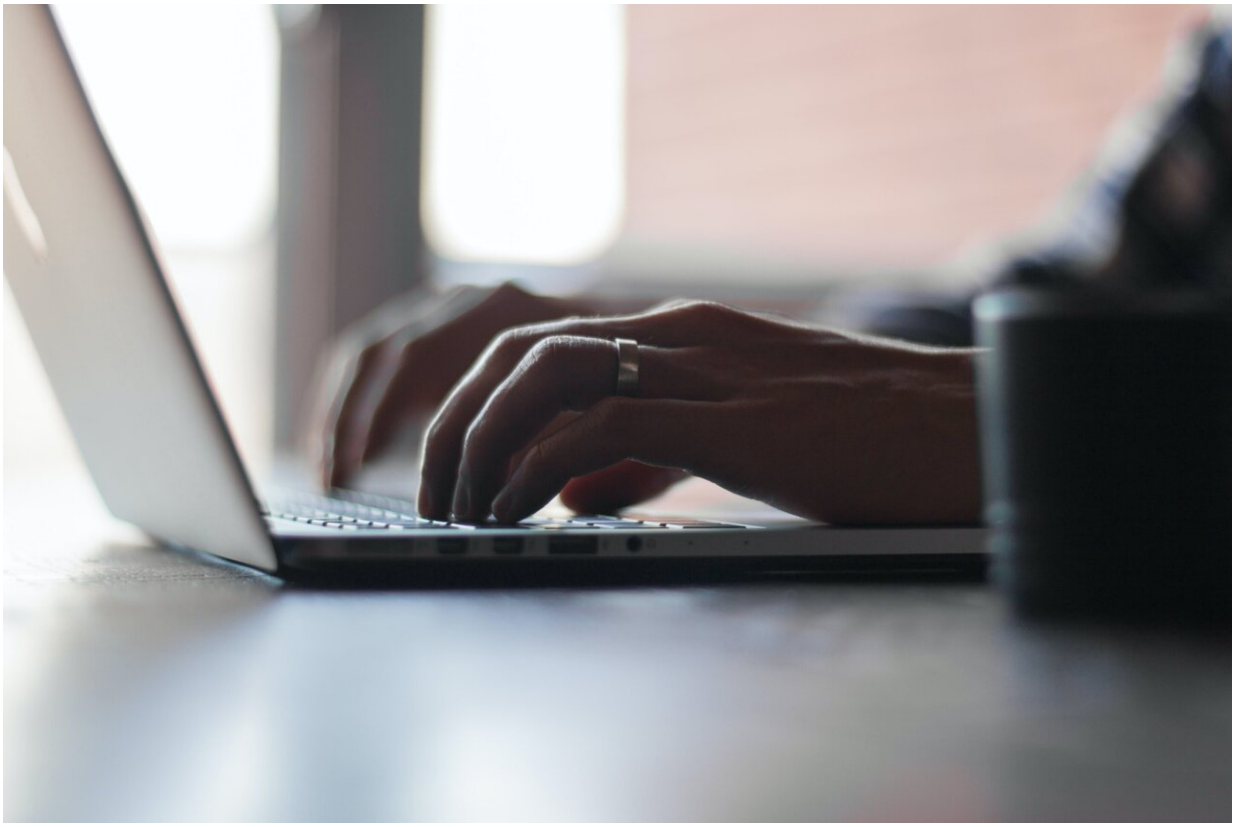
Stopping 'them' from spying on you: New AI can block rogue microphones

April 18 2022



Credit: Unsplash/CC0 Public Domain

Ever noticed online ads following you that are eerily close to something you've recently talked about with your friends and family? Microphones are embedded into nearly everything today, from our phones, watches,

and televisions to voice assistants, and they are always listening to you. Computers are constantly using neural networks and AI to process your speech, in order to gain information about you. If you wanted to prevent this from happening, how could you go about it?

Back in the day, as portrayed in the hit TV show "The Americans," you would play music with the volume way up or turn on the water in the bathroom. But what if you didn't want to constantly scream over the music to communicate? Columbia Engineering researchers have developed a new system that generates whisper-quiet sounds that you can play in any room, in any situation, to block smart devices from spying on you. And it's easy to implement on hardware like computers and smartphones, giving people agency over protecting the privacy of their voice.

"A key technical challenge to achieving this was to make it all work fast enough," said Carl Vondrick, assistant professor of computer science. "Our algorithm, which manages to block a rogue microphone from correctly hearing your words 80% of the time, is the fastest and the most accurate on our testbed. It works even when we don't know anything about the rogue microphone, such as the location of it, or even the computer software running on it. It basically camouflages a person's voice over-the-air, hiding it from these listening systems, and without inconveniencing the conversation between people in the room."

Staying ahead of conversations

While the team's results in corrupting automatic speech recognition systems have been theoretically known to be possible in AI for a while, achieving them fast enough to use in practical applications has remained a major bottleneck. The problem has been that a sound that breaks a person's speech now—at this specific moment—isn't a sound that will break speech a second later. As people talk, their voices constantly

change as they say different words and speak very fast. These alterations make it almost impossible for a machine to keep up with the fast pace of a person's speech.

"Our algorithm is able to keep up by predicting the characteristics of what a person will say next, giving it enough time to generate the right whisper to make," said Mia Chiquier, lead author of the study and a Ph.D. student in Vondrick's lab. "So far our method works for the majority of the English language vocabulary, and we plan to apply the algorithm on more languages, as well as eventually make the whisper sound completely imperceptible."

Launching 'predictive attacks'

The researchers needed to design an algorithm that could break neural networks in real time, that could be generated continuously as speech is spoken, and applicable to the majority of vocabulary in a language. While earlier work had successfully tackled at least one of these three requirements, none have achieved all three. Chiquier's new algorithm uses what she calls "predictive attacks"—a signal that can disrupt any word that automatic speech recognition models are trained to transcribe. In addition, when attack sounds are played over-the-air, they need to be loud enough to disrupt any rogue "listening-in" microphone that could be far away. The attack sound needs to carry the same distance as the voice.

The researchers' approach achieves real-time performance by forecasting an attack on the future of the signal, or word, conditioned on two seconds of input speech. The team optimized the attack so it has a volume similar to normal background noise, allowing people in a room to converse naturally and without being successfully monitored by an automatic speech recognition system. The group successfully demonstrated that their method works inside real-world rooms with natural ambient noise and complex scene geometries.

Ethical AI

"For many of us in the research community, ethical concerns of AI technology are an essential issue, but it seems to belong to a separate thought process. It is like we are so happy that we finally made a driving car but forgot to design a steering wheel and a brake," says Jianbo Shi, professor of computer and information science at the University of Pennsylvania and a leading researcher in machine learning. "As a community, we need to 'consciously' think about the human and societal impact of the AI technology we develop from the earliest research design phase. Mia Chiquier and Carl Vondrick's study poses the question 'How to use AI to protect us against unintended AI usages?' Their work makes many of us think in the following direction: Ask not what ethical AI can do for us, but what we can do for ethical AI? Once we believe in this direction, ethical AI research is just as fun and creative."

Chiquier will present her paper on April 25, 2022, at the International Conference for Learning Representations.

More information: Mia Chiquier et al, Real-Time Neural Voice Camouflage (2022). Available as a PDF at arXiv:2112.07076 [cs.SD] arxiv.org/abs/2112.07076

Provided by Columbia University School of Engineering and Applied Science