Voice Biometrics in Financial Services

By Clinton Alver
Senior Business Consultant, VeCommerce

While the added security of voice biometrics has made it a natural fit for the banking industry with current fears around fraud and identity theft, little coverage has been given to its applicability in other areas of the financial services industry. Several UK wealth management providers are considering the use of voice biometrics as a means of providing differentiated service to their customers. While the enhanced security of voice biometrics is clear, what is often ignored is the extra convenience the technology can provide to callers. For example, when Australian Health Management offer their members the option of being able to verify their identity just by speaking their membership number, over 95 per cent of them accept. This is perhaps hardly surprising when the alternative is the standard process of having to provide 4 pieces of personal information (member number, name, address and date of birth) to a customer service agent.

With the highly competitive nature of the wealth management industry, providing a high level of customer service is key in ensuring that Independent Financial Advisors (IFAs) recommend a company’s products to their clients. While hardly an issue for a policy holder calling every three years or less, questions asked in the contact centre to establish if the advisor is authorised on the policy can become frustrating for the IFA if they contact the organisation many times a week on behalf of the multiple clients they represent. With the introduction of voice biometrics, advisors are able to establish and confirm their identity just by speaking a single phrase such as a unique number or their company name. The proposition is clear: the advisor wins by gaining a level of service akin to the fast track now being offered for those who enrol a biometric at airport immigration; and the contact centre wins by saving time (up to 40 seconds) spent verifying the caller’s identity.

Identity-related fraud is one of the fastest growing areas of crime globally and it is of increasing concern to banking and finance sectors. The Australian Federal Police estimate that identity fraud costs businesses (and ultimately consumers) over $2 billion per year. Identity fraud is difficult to detect and has a significant emotional impact on its victims, comparable to that of violent crime. Most victims only find out about fraud through their own investigations and, on average spend around 600 hours resolving the issue.
INTRODUCTION

Biometrics (ancient Greek: bios = “life”, metron = “measure”) is the study of automated methods for uniquely recognising humans based upon one or more intrinsic physical or behavioural traits.

HISTORY AND BACKGROUND

Biometric technologies have been used for years since Chinese merchants first started stamping prints from palms and fingers as a way of identifying people. In the 19th century, efforts were made to develop a more scientific system than memory to identify criminals and biometrics became a field of study. Originally using various measurements of lengths, widths and characteristics of parts of the body, methods were standardised and fingerprinting became the method of choice. Recently, other biometric methods have become used more and the applications of biometrics have also widened from face to face identification to sophisticated security systems and consumer applications.

Voice Biometrics, developed in 1970 is a technology that enables callers’ identity to be authenticated using the unique characteristics of their voice. Used internally for years in high-security environments, such as government installations and financial institutions, voice biometrics can be used across multiple direct channels, is cost effective, requires little additional hardware, non-intrusive and can be used to verify remote transactions easily and conveniently. Voice verification relies on something truly unique that cannot be duplicated or stolen, an individual’s voice. Despite common misconceptions it is inherently more secure than most other verification.

According to Frost and Sullivan analyst Sapna Capoor, the global biometrics voice market is expected to grow rapidly over the next four years. “A primary driver for growth is the increased adoption of voice verification by financial institutions,” she says. “Key drivers will be the need for tighter security, enabling greater use of automated services, end-user convenience and cost reduction.” (source: www.frost.com)

HOW DOES IT WORK?

Typically, a customer enrols by speaking a unique identifier when they initially establish their identity. This is then used to confirm their identity on subsequent calls. The technology can be used to verify the identity information, as well as the unique way the caller speaks this information.

Once the caller has enrolled, this information is then converted to a numerical algorithm, and entered into a database. The algorithm creates a digital representation of the obtained biometric so the enrolled information is secure and cannot change. Each subsequent attempt to use the system, or authenticate, requires the user to make their identity claim and present their biometric again, this is then compared to those existing in the database to determine a match. This process is completed each time the user attempts to authenticate to the system. The comparison process involves the use of a Hamming distance, a measure of the percentage of dissimilar bits out of the number of comparisons made. The data from the caller is compared with the voiceprint retrieved by the application and with an imposter model; the resulting score from the verification engine represents the level to which the caller audio data matches the voiceprint better than the imposter model. When a valid user logs in, the match will be very close and, when someone tries to log in falsely, the voiceprint will not match closely and the system will not allow the person access.

In order to promote user convenience, commercial systems are usually designed to verify on a risk-based approach. Low-risk transactions proceed uninterrupted, while transactions deemed high-risk can be verified with additional layers of security, such as knowledge based or secret questions.

Current phone authentication techniques lack security, as they are generally ad hoc in nature, and are susceptible to social engineering attacks by increasingly intelligent fraudsters.

PERFORMANCE

There are three main parameters for measuring speaker verification performance, all related: the amount of speech, the False Reject Rate (FRR), and the False Accept Rate (FAR). The FAR measures the percent of false users who are incorrectly accepted as genuine users, while the FRR measures the percent of valid users who are rejected as impostors. Reporting one is technically meaningless as they are all related, and changing the acceptable FA rate will also change the FR rate or amount of speech required. (Usually we keep the amount of speech constant, and just vary FA/FR.) Text-independent speaker verification usually requires more speech than text-dependent speaker verification.

When a system is tuned, for a given amount of audio, the point at which the FA rate equals the FR rate is known as the equal error rate. The lower the EER, the more accurate the system is considered to be. A low EER is desirable for a biometric system because it balances the sensitivity and usability of the system.

Current technologies have widely varying Equal Error Rates, varying from as low as 60 per cent and as high as 99.9 per cent. Voiceprint verification is second in accuracy only to iris scanning and has considerable advantages in that it can be used remotely and is not invasive or confronting.

Figure 1. Measuring speaker verification performance
TYPES OF SPEAKER VERIFICATION

There are two types of speaker verification:

- **Text-dependent**, in which the enrolment and verification phrases must be identical (and often different for each user).
- **Text-independent**, in which the enrolment and verification phrases may be different (and often different from call to call).

As noted in speaker verification performance, there are three main parameters measuring the performance of speaker verification systems. Text-independent speaker verification requires a lot of audio to perform satisfactorily. Many research papers suggest that 30-60s of audio is required to enrol, and a similar amount to verify; however, this is dependent on the phonetic variability of the phrases used. It is expected that if the verification audio closely matches some of the enrolment audio, the amount required is less; further, if the verification audio covers more phonemes, less is required for good results.

TERMINOLOGY

Speaker verification is sometimes known as:

- **Speech verification**, which has the disadvantage that it is really the speaker, not the speech, that is being verified.
- **Voice Verification**, which is a somewhat more correct term than “speech verification”, since a voice is generally attributable to a particular person (debatably perhaps). This term is not widely used by industry, but does seem more commonly used by lay persons (it sounds less technical than “speaker verification” perhaps).
- **SV**, a commonly used acronym for Speaker Verification.

MARKET ACCEPTANCE

Speaker Verification has started seeing increasing awareness and acceptance, particularly in the banking and finance sectors, in a similar way to speech recognition. Often the two technologies are employed in a security solution, being combined on the same core technology to provide a convenient and secure voice interface to transaction services.

Head of ANZ’s Fraud and Risk Group, Lawrence Cox says, “For banking services where the user population is so widely dispersed, only speaker verification using the telephone system will work”.

CUSTOMER BENEFITS

- Customers like it
  - it uses the telephone, a device that they are familiar and comfortable using.
- It’s convenient
  - the customer controls the password
  - no PINs to remember; no PINs to forget.
  - it’s secure and privacy enhancing
  - it cannot be stolen or shared
  - cannot identify the customer
  - not even banking employees know the password.

BUSINESS BENEFITS

- Speaker verification now proven as a convenient and secure technology for authenticating identity.
- It is also extremely cost effective in terms of cost benefit and ROI, cost of deployment and cost of ownership.
- Because it uses the telephone, no investment in communication infrastructure needed.

WHO’S USING IT?

A wide range of organisations are now seeing the advantages of SV for not only authenticating customer identity but also authenticating identity for internal IT security.

Banking and Finance

- IT Security – PIN’s and password resets.
- Authentication and security.
- Authentication for professional and high value clients.
- Authentication in shopping and retail services.

Government services

- IT Security – PIN and password resets.
- Military and national security applications.
- Correctional services (prisons) – for authentication in telecommunication services.
- Border control – for authentication of ID in shipping.

Secure Communications

- Recharging pre-paid mobile telephones.
- Authentication for mobile telecommunications.
- Services – such as fax forwarding and access to voice mail.

Case Studies

- **ABN Amro** (Netherlands) initially applied to customers making balance enquiries, transfers and investment orders via the telephone. Verified on account number. Customers will not have to remember pass codes anymore – customers record their own secret question in addition to account number. The customer is first asked an open question: “How can we help you?” The option of voice verified access will be offered to customers on a voluntary basis.

- **AHM** (Australia) Verification of health insurance clients on their membership number – over 20,000 enrolled to date (Sept 2007.) What is interesting here is that the system can differentiate between multiple family members enrolled on the same membership number.

- **St George Bank** (Australia) Employee password reset. With total staff numbering 8,000 the current solution has been successfully rolled out across over 1,500 staff members to date (Sept 2007.) A key learning in this project is that success is more linked to people and process than technology. Why should people use this system? Do a small proportion of staff contribute to the majority of password resets? etc.

- **Department of Finance and Adminstration** (Australia) Classified.
Some organisations have chosen to introduce a second factor of verification in the form of keyring or SMS tokens. While these can be mildly effective, they can easily be lost or stolen, do not prevent spoofing or man-in-the-middle attacks and they are expensive to implement and maintain.

The deployment of speaker verification technology benefits both the end user and the organisation. It is convenient for the customer, eliminating the need to remember PINs and passwords; it does not require the intervention of a call centre agent, and thus the process is highly secure. What is more, as the process is performed automatically, speaker verification offers significant cost benefits in terms of call centre efficiency.

**Speaker verification offers more convenient services, more secure services and more effective and efficient use of business call centres.**

### A COMPARISON OF BIOMETRICS

All biometrics allow an additional factor of verification over PINs and passwords (something you know) and tokens and smart cards (something you have) by introducing the concept verifying your identity with something you are. All biometrics offer different benefits in different situations but voice is also the only biometric that does not require the physical presence of the person, making it ideal for telephone and web-based applications where remote authentication is required.

While all biometric technologies provide benefits, there are also some disadvantages. The most commonly deployed technologies include:

- Palm or Fingerprint
- Iris or Retina scan
- Facial Recognition
- Speaker (Voiceprint) Verification

The benefits of voiceprint over other biometrics is that it uses an existing infrastructure (the telephone) does not require any uncommon devices for access, can be accessed remotely, is not confronting or invasive and it is highly accurate. Given the network and technology to deliver is already in place, Speaker Verification solutions are also extremely cost effective and fast to deploy.

Table 1 compares the benefits of different authentication technologies.

### CURRENT AUTHENTICATION METHODS

In the modern world, consumers face the need to identify themselves every day. Some transactions are low risk such as frequent flyers and DVD rental, others such as funds transfers and health are sensitive and require high security. All of this increases the burden on customers to remember and keep secret passwords and PINs. Passwords and PINs are notoriously weak forms of security. When a password or PIN is lost or forgotten, the customer needs to re-establish his/her identity – usually by calling the call centre and answering a sequence of personal questions. As well as being costly and time consuming, this process does not necessarily establish that the caller is who they say they are. In addition, the call centre agent is also privy to identity information, providing yet another opportunity for fraudulent activity.

### USABILITY AND USER ACCEPTANCE

In the past, there have been questions regarding the performance of the technology. What happens if I have a cold? What happens if I have an accent? Can mimics break into my account? What happens if somebody records my voice? Can my password be decoded? Recent advances in technology have addressed all of these concerns.

Research by the British Government’s Communications Electronic Security Group and Edinburgh University has demonstrated that Speaker Verification accuracy results are superior to those of fingerprint, hand print and face recognition systems and are 100 times better security than PINs and passwords, with 97 per cent of callers successfully completing transactions without operator assistance.
Testing has also shown that accents do not cause problems as the requirement is that the original enrolment uses the same vocal and speech characteristics as subsequent verification, regardless of ethnicity or accent. Likewise, mimics are unable to recreate the physiology of the person they are impersonating, even if they sound identical. Usability is unaffected by all but the most serious cases of colds and flu (e.g. loss of voice). Recordings are difficult to obtain and require sophisticated equipment. All the same, to further improve security, solutions are usually set-up to ask random “liveliness” questions which will not allow pre-recorded responses. Secure data storage using encryption will prevent unauthorised access to voiceprints which cannot be decoded to reveal identity anyway.

With the rapid growth in identity fraud and the increasing cost for business and consumers, speaker verification is becoming a favoured option for improvements in convenience, security and privacy. Financial institutions are identifying significant ROI on the deployment of speaker verification through improved call centre efficiency and reduced agent time. Other measurable metrics include reduced staff turnover, increased customer satisfaction and subsequent success in upselling new products and services.

In deploying Speaker Verification, attention should be focussed on the user experience to ensure the solution is accepted. The benefits that can be explained to users are the reductions in their time spent, improvements in their service levels and personalisation, reduced costs and increased convenience. The advantages to the business are increased trust, customer confidence, privacy and compliance with regulatory obligations.

FUTURE IMPROVEMENTS

Core technology platform providers have plans for continued improvements in speaker verification, making it more accurate and more usable on more audio. As with other speech technologies (MRCP/ASR/TTS), the platforms are open architecture using standards such as SOAP interfaces and Java wrappers, allowing businesses to select solutions providers with demonstrated expertise. Recent improvements include the ability to verify against multiple voiceprints, being able to get an indication of gender and determination of line quality and handset type.

SPEAKER VERIFICATION – IMPOSTOR DATA

Impostor data is audio data that can be used to simulate an impostor attack against a verification system.

This usually means that it is an utterance or utterances spoken by one person (the impostor), saying things that would/could be spoken by another person (the target) in the process of verification.

A “liveliness” test is a mechanism used to establish whether a person is speaking on the end of the phone, as opposed to a recording of that person. It is almost always used with speaker verification, to ensure that the person whose identity is being used to access the system is present. (In other words, to prevent against recording attacks.)

IMPLEMENTATION

A number of liveness tests are possible; one is a random digit challenge (where the caller is asked to repeat back some arbitrary digits that differ from call to call). If using a random digit challenge, it helps to have the digits requested in ways that provide better matches to the enrollment data. For instance, if the enrollment data consists of users saying the digit string “123456789”, the random digits requested might include pairs of consecutive digits – such as “2378”.

GLOBAL TRENDS

- Youth Offenders (UK)
  - Telephone individuals at specific locations
  - Ensures offenders stay at home or work at specific times

- St. George Bank (Aus)
  - Internal employee password resets

- ABN Amro (Europe)
  - Verify by speaking account number
  - Balance Enquires, transfers and investments orders

- Bell Canada (North America)
  - Allows customers to verify and self serve
  - Balance Enquires, transfers and investments orders

- Australian Health Management (Aus)

Table 1. Comparison of security methods

<table>
<thead>
<tr>
<th></th>
<th>Speaker verification</th>
<th>PIN</th>
<th>Token devices</th>
<th>Fingerprint and hand geometry</th>
<th>Iris scan</th>
<th>Face recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of operating</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Access to security system</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Remote authentication</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Secure</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Intrusive</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Very</td>
<td>Very</td>
</tr>
<tr>
<td>Possible to lose or forget</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Note: SV is the most secure, most cost effective and most user-friendly technology for telephone use.
Business drivers

- Security and privacy.
- Servicing efficiency.
- Improved customer experience.

Benefits achieved

- Over 20,000 members registered since deployment in December 2006.
- Over 95 per cent opt in to enrol their voiceprint.
- Over 15 per cent of customer base enrolled (Aug 2007).
- Reduced cost of servicing – verified calls are 40 seconds shorter.
- Frees up staff to concentrate on value added opportunities.
- Improved customer service and enhanced security for customers.

SPAKER VERIFICATION AND IDENTIFICATION MANAGEMENT

Identity fraud, where someone uses your name, address, bank/credit card account number, or other identifying information without your knowledge to commit fraud or other crimes, is one of the fastest growing types of crime in the world. It has topped the US Government’s fastest growing crime list for the third year in a row and a growing number of reports suggest that the rate of identity fraud is even higher and is a greater concern outside the US.

Recently in Australia, a 73 year old woman was sentenced to at least three years in jail for social security fraud going back nearly 30 years. Using three different identities, she falsely claimed $380,000 in unemployment benefit, age and widow pensions between July 1975 and January 2004.

All it takes is 100 points...

Authenticating identity has always been problematic and subjective. Current methods rely on examining documents such as driving licences, passports and birth certificates. In most cases, these documents were simply not designed for the job. Many are readily available. With the advent of low cost, high quality reproduction technologies, creating forgeries is easier than ever. Google brings up over 63,000 results when searching on how to buy a fake ID.

Even today, an application for a passport requires a fairly basic proof identity.

So how do we know that a person is “who they say they are”?

With the rising trend towards call centres and internet-based customer service, establishing identity “remotely” is even more problematic, relying on customers to remember and keep secret passwords and PINs.

Passwords and PINs are notoriously weak forms of security. When a password or PIN is lost or forgotten, the customer needs to re-establish his/her identity, usually by calling the call centre and answering a sequence of personal questions. Not only is this costly and time consuming, this process does not necessarily establish the authenticity of the caller. In addition, the call centre agent is also privy to identity information, providing yet another opportunity for fraudulent activity.

Forward thinking organisations have engaged market leading solutions developers such as VeCommerce to integrate the latest developments in both speech recognition and speaker verification technology to enable organisations to confirm a claimed identity, ensuring that a person is “who they say they are”.

DIAL AN IDENTITY

Speaker verification (SV) is one of a range of biometric technologies that is starting to gain widespread appeal. Others include finger and palm print, iris and face recognition. While all these technologies have their place, the big benefit of SV is that it works over the world’s most common communications device: the telephone.

Speaker verification is a technology that relies on the unique characteristics of a person’s voice to create a distinct “voice identifier” that can be captured over the telephone, verified reliably, and appended permanently to an individual consumer’s personal ID credentials.

Using the telephone network, SV can be used to authenticate identity from anywhere in the world. The infrastructure to roll out SV is already in place. There is no need to invest in special sensors, data processing technology or networks. Given the technology to deliver SV is already in place, SV solutions are both extremely cost effective and fast to deploy.

As the authentication server can be centrally housed in a highly secure facility, with no connection to unsecured desktops, laptops and networks, SV provides an extremely secure means of identity management and authentication.

SIMPLE YET SECURE

Authentication of a person’s identity using the unique characteristics of their voice results in a verification process that is easy to use, fast, secure and much more convenient than passwords and PINs. Of all the verification technologies, Speaker verification is the only one that can be quickly and unobtrusively deployed to all customers as the infrastructure is already in place.

Before an authorised caller can use the system, they must go through an enrolment process where their voice is initially captured and digitised so that the system can authenticate their identity during live operation.

This crucial yet simple enrolment procedure involves each user speaking to the system for a short while until it forms a reliable “voiceprint” of the user’s voice. The system then encodes the voiceprint and stores it as an encrypted, mathematical algorithm that cannot be compromised, hacked or cracked. For additional security, the verification server can itself be located within a highly secure facility with no connection to unsecured computers or other access devices or equipment.

Subsequent calls are then compared to the stored voiceprint to verify the caller’s claimed identity. The voice is also checked against a model of other speakers’ voices to ascertain if the caller’s voice is significantly different than those of other people. Based on these comparisons, the caller is accepted or rejected.
STRENGTHENING SECURITY

SV can also add enhanced security functionality to new or existing voice or IT networks. For example, in many large organisations, an employee needing a new password is often required to telephone their IT help desk. The help desk agent confirms the caller’s identity by asking them a number of “personal information” questions. If the caller answers correctly, they are issued with a new password.

Disclosing personal or “identifying” information to a help desk agent for the purpose of reissuing a password is extremely insecure, yet many of us do this without questioning it. So now the help desk agent knows your personal identity information. This not only has issues relating to privacy obligations, but also has significant security implications, as now there is an opportunity for your personal information to be used in unauthorised circumstances. Unfortunately, criminal activity resulting from this scenario has already occurred around the world.

Identifying and verifying the caller by prompting for their personal information and then using speech recognition to ensure that the stated information matches the information on file, as well as confirming the caller’s voice characteristics against the registered voiceprint, means that in the event that an impostor may seek to obtain an unauthorised password while they may know an employee’s identity information, the system will reject their request as the “voiceprint” credential will not match.

MARKET GROWTH

Call centres are experiencing unprecedented growth as consumers are looking for access to more immediate and more personalised service. The customer services market is forecast to achieve a five-year compound annual growth rate (CAGR) of 12.7 per cent. The worldwide market for contact centres, help desks, customer care, customer relations management (CRM), and support centres, is growing rapidly. At the same time, new laws and regulations are coming into effect that could impede this projected growth. These new regulations, designed to protect national security and personal financial and health records, require increasing levels of authentication and the traditional challenge questions, PINS and passwords are no longer good enough. As well as asking someone to identify themselves, modern security requires that they can also prove who they claim to be, hence the rise of multi-factor verification.

Of the third factor biometrics, voiceprint is the only technology that can be utilised over the phone and is the ideal solution, providing higher levels of authentication and verification while reducing costs.