Open PhD position

"Leveraging hierarchical federated learning for monitoring and prediction in the Cloud-Edge-IoT continuum"

STACK team (IMT Atlantique, INRIA, LS2N), Nantes, France

Context

The current literature on using hierarchical federated learning (FL) for cloud-edge-IoT involves several innovative approaches aimed at enhancing efficiency, accuracy, and energy savings in distributed machine learning across IoT devices. Hierarchical FL frameworks, such as HELCHFL, have been developed to leverage the advantages of both cloud-centric and edge-centric FL paradigms, achieving high efficiency and low-cost hierarchical FL training. These frameworks address communication and user heterogeneity issues through utility-driven and heterogeneity-aware heuristic user selection strategies, significantly enhancing training performance and reducing energy costs [1].

In the Industrial Internet of Things (IIoT), hierarchical FL with mobile edge computing (MEC) servers has been proposed to reduce network burden and latency. Techniques like deep reinforcement learning (DRL)-based joint resource allocation and IIoT device orchestration policies have been shown to outperform other algorithms in achieving more accurate models while minimizing latency and energy consumption [2]. Similarly, intelligent and hierarchical resource facilitation frameworks, such as I-HARF, adapt to dynamic Edge-IoT situations, including service overloading and device mobility, by employing deep reinforcement learning designs for adaptive resource facilitation [3].

To address the heterogeneity of IoT environments, semi-synchronous communications in hierarchical FL frameworks have been proposed, reducing communication costs and overcoming unbalanced training problems in heterogeneous environments [4]. Semihierarchical federated analytics frameworks leverage multiple edge servers for aggregating updates from IoT devices, improving communication efficiency and providing robust, fault-tolerant data analytics to IoT networks [5].

Edge-assisted hierarchical FL schemes like FedEdge have been introduced to accelerate model training with asynchronous local federated training and adaptive model aggregation, addressing the challenges of model diversity and vulnerability to model poisoning attacks [6]. Furthermore, hierarchical FL with quantization has been explored to reduce communication overhead, with tighter convergence analysis providing practical guidelines for client-edge aggregation and edge-client association strategies [7].

These advancements collectively represent the state of the art in employing hierarchical FL for Cloud-Edge-IoT, focusing on efficiency, accuracy, and security.

PhD Subject

Our objective is to investigate the monitoring and prediction of the continuum Cloud-Edge IoT using hierarchical federated learning paradigm (HFL) [8], which has several advantages including (i) reduced communication cost (ii) improved security and privacy preservation (iii) reduced impact of heterogeneous environment, on the monitoring and detection/prediction.

- *Optimizing communication*: Given the large number of IoT devices participating in the FL training process and the increasing size of the (DL) models, the communication cost of FL often dominates the total cost of the FL system (which requires very little computational, only for weights aggregation). This demonstrates that traditional FL suffers from high overhead and latency for monitoring Cloud-Edge-IoT continuum. Therefore, we will study how HFL can be used to reduce this overhead without sacrificing accuracy. The high communication efficiency may come from the exploitation of the advantages from both cloud and edge servers to share the communication overhead as well as from sub-aggregations at the edge servers and fewer global aggregations at the cloud server.

- *Managing data/model heterogeneity*: Given numerous and geographically distributed locations of IoT devices, client training data are usually non-IID, which can degrade the global model performance. To solve the divergence issue and enhance the learning performance, we will study how HFL exploits the edge servers to group the IoT device models into smaller clusters based on the similarity of weight updates or other criteria (i.e., proximity). This attempts to reduce the variance of the updated weights and in turn, preserve the uniformity in the cluster and thus accelerate the convergence rate and adaptability of the monitoring system.

- *Handling privacy*: Despite that FL-based approaches exchange the model param- eters instead of the raw data, recent attacks (e.g., membership inference attack) demonstrate that such an approach does not provide a sufficient privacy guarantee. Methods like differential privacy or encryption have been applied; unfortunately, they introduce inherent trade-offs between protecting data privacy and the model performance or communication overhead. The two-stage aggregation of learning models in HFL can cope with this gap and hence offers a certain degree of privacy to be distributed throughout the hierarchy.

Expected skills

- Master or engineering student with experience in at least one of machine learning /deep learning and distributed programming/Cloud-Edge-IoT

- Solid programming skills in Python language. Familiarity with typical deep learning frameworks (TensorFlow/PyTorch) and models is a plus.

- Good English communication skills and teamwork abilities.

Application procedure

Interested applicants must submit the following documents (as a single PDF file) with the subject PhD Application: CEI-HFL to the contact below.

- Motivation letter

- Detailed CV (with link to GitHub or similar, if applicable)

- Past manuscript, projects, or other relevant works

- Copies of your engineer or MSc degrees and transcripts

- Names and contact information of at least one relevant reference.

Contact and setting

- Kandaraj Piamrat and Mario Sudholt: firstname.lastname@ls2n.fr
- PhD location: IMT Atlantique, Nantes, France
- Estimated start date and duration: ASAP, 36 months
- The PhD is part of the SPIREC and PEPR Cloud projects. The PhD features numerous contacts with their multiple partners.

References
----

[1] Y. Cui, K. Cao, J. Zhou, and T. Wei, "Optimizing training efficiency and cost of hierarchical federated learning in heterogeneous mobile-edge cloud computing," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2023.

[2] T. Zhao, F. Li, and L. He, "Drl-based joint resource allocation and device orchestration for hierarchical federated learning in noma-enabled industrial iot," IEEE Transactions on Industrial Informatics, 2023.

[3] I. AlQerm and J. Pan, "I-harf: Intelligent and hierarchical framework for adaptive resource facilitation in edge-iot systems," IEEE Internet of Things Journal, vol. 10, no. 5, pp. 3954–3967, 2023.

[4] M. G. Herabad, "Communication-efficient semi-synchronous hierarchical federated learning with balanced training in heterogeneous iot edge environments," Internet of things, 2022.

[5] L. Zhao, M. I. G. Valero, S. Pouriyeh, L. Li, and Q. Z. Sheng, "Communication-efficient semihierarchical federated analytics in iot networks," IEEE Internet of Things Journal, 2022.

[6] K. Wang, Q. He, F. Chen, H. Jin, and Y. Yang, "Fededge: Accelerating edge-assisted federated learning," in Proceedings of the ACM Web Conference 2023, WWW '23, (New York, NY, USA), p. 2895–2904, Association for Computing Machinery, 2023.

[7] L. Liu, J. Zhang, S. Song, and K. B. Letaief, "Hierarchical federated learning with quantization: Convergence analysis and system design," IEEE Transactions on Wire- less Communications, vol. 22, no. 1, pp. 2–18, 2023.

[8] O. Aouedi, K. Piamrat, and M. Suˆdholt, “HFedSNN: Efficient hierarchical federated learning using spiking neural networks,” in Proceedings of the Int’l ACM Symposium on Mobility Management and Wireless Access, pp. 53–60, 2023.