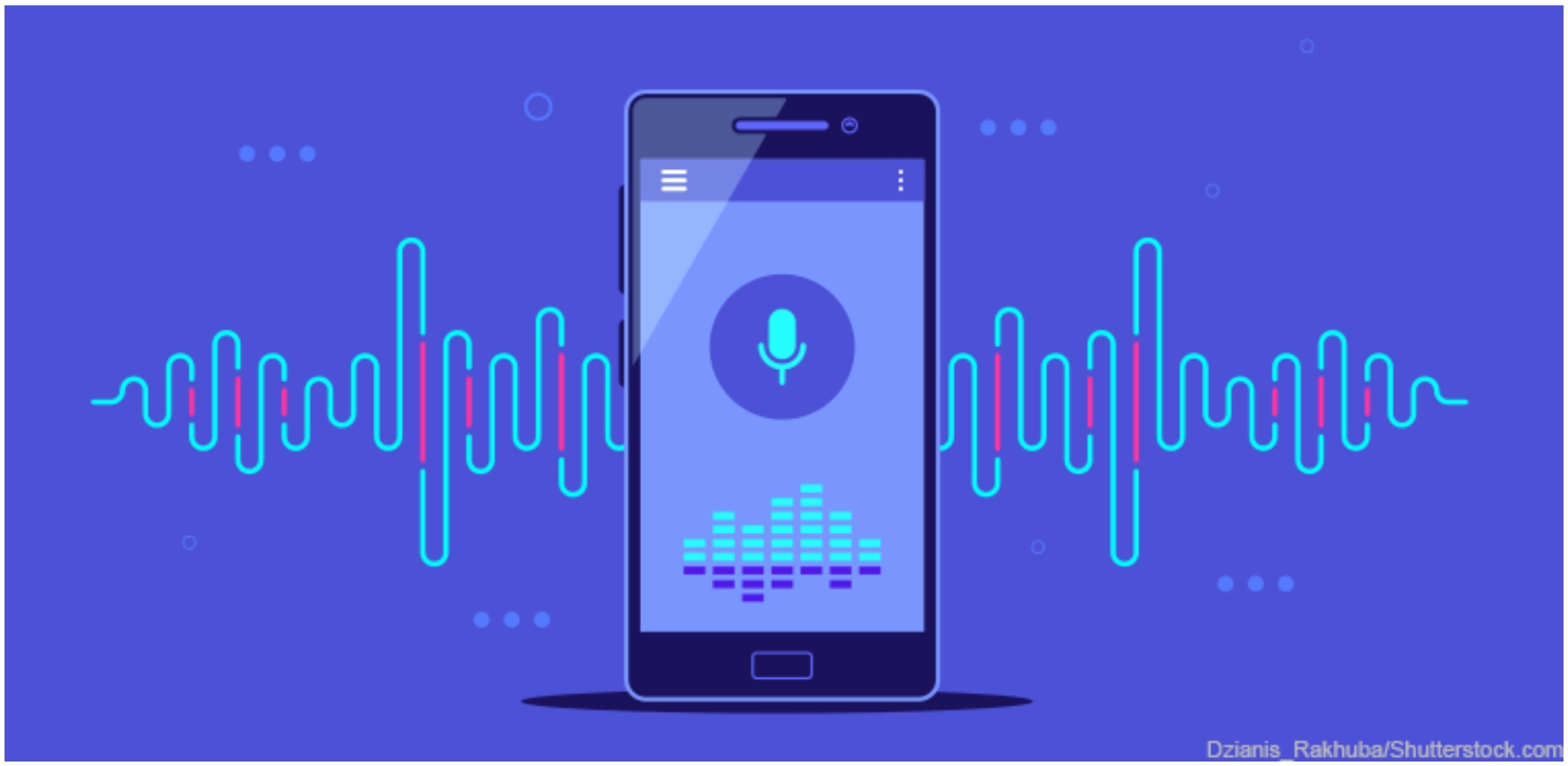

INDUSTRY INSIGHT

Masking the audio our devices hear is the best way to fight digital eavesdropping

BY MIKE FONG | AUG 01, 2019

Thanks in no small part to a barrage of news stories about smartphone spyware and misbehaving smart devices, there's a growing awareness that the cameras and microphones in our digital companions can be hijacked by sophisticated threat actors to remotely spy and eavesdrop on users. But while it's relatively straightforward for users of these devices to prevent illicit image capture by blocking built-in cameras, the complexities associated with blocking sound preclude a simple equivalent for microphones. To solve this gap in protection, sophisticated products have started to emerge that add noise near a device's microphones to mask sound in vicinity of the device.

This article will explore why on-device audio masking is being used to fight digital

eavesdropping and look at some considerations for maximizing the effectiveness of this technique.

Why audio data matters

Before discussing the ins and outs of preventing illicit audio capture, it's helpful to look at why threat actors seek audio data in the first place.

Sensitive and confidential conversations, of course, are digital eavesdropping's version of gold, as the information revealed in these discussions is often too sensitive or too fresh to be written down or recorded and therefore unavailable through other modes of hacking. Think of the foreign intelligence service learning about an adversary's strategic moves well in advance or the corporate spy learning about a competitor's product development while it's still in the early stages.

But it's not just long, uninterrupted sequences of dialogue that can be valuable to an eavesdropper. Random audio snippets, for example, can be pieced together to draw important conclusions, especially when the audio is accompanied with other data points. After the audio recordings of an Amazon Alexa user in Germany were mistakenly sent to another user, a neutral party listening to the audio files was able "to piece together a detailed picture of the customer concerned and his personal habits," Gizmodo reported.

And soon, eavesdroppers will be able to pick up information just from a person's vocal characteristics. A team at Carnegie Mellon University, for example, is developing machine learning-fueled voice forensics technology that can potentially yield a person's age, height, health status and much more from a voice sample. In fact, we're already seeing

artificial intelligence that reconstructs a person's face like based on a single sound bite. These kinds of inferences will prove to be invaluable to spies the world over.

Smartphone users might think, "I'll just tape the microphones on my smartphone and laptop and that should keep the bad guys from listening." Facebook CEO Mark Zuckerberg famously had a similar notion. Unfortunately, while providing a physical barrier over a microphone can muffle sound, the sound won't be stopped completely. Plus, if hackers have the tools to remotely install spyware on a smartphone or tap into a smart speaker, they likely have access to audio forensics technology for deciphering conversations distorted by a piece of tape.

Exacerbating the problem is the fact that turning off the microphones in devices is impossible or impractical. With today's popular smartphone models, users can disallow microphone access to any given app, but there's no way to completely disable the microphones from use, short of removing them from the device. And while some smart speakers -- like the Amazon Echo and Google Home -- have buttons or switches for shutting off their microphones, manually switching these back on requires touching the device, defeating the purpose of having voice-activated technology in the first place.

Bring on the noise

Enter audio masking. Similar to how white noise machines drown out the audio in a room, on-device audio masking -- whether delivered through an anti-surveillance smartphone case or smart speaker gatekeeper -- adds noise at the locations of the microphones. At the proper volume levels, the end result is that the device is essentially deafened, burying any valuable audio content beneath the noise floor of the masking signal and rendering recordings useless to eavesdroppers.

But not any noise will do. To prevent a conversation or audio snippet from being successfully deciphered using advanced audio forensics, the noise added to the audio mix must have the following three characteristics:

1. **Random:** The noise must be random, ideally generated using a true random number generator based on a source with high entropy. If the noise is repetitive or even pseudo-random, a recorded audio signal can be processed -- using two-channel adaptive filtering -- to essentially "subtract" the noise profile found in a reference file from the audio in a target file, delivering a version of the target file with significantly improved speech intelligibility. To help synchronize the timing of the two files, additional signal processing -- using fast Fourier transform analysis -- can help determine the repetition rate of the noise profile.

2. **Microphone-specific:** The audio masking must occur independently for each microphone, whether for the four microphones found in the iPhone XR or the seven found in the second-generation Amazon Echo. Doing so prevents an eavesdropper from using cross-correlation or known pseudo-random patterns to extract the audio from the masked output.

3. **Adaptive:** The level of audio masking must adapt to the volume of the audio being masked throughout the range of human speech (from a whisper to a shout). After all, if the level of audio masking were at maximum volume the entire time, the noise would be unbearable for the user, and if the level of audio masking stayed within a moderate range, it wouldn't be able to adequately mask louder speech and sounds.

With these three characteristics, a conversation's content (the words spoken) and context (accents, tones, number of participants, etc.) will be unidentifiable to an eavesdropper, as the final audio output will be indistinguishable from a recording of noise alone.

Of course, having the best audio masking in the world is meaningless if it gets in the user's way. Even though we may only want our microphones to be actively listening to us only a fraction of the time, we don't want any hiccups when it comes to making phone calls, recording audio messages or using a device's virtual assistant. For an audio-masking add-on for smartphones or other handheld devices, the solution is rather simple, as a physical mechanism can give the user control of when masking occurs. But a similar add-on for smart speakers requires some creativity; one novel solution allows the user to temporarily stop audio masking by speaking a custom wake word to the add-on, which then triggers the speaker by whispering the standard wake word.

Because our smartphones, smart speakers and other smart devices act as witnesses to our most private conversations and personal behaviors, limiting the exposure of these details while leveraging the powerful technology at our disposal is a challenging balancing act. On-device audio masking in an exciting development that promises to help us achieve that balance.

About the Author

Mike Fong is the founder and CEO of Privoro.

RELATED ARTICLES

- Facial recognition software prompts privacy, racism concerns in cities and states
- DARPA's disinformation detector
- Can government balance agency and commercial spectrum use?
- Stop chasing ghosts and build a threat hunting strategy

Phishing: The future is zero tolerance

MORE FROM PUBLIC SECTOR 360

Defense Systems

The Pentagon is looking for an AI ethicist

How the Army is advancing facial recognition

Pentagon awards $8 billion cloud business contract

Federal Soup

Trump announces 2020 raise in fed, locality pay

Understanding the biggest benefit of the TSP Modernization Act

Lawmakers pen letter urging Congress to protect feds' collective bargaining rights

GCN

About Us
Contact Us

Digital Edition
Advertise
Reprints
List Rental