

Voice Authentication Speaks to the Marketplace

Steven J. Vaughan-Nichols

Although the hype over biometric security has cooled down as users have gained exposure to the technology, one particular approach has started to become increasingly popular: voice authentication.

Individuals trying to make a purchase with a credit card, access a protected computer system, or retrieve account information from their bank can speak into a telephone and have their voice identified by a voice-authentication system to verify they are who they claim to be.

One US railroad uses ScanSoft's SpeechSecure voice-authentication software to ensure that the customer releasing a rail car after it's been unloaded is authorized to do so, noted ScanSoft spokesperson Marie Ruzzo.

Judith Markowitz, president and founder of J. Markowitz, Consultants, a voice-biometrics consultancy, said, "2004 has been an interesting and active year for security as a whole, including voice authentication, because the US government is distributing considerable homeland-security funding. And so, I'm seeing a lot of action not only for voice but for other advanced biometric security uses."

Jackie Fenn, vice president at Gartner Inc., a market research firm, said government will be the primary voice-authentication adopter to address its many security concerns. She said corporate adoption will continue to grow



slowly until biometric readers are routinely embedded in [private-branch-exchange phone systems].

A growing number of voice-authentication products are appearing, such as Courion's PasswordCourier, Nuance Communications' Nuance, Vocent Solutions' Confirmed Caller, Voice.Trust's Voice.Trust server, and Voice-vault's Voicevault.

However, concerns in such areas as security and accuracy may be significant hurdles to the technology's widespread adoption.

THE TECHNOLOGY

Essentially, voice-authentication systems capture and digitize speakers' voices. The basic equipment is a microphone or telephone to input speech, an analog-to-digital converter to digitize the spoken words, a high-powered computer, and a database to store voice characteristics.

Typically, these systems match a voice's harmonic and resonant frequencies, as well as the way the speaker

pronounces phonemes—a language's smallest distinctive sounds—against an authorized user's digital voiceprint. The voiceprint is created when the authorized user enrolls in the authentication system, and it is subsequently stored as a digital file in a database.

The system calculates a score that indicates how closely the spoken voice matches the stored voiceprint for the person the speaker claims to be.

Alvin F. Martin, a mathematician at the US National Institute of Standards and Technology and an organizer of NIST's annual evaluation of speech-recognition programs, noted, "The increase in processing power in modern computers has helped make voice authorization more effective."

In addition to chips that can quickly process the large amounts of information involved in voice authentication, the systems need huge memories to store the data and pattern-matching technologies to compare live speech with stored voiceprints.

SPEAKING OF ADVANTAGES

As a biometric identifier, voice authentication has much to offer, said Steve Bittner, vice president of development for Convergys, a business and call-center services company.

For example, the technology permits remote authentication, unlike other biometric approaches such as fingerprint or iris scans. A user can enroll in and work with a voice-authentication system from a remote location via a telephone.

Also, many users are more comfortable identifying themselves by speaking than by submitting to fingerprint or iris scans, which are frequently seen as invasive.

Voice authentication can also reduce the cost of handling customer-service calls, according to Bittner. Voice-authentication systems with speech recognition could verify a speaker's identity, determine the spoken reason for the call, and forward the call to the appropriate service. This would save money by reducing the number of call-center employees.

Industry Trends

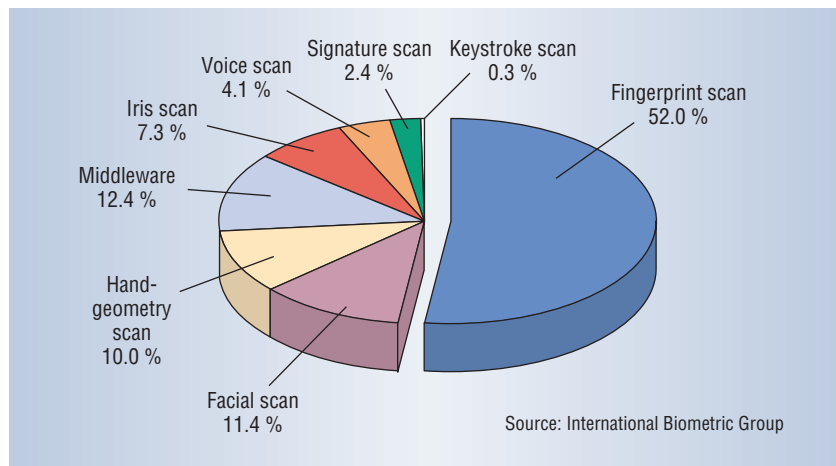


Figure 1. Voice authentication represents only a small part of the overall biometrics market.

Bank of America senior vice president Tim Wishon noted that his company used this approach to combine about 4,000 toll-free numbers into a much less expensive system using just one toll-free number.

USING VOICE AUTHENTICATION

Various government agencies use voice authentication for security purposes, such as ensuring that only authorized users have access to computer files or buildings.

There is also some commercial adoption of voice authentication, primarily by financial services companies. Markowitz said the technology is a good fit for the industry because of its security requirements and because customers like to perform many transactions and activities, such as verifying account balances and resetting passwords, via telephone.

In addition, merchants are using voice authentication for telephone-based, credit-card transactions, to reduce the risk of fraud by verifying that the voice on the line is that of the card's owner.

Voice authentication is particularly well suited for remote network and system access, employee timesheet record keeping, and other applications that require callers to use the same basic identification process, such as providing their mothers' maiden name, during authentication.

Voice authentication is also good for security processes that must identify many individuals because remote users need only a telephone and companies thus don't have to buy expensive equipment for them.

HURDLES

To inspire confidence and encourage more widespread adoption, voice authentication must overcome several obstacles.

For example, said Michelle M. Shen, consulting manager of ePolymath Consulting, a biometric consultancy, "The time [required] to verify a customer can be very long. Voice templates are so much larger than other kinds of biometric information. For example, data associated with a fingerprint may take up only 10 Kbytes, while a voiceprint typically takes up from 500 Kbytes to 1 Mbyte. This makes fast database servers and quick-filtering software a must."

Also, organizations may not feel comfortable adopting such a new technology yet. In addition, biometrics, after receiving so much hype a few years ago, has been shown to be less than completely reliable at times.

Security

As with any technology that allows access to sensitive systems, there are concerns about whether hackers could

compromise voice-authentication systems.

According to Markowitz, someone could play a recording of someone's voice to fool a low-end voice-recognition system. However, today's more sophisticated systems create detailed voiceprints that wouldn't match readily with a recorded voice.

Skilled human imitators, though, could still fool a pure voice-authentication system in many cases.

Consistent accuracy

Voice authentication is the least accurate biometric-security system, according to Gartner's Fenn.

In accuracy tests in lab settings, where environmental variables are controlled, voice-authentication systems compare favorably with other biometric approaches. In real-world use, though, behavioral and environmental factors such as background noise or changes in users' voices due to health, fatigue, or other causes reduce voice-authorization systems' accuracy.

"Voice characteristics vary with your age, your metabolic state, your emotional state, and all the ways you can say [various words]," said George Doddington, a speech-recognition expert and consultant to the US government. This makes relying on voice authentication alone as a security measure problematic.

"There are many breakdown points in voice authentication," explained ePolymath's Shen. A typical example occurs when people use a different type of phone for authentication than they did for enrollment. For example, a cellular phone used in traffic may produce somewhat different voiceprints than a high-quality wireline phone.

CLEARING THE HURDLES

Researchers are looking for ways to overcome voice authentication's obstacles.

Security

Some voice-authentication applications offer improved security via a two-

factor process, in which a user provides a voice sample along with another authenticating detail—such as a password or account number—in response to a question from the system. Voice authentication does the initial speaker identification and then speech recognition recognizes the user's answer to the contextual question.

The system rejects users who can't answer the questions correctly or refers them to a live agent.

Accuracy

Researchers are taking several approaches to improve voice authentication's accuracy.

For example, David Frogel, Courion's director of business development, said his company authenticates users more accurately based on a scoring system that compensates for a number of external factors that could change a speaker's voiceprint. These factors include whether a call is coming from an internal or external source or is affected by environmental variables such as background noise.

Although the voiceprint may be a bit different, the speaker's voice still must substantially match it for authentication to occur.

Speaker model synthesis. Nuance uses a speaker-model-synthesis approach to develop a machine-learning algorithm that identifies changes in a voice template—the stored master record of a voice—based on different equipment used. The system can recognize equipment by its transmission characteristics, such as cellular phones' narrower high and low voice-frequency ranges.

Over time, for all speakers, the system creates a transform voice template for each type of equipment used.

Model adaptation. Model adaptation is also a key to improving voice-authentication accuracy, said Kevin Farrell, ScanSoft's director of speaker verification. This approach creates a more accurate voice template by adjusting an individual's voiceprint parameters—such as harmonic and

resonant frequencies—over time, based on additional voice data received from encounters with the speaker after enrollment.

Analyzing new factors. NIST's Martin noted that modern voice-authentication systems are analyzing and classifying speech factors other than harmonic and resonant characteristics, such as word combinations, accents, and additional linguistic and idiomatic features.

At the Massachusetts Institute of Technology's Lincoln Labs, senior staff member Douglas Reynolds is working on ways to analyze and classify previously unexamined acoustic information such as voice pitch, pauses, and pronunciation style.

Meanwhile, at IBM's Thomas J. Watson Research Center, Ganesh Ramaswamy and other researchers are developing their *conversational biometrics* technique by analyzing and classifying multiple types of speech-related information, including pronunciation and how speakers use sounds like "uh" when thinking of what to say.

According to Courion's Frogel, "Voice-based authentication is already gaining traction in the marketplace." And recent enhancements, such as user-friendly interfaces, are increasing the technology's popularity. However, products still need additional improvements, such as faster voice recognition and elimination of the need for users to repeat phrases many times.

Markowitz said that more effectively combining voice authentication and speech recognition might help address these problems.

Another challenge is smoothly integrating voice authentication with other systems. Currently, organizations must use proprietary middleware or custom integration. Vendors are exploring ways to solve this problem via standards such as the Common Biometric Exchange File Format, the Voice Extensible Markup Language, and

programming interfaces such as BioAPI.

In the marketplace, said ePolymath's Shen, "Overall growth has been mediocre. If voice authentication is to grow, its real potential will be in financial services."

As Figure 1 shows, the International Biometric Group consultancy found that in 2003, voice authentication accounted for only 4.1 percent of the \$928 million biometrics market.

The IBG says voice authentication is just not accurate enough yet to increase its market share substantially in the near future. ■

Steven J. Vaughan-Nichols is a freelance technology writer based in Arden, North Carolina. Contact him at sjvn@vna1.com.

Editor: Lee Garber, *Computer*, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1314; l.garber@computer.org

Computer Wants You

Computer is always looking for interesting editorial content. In addition to our theme articles, we have other feature sections such as Perspectives, Computing Practices, and Research Features as well as numerous columns to which you can contribute. Check out our author guidelines at

www.computer.org/computer/author.htm

for more information about how to contribute to your magazine.

Innovative Technology for Computer Professionals
Computer