



Offre n°2024-07586

PhD Position F/M Towards a comprehensive speech anonymization framework

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

Contexte et atouts du poste

Cette offre s'inscrit dans le cadre du projet ANR SpeechPrivacy, qui vise à faire progresser les technologies d'anonymisation de la parole. Le doctorant sera co-encadré par [Emmanuel Vincent](#) et [Mickaël Rouvier](#). Il/elle aura l'opportunité de passer du temps à la fois dans [Multispeech](#) et au [LIA](#), et de contribuer aux objectifs généraux du projet, y compris l'organisation d'une campagne d'évaluation internationale [1].

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

Mission confiée

La collecte, le stockage et le traitement à grande échelle des données vocales posent de graves menaces pour la vie privée [2]. En effet, la parole renferme une multitude d'informations personnelles dans les traits du locuteur et dans le contenu verbal et non verbal (par exemple, l'identité du locuteur, son âge et son sexe, son origine ethnique, ses traits de personnalité, son état de santé et son statut socio-économique, etc.). Le respect des lois sur la protection des données telles que le RGPD nécessite des solutions de préservation de la vie privée adaptées aux technologies vocales. Cela a conduit au développement de méthodes d'anonymisation de la voix [3] qui visent à dissimuler la voix du locuteur dans la parole verbale et de méthodes d'évaluation [1] qui quantifient le risque de réidentification du locuteur par des attaques de corrélation et d'individualisation [4] sous des hypothèses simplificatrices. Cette thèse vise à étendre ces méthodes à un cadre complet d'anonymisation de la parole, qui quantifie le risque réel et prend en compte la parole non verbale et d'autres traits du locuteur.

[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

Principales activités

Les mesures existantes d'évaluation au cas moyen surestiment le risque de réidentification en raison des hypothèses irréalistes que l'attaquant est capable de fixer le seuil de décision optimal et sait quand la réidentification a réussi. En même temps, ces mesures sous-estiment le risque pour certains locuteurs ou phrases plus facilement réidentifiables [5, p.105]. Un premier objectif est donc de développer des mesures de fiabilité [6] permettant à l'attaquant de décider si la réidentification a réussi, et d'étudier ce qui rend certains locuteurs ou phrases plus facilement réidentifiables afin d'adapter les paramètres d'anonymisation associés sans effectuer d'expériences coûteuses de vérification du locuteur.

Les méthodes d'anonymisation et d'évaluation existantes ne tiennent pas non plus compte de la parole non verbale (rire, respiration, toux, cri, pleurs, etc.) et des caractéristiques du locuteur (âge, sexe, phonation, articulation, prosodie, accent, etc.). Bien que des classifieurs aient été développés pour plusieurs de ces traits [7,8], leur anonymisation et leur utilisation pour des attaques par inférence [4] restent à étudier. Un deuxième objectif sera donc d'étudier l'efficacité de transformations simples de la parole [9] ou de techniques de synthèse spécifiques [10] pour l'anonymisation de la parole non verbale, et de quantifier le risque de ré-identification induit par la combinaison de plusieurs traits du locuteur.

[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.

Compétences

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.

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

2100 € brut/mois (la 1ère année)

Informations générales

- **Thème/Domaine** : Langue, parole et audio
- **Ville** : Villers lès Nancy
- **Centre Inria** : [Centre Inria de l'Université de Lorraine](#)
- **Date de prise de fonction souhaitée** : 2024-10-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2024-05-26

Contacts

- **Équipe Inria** : [MULTISPEECH](#)
- **Directeur de thèse** :
Vincent Emmanuel / emmanuel.vincent@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.