PhD Position in Automatic Control

# Heterogeneity in federated learning: An open multi-agent systems approach

## General information

- **Location:** CRAN, Université de Lorraine, 2 avenue de la Fôret de Haye, Nancy, France.
- **Department:** Control Identification Diagnosis (CID)
- **Duration:** 01/10/2027 - 30/09/2030
- **Contact:** renato.vizuete@univ-lorraine.fr vineeth.satheeskumar-varma@univ-lorraine.fr
- **Keywords:** multi-agent systems, machine learning, distributed optimization, stochastic processes.

## Scientific context

Federated Learning is a recent area of machine learning whose aim is to train global AI models for multiple users while preserving the privacy of the data of each user [KMR+16]. The typic architecture consists of a central server that gathers information of several users without the need to collect private data. Federated Learning is naturally an open multi-agent system since the users are frequently offline such that the set of available agents at each computation round is time-varying. Furthermore, the connectivity of agents depends on their location and available infrastructure, introducing a clear level of heterogeneity. Unfortunately, mathematical analyses of the performance of federated learning algorithms have not considered the heterogeneity related to the connectivity of agents, which could degrade the performance of algorithms and even prevent convergence. The novel idea of this thesis is to analyze heterogeneity in federated learning through open multi-agent systems. Some preliminary results have been obtained for the case of replacements in distributed optimization, which is related to the distributed training of machine learning models [MVH+24]. The goal of the thesis is to consider different frameworks in open multi-agent systems to analyze the heterogeneity in federated learning where agents with different capabilities participate in the training process at different time instants.

## Research program

Federated learning can be classified into two types: cross-silo and cross-device [KM21]. In cross-silo, a small number of organizations participate in the training of the model. One of the main applications involve hospitals where several units want to collaborate in the training of a health-issue detection model without sharing the data of their patients. In cross-device, the number of agents is large (millions of devices), and they are usually spread around the world. Several applications include the training of models for text prediction, voice assistant or image recognition that require data from millions of users, or more local applications like power consumption [HZL+24]. The goal of this project is to use the different approaches in [VMF+24] to analyze federated learning algorithms in cross-silo and cross-device federated learning. In the cross-silo scenario, since the number of participating agents is known, we will use the framework of pseudo open multi-agent systems where the openness of the system is characterized by activations and deactivations in a superset of agents [VMN18]. The heterogeneity will be modeled by different distributions for the activation of the agents. In the cross-device scenario, since the number

of participating agents is unknown, we will consider a general open multi-agent systems where entirely new agents can join the system [VH25]. The heterogeneity will be modeled with different probabilities of arrivals and departures for the agents, resulting in different interaction time intervals. To address the challenges posed by these scenarios, tools from fields such as multi-agent systems, distributed optimization, machine learning, and stochastic processes—among others—will be applied.

## Lab

Established in 1980, CRAN is a Joint Research Unit (UMR 7039) shared by the University of Lorraine and the CNRS. It is part of the University of Lorraine's scientific cluster dedicated to Control Theory, Mathematics, Computer Science, and their Interactions (AM2I). It is primarily affiliated with the CNRS Institute for Information Sciences and Technologies (CNRS Sciences informatiques) and secondarily with the CNRS Institute of Biological Sciences (CNRS Biologie). The laboratory operates across eight sites in the Lorraine region: six in Nancy and two others in Longwy and Épinal. CRAN comprises 250 members, including:

- 120 researchers or faculty-researchers, including 9 CNRS researchers from CoNRS Section 03
- 5 emeritus researchers
- 27 research support staff (administrative, technical, and scientific services)
- Approximately 100 doctoral and post-doctoral researchers

The laboratory is internationally recognized for research activities spanning the following fields: control theory, signal and image processing, computer engineering, production engineering, and biology related to oncology and neuroscience. It conducts both theoretical and applied research, resulting in technology transfer, innovation, and the practical application of findings. The CID (Control, Identification, and Diagnosis) department encompasses all CRAN research activities focused on the control of continuous dynamic systems and cyber-physical systems. More broadly and concisely, the research conducted at CRAN lies "at the heart of systems and health"—the laboratory's signature tagline.

## Qualification requirements

- Applicants should hold, or be in the final year of, a master's degree or an engineering degree in automatic control, machine learning, optimization, or applied mathematics.
- Rigorous knowledge in optimization, networks, and multi-agent systems.
- Excellent programming skills (Python, MATLAB, etc).
- Proficiency in spoken and written English. French is not a requirement.

## To apply

Contact: Renato Vizuete (renato.vizuete@univ-lorraine.fr)
`https://sites.google.com/view/renato-vizuete`
Vineeth Varma (vineeth.satheeskumar-varma@univ-lorraine.fr)
`https://sites.google.com/site/vineethshome/home`

Please accompany the application form:

- detailed CV

- university transcripts
- letter of motivation for the subject
- letter(s) of recommendation or names of referents

## References


[KMR+16] Konečný, J., McMahan, H. B., Ramage, D., & Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence. arXiv preprint arXiv:1610.02527.

[MVH+24] Monnoyer de Galland, C., Vizuete, R., Hendrickx, J. M., Panteley, E., & Frasca, P. (2024). Random coordinate descent for resource allocation in open multiagent systems. IEEE Transactions on Automatic Control, 69(11), 7600-7613.

[KM21] Kairouz, P., & McMahan, H. B. (2021). Advances and open problems in federated learning. Foundations and trends in machine learning, 14(1-2), 1-210.

[HZL+24] Hassan, M., Zhang, C., Lasaulce, S., Varma, V. S., Debbah, M., & Ghosho, M. (2024, August). Strategic federated learning: Application to smart meter data clustering. In 2024 32nd European Signal Processing Conference (EUSIPCO) (pp. 1172-1176).

[VMF+24] Vizuete, R., Monnoyer de Galland, C., Frasca, P., Panteley, E., & Hendrickx, J. M. (2024). Trends and questions in open multi-agent systems. In Hybrid and Networked Dynamical Systems: Modeling, Analysis and Control (pp. 219-252). Cham: Springer Nature Switzerland.

[VMN18] Varma, V. S., Morărescu, I. C., & Nešić, D. (2018). Open multi-agent systems with discrete states and stochastic interactions. IEEE Control Systems Letters, 2(3), 375-380.

[VH25] Vizuete, R., & Hendrickx, J. M. (2025). Consensus on Open Multi-Agent Systems Over Graphs Sampled from Graphons. In 2025 IEEE 64th Conference on Decision and Control (CDC) (pp. 5282-5289).

Offre de thèse en Automatique

# Hétérogénéité dans l'apprentissage fédéré : une approche fondée sur les systèmes multi-agents ouverts

## Informations générales

- **Lieu :** CRAN, Université de Lorraine, 2 avenue de la Fôret de Haye, Nancy, France.
- **Département :** Contrôle Identification Diagnostic (CID)
- **Durée :** 01/10/2027 - 30/09/2030
- **Contacts :** renato.vizuete@univ-lorraine.fr vineeth.satheeskumar-varma@univ-lorraine.fr
- **Mots-clés :** systèmes multi-agents, apprentissage automatique, optimisation distribuée, processus stochastiques.

## Contexte scientifique

L'apprentissage fédéré (Federated Learning) est un domaine récent de l'apprentissage automatique dont l'objectif est d'entraîner des modèles globaux d'intelligence artificielle à partir des données de plusieurs utilisateurs, tout en préservant la confidentialité des données de chacun [KMR+16]. Son architecture typique repose sur un serveur central qui collecte des informations provenant de plusieurs utilisateurs sans nécessiter l'accès à leurs données privées. L'apprentissage fédéré constitue naturellement un système multi-agents ouvert, puisque les utilisateurs sont fréquemment déconnectés, ce qui implique que l'ensemble des agents disponibles à chaque itération de calcul varie au cours du temps. De plus, la connectivité des agents dépend de leur localisation et des infrastructures disponibles, introduisant ainsi une forme importante d'hétérogénéité. Malheureusement, les analyses mathématiques des performances des algorithmes d'apprentissage fédéré n'ont pas pris en compte l'hétérogénéité liée à la connectivité des agents, laquelle pourrait dégrader les performances de ces algorithmes, voire empêcher leur convergence. L'idée novatrice de cette thèse consiste à analyser l'hétérogénéité dans l'apprentissage fédéré à travers le cadre des systèmes multi-agents ouverts. Des résultats préliminaires ont été obtenus dans le cas du remplacement d'agents en optimisation distribuée, un problème étroitement lié à l'entraînement distribué des modèles d'apprentissage automatique [MVH+24]. L'objectif de cette thèse est d'explorer différents cadres théoriques des systèmes multi-agents ouverts afin d'analyser l'hétérogénéité dans l'apprentissage fédéré, où des agents possédant des capacités différentes participent au processus d'entraînement à différents instants.

## Programme de recherche

L'apprentissage fédéré peut être classé en deux catégories : l'apprentissage fédéré inter-organisations (cross-silo) et l'apprentissage fédéré inter-appareils (cross-device) [KM21]. Dans le cadre de l'apprentissage fédéré inter-organisations, un nombre limité d'organisations participe à l'entraînement du modèle. L'une des principales applications concerne les établissements hospitaliers, où plusieurs unités souhaitent collaborer à l'entraînement d'un modèle de détection de problèmes de santé sans partager les données de leurs patients. Dans le cadre de l'apprentissage fédéré inter-appareils, le nombre d'agents est très élevé (potentiellement des millions d'appareils), généralement répartis

à travers le monde. Parmi les applications figurent l'entraînement de modèles de prédiction de texte, d'assistants vocaux ou de reconnaissance d'images, qui nécessitent des données provenant de millions d'utilisateurs, ainsi que des applications plus locales, notamment dans le domaine de la consommation énergétique [HZL+24]. L'objectif de ce projet est d'utiliser les différentes approches présentées dans [VMF+24] afin d'analyser les algorithmes d'apprentissage fédéré dans ces deux contextes. Dans le scénario inter-organisations (cross-silo), puisque le nombre d'agents susceptibles de participer est connu, nous utiliserons le cadre des systèmes pseudo-multi-agents ouverts, dans lequel l'ouverture du système est caractérisée par l'activation et la désactivation d'agents appartenant à un ensemble prédéfini [VMN18]. L'hétérogénéité sera modélisée à l'aide de différentes distributions de probabilité régissant l'activation des agents. Dans le scénario inter-appareils (cross-device), puisque le nombre d'agents participants est inconnu, nous considérerons un cadre général de systèmes multi-agents ouverts dans lequel de nouveaux agents peuvent rejoindre le système [VH25]. L'hétérogénéité sera modélisée à l'aide de différentes probabilités d'arrivée et de départ des agents, conduisant à des intervalles de temps d'interaction variables. Pour relever les défis associés à ces scénarios, des outils issus de plusieurs domaines seront mobilisés, notamment les systèmes multi-agents, l'optimisation distribuée, l'apprentissage automatique et les processus stochastiques.

## Laboratoire d'accueil

Créé en 1980, le CRAN est une Unité Mixte de Recherche – UMR 7039 commune à l'Université de Lorraine et au CNRS. Il fait partie du pôle scientifique Automatique, Mathématiques, Informatique et leurs Interactions (AM2I) de l'Université de Lorraine. Il est rattaché principalement à CNRS Sciences informatiques et secondairement à CNRS Biologie. Le laboratoire est réparti sur 8 sites en Lorraine dont 6 sont situés à Nancy et les 2 autres à Longwy et Épinal. Le CRAN est composé de 250 membres au 1er septembre 2025 dont :

- 120 chercheurs ou enseignants-chercheurs dont 9 chercheurs CNRS de la section 03 du CoNRS.
- 5 chercheurs émérites.
- 27 personnels d'appui à la recherche.
- Environ 100 doctorants et post-doctorants.

Le laboratoire est reconnu à l'international pour ses activités de recherche couvrant les domaines suivants : automatique, traitement du signal et de l'image, génie informatique, productique, biologie en lien avec la cancérologie et les neurosciences. Il y développe à la fois des recherches théoriques et appliquées se concrétisant par du transfert, de l'innovation et de la valorisation. Le département CID (Contrôle Identification Diagnostic) regroupe l'ensemble des activités de recherche du CRAN portant sur l'automatique des systèmes dynamiques continus et cyber-physiques. Plus globalement et de manière synthétique, les recherches menées au CRAN sont « au cœur des systèmes et de la santé », signature rédactionnelle du laboratoire.

## Profil recherché

- Les candidats doivent être titulaires d'un diplôme de master ou d'ingénieur, ou être en dernière année de formation, dans les domaines de l'automatique, de l'apprentissage automatique, de l'optimisation ou des mathématiques appliquées.
- Solides connaissances en optimisation, réseaux et systèmes multi-agents.
- Excellentes compétences en programmation (Python, MATLAB, etc).
- Bonne maîtrise de l'anglais, à l'écrit comme à l'oral. La maîtrise du français n'est pas requise.

## Modalités de candidature

Contacts : Renato Vizuete (renato.vizuete@univ-lorraine.fr)
https://sites.google.com/view/renato-vizuete
Vineeth Varma (vineeth.satheeskumar-varma@univ-lorraine.fr)
https://sites.google.com/site/vineethshome/home

Les candidats sont invités à transmettre les documents suivants :

- un curriculum vitae détaillé
- les relevés de notes universitaires
- une lettre de motivation précisant leur intérêt pour le sujet de thèse
- une ou plusieurs lettres de recommandation, ou les coordonnées de personnes référentes

## Références

[KMR+16] Konečný, J., McMahan, H. B., Ramage, D., & Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence. arXiv preprint arXiv:1610.02527.

[MVH+24] Monnoyer de Galland, C., Vizuete, R., Hendrickx, J. M., Panteley, E., & Frasca, P. (2024). Random coordinate descent for resource allocation in open multiagent systems. IEEE Transactions on Automatic Control, 69(11), 7600-7613.

[KM21] Kairouz, P., & McMahan, H. B. (2021). Advances and open problems in federated learning. Foundations and trends in machine learning, 14(1-2), 1-210.

[HZL+24] Hassan, M., Zhang, C., Lasaulce, S., Varma, V. S., Debbah, M., & Ghosho, M. (2024, August). Strategic federated learning: Application to smart meter data clustering. In 2024 32nd European Signal Processing Conference (EUSIPCO) (pp. 1172-1176).

[VMF+24] Vizuete, R., Monnoyer de Galland, C., Frasca, P., Panteley, E., & Hendrickx, J. M. (2024). Trends and questions in open multi-agent systems. In Hybrid and Networked Dynamical Systems: Modeling, Analysis and Control (pp. 219-252). Cham: Springer Nature Switzerland.

[VMN18] Varma, V. S., Morărescu, I. C., & Nešić, D. (2018). Open multi-agent systems with discrete states and stochastic interactions. IEEE Control Systems Letters, 2(3), 375-380.

[VH25] Vizuete, R., & Hendrickx, J. M. (2025). Consensus on Open Multi-Agent Systems Over Graphs Sampled from Graphons. In 2025 IEEE 64th Conference on Decision and Control (CDC) (pp. 5282-5289).