Speaker Verification in a Multi-Vendor Environment

Mr Ross Summerfield, Dr Ted Dunstone, Dr Clive Summerfield

I. Introduction

A. Use of Speaker Verification in Government

AROUND the world, speaker verification is starting to gain increasing acceptance in both government and financial sectors as a method to facilitate quick and secure authentication of individuals. While there have been some prior deployments in the criminal management industry [1], the most ambitious deployment that we are aware of in government has been for the authentication of Welfare recipients undertaking telephone transactions at Centrelink in Australia.

Centrelink's use of speaker verification is specifically for authentication of its customer base. While Centrelink's fraud and compliance team may use speaker verification for assisting with prosecutions in the future, its primary purpose is to provide a two factor authentication facility. This allows the use of high value transactions over the telephone, where previously only transactions with lower security ratings have been permitted using PINs. It has been implemented so as not to require the use of a PIN, since Centrelink's experience is a high cost of re-registration due to forgotten PINs.

B. Centrelink Environment

With around 27,000 staff, as the second largest government agency within the Commonwealth of Australia, Centrelink serves 6.5 million customers through over 1,000 access points that are either operated by Centrelink or provided under an agency contract with Centrelink [2]. Centrelink delivers $66,300 million per year in payments on behalf of policy departments [2]. It conducts over 14 million self service transactions and handles more than 30 million telephone calls a year. More than 6,000 million transactions take place on customer records each year through self service and by Centrelink's staff.

Centrelink's telephony environment is wholly outsourced, with services being provided on a 'click charge' basis to a single prime vendor, Telstra. The provision of speaker verification is also provided as a managed service as a part of Centrelink's telephony contract. Telstra has supported Centrelink's speaker verification system through separate subcontractors for its IVR service and for its speaker verification authentication service. The speaker verification service provider, Kaz, supports its service through its own engine and an engine provided by Nuance. Measurement services for the purposes of verifying and supporting the service level agreement are provided through the biometric analysis software, Performix [3]. The IVR service is maintained by Information Technologies Australia (iTa), who are in turn supported by (and provide support to) Nuance for the Natural Language Speech Recognition components.

The output of the two engines is combined through linear discriminant analysis to yield a composite result. The service provides significant flexibility to grow through the addition of multiple different engines as well as accommodating both text dependent and text independent modes of authentication operating in tandem to enhance authentication accuracy.

Ross Summerfield is with Centrelink, Ted Dunstone is with Biometix Pty Ltd and Clive Summerfield is with Auraya Pty Ltd

C. Architecture

Figure 1 shows the implementation architecture used for speaker verification in Centrelink's system. The speaker verification components provided by Telstra interfaces to Centrelink's security services via Centrelink built J2EE based middle-ware.

In Figure 1, the call is received by the IVR via Telstra's telephone network. At that point the IVR requests the user identify themselves. A check is made on the identity to determine if they have a security entry (that is, they are enrolled). At the same time, environmental information about the claimed identity is obtained from Centrelink's Income Security Information System (ISIS) central data base via the SOA middle-ware layer. If they are not enrolled, but are set up to enrol, then an enrolment process is followed. Otherwise they are lead through an authentication process. In that process, the IVR requests speech samples, validates those speech samples using the speaker verification service via the Kaz provide Voice Authentication Control Module (VACM). The status of authentication, including the results generated by the speaker verification engine, are then passed to the security services via the SOA middle-ware layer.

To meet the requirements of privacy advocates, as well as assist in the overall security model around the vendor supplied components, the template data is encrypted within the data base drivers as it exits the speaker verification engines. Privacy advocates were particularly concerned about data in transit and even in memory, especially since the vendor supplied equipment is hosted in a Microsoft Windows operating system environment. Centrelink is also concerned, since verification of the vendor's patching levels by Centrelink's security specialists is difficult to achieve in practice. Centrelink has therefore given higher consideration to firewall arrangements around the equipment than it might have otherwise.

To meet fraud and compliance requirements, particularly the scores from the biometric engines are provided via Centrelink's security services to its logging service, "CRAM". There it may be used for analysis and, mixed with impostor information collected by Performix, used as a part of the likelihood calculations required for prosecutions.

In addition, security data (both identity and authentication data) travelling between the IVR Centrelink's security services

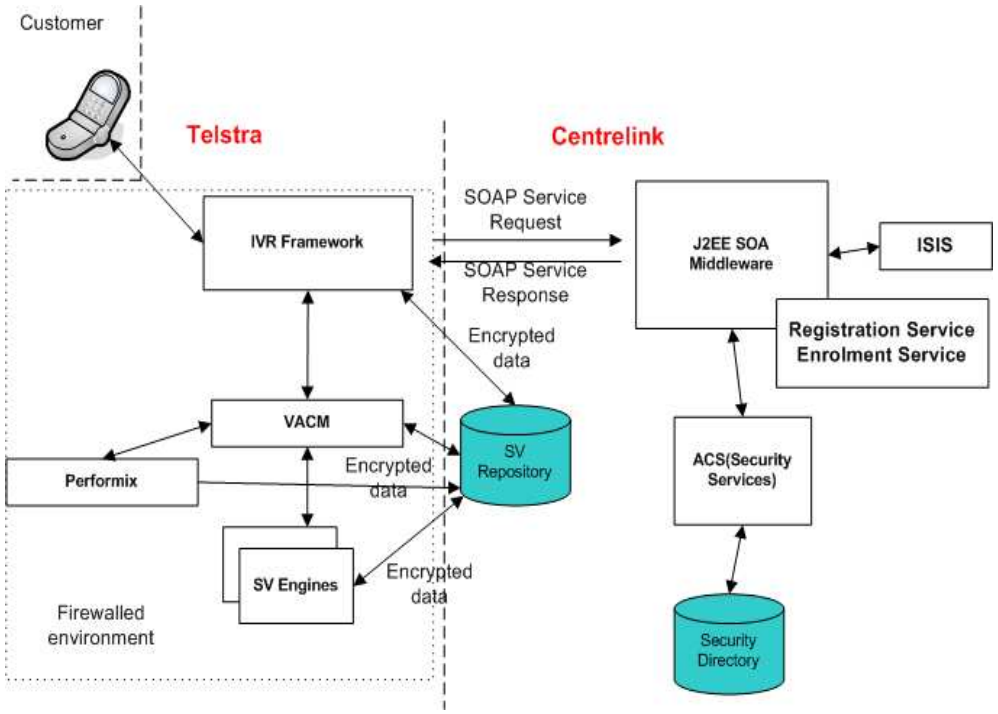

Fig. 1. Architecture for Centrelink's Speaker Verification system

(in both directions) is encrypted to AES 256 bit as it leaves the IVR. This protects the data from internal eves-droppers. It also helps enable Centrelink to provide assurances to courts that the information presented as evidence has not been tampered with during transit.

While the Centrelink system has its interface to the security services via the IVR, this is actually for historical reasons. If Centrelink had been a green field site, the security interface could just as easily have been via the VACM in Figure 1, even if Centrelink did elect to keep its PIN option for low risk transactions. From the perspective of overall authentication architecture, the speaker verification service is viewed as a back-office security service and the IVR is viewed simply as a user interface for the collection and delivery of the biometric sample. Centrelink is giving some consideration to a (at this early stage, admittedly distant) future possibility of using speaker verification over the World Wide Web, where no IVR and maybe no other voice service would be involved.

II. Discussion

A. Architecture Considerations

Because it would have been feasible for Centrelink to have deployed speaker verification using a model that did not have it directed by the IVR, and because such an architecture would likely exist in an all-web world, it is important to consider speaker verification (and perhaps, speaker identification) as a security service rather than as a voice service. Given the environment that it works in, it makes sense for it to be provided as a Web Service. This distinction is important for framing the approach to applying standards to the technology.

B. Issues with Communications

Kaz, being quite conscious of Centrelink's positive attitude towards open standards, built their components using Web Services based standards wherever possible. For IVR communications, this has worked well, since the IVR is primarily supported by Scansoft based software. However, they had to work with old Nuance (that is prior to the take-over by Scansoft) speaker verification engine and the relatively immature speaker verification technology from its internal voice technology arm.

Further, due to lack of standards and project scope restrictions, no standards relating to biometrics were adopted and a minimalist approach was taken to the interface between the voice authentication control component of the Telstra/Kaz solution and the IVR component of the Telstra/iTa solution. Essentially VoiceXML was adopted where it made sense.

C. Biometric Templates

There are no existing standards around speaker verification templates. Where there is a change vendors or where vendors change their technology, then re-enrolment may be required.

There would, however, be a serious cost to the implementation of a common template standard. In the case of a template that required a template that defined the feature set, that cost would be in stifling innovation in the design of the biometric system, an issue at this early stage of biometric development. In the case of the template essentially containing the biometric sample (similar to AFIS arrangements for fingerprints), then there are processing costs at each authentication (whereby the enrolment template needs to be rebuilt) as well as the obvious privacy risks.

Notwithstanding this, it would not be unreasonable for the evolution of a standard for the template wrapper. Further, it would not be unreasonable for such a standard to follow the standards already developed for biometrics (BioAPI and CBEFF).

D. Standards Issues and Opportunities

The real issue faced by the Centrelink service is the standards for communications between the speaker verification services and other systems. Those other systems may be IVRs, but may be other services as well.

The payloads that Centrelink's speaker verification service passes around are:

- biometric samples (unencrypted and encrypted);
- biometric templates (encrypted prior to entry to the transport layer);
- biometric results (that is, scores) from the speaker verification engines; and
- biometric decisions.

In any verification system where information may be required for legal purposes, all of the above information will be required to be transported around the system. Further, in such cases, the information will need to be transported securely so that courts can be assured that the information has not been tampered with, not only when stored, but also in transit.

III. Conclusion

The principal opportunity for standards in the biometric speaker identification and verification space at this point in time is for the transmission of processed and unprocessed biometric data between the biometric systems and other non data base systems. While it may be beyond the scope of these standards to specify encryption standards, they should allow for the requirement for data to be encrypted. Further, any guides on the use of the standards should advise on the use of encryption.

References

[1] T. Dunstone, "Summary report on biometric technology," 2003, not publically available.

[2] J. Whalan, *Centrelink Annual Report 2006-07*. Box 7788, Canberra Business Centre, ACT, 2610, Australia: Australian Government Printers, Sep 2007. [Online]. Available: http://www.centrelink.gov.au/internet/internet.nsf/filestores/ar0607pdf/$file/complete.pdf

[3] T. Dunstone and N. Yager. (2008) What is performix? Internet. Biometix Pty Ltd. Last accessed 18/12/2008. [Online]. Available: http://www.biometix.com/documents/Performix_introduction.pdf