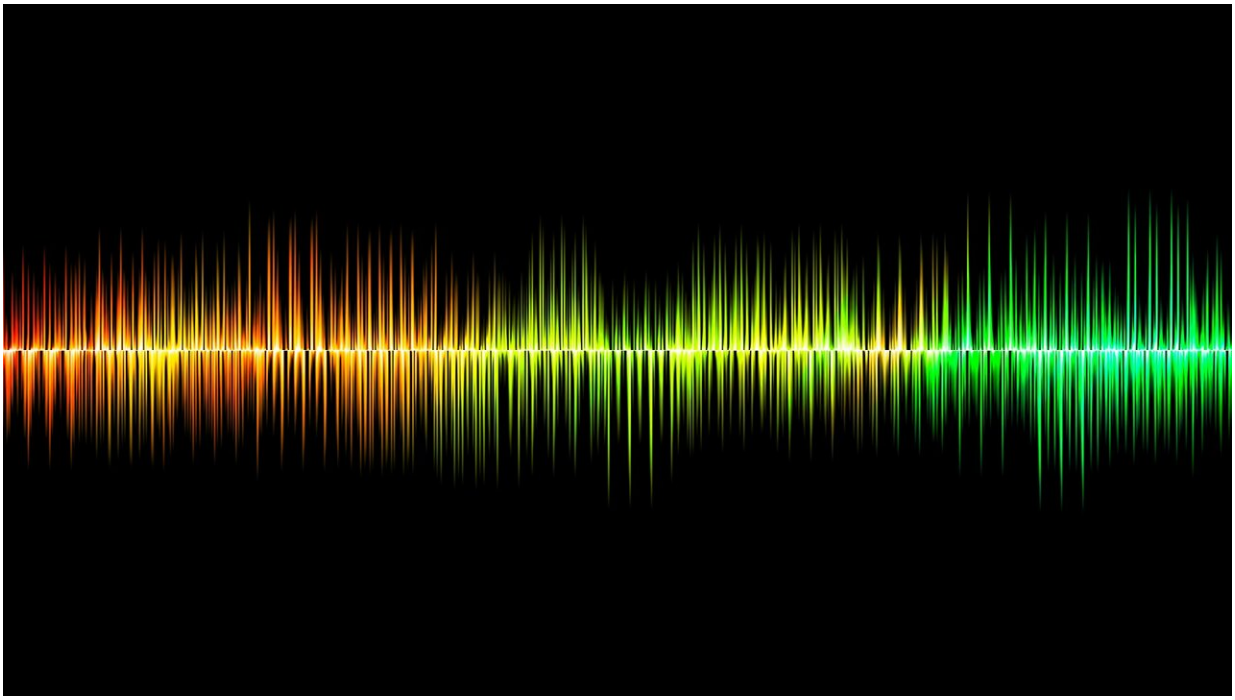
# DeFake tool protects voice recordings from cybercriminals

April 22 2024, by Shawn Ballard



Credit: Pixabay/CC0 Public Domain

In what has become a familiar refrain when discussing artificial intelligence (AI)-enabled technologies, voice cloning makes possible beneficial advances in accessibility and creativity while also enabling increasingly sophisticated scams and deepfakes. To combat the potential negative impacts of voice cloning technology, the U.S. Federal Trade Commission (FTC) challenged researchers and technology experts to

develop breakthrough ideas on preventing, monitoring and evaluating malicious voice cloning.

Ning Zhang, an assistant professor of computer science and engineering in the McKelvey School of Engineering at Washington University in St. Louis, was one of three winners of the FTC's Voice Cloning Challenge announced April 8. Zhang explained his winning project, DeFake, which deploys a kind of watermarking for voice recordings. DeFake embeds carefully crafted distortions that are imperceptible to the human ear into recordings, making criminal cloning more difficult by eliminating usable voice samples.

"DeFake uses a technique of adversarial AI that was originally part of the cybercriminals' toolbox, but now we're using it to defend against them," Zhang said. "Voice cloning relies on the use of pre-existing speech samples to clone a voice, which are generally collected from social media and other platforms. By perturbing the recorded audio signal just a little bit, just enough that it still sounds right to human listeners, but it's completely different to AI, DeFake obstructs cloning by making criminally synthesized speech sound like other voices, not the intended victim."

The project builds on Zhang's earlier work to thwart unauthorized speech synthesis before it happens. Zhang and the other two winners of the Voice Cloning Challenge, whose proposals focused on detection and authentication, illustrate the variety of approaches being developed to deter harmful practices and protect consumers from bad actors. The winners were selected by a panel of judges and will split $35,000 in prize money.

Provided by Washington University in St. Louis