



Offer #2024-07586

PhD Position F/M Towards a comprehensive speech anonymization framework

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

Context

This position is part of the ANR SpeechPrivacy project, which aims to advance speech anonymization technology. The PhD student will be co-supervised by [Emmanuel Vincent](#) and [Mickaël Rouvier](#). He/she will have the opportunity to spend time in both [Multispeech](#) and [LIA](#), and to contribute to the project's overall goals including the organization of an international evaluation challenge [1].

[1] <https://www.voiceprivacychallenge.org/>

Assignment

Large-scale collection, storage, and processing of speech data pose severe privacy threats [2]. Indeed, speech encapsulates a wealth of personal information in the speaker's traits and the verbal and non-verbal content (e.g., speaker's identity, age and gender, ethnic origin, personality traits, health and socio-economic status, etc.). Compliance with privacy laws such as the European general data protection regulation (GDPR) calls for privacy preservation solutions tailored to speech technology. This has led to the development of voice anonymization methods [3] which aim to conceal the speaker's voice in verbal speech, and evaluation methods [1] which quantify the resulting speaker re-identification risk against linkability and singling-out attacks [4] under simplifying assumptions. This PhD aims to extend these methods into a comprehensive speech anonymization framework, which quantifies the real-world risk and accounts for non-verbal speech and other speaker traits.

[2] A. Nautsch, A. Jimenez, A. Treiber, J. Kolberg, C. Jasserand, E. Kindt, H. Delgado, M. Todisco, M. A. Hmani, M. A. Mtibaa, A. Abdelraheem, A. Abad, F. Teixeira, M. Gomez-Barrero, D. Petrovska, N. Chollet, G. Evans, T. Schneider, J.-F. Bonastre, B. Raj, I. Trancoso, and C. Busch, "Preserving privacy in speaker and speech characterisation," *Computer Speech and Language*, vol. 58, pp. 441–480, 2019.

[3] B. M. L. Srivastava, M. Maouche, M. Sahidullah, E. Vincent, A. Bellet, M. Tommasi, N. Tomashenko, X. Wang, and J. Yamagishi, "Privacy and utility of x-vector based speaker anonymization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 2383–2395, 2022.

[4] https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

Main activities

Existing average-case evaluation metrics overestimate the re-identification risk due to the unrealistic assumptions that the attacker is able to set the optimal decision threshold and knows when re-identification has succeeded. At the same time, they underestimate the risk for certain speakers or utterances which are more easily re-identifiable [5, p.105]. Therefore, a first goal is to develop reliability measures [6] enabling the attacker to decide whether re-identification has succeeded, and to study what makes certain speakers or utterances more easily re-identifiable so as to adapt the anonymization parameters accordingly without running costly speaker verification experiments.

Existing anonymization and evaluation methods also do not account for non-verbal speech (laughter, breathing, cough, scream, crying, etc.) and speaker traits (age, gender, phonation, articulation, prosody, accent, etc.). While classifiers have been developed for several of these traits [7,8], their anonymization and their use for inference attacks [4] remain to be studied. Therefore, a second goal will be to study the effectiveness of simple speech transformations [9] or specific synthesis techniques [10] for non-verbal speech anonymization, and to quantify the re-identification risk resulting from the combination of several speaker traits.

[5] B. M. L. Srivastava, "Speaker anonymization : representation, evaluation and formal guarantees," PhD thesis, University of Lille, 2021.

[6] P.-M. Bousquet, M. Rouvier, and J.-F. Bonastre, "Reliability criterion based on learning-phase entropy for speaker recognition with neural network," in *Interspeech*, pp. 281–285, 2022.

[7] Y. Lin, X. Qin, N. Jiang, G. Zhao, and M. Li, "Haha-Pod: An attempt for laughter-based non-verbal speaker verification," in 2023 IEEE Automatic Speech Recognition and Understanding Workshop, pp. 1-7, 2023.

[8] N. Obin and A. Roebel, "Similarity search of acted voices for automatic voice casting," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no.9, pp. 1642-1651, 2016.

[9] J. Patino, N. Tomashenko, M. Todisco, A. Nautsch, and N. Evans, "Speaker anonymisation using the McAdams coefficient," in Interspeech, pp. 1099-1103, 2021.

[10] D. Xin, S. Takamichi, A. Morimatsu, and H. Saruwatari, "Laughter synthesis using pseudo phonetic tokens with a large-scale in-the-wild laughter corpus," in Interspeech, pp. 17-21, 2023.

Skills

MSc degree in speech processing, machine learning, or a related field.

Strong programming skills in Python/Pytorch.

Prior experience with speech processing is an asset.

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 (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

Remuneration

2100€ gross/month (the 1st year)

General Information

- **Theme/Domain** : Language, Speech and Audio
- **Town/city** : Villers lès Nancy
- **Inria Center** : [Centre Inria de l'Université de Lorraine](#)
- **Starting date** : 2024-10-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2024-05-26

Contacts

- **Inria Team** : [MULTISPEECH](#)
- **PhD Supervisor** :
Vincent Emmanuel / emmanuel.vincent@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

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.