# 1. Aim

Voice Authentication as a means of biometric authentication has been around for quite a while. Along with the technological progress and the improved accuracy, voice authentication has again attracted the market's attention. In this paper, we will look at one of the leading providers in voice authentication-Knuverse, and test its voice authentication solution "AudioPIN". Based on the experiment's result, we will advise on the status and maturity of the AudioPIN product, and how we can use voice authentication technology in the university to improve our security.

# 2. Voice Authentication Overview

## 2.1 What is Voice Authentication

Voice authentication is to use the voice biometrics as a method of determining the identity of a speaker for access control. Voice biometrics is a set of unique, distinguishable physiological characteristics and behavioural features of a person's voice, that can be used to identify that person. Voiceprint is a mathematical representation of the voice biometrics stored in electronic format. It is not a recording or file that can be played back or reverse engineered into original speech. During the voice authentication process, voice authentication software checks the user' voice against their storied "voiceprint", returning a verification result of the service software. If the result is positive, the user will be authenticated and get through to the service he requested.

## 2.2 Voice Authentication and Speech Recognition

It's easy to confuse voice authentication and speech recognition. Both technologies use recordings of the human voice, but they do different things with it. Speech recognition strips out the personal differences to detect the words, while voice authentication typically disregards the language and meaning to detect the physical person behind the speech. Speech recognition is language dependent, while voice authentication is independent of language.

Siri, Amazon Echo, Google Voice etc. are all examples of speech recognition because their technology recognize what is being said. These technologies are great advancements, but are not considered as voice biometrics, which identifies who someone is, and authenticates they are who. In situations where the identity of a speaker is required, including security and personalization use cases, voice biometrics are essential.

# 3. KnuVerse Solutions

KnuVerse voice authentication (https://www.knuverse.com/) is the re-branding of Knurld. KnuVerse voice authentication provides a few voice authentication solutions depending on the users' requirements. We will give an overview of the KnuVerse voice authentication and have a trial of their AudioPIN solution.

## 3.1 Key Features

KnuVerse voice authentication is a cloud-based solution that can be accessed via a set of RESTful APIs. All the product functionalities are residing in the cloud, requiring no servers, databases, or other systems provisioning and streamlining the development, test, and deployment process. Active Voice delivers secure, text dependent (TD) authentication with a current vocabulary of up to 42 customizable words. It can be integrated into the windows applications and mobile applications.

The KnuVerse voice authentication solution has two key components. AudioPass is a speaker authentication method that periodically prompts subjects to speak three or four words for identity verification. AudioPIN, meanwhile, is an authentication approach that lets users create PINs, and prompts them to say words that are dynamically generated and displayed.

KnuVerse delivers the following benefits:

- Streamlined user experience: as a multi-factor authentication method, users can authenticate using their voice prints with no other device needed;
- Military Grade Security: Knuverse is built on a military-grade biometric platform that is "battlefield proven". KnuVerse can extract user's voice from ambient noise, it not only works with standard microphones but also effective with mobile phones.
- Easy integration: Cloud-based platform enables integration of voice biometric APIs anywhere in the world.


Li Zhao, IT Innovation Specialist
itinnovation@contacts.bham.ac.uk

## 3.2 AudioPIN Solution

AudioPIN solution provides a combination of secure authentication using a passcode and voice prints in enterprise, mobile or web applications. It allows for highly secure authentication by providing multiple layers of protection against spoofing and other attacks. AudioPIN requires the client to be able to see the AudioPIN display on a computer or Android device and for the client's voice to be recorded either through a microphone or telephone.

The AudioPIN requires the client to be known, either as the primary user of a device such as a mobile phone or tablet, or otherwise identified through a login ID. Clients can then verify themselves by saying the words that are displayed in the positions corresponding to their AudioPIN number positions on the keypad. The AudioPIN consists of 4 or more digits. To use AudioPIN, users will enroll by reading a list of defined words three times. Currently, these words are a list of cities in the United States. The AudioPIN works as follows:

1. The client says the word displayed in the keypad location corresponding to the first digit of his/her AudioPIN;
2. The display will be updated with a new set of words randomly placed within each number on the keypad and the client will say the word corresponding to the second digit of his/her AudioPIN. This will continue for the third and fourth digits of the AudioPIN.
3. After the animation is complete, the device will send the four words of audio data to a AudioPIN server for analysis.
4. The AudioPIN server will return its decision whether or not the client has been verified.

## 3.3 License Models

Pricing for KnuVerse products is an annual subscription based upon the number of users. Education and reference discounts are applicable to the University of the Birmingham.

- The "AudioPass" voice authentication product, which user would add in front of the Shibboleth implementation, to add voice authentication beyond the user id and password authentication. The list price for this product, for 5,000 users with unlimited verifications, is $6.00 per user per year ($30,000 per year).
- The "AudioPIN" product, including the windows client application integrated with the Windows login process using both a PIN and the voice prints, is priced at $11.00 per user per year ($55,000 per year).

Please note there is a $10,000 annual subscription minimum.

# 4. AudioPIN Experiment

In order to test the AudioPIN solution, an experiment was carried out to test its functionality and reliability.

## 4.1 System Setup

The AudioPIN solution comprises of reusable and scalable components to fit users' needs. There are five major components of the AudioPIN systems:

- Front End Srvices: AudioPIN API, AudioPIN Console;
- Backend Servce: AudioPIN Engine, AudioPIN Phone;
- Database Service
- Messaging Service
- Redis Caching Service

AudioPIN is available as a Virtual Machine Applicance in Open Virtual Format. The appliance has all of the required components available and configuration tools to make it easy for user to instantiate a running AudioPIN service. In our testing scenario, Knuverse set up the test virtual machine appliance over the cloud for us in the trial period. We have been given the administrator right on this testing AudioPIN console to enrol and authenticate users.

## 4.2 User Enrolment and Authentication

Using the admin credentials provided by converse, user can log on the Audio PIN Console. Figure 1 shows the interface of the console. On the left navigation menu, it lists the main functions; the dashboard is displayed on the right side, which lists the statistics of the users, verification rates etc.

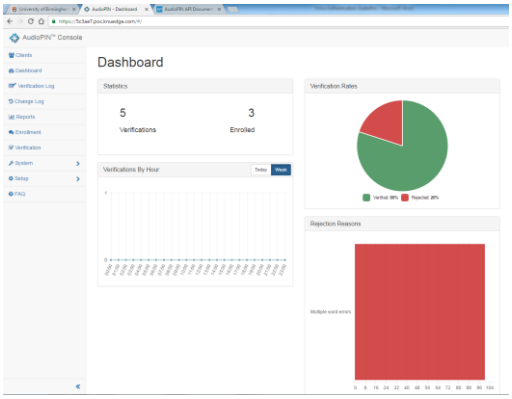


*Figure 1 AudioPIN Console*

### 4.2.1 User Enrolment

To use the AudioPIN service, first step is to enrol the users. Select "Enrolment" from the menu, there are two ways to enrol users: Random phase and AudioPIN.

- ***Random Phase Method:*** User will need to create a username, select the gender and the recording device. User clicks on "Continue" to enter the enrolment page. On the enrolment page, the system will display the instruction and show the words one by one, the user will need to concentrate and follow the word list in the exact order and read them clearly.

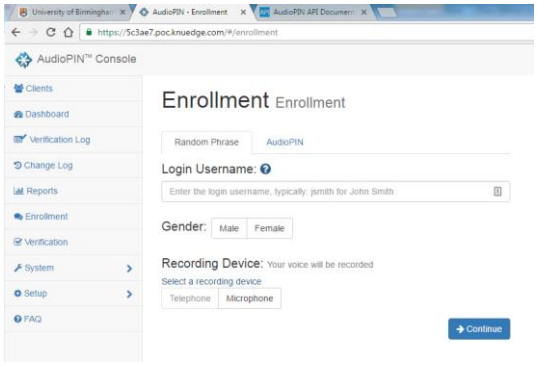


*Figure 2 Random Phrase Enrollment*

- ***AudioPIN Method:*** To enrol using the AudioPIN method, users will need to have a login username, select the gender, recording device and create a 4 digit PIN, see figure 3. Users will need to remember this 4 digit PIN for authentication purpose later. In the enrolment phase, words appear in the PinPad then fade away, users will follow the instruction to read the words at the corresponding position, see figure 4. To enrol a new user guided by the tutorial, the process takes about 4 minutes. Users will need to be very concentrated when doing the enrolment and follow every step of instruction clearly, otherwise users may fail to enrol. In our testing, it usually takes more than 1 attempt on average to enrol a new user successfully.

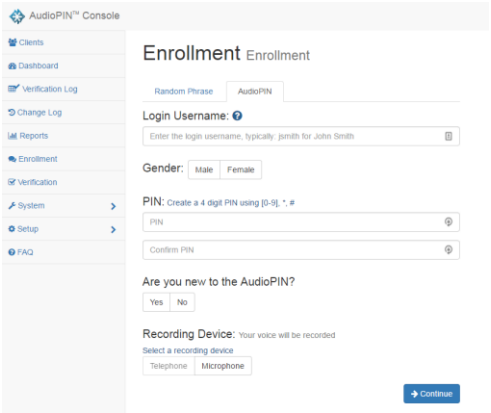


*Figure 3 AudioPIN Enrollment*

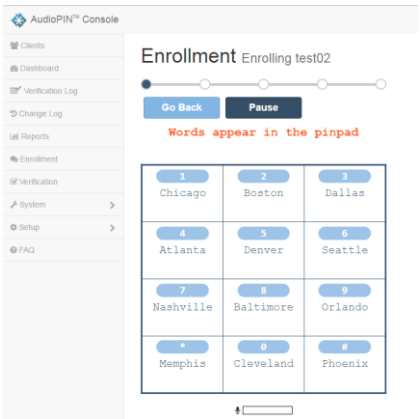


*Figure 4 AudioPIN Enrollment Process*

### 4.2.2 User Verification

Users can be verified using the same method when they are enrolled, either by "Random Phase" or "AudioPIN". To verify the user by random phase, the user just needs to read the words showing on the screen. The process is quite straight forward. To verify by "AudioPIN" method, the user will need to say the word displayed in the keypad location corresponding to the PIN code of his/her AudioPIN one by one, see Figure 5. This is more secure than just repeating the words, because the PIN code and the voice print information together act as a two factor authentication. The user uses their PIN to determine the correct sequence of words to read. The user's PIN code is protected by randomizing the word placement each login.

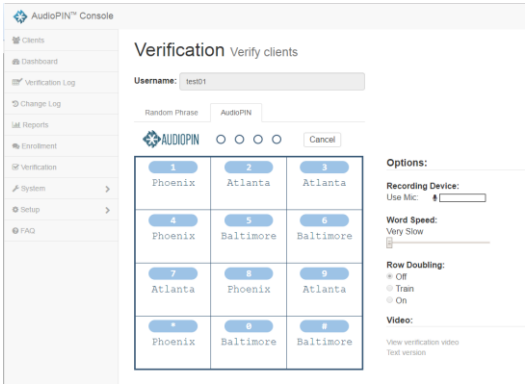


*Figure 5 Verification by AudioPIN*

## 4.3 System Management

The AudioPIN console provides quite a lot of functions including enrolling, authenticating, and removing users. One of the key functions is the system configuration. Under the system configuration, there is a function called "login Bypass" in the security tab, see figure 6. Login Bypass allows users to log on without voice authentication, administrators can set additional checks to make the bypass more secure. This is particularly useful in some special scenarios, such as equipment failure or user is having a very bad cold.

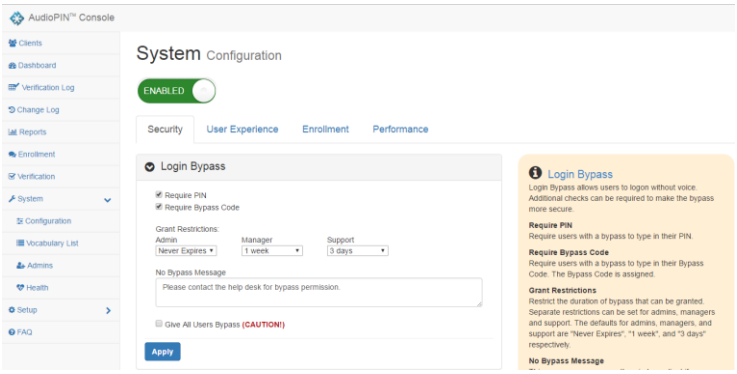


*Figure 6 Login Bypass*

Another function worth to mention here is the "Vocabulary List", see figure 7. In the testing environment, it currently has 42 words which are optimized for US English. The administrator can select which of the 42 words to be used, but cannot change these words. To redefine the vocabulary and include some specific UK words, it will require a collaborative effort from the Knuverse and us. We will need to conduct a data collection (getting 100+ people to say each of the new words) and Knuverse will use that data to create "background models' for each of the words. According to Knuverse, this process can be completed within a few weeks for a nominal fee.

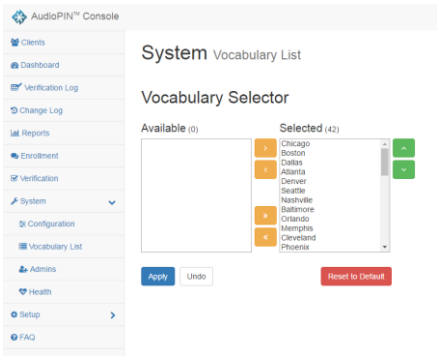


*Figure 7 Vocabulary List*


Li Zhao, IT Innovation Specialist
itinnovation@contacts.bham.ac.uk

### 4.4 Experiment Findings

Through the AudioPIN experiment, we find out:

1) AudioPIN solution, as a cloud based solution, is quicker to set up and certainly needs less in-house resources to maintain the services, however, it will need some development at the initial stage to identify a suitable vocabulary list and to integrate the AudioPIN solution into the university login system.
2) AudioPIN console is an easy to use administration tool, which provides the functionalities to configure the application and manage the users through a web interface.
3) One of the key benefits is that AudioPIN solution's capability of detecting speech in noisy environments. Testing has approved its reliability in the noisy environment, in the noisy music background, users can authenticate successfully to the system using a microphone.
4) We have also tried to test AudioPIN's reliability by rendering user's voice on purpose. When users have rendered their voice, AudioPIN failed to verify the users. So if users are having a cold or any other throat condition, there is a high chance that users will not be able to authenticate to the AudioPIN. It is essential to set up a bypass code in such scenarios.
5) The process of enrolling new users is not straight forward as expected. In the "Random Phase" method, users will repeat the listed words three times for the enrolment; while in the "AudioPIN" method, words appear in the PinPad then fade away, users will follow the instruction to read the words at the corresponding position. The whole process takes about 4 minutes and the users can not miss a word as it will affect the enrolment result.
6) Currently, the testing result is only based on a few users, which means the sample data is quite small. Testing from a larger number of users will improve the accuracy and precision of the result.
7) Further prototype of integrating AudioPIN into the university login system will help to understand how much the task involved and the degree of the users' acceptance.

## 5. Other Alternatives

There are quite a few contenders in the voice authentication technology, below are the leading ones:

- Microsoft launched Speaker Recognition API as part of their Cognitive Services (https://www.microsoft.com/cognitive-services) to identify speakers or use speech as a means of authentication. Developers can use it to power applications with a voice authentication tool. The plan starts free for up to 10K application calls per month, then various charges of the plan depend on the usage. Microsoft Speaker Recognition API gives developers the flexibility to embed voice recognition into applications, for example, it is possible for university to embed this service into university mobile apps.

- Voicevault (http://voicevault.com/) provides a voice biometric identity verification solution "ViGo" for mobile, on-device and telephony application. Amazon Web Services hosting is included with ViGo right out of the box, ViGo is configured out of the box for security and convenience. "ViGo" interface supports multi-language include English, Mandarin Chinese and Spanish. "ViGo" has the benefit of directl implementation on mobile, on-device and telephone, which potentially can be a complete voice authentication solution for the university.

- Nuance Communications (www.nuance.com ) provide a number of voice recognition biometric solutions. The FreeSpeech solution can verify the phone caller's identity simultaneously while conversing. The voice of the phone caller is matched with those stored in the database within a few seconds. This reduces the phone call length, which makes it efficient and cost-effective. The VocalPassword solution can verify the caller's identity against the database while interacting with a mobile app or interactive voice response (IVR). It allows the user to use his or her voice as password and avoid carrying any device or remember any information for login purposes. It also proactively detects fraudsters. This solution can be used for remote employee validation, IT helpdesk identity verification, web transaction, and mobile app authentication. Nuance Communications supply the voice biometrics technology for HSBC Voice ID banking.

## 6. Conclusion

KnuVerse voice authentication solution "AudioPIN" is a voice authentication application that lets users create PINs and prompts them to say words that are dynamically generated and displayed. Its key strengths are listed as follows:

- AudioPIN is a multi-factor authentication method, where users can authenticate using their voice prints and PIN code.

- KnuVerse is built on a military-grade biometric platform, which mean that AudioPIN can extract user's voice from ambient noise. Through the testing, users have been able to authenticate to the system in the loud music background.
- The Knuverse voice authentication solution is cloud-based, which enables easy integration of voice biometric APIs into applications anywhere in the world.

Through the testing of AudioPIN, we have identified some weakness which may need to improve:

- AudioPIN has failed to verify the users when users rendered their voice on purpose. In the situation when users are having cold or other throat condition, there is a high chance that users won't be able to authenticate with the AudioPIN. KnuVerse has provided the "bypass" mode, where administrator can set up a bypass code for the user to authenticate to the system.
- The new user enrolment process is relatively long. Especially in the AudioPIN method, it takes about 4 minutes to enrol a new user and the user can't miss a word. From our testing, it usually takes more than 1 time to enrol a new user.

Based on the experiment of AudioPIN solution, it randomizes the words to improve the security, however, it will need to improve its enrolment process to be more user friendly.

## 7. Recommendations

Voice authentication as one of the main areas in Biometric authentication has some key advantages over traditional authentication techniques.

1) The user does not need to remember any password or bring any device, which means it is more convenient.
2) Voice print is unique to every person, which means it is reasonably secure.
3) Unlike other biometric authentication, voice authentication works well for remote users. For example, finger printer authentication will need user to physically present in front of the verification equipment.

However, voice authentication also has its disadvantages:

1) In a noisy environment, the accuracy of voice authentication may be affected.
2) If the user's voice condition was affected by illness or other reasons, voice authentication may fail to verify the user.
3) It is possible to record the user's voice to do the security breach, so it is import to randomize the words phrase or integrate with other authentication methods as a second factor authentication.

Currently, voice authentication has entered the mainstream as a verification technology, more and more mobile-based e-commence applications and web-based voice portals become more prevalent in many sectors. Voice authentication technology has potentials in some scenarios in the university, for instance, authenticating user through mobile devices. However, adopting the voice authentication technology in the university successfully will highly depend on users' acceptance. The process of enrolling new users and authentication should be simple, easy to use and streamline the user experience.


Li Zhao, IT Innovation Specialist
itinnovation@contacts.bham.ac.uk