Offer #2024-08353

PhD Position F/M Robust Federated Learning

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

About the research centre or Inria department

The Inria center at Université Côte d'Azur includes 42 research teams and 9 support services. The center's staff (about 500 people) is made up of scientists of di?erent nationalities, engineers, technicians and administrative staff. The teams are mainly located on the university campuses of Sophia Antipolis and Nice as well as Montpellier, in close collaboration with research and higher education laboratories and establishments (Université Côte d'Azur, CNRS, INRAE, INSERM ...), but also with the regional economic players.

With a presence in the fields of computational neuroscience and biology, data science and modeling, software engineering and certification, as well as collaborative robotics, the Inria Centre at Université Côte d'Azur is a major player in terms of scientific excellence through its results and collaborations at both European and international levels.

Context

The position is part of a new Marie Curie Training Network called FINALITY, in which Inria joins forces with top universities and industries, including IMDEA, KTH, TU Delft, the University of Avignon (Project Leader), the Cyprus Institute, Nokia, Telefonica, Ericsson, Orange, and others. The PhD students will have opportunities for internships with other academic and industry partners and will be able to participate in thematic summer schools and workshops organized by the project.

Only people who have spent less than one year in France in the last 3 years are eligible.

The candidate will receive a monthly living allowance of about €2,735, a mobility allowance of €414, and, if applicable, a family allowance of €458 (gross amounts).

Assignment

Federated Learning (FL) empowers a multitude of IoT devices, including mobile phones and sensors, to collaboratively train a global machine learning model while retaining their data locally [1,2]. A prominent example of FL in action is Google's Gboard, which uses a FL-trained model to predict subsequent user inputs on smartphones [3].

Two primary challenges arise during the training phase of FL [4]:

Data Privacy: Ensuring user data remains confidential. Even though the data is kept locally by the devices, it has been shown that an honest-but-curious server can still reconstruct data samples [5,6], sensitive attributes [7,8], and the local model [9] of a targeted device. In addition, the server can conduct membership inference attacks [10] to identify whether a data sample is involved in the training or source inference attacks to determine which device stores a given data sample [11].

Security Against Malicious Participants: Ensuring the learning process is not derailed by harmful actors. Recent research has demonstrated that, in the absence of protective measures, a malicious agent can deteriorate the model performance by simply flipping the labels [12] and/or the sign of the gradient [13] and even inject backdoors into the model [14] (backdoors are hidden vulnerabilities, which can be exploited under certain conditions predefined by the attacker, like some specific inputs).

Differentially private algorithms [15] have been proposed to tackle the challenges of protecting user privacy. These algorithms work by clipping the gradients and adding noise to them before the transmission, ensuring that minor alterations in a user's training dataset will not be discernible to potential adversaries [16,17,18,19,20]. By leveraging the differentially private mechanisms, [19] shows that adversaries are unable to deduce the exact local information of vehicles for the applications such as Uber. Furthermore, [20] demonstrates that the quality of data reconstruction attack is significantly reduced when training a convolutional neural network on CIFAR-10 dataset.

To enhance system security against adversarial threats, Byzantine resilient mechanisms are implemented on the server side. These algorithms are designed to identify and mitigate potentially detrimental actions or inputs from users, ensuring that even if some components act maliciously or erratically, the overall system remains functional and secure [21,22,23,24]. Experiments [21] reveal that integrating these Byzantine resilient mechanisms sustains neural network accuracy at 90.7%, even when 10% of the agents maliciously flip the labels on the MNIST dataset. In contrast, without such protection, the accuracy of the neural network drops significantly to 77.3%.

Integrating differential privacy with Byzantine resilience presents a notable challenge. Recent research suggests that when these two security measures are combined in their current forms, the effectiveness of the resulting algorithm disproportionately depends on the number of parameters in the machine learning model (d) [25]. In particular, it requires either the batch size to grow linearly with the square root of d, or the proportion of the malicious agents in the system to decrease with rate inversely proportional to the square root of d. For a realistic model such as ResNet-50 (with around 25 million parameters), the batch size should be larger than 5000, which is clearly impractical. To tackle this problem, novel Byzantine resilient algorithms have been recently proposed [26,27]. However, these algorithms encounter significant computational complexity, with a rate of at least d^3 in each communication round. Hence, there is a pressing need for innovative methods that can seamlessly integrate differential privacy and Byzantine resilience with low computational complexity to train practical neural networks.

Project objective

The goal of this PhD is to propose novel FL algorithms to effectively tackle these two mutually linked challenges. In particular, we want to explore the potentialities of compression for FL training, as these techniques can highly reduce the model dimension d, which may provide a solution for a computation-efficient private and secure FL system.

Compression techniques were initially introduced to alleviate communication costs in distributed training processes, where only a proportion of model parameters are sent from the device to the server in each communication round [28,29,30]. The primary objective of compression design is to ensure a communication-efficient machine learning/FL system, by providing model parameters selection rules at the device side which optimize the trained model performance under a given communication budget. [31,32] combined Byzantine resilient methods with compression, to ensure a communication-efficient secure FL system. However, in these studies, even though devices transmit compressed models to the server, Byzantines resilient methods still operate on the full models of dimension d. Consequently, adopting their solutions to build a private and secure FL system still requires high computation load.

The goal of this PhD is to investigate the impact of compression strategies on the trade-offs among privacy, robustness, computational efficiency, and model performance, with the aim of designing novel compression techniques for a computationally efficient, private, and secure federated learning system.

References

[1] McMahan et al, Communication-Efficient Learning of Deep Networks from Decentralized Data, AISTATS 2017

[2] Li et al, Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, p.p. 50-60, 2020
[3] Hard, Andrew et al, Federated Learning for Mobile Keyboard Prediction. arxiv: 1811.03604, 2019
[4] Kairouz et al, Advances and Open Problems in Federated Learning. Now Foundations and Trends, 2021
[5] Geiping et al, Inverting gradients - how easy is it to break privacy in federated learning?, NeurIPS 2020
[6] Yin et al, See through gradients: Image batch recovery via gradinversion, CVPR 2021
[7] Lyu et al, A novel attribute reconstruction attack in federated learning, FTL-IJCAI 2021
[8] Driouich et al, A novel model-based attribute inference attack in federated learning, FL-NeurIPS22, 2022.
[9] Xu et al, What else is leaked when eavesdropping Federated Learning? PPML-CCS, 2021
[10] Zari et al, Efficient Passive Membership Inference Attack in Federated Learning, PriML-NeurIPS workshop, 2022
[11] Hu et al, Souce inference attacks in federated learning, ICDM 2021
[12] Fang et al, Local model poisoning attacks to Byzantine-robust federated learning, in 29th USENIX Security Symposium, 2020
[13] Wu et al, Federated variance-reduced stochastic gradient descent with robustness to byzantine attacks, IEEE Transactions on Signal Processing, vol. 68, pp. 4583–4596, 2020
[14] Wang et al, Attack of the tails: yes, you really can backdoor federated learning, NeurIPS 2020
[15] Dwork and Roth, A. The algorithmic foundations of differential privacy. Now Publishers Inc., 2013.
[16] Abadi, M et al, Deep learning with differential privacy. ACM CCS 2016
[17] Bellet et al, Personalized and Private Peer-to-Peer Machine Learning, AISTATS 2018
[18] Noble, M et al, 2022. Differentially Private Federated Learning on Heterogeneous Data, AISTATS 2022
[19] Zhao, Y et al. Local Differential Privacy based Federated Learning for Internet of Things. IEEE Internet of Things 2020.
[20] Balle, B et al. Reconstructing Training Data with Informed Adversaries. 2022 IEEE S&P
[21] Yin et al, Byzantine-Robust Distributed Learning: Towards Optimal Statistical Rates, ICML 2018
[22] Krishna Pillutla et al, Robust Aggregation for Federated Learning, in IEEE Transactions on Signal Processing, 2022.
[23] Blanchard et al, Machine Learning with Adversaries: Byzantine Tolerant Gradient Descent, NeurIPS 2017
[24] Guerraoui et al, Byzantine Machine Learning: A Primer. ACM Comput. Surv., August 2023
[25] Guerraoui et al, Differential Privacy and Byzantine Resilience in SGD: Do They Add Up?, PODC 2021.

[26] Zhu et al, Byzantine-Robust Federated Learning with Optimal Statistical Rates, AISTATS 2023
[27] Allouah et al, On the Privacy-Robustness-Utility Trilemma in Distributed Learning, ICML 2023.
[28] Alistarh et al, QSGD: Communication-efficient sgd via gradient quantization and encoding. NeurIPS 2017.
[29] Alistarh et al, The convergence of sparsified gradient methods. NeurIPS 2018.
[30] Haddadpour et al, Federated learning with compression: unified analysis and sharp guarantees, AISTATS 2021
[31] Gorbunov et al, Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top, ICLR 2023
[32] Zhu, H et al. Byzantine-Robust Distributed Learning With Compression. IEEE Trans. on Signal and Inf. Process. over Networks 9, 280–294, 2023.

Main activities

Research

Skills

We are looking for a candidate with coding experience in Python and good analytical skills.

We expect the candidate to be fluent in English.

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Contribution to mutual insurance (subject to conditions)

Remuneration

The candidate will receive a monthly living allowance of about €2,735, a mobility allowance of €414, and, if applicable, a family allowance of €458 (gross amounts)

General Information

- **Theme/Domain :** Optimization, machine learning and statistical methods
  System & Networks (BAP E)
- **Town/city :** Sophia Antipolis
- **Inria Center :** Centre Inria d'Université Côte d'Azur
- **Starting date :** 2025-03-01
- **Duration of contract :** 3 years
- **Deadline to apply :** 2025-05-31

Contacts

- **Inria Team :** NEO
- **PhD Supervisor :**
  Neglia Giovanni / Giovanni.Neglia@inria.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply

Applications must be submitted online on the Inria website. Collecting applications by other channels is not guaranteed.

Defence Security :
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical

potential (PPST).Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :
As part of its diversity policy, all Inria positions are accessible to people with disabilities.