ViGo Adaptation Processes

What they are and how they work

Title: ViGo Adaptation Processes

Part number: VV/VIGO/DOC/190/A

VoiceVault Inc.

400 Continental Blvd

6th Floor

El Segundo

CA 90245

USA

(310) 426 2792

info@voicevault.com

Table of Contents

Introduction

Voice biometric voice print adaptation is the process of updating a voice print to reflect changes in user behavior and environment. It is designed to ensure that a user has the best possible voice login / verification experience irrespective of their location or environment. The goal and purpose is to minimize the chance of a genuine user being falsely rejected by the voice biometric system.

Why is it needed?

When a user registers, their voice print is based on the 3 or 4 speech samples they were prompted for that will have been collected from the user at a specific time and environment / location. Subsequent login / verification attempts might take place in different locations, at different times of the day, or when the user is perhaps in a different mood. All of these factors can sometimes affect the ability of the user to voice verify their identity quickly and easily.

Adaptation is the process of adding extra samples of a users' voice to their voice print so as to improve the user experience. This is an ongoing process that keeps the voice print current and reflective of user behavior. For example, if a user registered their voice at home in the morning but typically verify themselves in the afternoon at the office, the voice print will adapt itself over time to provide a better user experience at the office (without adversely affecting their experience in other locations of course).

The end result of adaptation is that the advertised false reject rate (false negative) is maintained without compromising security thus improving user experience.

How does it work?

There are two types of adaption and they are generally both used in an application. One is automated and the other requires a third verification factor to be incorporated into the app or solution. This third factor is used to trigger 'supervised, or manual, adaptation.

At a very simple level, automated adaptation occurs when a user *successfully* verifies and supervised adaption is triggered if their verification attempt is *unsuccessful*.

Automated adaptation is configured into the voice biometric system and, as the name suggests, happens automatically. It is triggered if (a) the users' login / verification attempt is successful and (b) the system decides that their 'score' lies within a certain range that is below a 'Strong Accept' threshold level (see below). The levels at which it is triggered are not configurable.

Supervised adaptation is not automated and requires *supervision* by a third verification factor that provides the mechanism to re-established ground truth that is necessary because the user failed to authenticate themselves using voice biometrics. The app will engage the third factor if a users' 'score' lies within a certain range (see below).

Ground truth can be re-established in any number of ways and the mechanism used is defined by the security requirements of the application. It could be anything from entering a user-defined PIN or swipe (as is the case in the ViGo demo apps) or having the user speak to a call center agent and respond to knowledge based questions.

If ground truth is re-established, the speech sample in question is used to adapt the voice print so that next time the user finds themselves in a similar environment they will score higher as their voice print now contains speech representative of the new environment. If ground truth is not re-established, it can be assumed that the user is an imposter and suitable measures can be taken by the app.

How and when is adaptation triggered?

VoiceVault voice biometrics, including ViGo, provides a simple mechanism that informs the app logic when a user has scored in one of 4 possible bands:

- Strong accept
- Accept
- Reject
- Strong reject

These bands are implemented in the VoiceVault voice biometric system so as to provide an abstraction layer away from the underlying thresholds for acceptance and rejection. This is because these thresholds can be modified over time as tuning and optimization processes are used; and also because individual thresholds, and the widths of the bands, can vary depending on the verification mode used (whether digits or phrases) and between phrases themselves, and with different phrase languages.

Looking at each of the bands in turn:

Strong Accept

If a user scores in this band with their first utterance, they will be automatically accepted. Automatic adaptation will not be triggered, as it is not necessary / there would be no benefit.

The operating point for scoring in this band is configured to ensure that there is a near-negligible chance that the user is an imposter. As a result, the reject rate is relatively high which has the effect that if a user does not quite reach this level, they will fall into the next lowest band, Accept.

Accept

If a user scores into this band with their first utterance, they will automatically be prompted for another one. This is because the system has a lower confidence level than if they scored in the Strong Accept band on their first utterance.

If after providing a further utterance, the user scores in this band with the average of the two, they will be accepted, and automatic adaptation will be triggered.

If after providing a further utterance however, the average of the two scores falls in the next lower band (Reject), they will be rejected; automatic adaptation will not occur; and they will be prompted for the third factor (see below).

Reject

If a user scores into this band with their first utterance, they will automatically be prompted for another one.

If after providing a further utterance, the user scores in this band with the average of the two, the mandatory (for a ViGo app) third factor / supervised adaptation mechanism will be triggered by the app logic. This is used to determine if they are the genuine person or not, and therefore whether supervised adaption can take place. Once a user has scored into this band, it is up to the app logic to prompt the user for whatever the third factor mechanism is. It is also the responsibility of the app to programmatically call the supervised adaptation process if the user passes the third factor verification process successfully.

If the user fails to satisfy / pass the third factor successfully, it is the responsibility of the app to decide what to do next. At this point, the user's identity will not have been verified and the app may let the user have another go (after a time delay perhaps); to lock them out, etc.

The supervised adaptation process is more likely to be seen in the early stages of the use of the app by a user. Over time, unless their usage patterns change radically, it will become less and less necessary. It can be considered as being one of the mechanisms by which the authentication system as a whole learns about the user and their app usage patterns.

If after providing a further utterance however, the user scores in the Accept band with the average of the two, automatic adaptation *will* occur.

Strong Reject

If a user scores in this band with their first utterance, they will be automatically rejected. It is the responsibility of the app to decide what to do with this user who's verification

attempt will have been rejected outright. The system will have decided that it is very unlikely that this person is genuine / who they claim to be.