

VoiSentry speaker verification

Datasheet



Identity verification for any application

Aculab's VoiSentry is designed to enable OEM partners to add voice authentication to any telephony-based solution. Now, your business customers can replace frustrating, time-consuming processes simply by allowing their customers and/or employees to verify with their voice.

Across all market sectors, businesses compete in terms of customer experience. All seek efficiency and productivity gains. The simplicity and convenience of verifying or identifying a caller by their voice, as opposed to agent-led ID&V or knowledge- and token-based authentication (KBA or TBA), provides seamless user interaction, while also offering time and cost savings for the end user organisation.

With the escalation in identity theft, fraud, and social engineering attacks, businesses have a compelling imperative to provide additional security in terms of access to user data, accounts and services, for both employees and customers. As well as verification, VoiSentry offers speaker identification, whether to help detect known fraudsters or to differentiate between trusted speakers.

Speaker verification	Verify callers against an identity claim
Speaker identification	Identify a speaker from a list of known candidates
Multiple verification modes	Text-dependent, text-independent, and text-prompted modes, all using the same speaker models
Autonomous passphrases	Individual passphrases provide an additional level of security
Multi-factor authentication	Integrated, multi-lingual spoken number and DTMF recognition
Presentation attack detection	Anti-spoofing technology mitigates risks from presentation attacks, including synthesised and replayed speech
Algorithmic adaptation	Per-speaker adaptation and model updating track changes in the user's voice and improve accuracy
Machine learning and artificial intelligence (AI)	Combines machine learning and AI with hand-crafted algorithms and advanced signal processing for robustness and accuracy
Easy integration	A 'virtual appliance' that runs as a VM on the developer's platform of choice; on-premise, data centre, or hosted cloud
Web-based UI	All nodes accessible from a log-in at a single node; remote, web-services administration
Multi-tenant	Runs multiple, distinct applications, and hosted solutions
Scalable and redundant	Add capacity and redundancy by creating a multi-node system
Load balancing	Optimise system node throughput, loading and response time

Key features

Product benefits

Self-service assist (IVR containment) – simple, convenient, intuitive verification by voice

Improves the customer experience (CX) – reduces irritating security interrogations

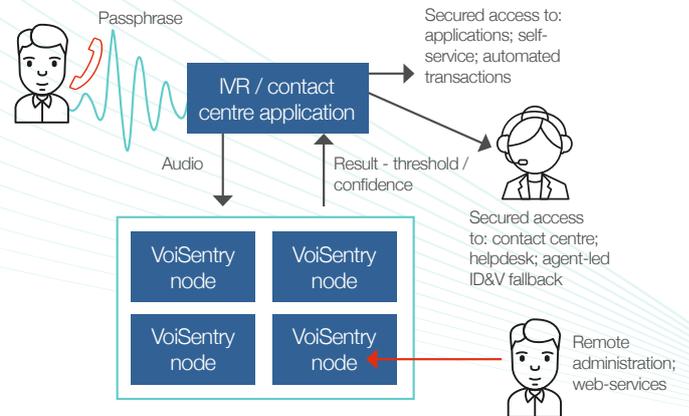
Cost-effective and efficient – dramatically reduces time/cost associated with manual ID&V

Multi-factor authentication – enables enhanced security and fraud prevention

Flexible deployment options – ideal for multiple, distinct applications, and hosted solutions

For developers of solutions requiring secure identify verification, or providers offering services to multiple clients from their hosted, cloud-based platforms, VoiSentry presents an ideal solution.

A multi-tenant customer interaction solution, for example, can offer caller ID&V, simply and conveniently, according to individual businesses' needs. Each has full, independent control over the creation of datasets against which enrolments and verification or identification attempts are performed.



Technical summary

Minimum hardware configuration	
Processor	8 CPU cores
Memory	20 GB RAM (including 4GB system headroom)
Storage	20 GB of VM disk storage (when using the integrated volume)
Software environment	
Deployment	Supplied as a 'virtual appliance'; for deployment as a VM (ideally the sole VM) onto a hardware platform with an installed virtualisation hypervisor; alternatively, as a 'Docker image'
Hypervisor	VMWare vSphere ESXi 6.0, or ESXi 6.5; Microsoft Hyper-V; Docker (Linux host)
Application Programming Interface (API)	Web-services, REST-based API (provides enrolment, verification, etc.)
User interface (UI)	HTML-based administration UI (allows administrative control and status visibility over an entire cluster)
Web-services (WS)	A WS-based administration REST interface (enables VoiSentry to be remotely administered from 3rd party applications)
Data storage	Self-contained volume (within the virtual appliance); alternatively, greater external capacity may be assigned
Additional specifications	
Voiceprints	Derived biometric metadata for verification and identification; datasets not externally accessible; proprietary format; cannot be used elsewhere
Passphrase	Fixed, or autonomous selection (applies to text-dependent)
Verification modes (API configurable)	Text-dependent (default; with pattern matching); text-independent; text-prompted

Identification	Identifies a speaker from a list of known candidates
Liveness detection	Prompted, random text/digits; enhanced, multi-factor, token-based authentication
Similarity (anti-spoofing)	Performs a similarity (duplicate/copy) test on audio data
Presentation attack detection (PAD)	Selectable; duplicate, replayed or synthesised speech (ref: ISO 30107-3)
Languages	Language independent
Minimum input audio (min)	2 seconds of speech
Recommended enrolment audio (min)	3 repetitions of speech
Recommended verification audio (min)	1 repetition of speech
Enrolment time	Sub-second (for 3 repetitions of a passphrase of 3 seconds)
Verification latency	Sub-second (for a passphrase of 3 seconds)
Identification latency	Dependent on comparison ratio
Audio format	.wav format; G.711 (A-law or μ -law) or 8kHz, 16-bit linear (mono) PCM
Feature analysis	Designed for real-world telephone speech, sampled at 8kHz; noise robust
Verification performance	A standard configuration ¹ has been proven at up to 40,000 verifications per hour
Identification performance	In excess of 40,000 comparisons per hour; dependent on the ratio
Verification accuracy ²	A system ¹ has been shown to achieve an imposter detection rate of 99.65%
Equal error rate (EER) ^{2 & 3}	A system ¹ has been shown to have an EER of <2%
Algorithmic adaptation	Dynamic, per-speaker adaptation and model updating
Sensitivity threshold	API configurable
Automatic speech recognition	Selectable (digits 0-9; 'yes' and 'no' – English; French; German; Spanish; Italian)
DTMF detection	Selectable
Scalability	Via node clustering (adding nodes to a cluster can be expected to scale throughput approximately linearly)
Resilience and redundancy	Fail-over protection via clustering nodes
Load balancing	Reverse proxy; node polling
Management	May be monitored/managed from administrator log-in at a single node
Licensing	Embedded licence server

Notes:

¹A cluster consisting of a single node running as a sole guest under VMWare ESXi 6.0 on an Intel i7 at 3.2GHz, and assigned all 4 cores (8 hyperthreaded cores) and 16GB RAM (an additional 4GB is used by the hypervisor) with a local, commodity SSD.

²VoiSentry has been evaluated using an industry standard SpeechDat database, supplemented with data collected by Aculab, covering a wide range of telephone equipment and acoustical environments.

³The number of false acceptances being equal to the number of false rejections.

Owing to the dynamic nature of our business, specifications are constantly being changed and therefore this product overview is for informational purposes only. Aculab makes no warranties, express or implied, in this document. E&OE.

Follow VoiSentry:  Twitter |  Blog |  LinkedIn

+44 (0) 1908 27 38 38 | www.voisentry.com

VoiSentry