# New voice verification technology prevents impersonators from obtaining voiceprints

September 17 2012

Computer users have learned to preserve their privacy by safeguarding passwords, but with the rise of voice authentication systems, they also need to protect unique voice characteristics. Researchers at Carnegie Mellon University's Language Technologies Institute (LTI) say that is possible with a system they developed that converts a user's voiceprint into something akin to passwords.

The system would enable people to register or check in on a voice authentication system, without their actual voice ever leaving their smartphone. This reduces the risk that a fraudster will obtain the person's voice biometric data, which could subsequently be used to access bank, health care or other personal accounts.

"When you use a speaker authentication system, you're placing a lot of faith in the system," said Bhiksha Raj, an associate professor of language technologies. "It's not just that your voiceprint might be stolen from the system and used to impersonate you elsewhere. Your voice also carries a lot of information—your gender, your emotional state, your ethnicity. To preserve privacy, we need systems that can identify you without actually hearing your voice or even keeping an encrypted record of your voice."

Raj and Manas Pathak, a recent Ph.D. graduate of the LTI, have devised a method for converting a voiceprint—a spectrogram that represents the acoustic qualities of speech—into alphanumeric strings that can serve as passwords. They will present the work as a keynote address Sept. 21 at the Information Security Conference in Passau, Germany.

Because a person's voice never sends the same signal twice, even when repeating the same word or phrase, converting the voiceprint into a single password won't do. Instead, the CMU system uses different mathematical functions to generate hundreds of alphanumeric strings. To authenticate the user, the system compares all of the strings with those that the system has on file from the initial registration; if enough of the strings match, the user is authenticated.

The system also adds what the researchers call "salt"—a random string of digits unique to each smartphone—to the alphanumeric strings to provide an additional level of security.

In tests using standardized speech datasets, Raj and Pathak found that their system was accurate 95 percent of the time. The privacy-preserving method is computationally efficient, so it could be used with most smartphones, they noted.

But Raj also warned that improving the security of voice authentication systems would be just a first step to protecting privacy overall. "With increasing use of speech-based services, such as the iPhone's Siri assistant or personal videos uploaded to YouTube, the issue of the privacy of users' speech data is only just beginning to be considered," he said.

Provided by Carnegie Mellon University