Internship

Secure Decentralized Learning

Context: Machine learning plays a central role in many applications, and the increasing adoption of decentralized solutions, combined with dependability requirements, necessitates that learning tasks be carried out in a decentralized manner. In such settings, nodes in the system can assume multiple roles: performing learning on their local data while also aggregating the computational results of other nodes.

In this context, we aim to achieve **confidentiality**, ensuring that private local data is not leaked while providing a practical solution.

Objective: The goal of this internship is to explore the combination of **Multi-Party Computation (MPC)** and **Differential Privacy (DP)** to assess the feasibility and effectiveness of these approaches in ensuring confidentiality.

Our team has previously conducted an exploratory study on MPC in the Federated Learning context, which provides a strong foundation for studying the integration of Differential Privacy techniques.

The successful candidate will join the **Laboratory for Trustworthy, Smart, and Self-Organizing Information Systems (LICIA)** at CEA LIST, working in a multicultural, multidisciplinary environment with the opportunity to collaborate with external researchers.

Methodology: The intern will be responsible for the following tasks:

1. Become familiar with the Differential Privacy principles.
2. Conduct a state-of-the-art review of Differential Privacy in Federated Learning and its combination with Multi-Party Computation.
3. Become familiar with the MPC solution developed in the laboratory.
4. Select a Differential Privacy approach and design a solution to integrate it with the existing MPC solution.
5. Implement the integrated solution.
6. Evaluate the performance of the solution.

Requirements:

- **Background in computer science or a related field**, with a focus or strong interest in distributed systems, cryptography, and machine learning.
- **Programming skills** in languages commonly used for cryptographic or machine learning tasks (e.g., Python, C++, or Rust).
- Comfortable working in **English**, both for communication and documentation purposes.

Domaine de spécialité requis : Informatique

Moyens mis en œuvre (expériences, méthodes d'analyses, autres...): distributed systems, programming languages, machine learning, data privacy.

Moyens informatiques mis en œuvre :
Langages: C++, Python, Java

Niveau souhaité : Bac + 4/5

Durée: 6 mois

Niveau d'habilitation défense (AS au minimum) : AS

Formation souhaitée : Ingénieur/Master

Possibilité de poursuite en thèse : Oui

Lieu du stage : CEA, Centre de Saclay Nano-Innov, 91191 Gif sur Yvette

Contacts : Antonella Del Pozzo antonella.delpozzo@cea.fr +33 1 69 08 04 69