Robust speaker recognition for smart assistant technology

**Starting date:** Earliest Fall 2019  
**Duration:** 3-4 years  
**Supervisors:** Md Sahidullah (Inria, France) and Tomi Kinnunen (University of Eastern Finland, Finland)

We are looking for an outstanding, highly motivated PhD candidate who will be working on the application of deep learning in speech processing area, more specifically on speaker recognition. The PhD position offer fits within the scope of a collaboration between the Multispeech team of Inria Nancy - Grand Est, France and Computational Speech Group, School of Computing, University of Eastern Finland. The PhD student will be jointly funded by both the institutes and he/she will spend time at both locations, to be detailed upon preparing PhD plan.

**Ideal candidate profile**
- Highly self-motivated and solution-oriented individual
- Master’s degree in computer science, electrical engineering or a closely related field
- Prior knowledge in signal processing, statistics, probabilistic modeling
- Experience with programming in Python and/or MATLAB
- Experience with PyTorch/TensorFlow is a plus
- Experience with speech processing is a plus
- Excellent written and oral English language skills

**Thesis objective**
Artificial intelligence systems, such as smart assistants, have become part of our everyday life. Recent advances in machine learning, including speech technology, have enabled us to interact naturally with such devices. Nonetheless, with an increasing demand for more intuitive user experience, the underlying speech technology is expected to work flawlessly. Due to large variety of the operational environments and intrinsic variation of human voice this is challenging. Automatic identification of persons using their voice, speaker recognition, can be easily fitted with smart assistant technology to deliver individual user experience. While speaker recognition system with deep learning based speaker embeddings are promising under controlled conditions, real-world conditions remain challenging due to inevitable variations in *acoustic conditions, channel, and distance between the user and microphone* to name a few.

The state-of-the-art speaker recognition system mostly uses short-term magnitude information where other important information, such as *prosody, phase, and long-term speech characteristics*, are not taken into consideration during embedding learning. This PhD work aims at improved understanding and modeling of these overlooked speech attributes and their implications to speaker characterization. To this end the selected candidate is expected to explore advanced signal processing and deep learning algorithms. Exact PhD plan is to be detailed and agreed together with the supervisors and the successful candidate.
About the labs
The Inria is a French national research institution focusing on computer science and applied mathematics. Inria has eight research centers. Inria Nancy – Grand Est is hosting more than 150 PhD students. Inria is a worldwide recognized research institution. Research at Inria is organised in “project teams” which bring together researchers with complementary skills to focus on specific scientific projects. With this model, Inria is able to explore original approaches with its partners in industry and academia and provide an efficient response to the multidisciplinary and application challenges of the digital transformation. The MULTISPEECH is a joint research team in at Inria Nancy – Grand Est between the Université of Lorraine, Inria, and CNRS. The goal of the MULTISPEECH is the modeling of speech for facilitating oral-based communication. More details about MULTISPEECH: https://team.inria.fr/multispeech/

University of Eastern Finland (UEF) is one of the largest universities in Finland with a multidisciplinary research profile. The Computational Speech Group of the School of Computing, physically in Joensuu, focuses on machine learning for speaker and language recognition. We provide modern facilities and research infrastructure with a strongly international environment. We hosted the Odyssey 2014 conference, have received funding both from the EU and Academy of Finland, and are known as a co-founder of the Automatic Speaker Verification and Countermeasures (ASVspoof) challenge series. We have also organized a number of international summer schools. More details: https://www.uef.fi/en/web/speech/

Inquiries
Informal inquiries with a CV can be addressed to Dr. Md Sahidullah <md.sahidullah@inria.fr> and Dr. Tomi Kinnunen <tkinu@cs.joensuu.fi>.

References