Offre n°2020-02918

PhD Position F/M Differential Privacy for Machine Learning and Fairness

Le descriptif de l'offre ci-dessous est en Anglais

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

The Inria Sophia Antipolis - Méditerranée center counts 34 research teams as well as 8 support departments. The center's staff (about 500 people including 320 Inria employees) is made up of scientists of different nationalities (250 foreigners of 50 nationalities), engineers, technicians and administrative staff. 1/3 of the staff are civil servants, the others are contractual agents. The majority of the center's research teams are located in Sophia Antipolis and Nice in the Alpes-Maritimes. Four teams are based in Montpellier and two teams are hosted in Bologna in Italy and Athens. The Center is a founding member of Université Côte d'Azur and partner of the I-site MUSE supported by the University of Montpellier.

Contexte et atouts du poste

This is an industrial PhD thesis in collaboration with SAP Security Research.

SAP Security Research identifies and solves research problems relating to software and operational security, and shares its insights with others, particularly product groups at SAP, to drive advances in security and privacy.

The PhD candidate will also join NEO project-team https://team.inria.fr/neo/.
NEO is positioned at the intersection of Operations Research and Network Science. By using the tools of Stochastic Operations Research, the team members model situations arising in several application domains, involving networking in one way or the other. The aim is to understand the rules and the effects in order to influence and control them so as to engineer the creation and the evolution of complex networks.

The research activity will be supervised by

- Michele Bezzi, SAP
- Anderson Santana de Oliveira, SAP
- Giovanni Neglia, Inria, http://www-sop.inria.fr/members/Giovanni.Neglia/

Mission confiée

The student will work on applications of differential privacy [1] to machine learning.

There are increasing concerns among scholars and the public about bias, discrimination, and fairness in AI and machine learning. Decision support systems may present biases, leading to unfair treatment of some categories of individuals, for instance, systematically assigning high risk of recidivism in a criminal offense analysis system. Essentially, the analysis of whether an algorithm's output is fair (e.g. does not disadvantages a group with respect to others) depends on substantial contextual information that often requires human intervention. There are though several metrics for fairness that have been developed [2], which may rely on collecting additional sensitive attributes (e.g., ethnicity) before imposing strong privacy guarantees to be used in any situation. It is known that differential privacy has the effect of hiding outliers from the data analysis, perhaps compounding existing bias in certain situations [3]. This proposal encompasses the search for a mitigating strategy.

[1] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3–4), 211-407.

[2] Beutel, Alex, Chen, Jilin, Doshi, Tulsee, et al. Putting fairness principles into practice: Challenges, metrics, and improvements. In : Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. 2019. p. 453-459.

[3] Bagdasaryan, Eugene, Omid Poursaeed, and Vitaly Shmatikov. Differential privacy has disparate impact on model accuracy. Advances in Neural Information Processing Systems. 2019.

Principales activités

Research.

Compétences

- Completed MSc or equivalent degree in Data Science, Computer Science, Mathematics or other relevant field combined with a scientific interest in privacy and data protection
- Technical knowledge about Algorithms & Data Structures combined with sound programming skills (e.g. Python, JavaScript, JAVA or C++)
- Knowledge about, and hands-on experience with, machine learning, as well as a willingness to dive deep into current research aspects of the field (DNNs, VAEs, GANs, Tensorflow, Keras, PyTorch)
- Good communication & writing skills (We expect the successful candidate to be fluent in English.)
- Team Player, flexible & dedicated

Avantages

- 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 (after 6 months of employment) 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
- Social security coverage

Rémunération

Duration: 36 months
Location: Sophia Antipolis, France
Gross Salary per month: 1982€brut per month (year 1 & 2) and 2085€ brut/month (year 3)

Informations générales

- **Thème/Domaine :** Réseaux et télécommunications
- **Ville :** Sophia Antipolis
- **Centre Inria :** Centre Inria d'Université Côte d'Azur
- **Date de prise de fonction souhaitée :** 2020-10-01
- **Durée de contrat :** 3 ans
- **Date limite pour postuler :** 2020-10-04

Contacts

- **Équipe Inria :** NEO
- **Directeur de thèse :**
  Neglia Giovanni / Giovanni.Neglia@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Sécurité défense :
Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :
Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.