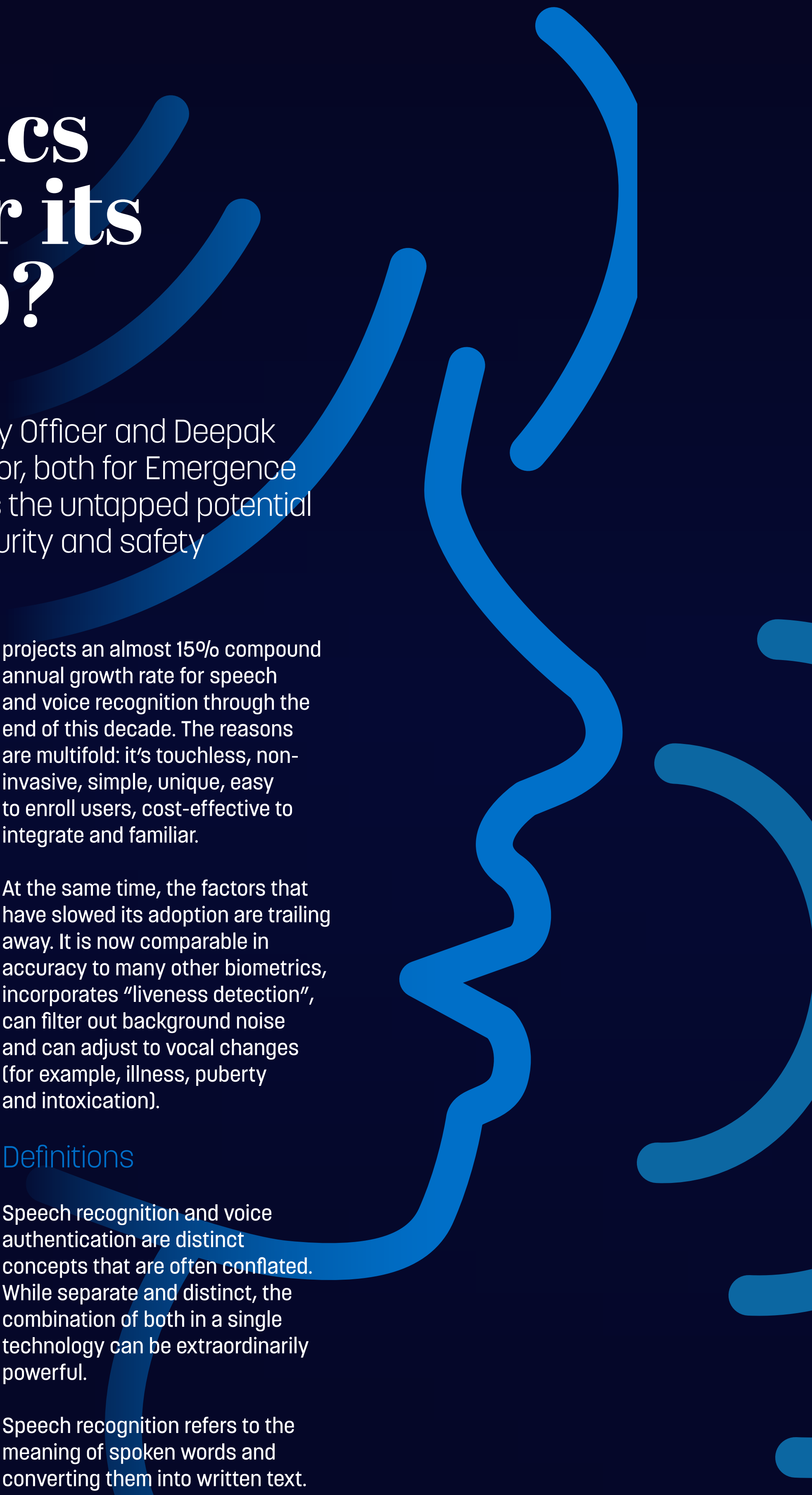
Is voice biometrics ready for its mic drop?



Michael Gips, Chief Strategy Officer and Deepak Chandran, Technical Advisor, both for Emergence Technology Group, discuss the untapped potential of voice biometrics for security and safety

Tens of millions of people have facial recognition on their smartphones. Fingerprints, which can also unlock smartphones, have been used for access control for decades. Iris recognition is deployed for national ID programs, high-security access control and border management. Among other qualities that can be used for biometric technologies identified by The Biometrics Institute are ear shape, scleral vein, gait, hand geometry, odor and keystrokes.

Voice is a well-recognized biometric, but it is overwhelmingly used for non-security purposes – think of your everyday commands to Siri or Alexa. To be sure, banks and some customer support centers use limited voice recognition for authentication purposes and to detect fraud, but security uses are not widespread.

Is voice biometrics for security due for a breakthrough? Many experts think so. Nacho de Marco of the Forbes Technology Council has dubbed voice biometrics "the next big thing". Grand View Research projects an almost 15% compound annual growth rate for speech and voice recognition through the end of this decade. The reasons are multifold: it's touchless, non-invasive, simple, unique, easy to enroll users, cost-effective to integrate and familiar.

At the same time, the factors that have slowed its adoption are trailing away. It is now comparable in accuracy to many other biometrics, incorporates "liveness detection", can filter out background noise and can adjust to vocal changes (for example, illness, puberty and intoxication).

Definitions

Speech recognition and voice authentication are distinct concepts that are often conflated. While separate and distinct, the combination of both in a single technology can be extraordinarily powerful.

Speech recognition refers to the meaning of spoken words and converting them into written text. It is commonly used in applications such as dictation software, voice assistants and call transcription services. Voice recognition does not identify the speaker.

Voice authentication verifies the identity of the speaker. It analyzes characteristics such as pitch, rhythm and tone to create a "voiceprint" that can be used for secure access control in various applications to include phone banking and smart home devices. Unlike speech recognition, voice authentication does not necessarily comprehend the content of what's being spoken.

Voice recognition can be compared to understanding the language of a book you're reading. You can grasp the meaning of the words and follow the story without knowing the identity of the author. Voice authentication is akin to recognizing a friend's handwriting in a letter before reading its content – you can identify their penmanship without knowing what they wrote.

Use cases for voice and speech technologies

Intelligent assistants have become an integral part of our lives, providing a hands-free way to control our devices, search for information and manage our tasks. However, while these voice-activated/speech-recognizing assistants like Alexa and Siri are convenient, they also come with significant concerns – they store data in the cloud and cannot recognize individual users accurately. New voice biometrics technology is emerging that fills the gap, offering a secure and reliable solution to these issues.

This technology employs an advanced on-edge voice capture algorithm to create a unique voice print for each user. This voice print is a distinctive audio pattern the system uses to identify and authenticate the user. The technology can store multiple prints of the same user to accommodate changes in voice per time of day, per environment, over time and so on.

New developments allow the technology to accurately identify a specific user's voice even in noisy environments, regardless of their emotional state or physical health. Coupled with speech recognition of particular words and phrases, the technology can verify that a specific person uttered a particular word or phrase.

> “NEW DEVELOPMENTS ALLOW THE TECHNOLOGY TO ACCURATELY IDENTIFY A SPECIFIC USER'S VOICE EVEN IN NOISY ENVIRONMENTS.”

This dual technology unlocks multiple potential uses.

Home automation: smart home devices can recognize who is giving the command, allowing personalized experiences. For instance, when a user says, "Play my playlist," the system will play the specific user's collection of songs, not a generic one.

Security systems: technology can be incorporated into security systems for voice-based authentication, thus adding an extra layer of security, as the system will only grant access to specific phrases and registered voice prints.

Customer service: companies can authenticate customers during phone-based customer service, thus streamlining the process, as customers won't have to answer multiple security questions. ▸

Health monitoring: technology can be used in health apps to monitor patients' health through their voice. Any significant changes in the voiceprint could alert medical professionals to potential health issues.

Personalized learning: in education, voice biometrics can be used to create personalized learning experiences. For instance, an educational app could adapt its content based on who it recognizes is using the app.

Life safety: voice biometrics have the potential to enhance safety and security. In terms of life safety, it could be used to activate emergency response systems when individuals are unable to trigger an alarm. For instance, if a senior citizen falls, they could call for help using their voice as a means of identification. Or, in the case of a home invasion, a trigger word could notify emergency response personnel, family members, friends, neighbors and others, geolocate where the victim is and send help on the way.

Hospitality: hotels could implement voice biometrics for check-ins and room access instead of dispensing key cards and codes. This would allow guests to unlock doors, order room service or request hotel information by using their unique voice. For example, simple oral commands could allow a guest to tune to a specific TV channel, turn on a particular light, close the shades or set a wakeup call.

Banks: banks benefit from voice biometrics by authenticating customer identities and authorizing transactions. The technology provides a layer of security beyond frequently defeated passwords or PINs.

> “BANKS BENEFIT FROM VOICE BIOMETRICS BY AUTHENTICATING CUSTOMER IDENTITIES AND AUTHORIZING TRANSACTIONS.”

Additive manufacturing: voice commands can play a role in controlling industrial 3D printers and other equipment. By recognizing authorized users through their voices this technology can control and monitor usage.

Call centers: call centers can streamline their verification processes by implementing voice biometrics. Use of knowledge-based questions can help call centers identify customers through their responses and their voices.

Automotive: motor vehicles can leverage voice recognition technology to personalize settings and experiences for drivers. Driver preferences, entertainment options and other features would be based on vocal characteristics. This application provides an important safety function because it keeps the driver’s eyes on the road instead of on dashboard screens, knobs and buttons.

Considerations

When considering potential applications for voice biometrics, users should ask themselves the following questions.

1. Can the system recognize both vocal patterns and specific words? The combination of understanding words and recognizing an individual’s vocal patterns provides more robust security

2. Is the system language-dependent? This question is only relevant for speech recognition. Is the technology sophisticated enough for applications in which system users will be using multiple languages?

3. Is your voice data stored in the cloud or on much less intrusive edge devices? Exactly what data? Cloud storage raises privacy concerns. Amazon’s Echo device has been known to record private conversations and send them to so-called “graders,” who evaluate the system to improve its voice recognition. On occasion, private conversations have been recorded and sent to random people on the user’s contact list, according to reports

4. Can the system operate effectively in a loud environment? Many systems are badly degraded by high-volume ambient noise such as vehicle traffic, conversations, music, air conditioners, television or wind. An effective system should have noise-cancelling

software to screen out or eliminate background sounds, short of extremely close and loud noises such as nearby blenders on high settings

5. Can the system be spoofed by your recorded voice or a soundalike? Technically, any voice can be spoofed. However, so can any biometric – fingerprints, irises, passwords and so on. Voice biometrics that use specialized algorithms for voice-liveness detection and other features prove difficult to spoof

> “SPEECH RECOGNITION AND VOICE AUTHENTICATION ARE DISTINCT CONCEPTS THAT ARE OFTEN CONFLATED.”

6. Can the system automatically call 911 without the need for a callback (i.e., it will work with NextGen 911)? For decades, the 911 system has been based on analog technology. Next Generation 911 (NG911) will allow users to communicate with public safety answering points via voice, photos, images and text messages. When NG911 is operational, 911 centers will be able to receive and accept calls – and send help – without requiring a callback for confirmation of location or specific individual details. This will greatly speed up emergency response and avoid allowing an intruder to cancel an emergency call

8. Can the technology be integrated with other systems and devices (alarm panels, sensors, routers, access control systems, video cameras, anything IP-enabled)? End users already have dozens of internet-connected devices. Few customers are looking for another box to affix to the wall or another assistant to take up room on a desk. The voice technology should be able to integrate via application programming interfaces (APIs) with other original equipment manufacturer (OEM) systems

9. Can it be incorporated into/integrated with wearables such as pendants, bracelets, watches and badge holders? Forthcoming uses of voice technology for personal safety include coded distress calls in cases such as assaults on campus, abductions, armed robberies and missing persons. Voice technology embedded into jewelry, clothing and other wearable accessories, when connected to a mesh network, can summon immediate assistance no matter where the user is

10. Can it add business value to the organization? Many organizations see security investments as expenditures, rather than a value add. Voice biometrics that generate revenue as well as provide security or life safety become a much more attractive purchase. For example, a system that can be used to summon emergency assistance may also be used in a hotel for bespoke concierge services. It could provide an elite level of service to loyalty-program members by, say, enabling them to check in at the entrance vestibule and avoid the long check-in line

Speech and voice are exquisitely powerful tools with the potential to unlock doors, trigger alarms or call for help. Exciting new developments are transforming the world of speech and voice-related biometrics. While other biometric technologies will likely always have a place in security, the world is just starting to learn the power of its own voice. ▪

About the authors

Michael Gips is the Chief Strategy Officer at Emergence Technology Group, which develops voice recognition and authentication technology. He is the former Chief Knowledge and Learning Officer and Chief Security Officer for ASIS International.

Deepak Chandran is Technical Advisor for Emergence Technology Group. He has developed cutting-edge technologies for the smart grid, aquaculture, electric vehicles, lighting and many other applications.