An Exploration of Voice Biometrics

Lisa Myers

An Exploration of Voice Biometrics

By Lisa Myers

GSEC Practical Assignment version 1.4b Option 1

Date Submitted: Monday April 19, 2004

Table Of Contents

Summary

Biometrics is, in the simplest definition, something you are. It is a physical characteristic unique to each individual. Using biometrics to identify individuals is a practice as old as ancient Egypt. Today, it is becoming more and more popular to use biometrics to identify people and authenticate them for access to secure areas and systems.

This paper will briefly define biometrics and enumerate the various kinds. I will examine in more depth voice biometrics, specifically how voice biometrics compares with other biometric technologies, how accurate it is, and how it is used to accomplish identity authentication. I will discuss privacy issues with the technology. Finally, I will explore how the technology has evolved, and some current and future applications of voice biometrics in our daily lives, illustrating why voice biometrics have significant future potential.

What is biometrics

A biometric is a physical characteristic, a measure of a biological trait such as a fingerprint. Biometrics has a very useful application in security; it can be used to authenticate a person's identity and control access to a restricted area or electronic system, based on the premise that certain of these physical characteristics can be used to uniquely identify individuals. All security systems that use user-based authorization require users to be accurately identified to ensure that the correct access privileges are granted. Biometrics as an authentication tool is very powerful because unlike other techniques currently used to authenticate people, such as passwords or access control badges, it cannot be easily taken away, lost, counterfeited, or forgotten. There are several categories of biometrics: fingerprints, hand geometry, retina, iris, face, handwriting, and voice.

Fingerprint biometrics compares points or patterns on the fingertip to establish identity. Dirt, aging, and wearing of the skin on the finger affect the performance of this biometric.

Retinal scanners compare the blood vessels in the eye. A scanning device that uses low light compares unique patterns on the retina. The presence of glasses adversely affects retinal scanning.

Iris scanners look at the colored ring around the pupil of the eye, and even work when the individual is wearing glasses. Low light and movement when scanning will reduce the effectiveness.

Hand geometry compares hand shape and size. Scarring, skin changes with age, and jewelry can all cause problems with this biometric.

Facial biometrics uses the size and temperature of facial features. A digital camera is used to create the facial image. Facial scans can be different from day to day and year to year due to changes in age, scars, glasses, hairstyle, and lighting.

Handwriting biometrics analyzes patterns in writing. The speed, pressure and velocity used when writing, and the shape of the finished writing is used in the comparison. Variations in writing from day to day can change a person's handwriting and affect this biometric.

Finally, voice biometrics uses the pitch, tone, and rhythm of speech. Background noise, illness, age, and differences in telephones and microphones can cause problems with voice identification and authorization.

How voice biometrics measures up

Advantages

Voice authentication has a number of advantages. The cost of implementation is low because there is no special hardware required. A simple telephone or microphone is all that a user needs to authenticate using her voice. Other methods of biometric authentication like fingerprinting and retinal scans require special devices.

Voice authentication is easy to use and easily accepted by users. It is quite natural to speak. It is not as natural to put an eye up to a reader. The concept of identifying people by voices is also quite natural. Every time someone answers a telephone call, the natural instinct is to try to identify the caller by his voice.

Perhaps most important to the future of voice biometrics is that it is the only biometric that allows users to authenticate remotely. Allowing a user to call a phone number and authenticate with her bank vocally to perform a transaction is much easier than asking the user to go to the bank in person and authenticate via fingerprint.

It is quick to enroll in a voice authentication system. The user is asked to speak a certain set of words or phrases, or to speak for a certain length of time. From that sample, a digital representation of the voice, called a voiceprint, is created. A good voiceprint is between 2-8 seconds of speech.[1]

Authenticating a user is accomplished by comparing the voiceprint that was created at enrollment to a sample given when the user wants to enter the restricted area or system. Authentication is very fast; it can be completed in 0.5 seconds. [2]

Another advantage is that the storage size of the voiceprint is small. How small it is will be vendor specific, but one vendor, Voice Security Systems, states that a user's voiceprint is less than 1K in size. This is so small that it can be stored almost anywhere: smart cards, floppy disks, databases, even on cell phones.[3]

Disadvantages

Voice biometrics is not the most secure of the biometric technologies, as indicated in Table 1 below. For this reason, it is not appropriate to use them independently for authorization to systems that require high security. They become more powerful when used in conjunction with another form of authorization, such as a password.

---

[1] Schmidt, Regina. http://www.biometritech.com/features/090303nu.htm
[2] Schmidt, Regina. http://www.biometritech.com/features/090303nu.htm
[3] http://www.voice-security.com/NewVoice.html

In addition, the human voice changes over time. A voiceprint taken when a user is young may not match the same user when she is older. The longevity of the sample is not as good as with some of the other biometric technologies. The table below shows a comparison of the current biometric technologies in a variety of categories.

Table 1 – A comparison of biometric technologies [4]

Characteristic	Fingerprints	Hand Geometry	Retina	Iris	Face	Signature	Voice
Ease of Use	High	High	Low	Medium	Medium	High	High
Error incidence	Dryness, dirt, age	Hand injury, age	Glasses	Poor Lighting	Lighting, age, glasses, hair	Changing signatures	Noise, colds, weather
Accuracy	High	High	Very High	Very High	High	High	High
Cost	*	*	*	*	*	*	*
User acceptance	Medium	Medium	Medium	Medium	Medium	Medium	High
Required security level	High	Medium	High	Very High	Medium	Medium	Medium
Long-term stability	High	Medium	High	High	Medium	Medium	Medium

Accuracy

The accuracy of voice authentication is affected by several factors. The difference in telephones can be problematic. For example, enrolling with a home telephone and then trying to authenticate on a cell phone can be enough of a difference to cause a false rejection. Background noise, illness, and vocal changes from age can all affect accuracy. This type of problem is not unique to voice authentication. Fingerprint scans are affected by environmental factors such as dirt on the finger or device. Face scans can fail due to aging of the skin or scarring.

Accuracy of biometrics is measured in three categories:

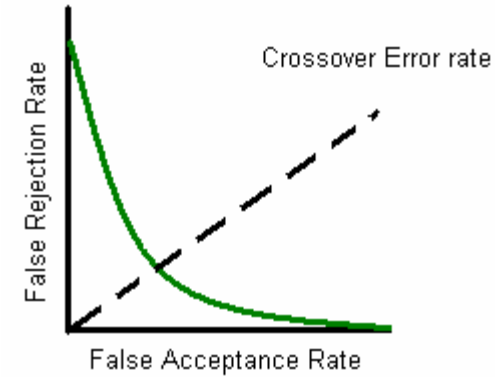
1. Failure to enroll – user's registration into the system is not successful
2. False acceptance – user is authenticated when she should not be
3. False rejection – user is not authenticated when she should be

Systems that analyze biometric data must balance the false acceptance rate and the false rejection rate. Tightening the requirements for a match can decrease the false acceptance rate (FAR), but it is likely to increase the false

[4] Liu, Simon; Silverman, Mark. http://www.computer.org/itpro/homepage/jan_feb01/security3b.htm

rejection rate (FRR). If you plot FAR and FRR against each other, the point at which they intersect is called the crossover error rate (CER). The lower the CER, the better the system is performing.

Figure 1 – FAR, FRR, and CER [5]



The table below illustrates the crossover rates for the various biometrics.

Table 2 – Crossover rates [6]

Biometric	Crossover Accuracy %
Retinal Scan	.0000001%
Iris Scan	.000763%
Fingerprints	.2%
Hand Geometry	.2%
Signature Dynamics	2%
Voice Dynamics	2%

As illustrated in the table above, iris and retinal scanning are considered to have the highest accuracy rate of the current biometric technologies. Among the

[5] Liu, Simon; Silverman, Mark. http://www.computer.org/itpro/homepage/jan_feb01/security3b.htm
[6] http://www.idsmartech.com/english/Biometric.htm

remaining technologies, voice performs comparably. Voice error rates are typically between 2 and 5%.[7] It is difficult to gather concrete data on accuracy of voice biometrics because it varies by vendor implementation. According to Samir Nanavati, a founding partner of International Biometric Group, "if you look at three key metrics – failure to enroll, false acceptance and false rejection – the best voice system will outperform the worst fingerprint system"[8]

How the technology works

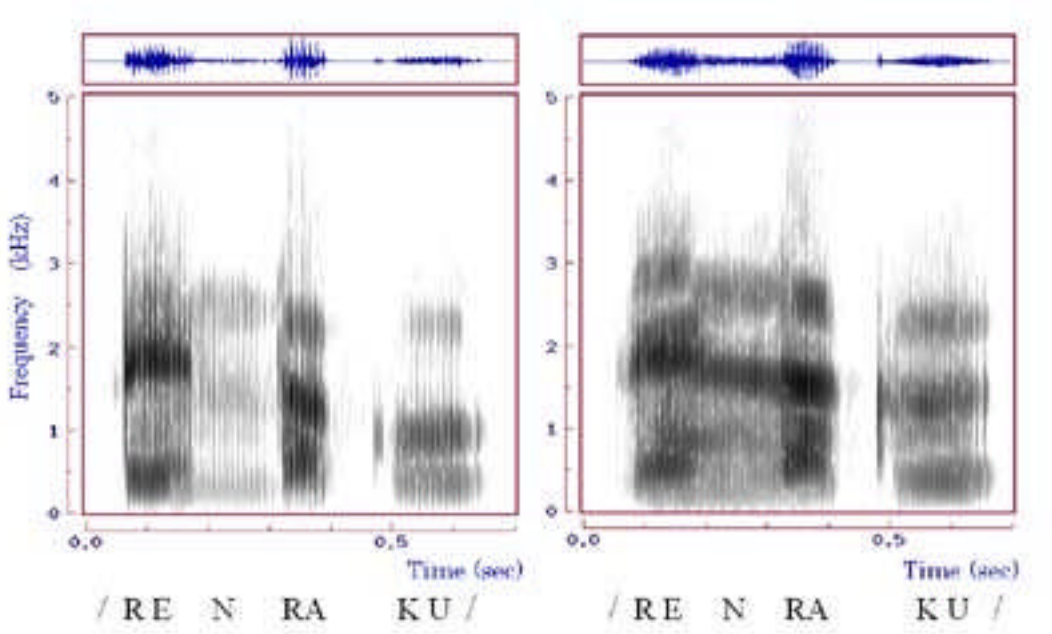
The underlying premise for voice authentication is that each person's voice differs in pitch, tone, and volume enough to make it uniquely distinguishable. Several factors contribute to this uniqueness: size and shape of the mouth, throat, nose, and teeth, which are called the articulators, and the size, shape, and tension of the vocal cords. The chance that all of these are exactly the same in any two people is low.

The manner of vocalizing further distinguishes a person's speech: how the muscles are used in the lips, tongue and jaw. Speech is produced by air passing from the lungs through the throat and vocal cords, then through the articulators. Different positions of the articulators create different sounds. This produces a vocal pattern that is used in the analysis.

A visual representation of the voice can be made to help the analysis. This is called a spectrogram. A spectrogram displays the time, frequency of vibration of the vocal cords (pitch), and amplitude (volume). Pitch is higher for females than for males.

Figure 2 – Examples of Voiceprints[9]

[7] Lewis, Scott; Steigerwalt, Todd. http://www.emory.edu/BUSINESS/et/biometric/Voice.htm
[8] Gilhooly, Kim. http://www.computerworld.com/securitytopics/security/story/0,10801,86900,00.html
[9] http://www.nrips.go.jp/org/fourth/info3/index-e.html



These voiceprints are a visual representation of two different speakers saying “RENRAKU”.

There are three broad categories of sounds:

1. Voiced sounds – produced when the vocal cords vibrate. This includes vowel sounds. The resonance frequencies for these sounds are called formants.
2. Unvoiced sounds – produced when the vocal cords do not vibrate. An example of an unvoiced sound is “Sh”. These sounds do not produce resonant peaks.
3. Plosive sounds – produced when a build up of pressure is released, as in the ‘b’ in ‘butter’. Plosives can be voiced or unvoiced.

One model of speech analysis is the source-filter model. This model has two sources of human vocal sounds: the glottal pulse generator and the random noise generator. The glottal pulse generator creates voiced sounds. This source generates one of the measurable attributes used in voice analysis: the pitch period. The random noise generator produces the unvoiced sounds. The vocal tract is the filter in the model. It filters both the voiced and unvoiced sounds. The filter changes every 5-20 milliseconds during speech. The filter produces intensification at specific formants. These frequency peaks corresponding to certain formants are a second measurable attribute.

Voiced sounds such as vowels are the easiest to analyze because they have a consistent pitch period and formant structure. The frequency of unvoiced sounds is nearly flat, so they represent the horizontal areas on a voiceprint. See the “N” portion of the voiceprints in Figure 2 for an example of this.

Various algorithms exist for processing this data. All of them use some combination of time and frequency to determine whether two voice samples match.

There are two approaches to doing the comparison. The first is text dependent, meaning that the two voiceprints being compared must be the same words or phrases. This is the more simple of the two because the patterns being compared will match more closely. The second approach is text independent, meaning that the two voiceprints are not necessarily of the same words or phrases. In this case, the analysis attempts to find common patterns between the two samples within the words. Text independent analysis requires a longer sample to compare.

It is worth noting that this technology is not the same as speech recognition. Speech recognition is the process of identifying what words have been spoken, not identifying who spoke them. Speech recognition can be used in combination with voice authentication to make it more powerful. For example, if the user is required to speak a password, the voice biometric can confirm the person's identity, and speech recognition can be used to validate that the password given is correct. The combination of the two creates a highly accurate authentication mechanism.

Preventing Hacking

Hackers might attempt to gain unauthorized access to a voice-authenticated system by playing back a pre-recorded voice sample from an authorized user. One way to thwart this sort of attack is to use a challenge-response system. The system can prompt the user to repeat a random set of words or phrases in a specified order. Then the system verifies that the voice sample matches, and that the sample contains the requested words and phrases in the correct order. This makes it difficult for anyone to use a prerecorded voice sample for authentication.

Privacy Issue

A major issue facing all biometric technologies that store data is maintaining the privacy of that data. As soon as a user registers with a voice biometric system, that voiceprint is stored somewhere just like an address or a phone number. What if companies decide to sell voiceprints like addresses? Will we need a public "opt out" registry like we now have for telephone numbers to prevent the sharing of biometric data?

Furthermore, even if the data is encrypted in storage and in transport, there is always the possibility of cracking the encryption and stealing the data. Biometric data is unique in that once it has been compromised, a user cannot merely request a new one like one can with a password reset or a new credit card number. Each person only has one voice.

History of the technology

Analyzing speech and creating visual representations of speech predates computers. As early as 1867, Melville Bell, father to Alexander Graham Bell,

began work on translating speech to visual forms.[10] In 1941, Bell Telephone Laboratories began work on early spectrograms.[11] The variation in voices, the very thing which makes voice authentication possible, made it difficult for researchers to create automated ways of analyzing voices.

In these early days, the research was driven by the potential to use voice technology to track enemy movement via radio traffic. The idea was to identify a specific voice, note the location of the transmission, and then try to identify the same voice at a later point in time in a new location to determine troop movements. At that point, the technology was not advanced enough to support this use case, and efforts were frustrating. As time went on, applications in law enforcement began to drive the research.

As spectrograph technology improved, enhanced automated analysis began to emerge. Today there is an array of commercial software available that utilizes automated analysis successfully to accomplish voice authentication.

Current applications

Voice biometrics is already being used in some interesting ways. Many of us have probably heard of or already used voice authentication in some of the ways described below. Here are some of the current applications.

Union Pacific Railroad

Union Pacific moves railcars back and forth across the United States every day. The railcars travel loaded in one direction and empty on the way back. When the loaded railcar arrives, the customer is notified to come pick up the contents. Once emptied, the customer needs to alert Union Pacific to put the railcar back to work. Union Pacific now has an automated system that utilizes voice authentication to allow a customer to release empty railcars. Customers enroll in the voice authentication system over the phone. When they call back to release an empty railcar, the system authenticates them and allows them to release their railcars. In this case, voice authentication has allowed customers to get off the phone faster, and Union Pacific to guarantee that a customer is not releasing a railcar that doesn't belong to him.

New York Town Manor

New York Town Manor is a residential community in Pennsylvania designed for senior citizens with technologically advanced features. The residents no longer have to remember passwords. They do carry ID cards that are used in conjunction with voice authentication to allow access to the complex. To enter their apartments, they speak for a few seconds while the system authenticates them. With this approach, voice authentication provides an extra measure of security.

[10] http://www.knr.net/geninfovoxid.htm
[11] http://www.knr.net/geninfovoxid.htm

Bell Canada

Technicians for Bell Canada used to have to carry laptops on the job with them. A technician would dial up using a modem to report the current job as finished and to get the next job. Bell Canada has rolled out a new system that uses voice authentication to verify the identity of the technician through a phone call and give him access to the data. This eliminates the need for a laptop.

Password Journal

Anyone who has ever had a diary has probably worried that someone would read it without permission. One company has solved this problem by adding voice authentication as a privacy measure to their Password Journal product. The journal has its own speaker, raises an alarm if an unauthorized person attempts to access it, and keeps track of how many failed attempts there have been.

Password Reset

Some companies are allowing users to reset passwords themselves. Users dial an automated system. The system asks questions. When the user answers, the system authenticates his voice and allows him to reset his own password. This saves companies time and money in support costs, and users need not spend time on hold waiting for the next available support person.

Banking

Reducing crime at Automated Teller Machines is an ongoing struggle. Banks have started using biometrics to authenticate users before allowing ATM transactions. Users generally must provide a pin number and a voice sample to be allowed access.

Royal Canadian Bank is using voice authentication to allow access to telephone banking.

US Social Security Administration

The United States Social Security Administration is using voice authentication to allow employers to report W-2 wages online. Used in combination with a pin number, the voice authentication provides system security and user convenience.

Law Enforcement

In Louisiana, criminals are kept on a short leash with voice biometrics. This inexpensive approach allows law enforcement to check in with offenders at

random times of the day. The offender must answer the phone and speak a phrase that is used for authentication. This system guarantees that they are where they are supposed to be!

Voice authentication has also been used in criminal cases, such as rape and murder cases, to verify the identity of an individual in a recorded conversation. There is a terrorism application also. Voice authentication is frequently used to validate the identity of terrorists such as Osama Bin Laden on recorded conversations. Hopefully these clues will one day assist in his capture.

Future applications: What If?

Where will this technology lead us? Here are a few thoughts.

VISA

Visa is one of the many companies who are utilizing voice authentication to allow users to reset their own passwords through automated systems. The company has also envisioned allowing customers to verify online purchases though voice authentication. What if someone got your VISA card number and went on a shopping spree online? Voice authentication would provide an extra measure of security against credit card fraud in online transactions by requiring users to speak into a microphone on their PC to authenticate before making online purchases.

EBay

If a registered EBay user has enough negative feedback, EBay will revoke privileges to that account. Unfortunately, what often happens is that individual simply opens a new account under a different username and continues to buy and sell until that account accumulates enough negative feedback to be shutdown. What if users were required to provide a voice sample when registering? If the sample matched an existing one in the database for an account that was terminated, EBay could refuse to allow the new account, effectively creating a lifetime EBay ban for violators.

Personal Devices

Anyone who has carried a cell phone has probably misplaced it at some point. If an unsavory individual finds it, he could run up thousands of dollars in phone charges quickly. Many cell phones have the option to secure the keypad with a pass code. Typing that code in every time you want to make a call is inconvenient. What if you could just flip open your cell phone and speak a phrase to authenticate yourself as a valid user of this cell phone? If the phone were stolen, the thief would not be able to use it.

This could extend to securing laptops, PDA's, and any personal device that could benefit from an extra measure of security and that could embed a microphone.

Parking Garages

Most urban commuters park in daily parking garages. These parking garages almost universally require some sort of key card for entrance, and often for exit as well. Sometimes these key cards melt in the hot car when parked in the sun. Sometimes they fall between the seat, and the anxious driver must search for it as the drivers queue up in line impatiently behind them waiting to get into the garage. Some drivers can't reach the readers easily unless the car is positioned just right. What if drivers could simply drive up, lower the window, and speak a phrase into a microphone to authenticate for entry and exit via voice? There would be no card to lose or get destroyed, no pin to remember, and it would be fast.

How to implement this technology in your company

If using voice biometrics for authentication sounds appealing, there are several commercial software packages available. There are consultant companies such as J. Markowitz, Consultants who rank voice vendor software. These reports are often available for purchase online. One vendor's implementation may be more suitable than another depending on the use case. Nuance (www.nuance.com) and Vocent (www.vocent.com) are two vendors who provide voice authentication solutions and who came up repeatedly in the research for the "Current Applications" section of this paper.

Nuance has been rated the top voice authentication technology vendor by Celent Communications.[12] Nuance claims a 96% or higher accuracy rate for speech recognition[13], and Nuance's Verifier product claims a "high" accuracy rate.[14]

Conclusion

Voice biometric authentication has a number of advantages over other biometric technologies, and will continue to grow in popularity due to the ease of use, user acceptance, and remote authentication capabilities. It is already enhancing our daily lives in some very practical applications. The future uses are only limited by the scope of our imagination.

[12] http://www.celent.com/PressReleases/20020430(2)/BioVendors.htm
[13] http://www.nuance.com/learn/speechaccuracy.html
[14] http://www.nuance.com/prodserv/prodverifier.html

References

1. Markowitz, Judith. “Voice Biometrics - Are You Who You Say You Are?” December 2003. http://www.speechtechmag.com/issues/8_6/cover/2751-1.html
2. Markowitz, Judith. “The For-Real Story” December 2002. http://www.speechtechmag.com/issues/7_6/voiceideas/1468-1.html
3. Gilhooly, Kim. “Q&A: Where Voice Authentication Fits” November 10, 2003. http://www.computerworld.com/securitytopics/security/story/0,10801,86900,00.html
4. Gilhooly, Kim. “Voice Authentication: Making Access a Figure Of Speech” November 11, 2003. http://www.computerworld.com/securitytopics/security/story/0,10801,86897,00.html
5. “A Business Case for Voice Authentication” http://www.atio.co.za/news/2002/A%20Business%20Case%20for%20Voice%20Authentication.htm
6. Schmidt, Regina. “Identity Confirmed, Access Permitted: The Basics on Voice Authentication, Security And Consumer Use Of An Emerging Biometric” September 3, 2003. http://www.biometritech.com/features/090303nu.htm
7. Liu, Simon; Silverman, Mark. “A Practical Guide to Biometric Security Technology” February 2001. http://www.computer.org/itpro/homepage/jan_feb01/security3.htm
8. April 30, 2002. http://www.celent.com/PressReleases/20020430(2)/BioVendors.htm
9. http://www.nuance.com/learn/speechaccuracy.html
10. http://www.nuance.com/prodserv/prodverifier.html
11. Lewis, Scott; Steigerwalt, Todd. http://www.emory.edu/BUSINESS/et/biometric/Index.htm
12. http://www.voice-security.com/NewVoice.html
13. Mearian, Lucas. “VISA Eyes Voice Recognition for Online Purchases” November 4, 2002. http://www.computerworld.com/securitytopics/security/story/0,10801,75553,00.html
14. Buckler, Grant. “Voice ID Sets the Tone For Security” January 8, 2004. http://www.globeandmail.com/servlet/ArticleNews/freeheadlines/LAC/20040108/TWVOIC08/technology/Technology
15. Cain,Steve; Smrkovski, Lonnie; Wilson, Mindy. “Voiceprint Identification” http://expertpages.com/news/voiceprint_identification.htm
16. “History of Voice ID” http://www.knr.net/geninfovoxid.htm
17. “Speaker Identification by Voiceprint” http://www.nrips.go.jp/org/fourth/info3/index-e.html
18. Sankaranarayanan, A. “A Text-Dependent Approach to Speaker Identification” September 16, 2002.

http://www.techonline.com/community/ed_resource/feature_article/21068_JD7349406658EL
19. "Biometric Selection: Body Parts Online" http://www.idsmartech.com/english/Biometric.htm
20. "Biometrics and Security" http://www.cesg.gov.uk/site/ast/index.cfm?menuSelected=4&subMenu=4&displayPage=404
21. Cole,Eric, Fossen,Jason; Northcutt,Stephen;Pomeranz,Hal. SANS Security Essentials With CISSP Volume 1. "Physical Security". USA: SANS Press, 4/1/2003. 277-283.