and compare the state-of-the-art
    b. build the testbeds to host the experiments