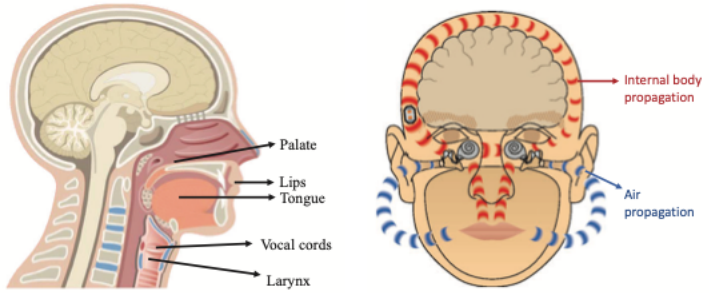
2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)

Enabling Secure Voice Input on Augmented Reality Headsets using Internal Body Voice

Jiacheng Shang and Jie Wu
Center for Network Computing, Temple University, Philadelphia, PA 19121

***Abstract*—Voice-based input is usually used as the primary input method for augmented reality (AR) headsets due to immersive AR experience and good recognition performance. However, recent researches have shown that an attacker can inject inaudible voice commands to the devices that lack voice verification. Even if we secure voice input with voice verification techniques, an attacker can easily steal the victim's voice using low-cast handy recorders and replay it to voice-based applications. To defend against voice-spoofing attacks, AR headsets should be able to determine whether the voice is from the person who is using the AR headsets. Existing voice-spoofing defense systems are designed for smartphone platforms. Due to the special locations of microphones and loudspeakers on AR headsets, existing solutions are hard to be implemented on AR headsets. To address this challenge, in this paper, we propose a voice-spoofing defense system for AR headsets by leveraging both the internal body propagation and the air propagation of human voices. Experimental results show that our system can successfully accept normal users with average accuracy of 97% and defend against two types of attacks with average accuracy of at least 98%.**

***Index Terms*—AR headsets, voice spoofing attack, liveness detection.**



(a) The human vocal system. (b) Two propagation paths.
Fig. 1. Human vocal system and two propagation paths of the voice.

devices that lack voice verification. Moreover, unlike other human biometrics, the human voice is often exposed to the public in many different scenarios, e.g., people making a presentation in public. Even if we secure devices with voice verification techniques, an attacker can easily steal the victim's voice using low-cast handy recorders and attack voice-based applications with the help of state-of-the-art voice synthesis/conversion software. Several security issues are, therefore, caused by the leakage of people's voices and pose a severe