IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CHAIRE CYBER CNI
Sécurité des infrastructures critiques

cyber-cni.fr/

PhD Thesis

Federated Approaches for Defending Cyber-Attacks

Author

Léo LAVAUR

Advisors

Marc-Oliver PAHL
Yann BUSNEL
Fabien AUTREL

School

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Partners

AIRBUS DEFENCE & SPACE

BNP PARIBAS La banque d'un monde qui change

EDF

NOKIA Bell Labs

SNCF

I. Context and Aims

- In 2016, Google introduced the concept of **Federated Learning (FL)**, enabling collaborative Machine Learning (ML). FL does not share local data but ML models, offering applications in diverse domains. FL has been studied to overcome challenges of **collaborative intrusion detection** and mitigation systems, such as communication overhead and information disclosure.
- This thesis addresses current limitations of Federated-learning Intrusion Detection and mitigation Systems (FIDS) in terms of **transferability, adaptability and scalability**. The chair's realistic test beds [1] will be used to host experiments and validate our hypotheses. The long-term objective is to build a distributed collaborative observatory of cyber-threats that would feed the detection systems of organizations.
- While our work on the literature answered multiple questions already, the following research questions are open:

RQ1: What are the relevant features to train FIDSs?

RQ2: How can we federated knowledge between parties with different use cases?

RQ3: Is there a trade-off between model specialization and generalization for FIDSs?

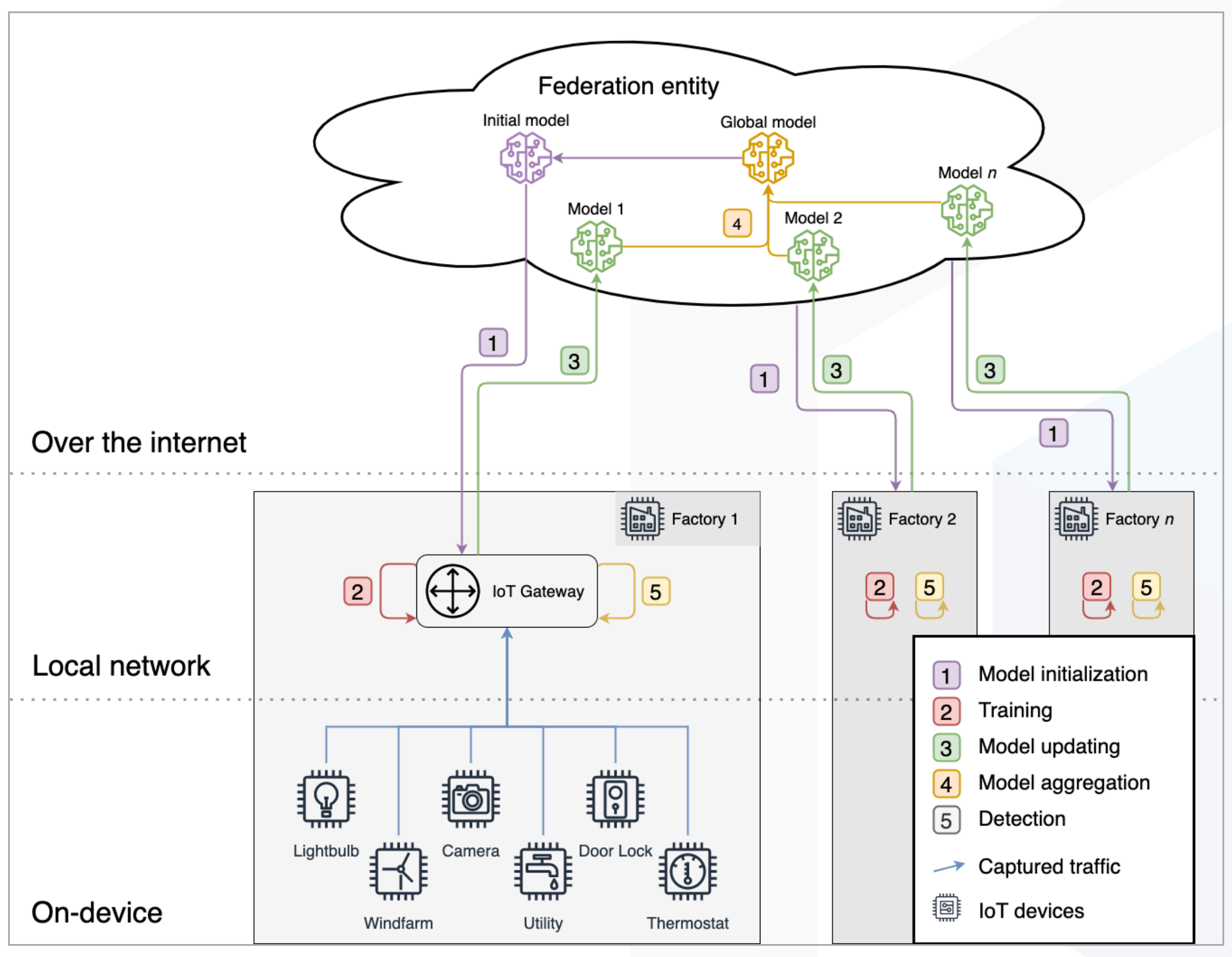


Fig 1: Federated Learning for intrusion detection in Industry 4.0 [2]

II. State of the Art of FIDSs

- We focus on FL-based intrusion detection systems (or FIDSs), which has become the state of the art for Collaborative IDSs (CIDSs). A survey paper [2] has been submitted to TNSM in 2021 and is currently under review. In particular, this **Systematic Literature Review (SLR)** shows: (a) how FIDSs are used in different domains; (b) what differences exist between architectures; (c) the state of the art of FIDSs.
- FIDSs are a *trending topic* whose evolution is following the one of FL. Publications are heterogeneous in term of venues and research groups. Most publications are use-case–based.

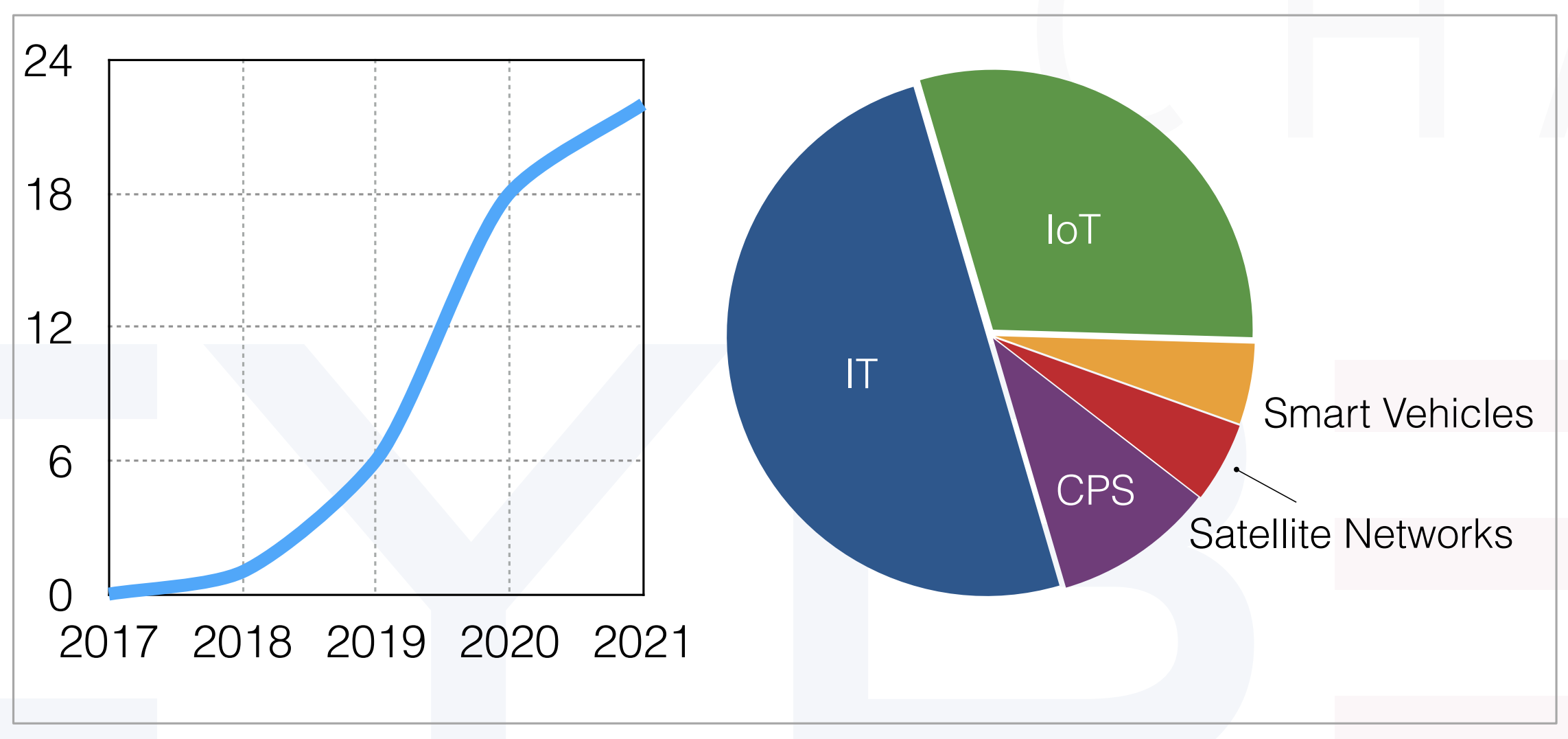
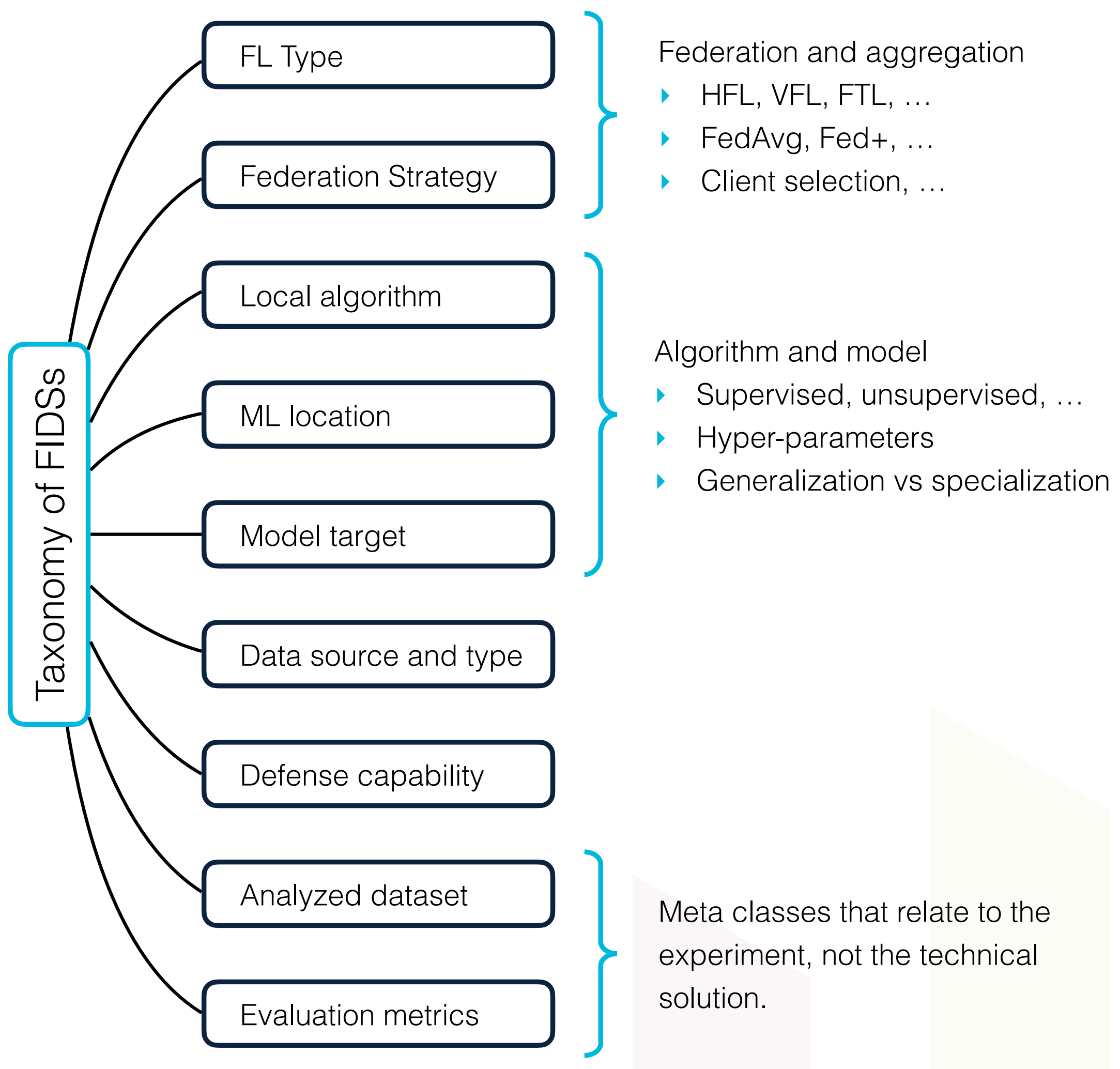


Fig 2: Evolution and repartition of FIDSs publications [2]



Taxonomy of FIDSs

FL Type
Federation Strategy
Local algorithm
ML location
Model target
Data source and type
Defense capability
Analyzed dataset
Evaluation metrics

Federation and aggregation
- HFL, VFL, FTL, …
- FedAvg, Fed+, …
- Client selection, …

Algorithm and model
- Supervised, unsupervised, …
- Hyper-parameters
- Generalization vs specialization

Meta classes that relate to the experiment, not the technical solution.

Fig 3. Taxonomy of FIDSs [2]

- The survey highlighted research directions for the community to follow. FIDSs have limited adaptability when dealing with architectures that are too different. Therefore, we will first work on **knowledge transfer in heterogeneous federations**: heterogeneous data, models, or features are considered.
- Other relevant research directions include **adaptability**—eg. dealing with data changes or clients with different distributions—, and **scalability**—eg. high number of clients, hierarchical federation aspects.
- The community identified other open issues, mostly **performance**– and **security**-wise.

III. Future work

- The chair builds and hosts realistic test beds to perform real-life reproducible experiments. Three use cases are considered in this thesis: IT infrastructures, industry 4.0, and smart buildings. The three use cases are covered by the chair's test beds and its partners. Two projects will start to address FIDSs limitation:

P1: FIDS-EP: an evaluation platform for reproducible experiments

P2: Cross-Silo and Hierarchical FL for FIDSs

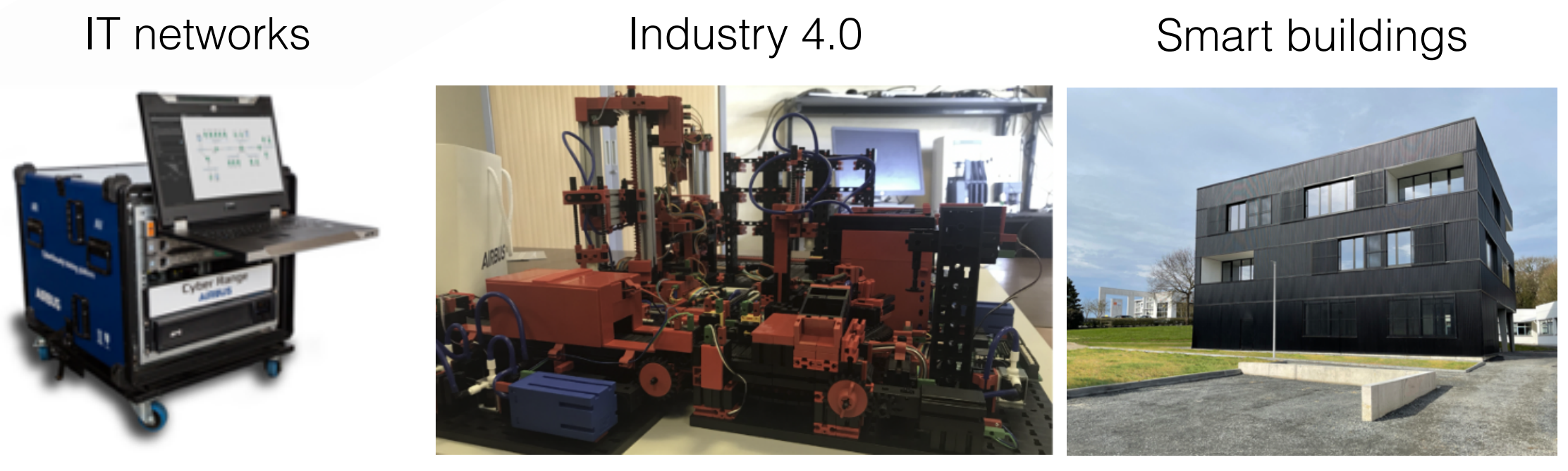


Fig 4: Test beds of the chair Cyber CNI: Airbus Cyberrange, Fischertechnics models, Cencyble building (Rennes campus)

- One of the major caveats of the literature review is the **inability to compare the performance of existing approach**, due to the differences in term of dataset, algorithms, participants, and use case. Therefore, we will develop an **evaluation platform for reproducible experiments**. This will allow us to study the impact of existing FL strategies on performance, and eventually provide objective insights on FIDSs design.
- The second project addresses the transferability and adaptability aspects of FIDSs to match the needs of organisations. Cross-silo FL enables privacy-preserving **training of heterogeneous models**, so that each organization can train a model of its own, while benefiting of the experience of the other participants. This is extremely relevant for IDS tasks. Organisation could train a model locally with FL among agents, and collaborate externally, using a **hierarchical approach**.

References

[1] M.-O. Pahl, A. Kabil, E. Bourget, M. Gay, and P.-e. Brun, "A Mixed-Interaction Critical Infrastructure Honeypot," *European Cyber Week C&ESAR Conference*, 2020.

[2] L. Lavaur, M.-O. Pahl, Y. Busnel, and F. Autrel, "The Evolution of Federated Learning-based Intrusion Detection and Mitigation: a Survey," under review *in IEEE TNSM Special Issue on Advances in Network Security Management*, 2022

Contact: leo.lavaur@imt-atlantique.fr
Twitter: @phdcybersec
Github: @phdcybersec
LinkedIn: linkedin.com/in/leo-lavaur

This poster was made at the chaire Cybersecurity for Critical Networked Infrastructures (Cyber CNI) | https://cyber-cni.fr/ | Our monthly cybersecurity speaker series https://talk.cybercni.fr/ | Our science blog https://future-iot.org/

Institut Mines-Télécom
IMT Atlantique Bretagne-Pays de la Loire École Mines-Télécom
TELECOM Paris
TELECOM SudParis
AIRBUS
AMOSSYS
BNP PARIBAS La banque d'un monde qui change
EDF
NOKIA Bell Labs
SNCF
PÔLE D'EXCELLENCE CYBER
Région Bretagne