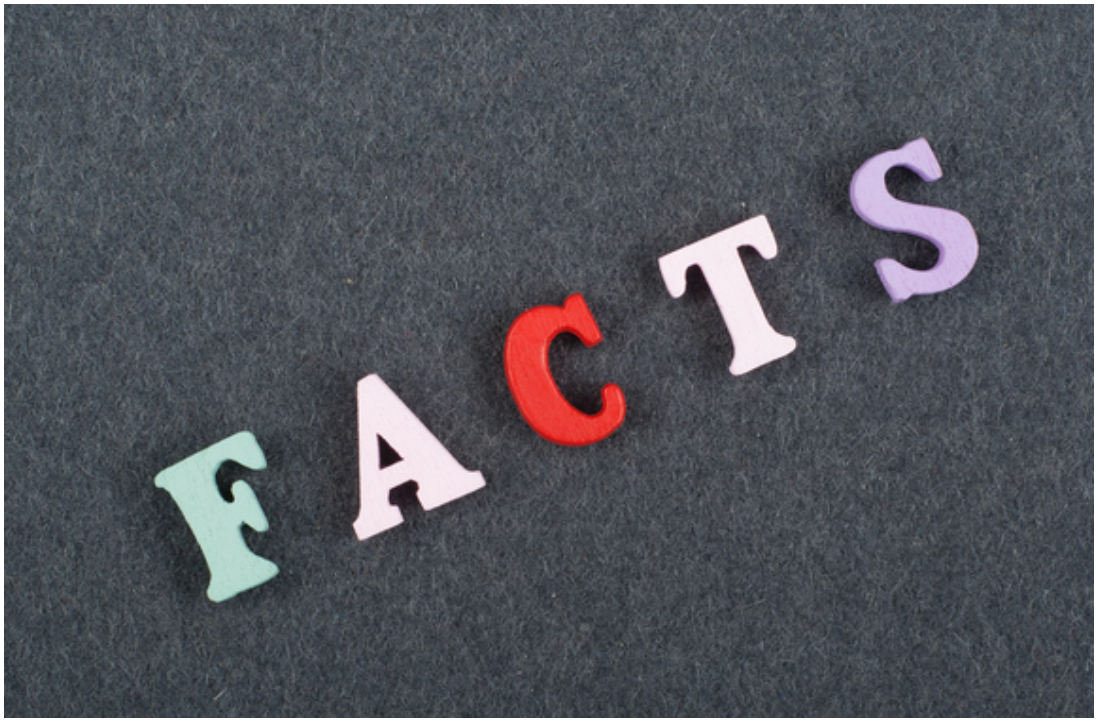
NUANCE | WHAT'S NEXT BLOG

Customer engagement, Authentication & fraud prevention

Can cyber criminals “compromise speech recognition systems with ease”?

Brett Beranek | Vice President & General Manager, Security & Biometrics

January 9, 2018



Not all scientific studies produce the voice of authority on their subject matter. Brett Beranek takes a closer look at why news stories on voice biometrics are missing the mark. What he finds will make you reconsider…

A recent Finnish university study on voice biometrics has been making headlines – and most of those news stories have been inaccurately summarizing the results with concerns as in our title above, leading many to believe that cyber crooks can compromise even the best speech recognition systems.

Before commenting on the article and the study, I feel it is important to highlight that Nuance's voice biometrics solutions have secured over five billion transactions to date, ***and not once has an impersonation attack been reported***. We have conducted several voice impersonation attacks with famous voice impersonators in the US and the UK, and none proved successful.

So why are the news stories missing the mark? The real story? Let's start with the conclusion.

> *“The results indicate a **slight increase** in the equal error rate (EER). Speaker-by-speaker analysis suggests that the impersonations scores increase towards some of the target speakers, but **that effect is not consistent**.”*

So how could the researchers write that “Voice impersonators can fool speaker recognition systems”? To understand that, you need to dig deeper into the study. Here are the actual data points:

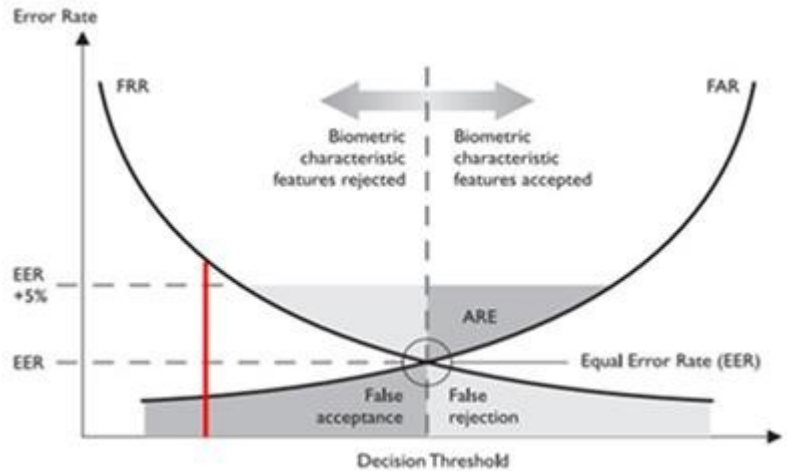
Table 5.3: Effect of impersonation on the pooled data from both impersonators in terms of *equal error rate* (EER %) for *Text-independent* and *Same text* test segments of 20 seconds duration.

Test	Case	GMM-UBM	i-vector Cosine	i-vector PLDA
Text independent	Baseline	10.83	6.80	4.36
	Impersonation	10.31	13.76	7.38
Same text	Baseline	10.38	13.73	7.16
	Impersonation	11.21	15.76	7.74

So what does this data mean? Let’s start with some definitions.

Text Independent– This is passive voice biometrics where a voiceprint is created from listening in on a normal conversation and that voiceprint is compared to a voiceprint on file.

Same Text– This is active voice biometrics where the user is given a specific phrase to repeat. (Often it is “My voice is my password”.) Once enrolled the user is asked to speak the phrase and then this new voiceprint is compared to the voiceprint on file.



False Accept Rate– This is the percentage of times a system incorrectly matches to another individual’s existing biometric. Example: fraudster calls claiming to be a customer and is authenticated.

False Reject Rate– This is the percentage of times the system does not match the individual to his/her own existing biometric template. Example: customer claiming to be themselves is rejected.

Equal Error Rate or EER– The EER is the location on the graph curve where the false accept rate and false reject rate are equal. In general, the lower the equal error rate value, the higher the accuracy of the biometric system. Note, however, that most operational systems are not set to operate at the “equal error rate”, so the measure’s true usefulness is limited to comparing biometric system performance.

GMM-UBM; i-vector Cosine; & i-vector PLDA– These are three different algorithmic approaches to voice biometrics. Notice that the latest technology, Deep Neural Networks, is not tested.

Now that we have that, the data showcases the following:

1. In one instance (text-independent GMM-UBM) the EER is decreased with impersonation – meaning that the **imposters were less successful** at generating a false accept than a random individual not attempting any voice mimicry.
2. In another instance (same text i-vector PLDA) the EER is virtually identical between the impersonation testing and random attacks. In other words, **imposters have the same performance via mimicry as a random individual** not attempting to modify their voice.
3. In four instances, there is an increase to the EER rate, but given the small sample size (60 voices) **the results are not statistically relevant**. In other words, a test performed with a larger sample may showcase opposite results.

Finally, and maybe most importantly, the researchers did not perform the tests with Nuance voice biometric technology. This is evident by the very high EER rates reported by the study as a “baseline”

result, ranging from 4.26% EER to 10.83% EER. No tests were conducted on deep-neural-network based voice biometric algorithms, the technology used by Nuance and deployed through scores of enterprises worldwide.

In conclusion, although this topic does merit additional research, Nuance will continue its focus on improving our ability to address actual fraud attack vectors, including brute force attacks, voice imposters, and recording attacks while continuously improving the voiceprint and also improving mitigating strategies for future attack vectors that we believe will eventually be used by fraudsters such as synthetic speech attacks.

Contact us if you would like to learn more about the great strides Nuance has made in Voice Biometrics.

Tags: Voice biometrics, Speech recognition, Voice recognition

More Information

About Brett Beranek

Brett Beranek is responsible for overseeing the security and biometric line of business at Nuance, a Microsoft company. In this role for the past 12 years, Beranek has brought Nuance to a leadership position in the biometric authentication and biometric fraud prevention space. A thought leader in the field of biometrics, Beranek is a frequent contributor in industry events and the media on the topic of AI technology and it’s use by the fraud community, and how society can mitigate against these evolving threats. Prior to Nuance, he held various leadership positions in the biometrics and security industry. He has earned a Bachelor of Commerce, Information Systems Major, from McGill University as well as an Executive Marketing certificate from Massachusetts Institute of Technology’s Sloan School of Management. Beranek is also a certified Master Fraud Prevention Black Belt professional.

View all posts by Brett Beranek