PROPOSED METHOD FOR EVALUATING VOICE AUTHENTICATION SYSTEMS

By

LESLIE NELSON

Bachelor of Science in Electrical Engineering
Oklahoma State University
Stillwater, OK
2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
DECEMBER, 2012

PROPOSED METHOD FOR EVALUATING VOICE AUTHENTICATION SYSTEMS

Thesis Approved:

Dr. John A. Acken

Thesis Adviser

Dr. L. Johnson

Dr. S. Sohoni

Name: LESLIE NELSON

Date of Degree: DECEMBER, 2012

Title of Study: PROPOSED METHOD FOR EVALUATING VOICE AUTHENTICATION SYSTEMS

Major Field: ELECTRICAL ENGINEERING

ABSTRACT: Security is an important part of our daily lives. This thesis focuses on the biometric component of security. Identity authentication for security is improved when voice authentication is augmented by other identity authentication components. This thesis presents a method for evaluating existing algorithms. This thesis details a method to compare pass/fail rates of various voice authentication algorithms. The method provides a formal mechanism to evaluate algorithms for different security levels. Different security levels have different security requirements. The application sets the security level. The level for entering a level for entering a high security research facility is more stringent for imposters at the risk of inconveniencing employees. The method also can be used to optimize setting the threshold of the algorithm. The developed method works even when the algorithm is less than ideal.

ACKNOWLEDGMENTS

I would first like to acknowledge my parents Robert and Judith Howard. Without their support I would not have been able to succeed. I want them to know how much their assistance has meant to me.

Next I would like to acknowledge my children, Daniel, John and Elizabeth. They have been an inspiration to me. They have been a major driving force in my desire to make a better life for myself and them.

I would like to acknowledge my husband Carl Butts. He has been there to encourage me to complete my thesis with phrases like, “Aren’t you going to work on that.”

I would like to acknowledge the time and effort and not always gently prodding that Dr. John M. Acken has provided over the last several years. It was through his guidance that this thesis has been written.

I also want to acknowledge my brother John Howard, who gave me my love of computers and music at the age of ten. This thesis is a tribute to both of these passions.

I would like to thank my committee members Dr. Johnson and Dr. Sohoni for taking the time to read my thesis and give their feedback.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Security is an important part of our world. There are many aspects to consider when securing a system. One of the most difficult situations to defend against is insider threat (10, 11). Insider threat is nearly impossible to guard against. Security systems need to be able to identify users. Using more than one aspect of identity authentication helps protect from other types of security circumvention. Having a reliable system to authenticate a person's identity is paramount to any secure system.

There is a trade off when allowing a person access to a secured system. A bank would not want to give access of their customer's bank accounts to a thief, yet they would like to allow their customers to access their own bank accounts. You would not want someone to be able to acquire your driver's license and create a forgery, but you would like to be able to replace your license easily if your license became lost or stolen. To identify an individual, different components are used.

There are three components to authenticating an individual's identity. The first component is knowledge of specific information. Knowledge is referred to as what they know. The second component is having physical possession of an item. The possession is referred to as what they have. The third component is physical characteristics of a person. The physical characteristics (also known as biometrics) refer to as what they are (3). Two examples of a person's knowledge are pin numbers and passwords. Three examples of a person's possessions are a smart card, a credit card or a house key. Two examples of physical characteristics are measurements of an individual's height or weight (8). The most effective way to authenticate an individual is to use all three components of identification. In secured systems there are several players involved.

Authentication of an identity is necessary for controlling access into a secure system. A secure system usually consists of a transaction between two or more parties. These transactions can have several players. There are typical symbols and names used to represent these players. The players symbols are A, B, C, D, E, I, M, P, T, R, V and W. Their nick names and descriptions are shown in Table 1 below (12, 16, 17). (A)lice is the initiator of the communication. Alice will be required to authenticate her identity to gain access into the system. (B)ob, (C)arol and (D)ave are the responders to Alice. Bob will communicate with Alice after she has been verified. Carol and Dave are other authorized users that can respond to Alice. These four players are the authorized participants. There are others that try to disrupt the transaction or steal information. (E)ve , (I)melda and (M)allory are players trying to get unauthorized access. Eve is an eavesdropper attempting to get information by listening. The information Eve collects could be used in identity theft and/or profit. Imelda is an imposter that will pretend to be Alice and try to gain access to the system. Mallory will maliciously try to disrupt the access attempts by Alice. Mallory will disrupt communication by intercepting, stealing, corrupting, or changing the information Alice is trying to send and receive. (P)eggy proves the identity of the initiator, Alice. (R)andy is a random innocent bystander who tried to access the system on accident. (T)rent is the trusted arbiter who confirms Alice's identity. (V)ictor verifies that Alice has passed the test. (W)alter is the warden who monitors communication, and may also need to prevent some communication from occurring. These players can be used in biometric security systems.

Symbol	Nick Name	Description
A	Alice	Initiator of transaction, is the person under test
B	Bob	Responder to Alice
C	Carol	Responder 2
D	Dave	Responder 3
E	Eve	Eavesdropper who is recording and analyzing information being sent to Alice.
I	Imelda	Imposter that is pretending to be Alice
M	Mallory	Malicious participant that will attempt to disrupt or damage the secure system
P	Peggy	Prover who provides evidence identity without revealing any data
R	Randy	Random individual attempting to access the system by accident
T	Trent	Trusted arbitrator whose participation is to confirm Alice's identity
V	Victor	Verifier will verify that Peggy had passed the identity authentication and is allowed that access level
W	Walter	The Warden prevents all unauthorized transactions

Table 1 Security Players

Biometrics are the measurement of an individual's physical characteristics. These characteristics are used for identification. Biometric characteristics are an important part of securing a system. The biometrics are used to authenticate the identity of an individual. Instead of only looking at biometrics as a replacement to an existing system, biometrics can be used to augment an existing system. An example is when a pin number is used in conjunction with a voice authentication (also known as speaker recognition) system. When used together the false matching rate (a match rate that allows an imposter into a system) changes from 1:100 to 1:1,000,000 (3). Voice authentication can be used with an ATM card to help secure a system. A voice system was used to verify the demonstrated evaluation method.

Current research focuses on improving algorithms. This research has created a method to evaluated voice authentication systems in relationship to its application. This is done by creating a database of user sound files. The sound files are then converted to mathematical files using fast Fourier transforming or other means. Then the algorithm(s) under test is/are used to generate pass/fail rates with various thresholds. The final steps are to evaluate the pass/fail rates for the application and draw a conclusion of what algorithm (if more than one is being considered) and what threshold should be used.

The evaluation research has concentrated on voice print biometrics. The voice authentication system (VAS) shows how the evaluation method can be used. The VAS tests different voice authentication algorithms. The steps for identity authentication based on speech are as follows. Voice prints are created from known individuals. When a user wants to access the system they will provide a new voice print. The known voice print from the data base is compared to the access user's voice print. If the access user's voice print passes the comparison test, then the user is authenticated. If the voice prints fail the comparison test, then the user is rejected.

In order to evaluate the algorithms it is imperative to know when an authentic user or an imposter is trying to gain access. Knowing the true identity of the user is required to generate pass/fail rates. These rates will be used to weigh the effectiveness of the algorithm. The focus of the research is to propose a method for evaluating different algorithms using the voice print biometric.

Examples of situations in which a person's identity would need to be authenticated are when goods, services, information and/or currency are changing hands. Originally bartering would take place face to face. If the people had met before they could use biometrics such as voice and face recognition to confirm identity. If they had not met before, then another method of identification would have been used. The other method could be an introduction by a mutual acquaintance or a token identifying the person. Biometrics are only a part of how systems are secured.

Security systems can be used to protect communication. An essential part of protecting communication is to ensure unauthorized people cannot use it. One component to secure a system is cryptography. Cryptography is used to provide protection, verification, and non-repudiations. Cryptography transforms plaintext into ciphertext using encryption and then ciphertext is decrypted back to the plaintext message using special keys (15). Special keys are used to thwart imposters, because they will not be able to decipher the message without the correct key. Steganography is another method of protection. Stenography is used to hide messages within another message, such as text in pictures or music. Stenography is done today with watermarks that are not visible to the naked eye. The watermarks give information about the hardware that printed the document.

I.1 BACKGROUND FOR BIOMETRICS

Biometrics are measurable physical characteristics of an individual. These characteristics can be used to distinguish one individual from another. Examples of these characteristics are face dimensions, finger prints, iris scan, DNA and voice print (6, 19). Some of the parameters of biometrics are described in the following Table 2. Universality describes a parameter that the majority of people possess. Acceptability is how comfortable a person is to providing the sample. An example is that a person may choose to change banks if an iris scan is required to access their bank account. Collectability is the difficulty of acquiring the biometric. Recording a voice is easier than taking a blood sample. The technology parameter is the ease of obtaining equipment for the quality of data collection needed.

Parameter	Description
Universality	Will most individuals be able to meet the requirements of the system?
Distinctiveness	How different will the measurement of one individual be from another?
Permanence	Will the biometric pattern change over time, when ill, environmental conditions with the quality of equipment?
Collectability	How easy is it to get samples of the biometric?
Acceptability	Will a person feel comfortable giving the sample?
Circumvention	How easy is it to trick the system?
Cost	How much time, money and data storage space is required?
Accuracy	Will a person be correctly authenticated and imposters correctly rejected?
Time	How long it takes an individual to state a phrase?
Repeatability	If an individual is asked to singing a response can the person always hit the same note?
Storage of Data	How much space is available for sample database?
Technology	Is the technology available to adequately take the sample?

Table 2 Parameters of Biometric Authentication Systems

Evaluations of these parameters are shown in Table 3. The universality of DNA is great because every person has DNA. Collectability of DNA samples is terrible because of the complexity of the equipment to test the blood samples, the drawing of blood can cause issues of safety and the analysis of DNA is expensive. Acceptability is terrible because the majority of people do not want to give blood samples. Universality of fingerprints is great because barring an accident or purposeful disfiguration most people have fingerprints. Collectability of fingerprints is great because you just need a piece of paper and some ink. Acceptability of fingerprinting is bad, because people associate fingerprints with criminal investigations and think fingerprinting is intrusive. Universality of hand dimensions is good because most persons have five fingers on each hand. Collectability of hand dimensions is medium because the equipment is cumbersome. The acceptability of hand dimensions is great because people do not see hand dimension sampling as intrusive. Universality of height and weight are great because everyone has both. Collectability is great for height and weight because methods to measure them are readily available. Acceptability for taking a height measurement is great, because most people do not mind others knowing how tall they are. Acceptability for taking a weight measurement is medium, because many people are self-conscious about their weight. Universality for voice prints is great because most people can speak. Collectability of voice prints is great because recording devices are inexpensive and readily available. Acceptability for using a voice biometric is great, because people do not have a negative connotation with voice authentication (18). Universality for eye iris scans is great because most people have eyes. Collectability for iris scans is terrible because of the high expense and low availability of equipment. Acceptability is low because people do not like having their eyeballs scanned. Universality for face dimension is medium because people can wear scarves, change hair styles or grow beards. Collectability of face dimensions is medium due to equipment availability. Acceptability of face dimensions is medium because people are uncomfortable with having their face measured. Although the biometric comparison analysis of voice prints is not exhaustive the comparison is sufficient for this evaluation. The voice print biometric was also chosen for this study because of personal interest and the ranking of good in a majority of parameters.

Parameter	DNA	Fingerprints	Hand Dimension	Height	Voice	Weight	Eye Iris Scan	Face dimensions
Cost	--	-	-	++	++	++	-	-
Time	--	-	-	++	++	++	-	-
Universality	++	++	+	++	++	++	++	M
Distinctiveness	++	++	+	-	-	-	++	M
Permanence	++	++	+	M	+	-	++	-
Collectability	--	++	M	++	++	++	--	M
Acceptability	--	-	+	++	++	M	--	M
Circumvention	++	+	-	-	M	--	++	-
Accuracy	++	++	+	-	-	--	++	-
Repeatability	++	++	+	M	+	--	+	-
Storage Requirements	--	--	++	++	M	++	-	++
Availability of Technology	--	+	M	+	++	++	-	M

++ = Great (or cost is low, time is short; hard to circumvent) ; + = Good;
M = Medium; - = Bad; --= Terrible (cost is high)

Table 3 Evaluation of Parameters of Biometrics

As stated previously voice patterns will be used in this thesis to prove the proposed evaluation method. Voice authentication is a combination of behavioral and physiological characteristics (3). The average speech spectrum energy amplitude is from 50 to 10000 Hz (14). These energies are greatest at 100 to 600 Hz, where the first formant is located (14). There are discrepancies on where normal speech takes place. According to the previous source normal speech is from 60 to 350 Hz and others state 20 – 4000 Hz (13). This research focuses on 40 to

2000Hz. The physiological characteristics are spoken of in the speech pathology section later in this thesis. The behavioral characteristics would be timing or accents. Some strengths of a voice authentication system are acceptability of the users and the availability of the existing telecommunication system (4). A person will usually be more willing to use a microphone then be fingerprinted (4). One drawback to voice authentication is that people perceive voiceprints as easy to circumvent (i.e. a tape recording). Research that has already been conducted has dispelled the recording circumvention myth (4), but there are still some challenges with using voice authentication.

There are several difficulties with using voice for identification. Passage of time can cause changes to the voice physiology. Therefore if the database is not periodically updated then a person can be rejected when they should not be. There can be changes with the quality of communication (i.e. background noise, illness, inconsistent voice sample, drunkenness, duress or hardware changes) causing a false results. The difficulty with background noise became apparent during the course of this research. The samples that were taken on different days and in different location had different sound recording environments. The noise and signal quality changed between the different environments. Noise and signal quality can be eliminated by having a standard system setup and location. Another difficulty is the size of the voice samples 41-140kbytes. Storage was not an issue for this research because there were not many samples of data. The file size would be a larger issue with large databases such as a bank account access system. The sample names and sizes are shown in the appendices. A tradeoff may need to take place for amount of accuracy verses cost of space. A voice authentication system requires a method to collect voice samples, such as a microphone hooked up to a computer. Next the voice will be compared to the known data samples with an algorithm. Once the algorithm comparison has been made a pass or fail message will be displayed. The comparison process is shown below in Figure 1.

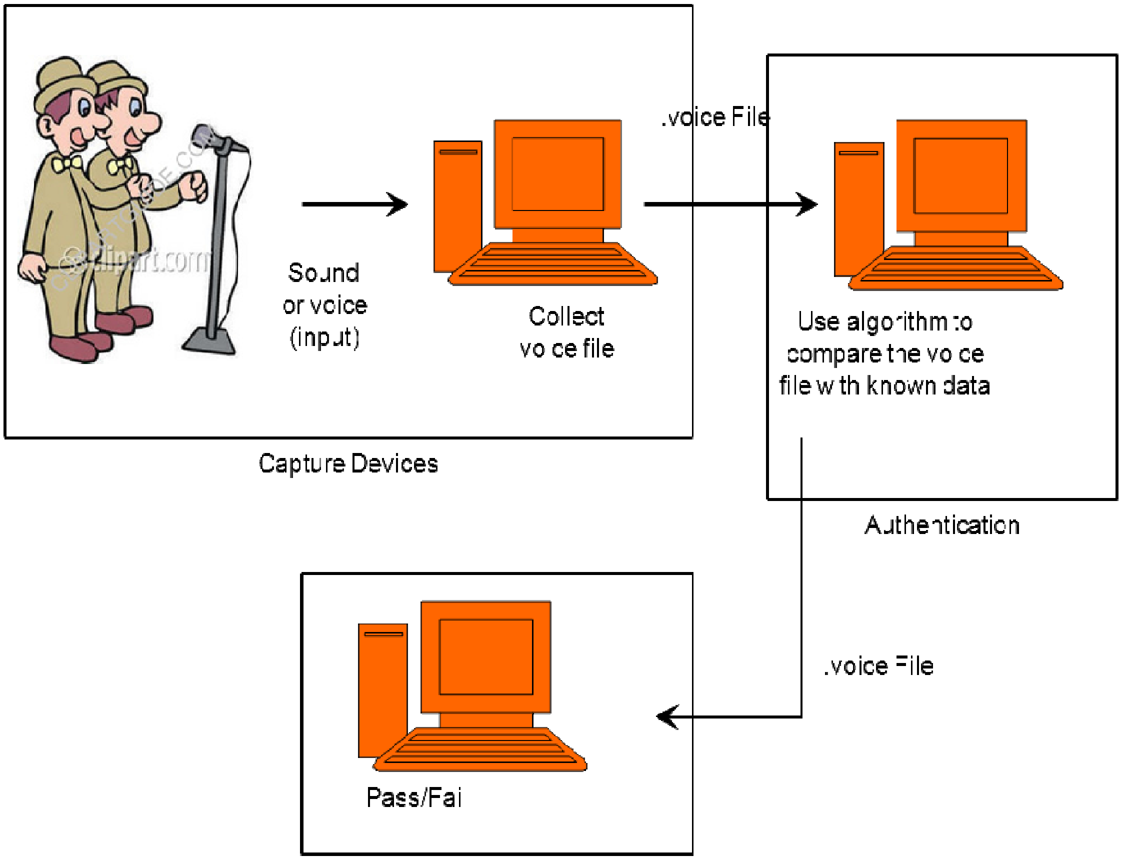


Figure 1 Basic Diagram of a Voice Authentication System

The testing of statistical hypotheses is an important decision making tool (9). The hypothesis is that if the user passes the test they are the person. There are four possible outcomes to the testing. If the person under test is the authentic person they can be allowed access, or rejected. When the authentic person is authenticated (passes) then (s)he is allowed access to the system. When the authentic person is rejected an error has occurred. If the person under test is an imposter and fails the test then (s)he is blocked from the system. When the imposter passes the test an error has occurred. The four outcomes are detailed in Table 4, and will be described in more detail in the evaluation methods section.

	Accept	Reject
Authentic Person	User is allowed access when (s)he (pass)	User is not allowed access when (s)he should be allowed access (fail)
Imposter	User is allowed access when (s)he should not have access (fail)	User is not allowed access when (s)he should not have access (pass)

Table 4 Pass/Fail Test

I.2 EVALUATION METHOD TO DETERMINE BEST ALGORITHM FOR A SYSTEM

Once the data has been recorded the next step is to convert the files from time based sound waves to frequency based sound waves. The time to frequency conversion allows for mathematical matching of the algorithms. The results of the mathematical comparisons will determine which algorithm is more effective for a given application.

This thesis created a method for evaluating algorithms. The algorithms chosen to display the evaluation method are described below. There were a few algorithms that were chosen to demonstrate the evaluation method. The first algorithm uses the maximum amplitude (peaks) of sound at given range of frequencies. These maximums are the peaks of the fast Fourier transforms. The second algorithm uses minimum amplitudes (i.e. valleys) matching to cancel a peak match. The third algorithm compares the time an individual takes to say a given phrase. The fourth algorithm uses the relative amplitude. Other algorithms that can be used are frequency range, minimum amplitude of the speaker (valleys), and patterns of maximums (peaks) and minimums (valleys) as seen in Table 5.

Measurements	Descriptions
Frequency range	The range of frequency for a specific persons voice (Men tend to have a lower frequency the women)
Valleys (Minimums)	The local minimum amplitude is a sum of amplitudes within a range of frequencies
Peaks (Maximums)	The local maximum amplitude is a sum of amplitudes within a range of frequencies
Peaks using a peak valley match to change the peak fail rate	The minimum amplitude amplitudes compared to maximum amplitude amplitudes
Patterns of Valleys and Peaks	If there is a certain amount of maximum or minimum amplitudes at specific frequencies
Timing	The time a person takes to say a given phrase
Relative Amplitude to Average	The ratio of the maximum amplitude to the average of the amplitude

Table 5 Measurements for Voice Authentication

The number of matches for valleys and peaks used in the algorithm are adjustable. These adjustments are made on a need basis. When the system is not critical the number of matches for valleys and peaks can be selected to give a larger overall matching rate even if the number of matches allows more imposters system access. The valleys have been used to cancel between peak matches between two samples. Timing is the length of time a person takes to speak a phrase. The relative amplitude is the proportion of the highest peak to the sound file mean.

The number of valleys and/or peaks to consider for matching is adjustable. In the next three figures, five points were selected as a simple example. The more peaks and valleys used the larger the match count. The match count is adjusted per application because a larger match rate also means more imposters get into the system. The below peaks show five maximum points with arrows at the highest part of the peaks. See Figures 2, 3, and 4 for simple examples of comparison points.

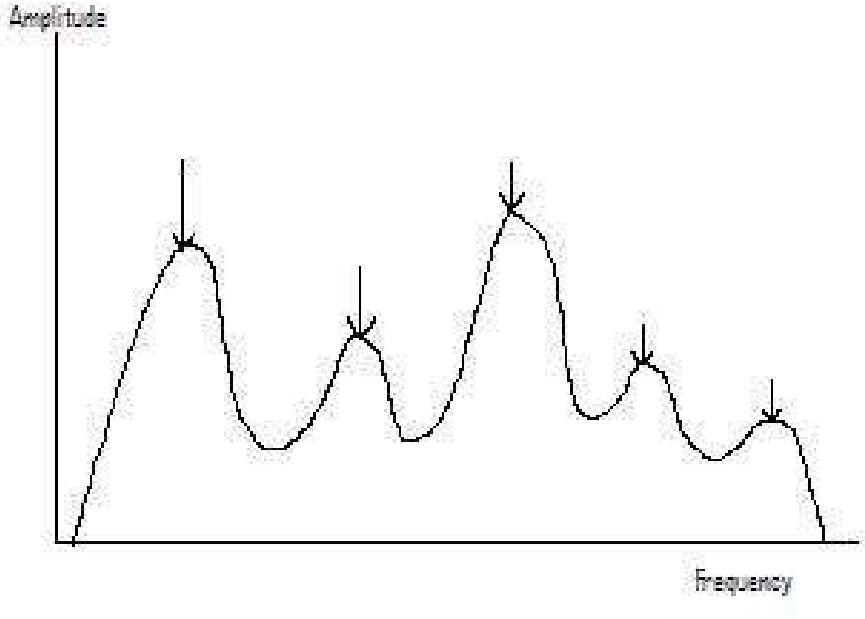


Figure 2 Peaks of Maximum Amplitude

The amplitude peaks will be compared between the existing database and the sample taken to attempt system access. The amplitude peaks are the local maximum point amplitude sums at certain frequency ranges. These amplitudes will be described in more detail.

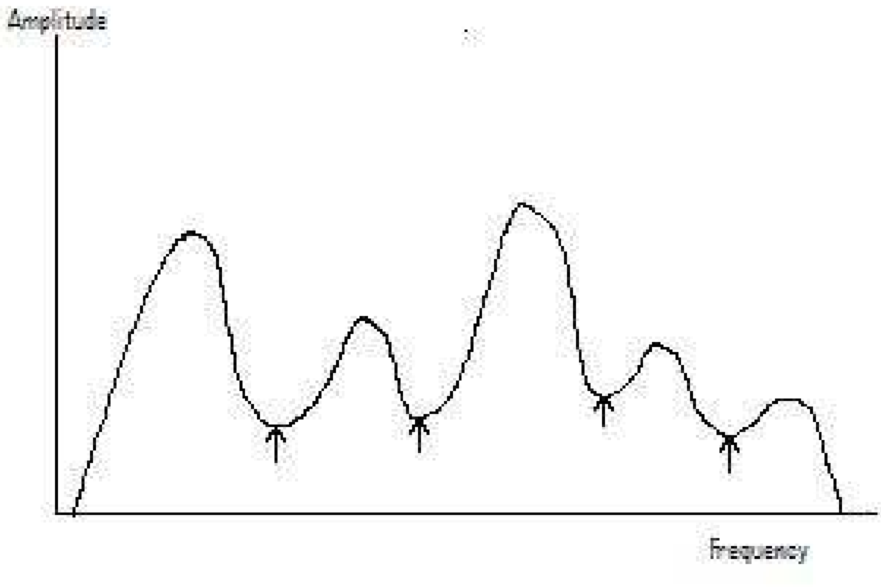


Figure 3 Valleys of Minimum Amplitude

The amplitude valleys will be compared between the existing database and the sample taken to attempt system access. The amplitude valleys are the local minimum point amplitude sums at certain frequency ranges. These amplitudes will be described in more detail.

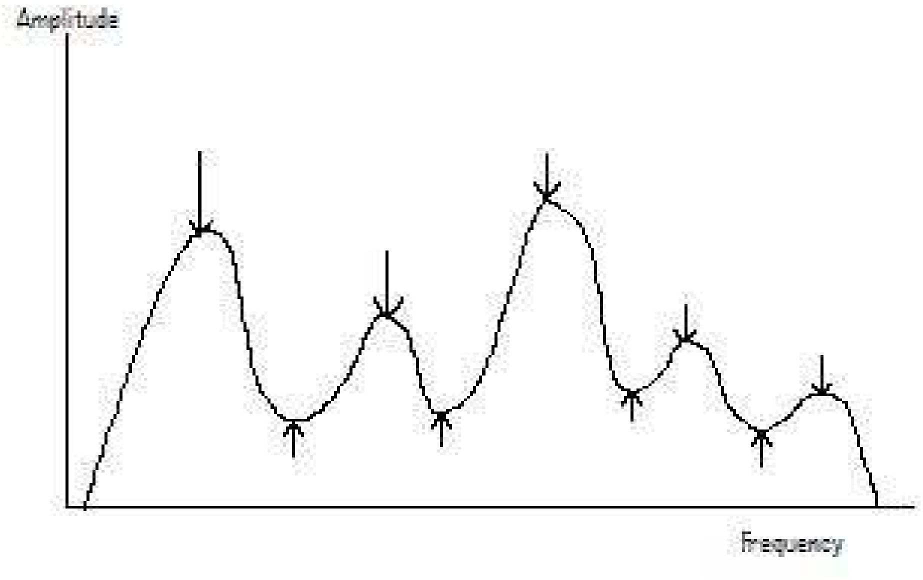
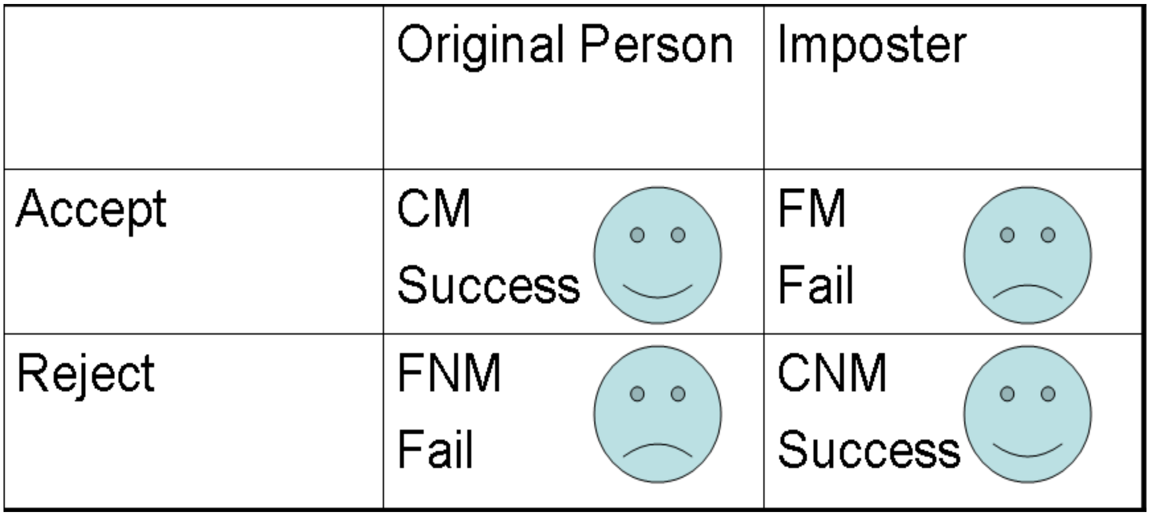


Figure 4 Mixed Amplitude

The peaks and valleys will be compared between the existing database and the sample taken to attempt system access. The peaks and valleys are the local maximum and minimum point amplitude sums at certain frequency ranges. These amplitudes will be described in more detail.

The needs of applications differ because there is a balance between if the application would rather grant access to an imposter or exclude the authentic person from the secured transaction. An example of the balance tradeoff is that access to a fast food restaurant would prefer to have less rejection of an authentic person than a nuclear power plant. The nuclear power plant would prefer to have a smaller incidence of allowing access to imposters at the cost of rejecting authentic people. The pass/fail rate of the comparisons of the system under test determines if the algorithm is acceptable for its proposed use. There are two ways that a test for matching can be successful. The first is that the authentic individual is accepted as a correct match (CM), which provides a correct match rate (CMR) and allows authentic people into the system. The second is that an imposter is rejected as a correct non-match (CNM), which provides a correct non-match rate (CNMR). There are two ways that a comparison can fail. The first is false non-match (FNM) when the authentic person is denied access. The denial provides a false non-match rate (FNMR). The second is a False Match (FM) which provides a false match rate (FMR). FM is when an imposter is accepted. The rates are detailed in Figure 5 below.



	Original Person	Imposter
Accept	CM Success ☺	FM Fail ☹
Reject	FNM Fail ☹	CNM Success ☺

Figure 5 Matching Pass and Fail Types

Below are the match rate equations:

TOTAL TESTS = M+NM

M = CM + FM

NM = CNM + FNM

Correct Rate = (CM+CNM)/TOTAL

Fail Rate = (FM+FNM)/TOTAL

FNMR = FNM/NM

FMR = FM/M

CMR = CM/M

CNMR = CNM/NM

Total test is the count of every comparison for the sample. The match (M) is the total count for a match. The non-match (NM) is the total count for a non-match. The correct rate is the percent of correct matches and correct non-matches. The fail rate is the present of false matches and false non-matches. The false non-match rate (FNMR) is the false non-match count divided by the non-match count. The false match rate (FMR) is the false match count divided by the non-match count. The correct match rate (CMR) is the correct match count divided by the match count. The correct non-match rate (CNMR) is the correct non-match count divided by the match count.

Example 1: There are 25 authentic samples of 25 individuals who work in an office. They are required to speak into a microphone to access their work computer. Two of the persons under test have colds and are rejected from their computers. Two persons (imposters) do not work in the office, but try to access a computer while the authentic person is at lunch, one succeeds. See Table 6 for details.

TERM	EQUATION	VALUE
CM		23
FM		1
CNM		1
FNM		2
Total	M+NM	27
Correct Count	CM+CNM	24
Fail Count	FM+FNM	3
FNMR	FNM/Fail Rate	66.67%
FMR	FM/Correct Rate	4.167%
CMR	CM/Correct Rate	95.83%
CNMR	CNM/Fail Rate	33.33%

Table 6 Example 1 Summary

The match rates depend upon the specific comparison algorithm used in the office's secured system. If the test becomes easier to pass then the two authentic individuals with false non-matches may be able to gain computer access, but the imposters are more likely to gain access as well. This offset is known as the threshold.

Important consideration should be used in the comparison of the authentic person and an imposter trying to access the system. The FNM error is when the authentic person/individual has been rejected. The FM error is when the imposter is accepted into the system. Both of these situations are undesirable, but the ways to reduce the errors are different. The threshold is an adjustable set point based on the importance of allowing the authentic person/individual into the system, or to keep an imposter out of the system. A graphical example of threshold is shown in Figure 6(6).

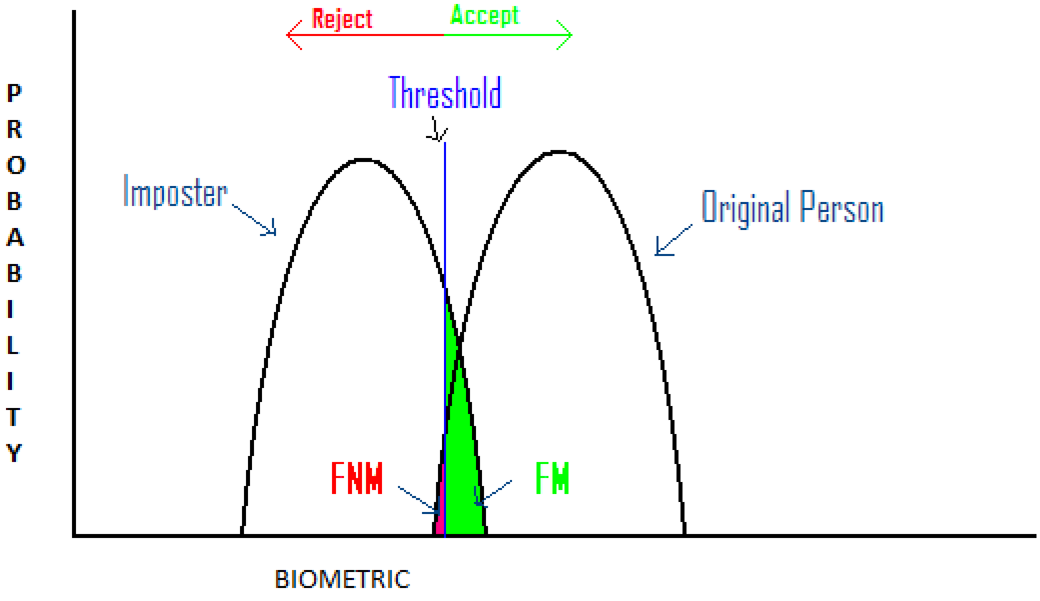


Figure 6 Representation of Threshold Position

The threshold is positioned by determining the requirements of the system. There is a sliding scale between ease of access and the security of the data that is being protected. Each application must be evaluated, so that the threshold can be determined.

I.3 PROJECT HISTORY

The precursor of this thesis was the investigation of a voice authentication system. The first task was to develop a system for voice authentication. The first task involved creating a database of voice files. To create the database voice prints were captured and analyzed. The data was collected using a microphone, a computer, and the standard Microsoft .wav recording software. The .wav files sample at 22000Hz rate. Nyquist-Shannon sampling theorem states: if a function x(t) contains no frequency higher than B hertz, the function is determined by time 1/(2B) seconds apart (5).

Because of Nyquist-Shannon sampling theorem the data after 11000Hz may be ignored (5). The wave files were recorded with a simple and widely used Microsoft software program called Wav. Wav is a program available on windows based PCs and has several resources describing how to use and apply it. All the samples taken for this thesis were saved as .wav files. The database samples are seven individuals speaking their name and four different greetings plus three additional files stating names and ten additional files stating English greetings. The greetings are in English, Spanish, Russian, and Farsi. A program was written by my brother, John Howard, with my direction to breakdown the .wav files, so that they could be analyzed. The program was written to display the numerical values of the waveforms. Observing the numerical values added an understanding about the data. This understanding allowed a more detailed system design.

The detailed study involved a microphone, computer, and the database of .wav files. When a person under test speaks into the microphone their voice creates a .wav file. The .wav file will be converted, and then the database file will be retrieved. The algorithm under test will be used to compare the access wave file to the database wave file. The comparison generates a pass/fail decision. The decision of multiple access attempts will be used to generate a pass/fail rate to make a qualitative decision of which algorithm is most effective for the application. The pass/fail rate will allow for comparisons to be made and a pass/fail rate determined as seen in Figure 7.

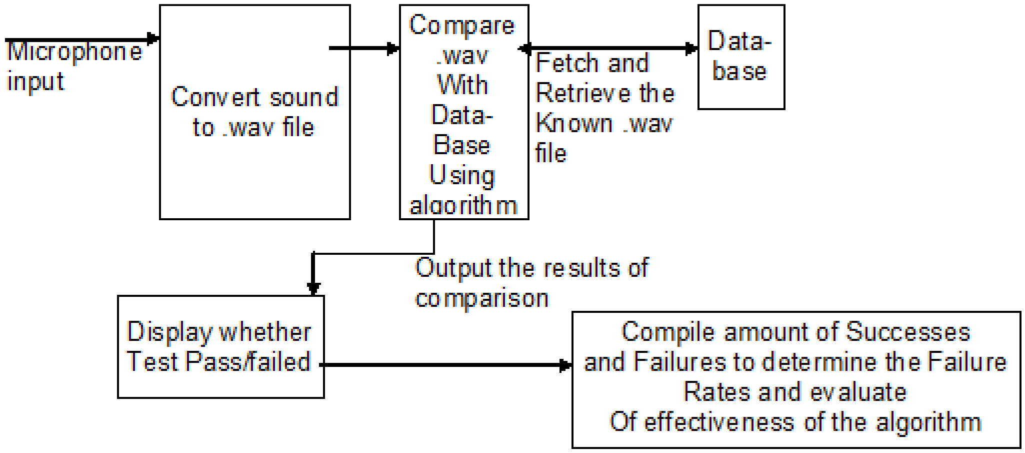

Microphone input
Convert sound to .wav file
Compare .wav With Data-Base Using algorithm
Fetch and Retrieve the Known .wav file
Data-base
Output the results of comparison
Display whether Test Pass/failed
Compile amount of Successes and Failures to determine the Failure Rates and evaluate Of effectiveness of the algorithm

Figure 7 Process to Obtain Pass/Fail Rate of a Single System

CHAPTER II

METHODOLOGY FOR COMPARISON

Voiceprint biometric comparisons were used to display the evaluation method of this thesis. Several different algorithms were used to impart how the method can be used. The first step to accomplish showing how the method works was to design a simple voice authentication system.

The following figure is a detailed flow chart of a system using this thesis' evaluation methodology as a tool to select an algorithm and set limits. The voice authentication system analyzes the effectiveness of an algorithm. To do the analysis the users need to speak a phrase. The phrase is converted to a digital file. The user's voice pattern is compared to the authentic person's voice pattern (in the database). This system is more detailed version of the testing block in Figure 7 and shows what will happen if a mismatch is made and how many times a person will be given to get a correct match Figure 8. The number of attempts is adjustable.

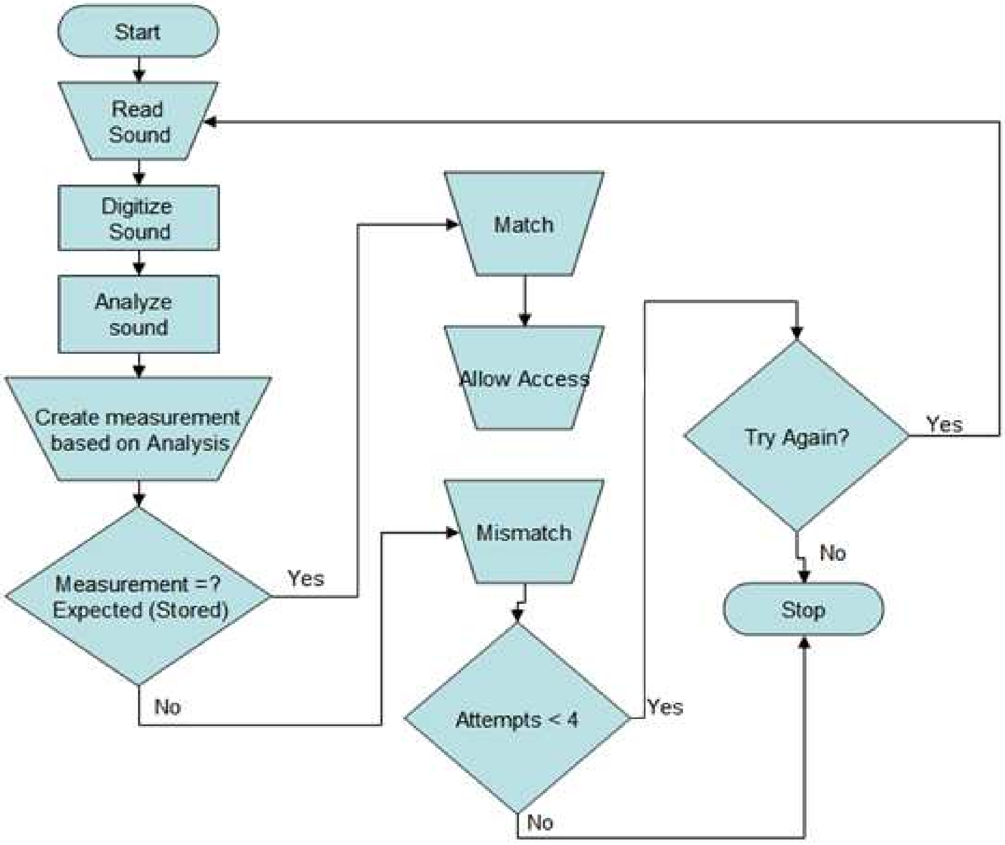


Figure 8 Voice Authentication System

Examples of different algorithm possibilities are shown in Table 7.

Measured Parameters
Frequency Range
Valleys (n deepest)
Peaks (n highest)
Combination of Valleys and Peaks
Patterns of Valleys and Peaks
Timing
Low Frequency Peak
High Frequency Peak
Low Frequency Valley
High Frequency Valley
Relative Amplitude

Table 7 Algorithms

For this research the maximum amplitudes (i.e. peaks) with the respective frequencies were initially chosen to be compared. The numbers or peaks chosen were 5, 10, and 14. After the first analysis was done peaks of n = 14 was chosen to test because it had the highest match rate without saturating the comparisons. After the maximum amplitude algorithms of all the data files were analyzed; time, maximums with minimum cancelations and relative amplitude algorithms were also evaluated for the English greeting. Once the evaluations were accomplished the maximum amplitude algorithm's pass/fail rate was quantified numerically and visually.

Figure 9 shows a single algorithm that is being tested. The test data is collected. Then the database file is retrieved. After that the algorithm is used to make a comparison between the two sound files. Lastly a pass/fail evaluation is made of the algorithm. Figure 10 shows how the comparison of different voice authentication systems (VAS) will be made. The VAS represents different algorithms. They are evaluated for use in an application. The result will be a choice of which algorithm to use.

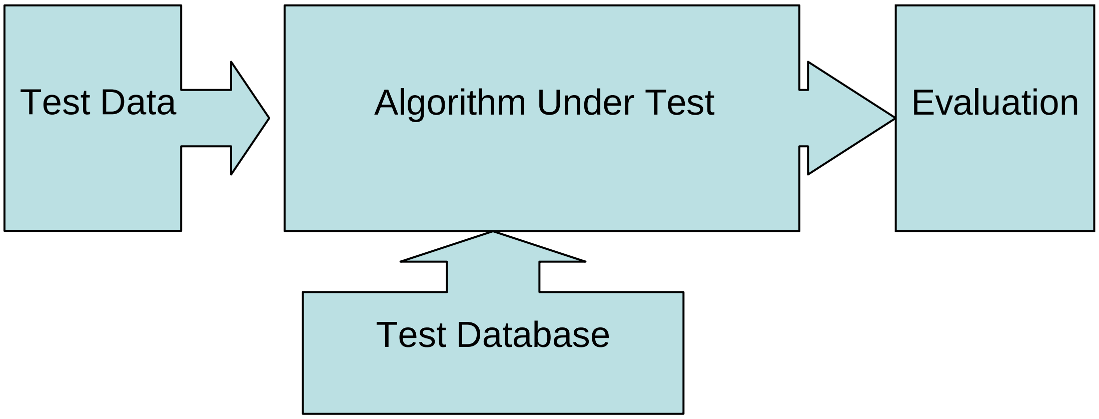


Figure 9 Evaluating a specific VAS

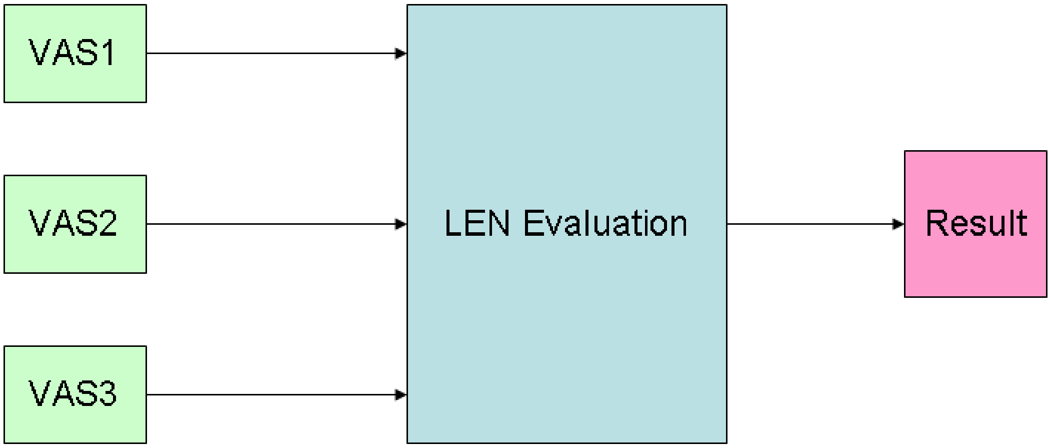


Figure 10 Comparing Voice Authentication Systems (VAS) to Pick One as a Result

II.1 VOICE CHARACTERISTICS AND SPEECH PATHOLOGY

Speech pathology studies the physics and physiology of how humans produce speech. The speech pathology field helps people who have a break down in their speech communication (1). Speech pathology is done through understanding of the principle organs used in speech production. These organs are the lungs, the trachea, the larynx, the pharynx, the nose, the jaw and the mouth. These organs make a ‘tube’ from the lungs to the mouth. Within the mouth the articulators: the soft palate, the tongue, the lips and the jaw are used to shape the air into sounds. Each speaker has a unique natural (resonance) frequency in both their chest cavity and mouth. The dividing point between the two cavities is the larynx. Physics uses mathematics to analyze sound production. The analysis helps isolate physical problems associated with speech production. Understanding speech helps to realize the characteristic and limitations of voice analysis. The understanding allows for a starting point for the spectrum of voice data files.

Human speech spectrum or amplitude of speech is within the range of 50-10000 Hz (14, 2). The human voices fundamental frequency F_0 and first formant F_1 are clustered around 100 and 600 Hz. Formants are a concentration of acoustic energy around a particular frequency in a speech wave. These distinguishing frequencies are components of human speech. Each vowel has a distinct formant pattern. Each person’s voice has unique formant qualities. Appling knowledge of speech science to the study of voice authentication has led to a better understanding of both.

II.2 RESEARCH PLAN

The goal of this research was to devise a method to choose between different voice authentication algorithms. The method was displayed by using voice authentication. This method uses different algorithms to find which one will be the best for a given situation.

The first step was to capture the sound files using a microphone and a laptop. After the files were collected the second step was to perform spectrum analysis on the data samples. The spectrum analysis allowed a comparison of the amplitudes versus frequencies of the data files. Examples of spectrum relying algorithms are n-valleys, n-peaks, pattern of valleys and peaks and n-valleys and peaks. Examples of non-spectrum algorithms are file length and file size.

Once the spectrum analysis was created for the all the data samples the samples were put into frequency 'buckets'. A 'bucket' used was 0-20Hz created for all samples from 0-11000Hz. The 'bucket' correlates to the amplitude sums at those frequencies. The 20 Hz size of 'buckets' allowed for a more manageable amount of data. After the 'buckets' were created the sums of the amplitudes were compared with the chosen algorithms. For the initial test of the evaluation method n-peaks and n-peaks using n-valley exclusion was used. After the test occurred then the pass fail rates can be compared. The comparison will display how adjusting the threshold can make the most effective match for the given application.

CHAPTER III

FINDINGS

To achieve the goal set forth by this research a database of recordings was required. There were 48 sound files used in the voice authentication systems. There were several different algorithms chosen as well. The reason different algorithms were used was to show that the evaluation method can be used on any voice authentication algorithm.

To show how the evaluation method works in practical application, voice samples were created. The voice samples were then converted from time domain to frequency domain using fast Fourier transforms. After the conversion the samples were compared and pass/fail rates were generated. If an organization asks for their algorithm to be tested the pass/fail rates can be used to determine if an algorithm is adequate for system access (CMR and FMR) versus system denial (CNMR and FNMR).

Figure 11 is an example of the raw sound wave displayed in Matlab. The LEN1 designates the file as Leslie E. Nelson stating her name. The appendix has a table that designates the file names with the person recorded and what they are saying. Once the samples were taken they had to be changed from time based domain to frequency based domain. The domain change was done using Matlab to do fast Fourier transforms (FFT) (15).

$$\int_{-\infty}^{\infty} f(x)e^{-i2\pi xs}\,dx = \int_{-\infty}^{\infty} F(s)e^{-i2\pi wd}\,ds$$

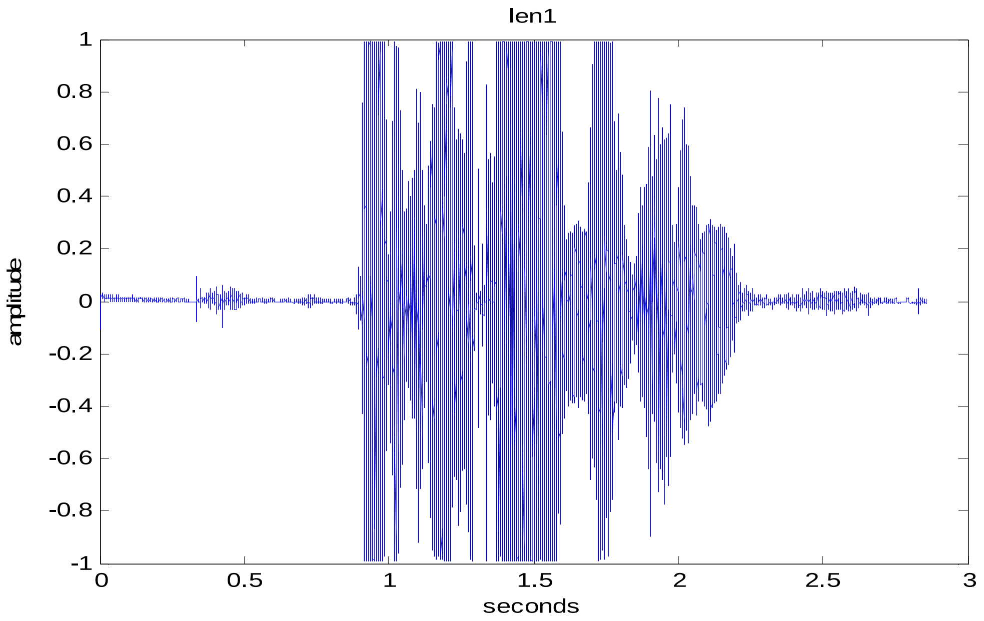


Figure 11 Wave File for Sample LEN1

The FFTs gave the data wave a numerical representation of amplitude at given frequencies shown in the below example of Figure 12. (The frequencies for the following FFT graphs are for 0-4000 Hz).

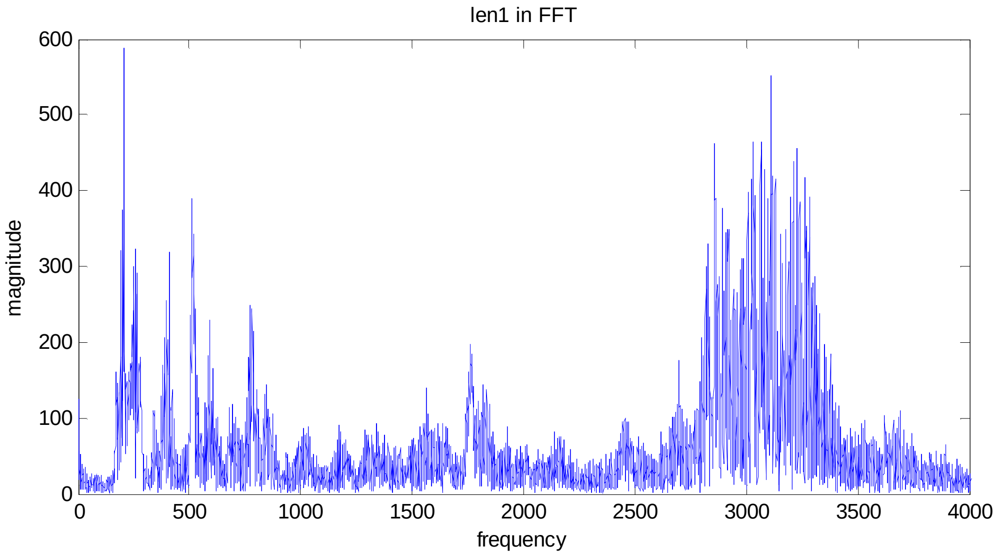


Figure 12 FFT Analysis of LEN1 to 4000 Hz

After the FFT were created the amplitudes observed at each frequency were small and there were 11000 different amplitudes per sample. To make the data manageable 'buckets' were created. These 'buckets' are amplitude sums for a given frequency range. Three frequency ranges were chosen 500 Hz, 100 Hz, and 20 Hz frequency ranges. Upon doing initial comparison testing and through research into the sound properties of voice 20 Hz frequency 'buckets' were chosen. The 20 Hz frequency 'buckets' had better matching capabilities, and yet were still of a manageable amount of data. See a transformation example below in Figure 13.

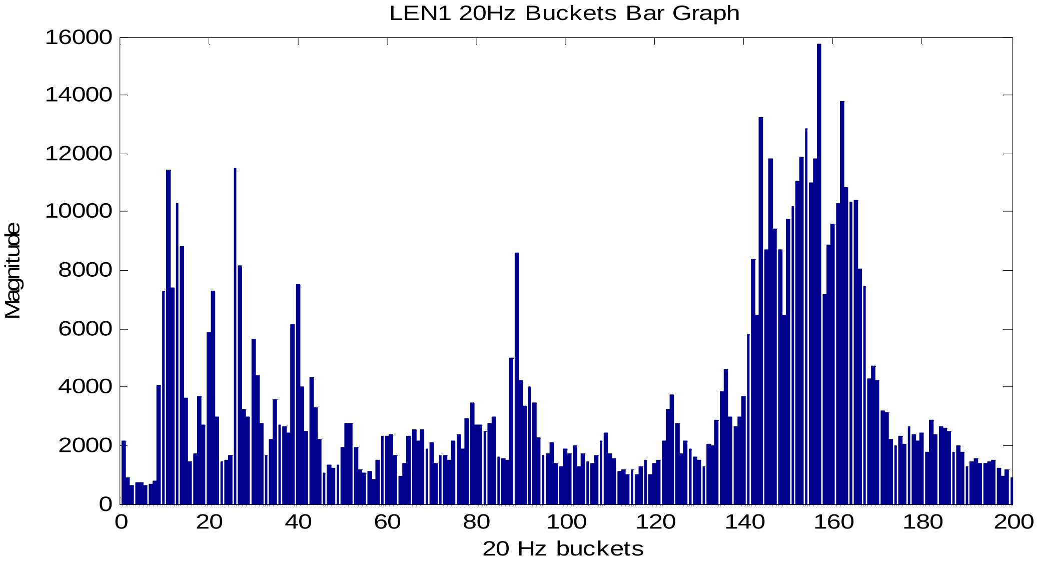


Figure 13 FFT Summarize in 20 Hz 'Buckets' for LEN1

Figure 13 represents the frequencies from 0 to 4000 Hz so there are two hundred 20 Hz 'buckets'. The name files represent text independent files and the greeting files represent text dependent files. These files are listed with the database files in the appendix.

Once all forty-eight (48) files were converted, the 20 Hz 'bucket' sums were taken from Matlab to excel to be compared. The 20 Hz 'buckets' are frequency values such as 21-40 Hz are represented with the value 40 in the 20 Hz 'buckets' HZ frequency column. The BK2-20 Hz ‘bucket’ sum is the amplitude values of the 21-40 Hz frequencies in the previous column for the BK-2 sound file valued at 6170.831. Below shows a truncated example of the excel file Table 8:

20 HZ buckets HZ frequency	BK2-20 HZ bucket sum	20 HZ buckets HZ frequency	JMA2-20 HZ bucket sum	20 HZ buckets HZ frequency	LEN2-20 HZ bucket sum
20	2891.437	20	5560.536	20	1374.166
40	6170.831	40	8177.412	40	355.7565
60	6904.46	60	9936.122	60	226.5954
80	6744.04	80	4941.406	80	180.675
100	2620.522	100	5253.837	100	121.6085

Table 8 Truncated English 20 Hz ‘bucket’ valves

III.1 ENGLISH SAMPLE FILES

The main body of analysis was done upon the English greeting files. The English greeting spreadsheet is for the English greeting, and the spreadsheet has 17 voice sample files with frequencies from 20 – 11000 Hz ‘buckets’. The next step was to compare maximum amplitudes at given frequency ‘buckets’. The range selected to make the comparisons was 40-2000Hz. The 40-2000Hz frequencies were put onto the next spread sheet and sorted into maximum amplitudes. During initial trials the top 5, 10, and 14 values were chosen to make comparisons. The maximum 14-peak amplitudes were selected for the comparisons because more amplitude values gave better results in initial testing. See truncated English greeting example in Table 9. Teal is a match between BK2 and JMA2. One of the values is 420 Hz. Green is a match between JMA2 and LEN2. One of the values is 440 Hz. Purple is a match between BK2 and LEN2. One of the values is 220 Hz. **Bold yellow** is a match between all three files. One of the values is 400 Hz.

20 HZ buckets HZ frequency	BK2-20 HZ bucket sum	20 HZ buckets HZ frequency	JMA2-20 HZ bucket sum	20 HZ buckets HZ frequency	LEN2-20 HZ bucket sum
420	18296.63	380	10737.49	220	8683.922
220	17072.63	60	9936.122	200	6530.573
600	16351.47	**400**	8337.78	800	6241.05
480	15579.07	40	8177.412	440	5467.849
240	14354.41	200	7130.427	1000	4942.6
400	11732.45	360	6759.238	300	4818.196
620	11038.11	140	6355.332	980	4568.269
200	10958.82	420	5921.535	280	4488.297
580	10608.41	460	5580.642	1180	4457.795
820	8894.832	440	5286.589	240	4305.509
60	6904.46	100	5253.837	820	4199.796
840	6901.262	480	5020.54	780	4196.163
80	6744.04	80	4941.406	**400**	3832.825
460	6604.634	180	4837.798	960	3624.737

Table 9 Maximum amplitudes top 14

Also the minimum 50%-valley amplitudes were placed in a table. See below for the truncated English greeting Table 10. The first column second row is range of 1101 – 1120 Hz. The second column second row value is 393.8151 which are the sum of amplitudes for the 1101 – 1120 Hz range. The lowest amplitude value is for the BK2 file.

20 HZ buckets HZ frequency	BK2-20 HZ bucket sum	20 HZ buckets HZ frequency	JMA2-20 HZ bucket sum	20 HZ buckets HZ frequency	LEN2-20 HZ bucket sum
1120	392.8151	1600	462.4067	100	121.6085
900	463.5154	1620	490.4632	120	122.3566
1500	489.7188	1560	642.3854	140	137.9888
1300	620.5594	1640	655.4751	160	152.8198
1480	654.0862	800	663.5777	1640	174.8138
1320	667.7239	1860	680.7095	80	180.675
1980	707.4544	820	684.9521	1620	184.1732
1600	728.4031	1740	688.7146	60	226.5954
1520	741.3775	1760	706.438	1460	242.455
1140	761.9576	840	742.9762	1240	262.1259
1100	767.1722	1580	800.8837	1280	265.2008
1680	768.0386	1480	869.8219	1440	269.8197
1340	779.6717	1500	885.6831	1260	272.6457
1820	799.5634	1520	896.0801	1660	279.6975

Table 10 Minimum amplitudes bottom 14

After the 14-peak maximum and 50 percent valley minimum amplitudes were found, two additional spreadsheets were used. The spreadsheet for the maximum amplitudes has columns of frequency and columns with a 1 assigned when the amplitude at that frequency ‘bucket’ was equal to or higher than the smallest top 14-peak amplitude sum (Table 11 shows a truncated table for the English greeting). The code used for the spreadsheet was automated to allow for the n to be adjusted from 14 for future ease of use.   The same technique was used to create the minimum amplitude spread sheet, but the spreadsheet used the minimum 50 percent of the amplitudes.

20 HZ buckets HZ frequency	BK2-20 HZ bucket sum	20 HZ buckets HZ frequency	JMA2-20 HZ bucket sum	20 HZ buckets HZ frequency	LEN2-20 HZ bucket sum
40	0	40	1	40	0
60	1	60	1	60	0
80	1	80	1	80	0
100	0	100	1	100	0
120	0	120	0	120	0
140	0	140	1	140	0
160	0	160	0	160	0
180	0	180	1	180	0
200	1	200	1	200	1
220	1	220	0	220	1
240	1	240	0	240	1

Table 11 Automated Sheet for Maximum Frequencies

After the columns in Table 11 were created the following table incorporated a formula that was used to compare how many frequencies had ones in both columns. The sum products were put into Table 12. Five thresholds were selected based on initial research. Threshold 6, 5, 4, 3 and 2 were used. When the sum product was equal to the threshold or higher the sum was highlighted in yellow. The boxed portions in Table 12 were when the amplitudes were for the same person, but a different sound file. An example of multiple files for the same person is six sound files by Dr. Acken and six sound files by Leslie Nelson. Since these files are by the same people then by theory they should match.

Threshold 6 values

	bk2	jma2	len2	jl2	et2	mdc2	na2	jma2b	jma2c	jma2d	jma2e	jma2f	len2b	len2c
bk2														
jma2	7													
len2	5	3												
jl2	2	2	3											
et2	6	9	3	5										
mdc2	5	5	3	3	5									

na2	3	4	2	4	4	4								
jma2b	6	9	4	2	6	4	5							
jma2c	4	8	2	1	7	4	4	10						
jma2d	4	6	2	3	5	5	4	5	4					
jma2e	5	1	5	4	4	5	5	4	2	4				
jma2f	5	10	4	4	8	4	7	9	7	7	5			
len2b	2	2	4	2	4	3	3	5	5	4	7	5		
len2c	3	0	3	3	1	3	2	1	0	1	5	2	4	
len2d	1	0	2	1	2	1	1	0	1	2	4	1	5	4
len2e	2	0	3	2	2	1	1	0	2	0	4	1	4	6
len2f	8	7	6	4	6	5	4	6	4	5	5	6	3	2

Table 12 Comparisons for a threshold of 6

Below are the pass/fail rates Table 13 and Figure 14 for all the max files using the above process. The boxed portions in table 12 are represented in the CMR (the yellow ones in the boxes) and the FNMR (the non-yellow ones in the boxes).

English pass fail rates:

Threshold 6	CMR	28/41	68.3	CNMR	93/112	83.0	FMR	13/41	31.7	FNMR	19/112	17.0
Threshold 5	CMR	29/72	40.3	CNMR	65/81	80.2	FMR	43/72	59.7	FNMR	16/81	19.8
Threshold 4	CMR	38/93	40.9	CNMR	51/60	85.0	FMR	55/93	59.1	FNMR	9/60	15.0
Threshold 3	CMR	41/111	36.9	CNMR	36/42	85.7	FMR	70/111	63.1	FNMR	6/42	14.3
Threshold 2	CMR	45/131	34.4	CNMR	20/22	90.9	FMR	86/132	65.6	FNMR	2/22	9.1

Table 13 Pass/Fail Rates for English Maximum Algorithm

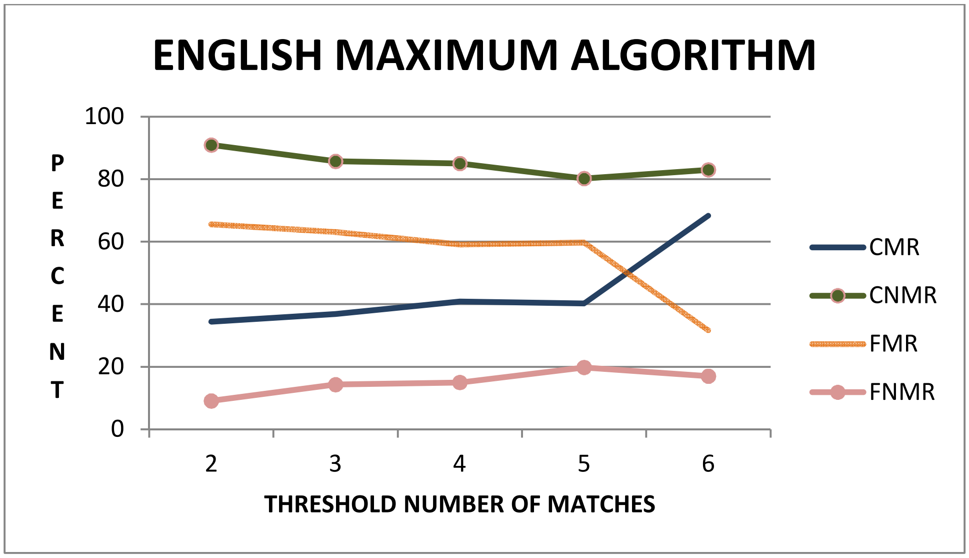


Figure 14 Graph of Pass/Fail Rates for English Maximum Algorithm

When the threshold value is lower the match rates are higher, but the correct match rate is lower and the correct non-match rate is higher. The lower threshold value allows for many people to enter the system and rejects very few. This could be used to allow assess into a system such as a convenience store. The median threshold can be used when you have an application for medium security such as an office building. The higher threshold would be where you would apply higher level security such as a nuclear power plant. The thresholds used make a strong difference in this algorithm because the CMR and FMR rate lines cross. The rates trend as expected.

To investigate the rejection rate of imposters another aspect was added using the minimum amplitudes in Table 14. Adding the minimum amplitudes decreased the amount of matches. A truncated table is shown below.

20 HZ buckets HZ frequency	BK2-20 HZ bucket sum	20 HZ buckets HZ frequency	JMA2-20 HZ bucket sum	20 HZ buckets HZ frequency	LEN2-20 HZ bucket sum
40	0	40	0	40	0
60	0	60	0	60	1
80	0	80	0	80	1
100	0	100	0	100	1
120	0	120	0	120	1
140	0	140	0	140	1
160	0	160	0	160	1
180	0	180	0	180	0

Table 14 Automated Sheet for Minimum Frequencies

Threshold minimum 4 was used to display the research method because threshold 4 was the median of the maximum thresholds. The truncated table below shows how the adjustments were made in Table 15.

Threshold 4 value match vs. 4 max matches

	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
	bk2	bk2	jma2	jma2	len2	len2	jl2	jl2	et2	et2
bk2										
jma2	0	7								
len2	0	5	2	3						
jl2	2	2	1	2	3	3				
et2	1	6	0	9	3	3	2	5		
mdc2	0	5	0	5	1	3	3	3	***4***	***5***
na2	1	3	0	4	2	2	2	4	2	4
jma2b	0	6	0	9	***6***	***4***	5	2	1	6
jma2c	1	4	0	8	5	2	4	1	0	7

Table 15 Comparisons for a Max Threshold of 4 and Min Threshold of 4

The min columns are number of minimum amplitudes (for the column file) that also appear in the maximums of the row file. The maximum columns are the number of maximum amplitude matches between the two files. When the thresholds are met for both columns then the minimum match is considered a cancellation of a maximum match. The changes for the match rates for the English greeting are listed below.

English pass fail rates with the valley exclusion using 4 minimum to maximum amplitude cancellations are shown in Table 16 and Figure 15:

Threshold 6	CMR	27/40	65.0	CNMR	93/113	83.2	FMR	14/40	35.0	FNMR	19/113	16.8
Threshold 5	CMR	27/60	45.0	CNMR	75/93	80.6	FMR	33/60	55.0	FNMR	18/93	19.4
Threshold 4	CMR	31/62	50.0	CNMR	75/91	82.4	FMR	31/62	50.0	FNMR	16/91	17.6
Threshold 3	CMR	33/64	51.6	CNMR	75/89	84.3	FMR	31/64	48.4	FNMR	14/89	15.7
Threshold 2	CMR	34/64	53.1	CNMR	76/89	85.4	FMR	30/64	46.9	FNMR	13/89	14.6

Table 16 Pass/Fail Rates for English Minimum Algorithm

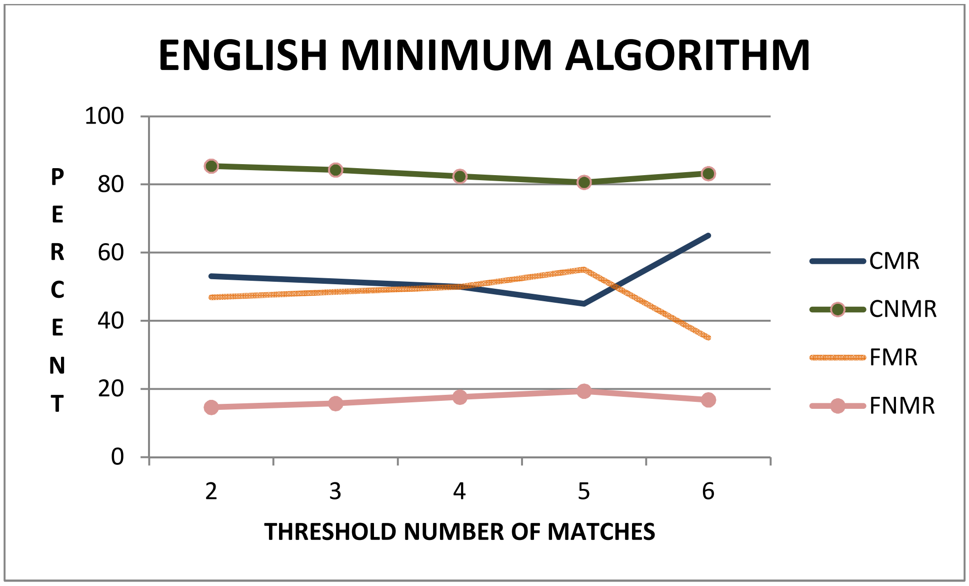


Figure 15 Graph of Pass/Fail Rates for English Minimum Algorithm

Using the minimum matches to cancel the maximum matches lowered the matching rate of all scenarios. The rates trending had some unexpected results expected. The CMR increased for threshold 2, 3 and 5 and decreased for 4 and 6, while the CNMR decreased for threshold 2, 3 and 5 and decreased for 4 and 6. The trends were expected to adjust the same.

III.2 NAME PASS/FAIL RATES

The text independent pass/fail rates for the name files are in Table 17. There were 10 sample files. One was John M. Acken saying Leslie E. Nelson's name. The rest of the files are people saying their own names as detailed in the appendix.

Name pass fail rates shown in Table 17 and Figure 16:

Threshold 6	CMR	11/17	65.0	CNMR	37/38	97.4	FMR	6/17	35.0	FNMR	1/38	2.6
Threshold 5	CMR	12/23	52.2	CNMR	32/32	100.0	FMR	11/23	47.8	FNMR	0/32	0.0
Threshold 4	CMR	12/31	38.7	CNMR	24/24	100.0	FMR	19/31	61.3	FNMR	0/24	0.0
Threshold 3	CMR	25/37	67.6	CNMR	18/18	100.0	FMR	12/37	32.4	FNMR	0/18	0.0
Threshold 2	CMR	12/49	24.5	CNMR	6/6	100.0	FMR	37/49	75.5	FNMR	0/6	0.0

Table 17 Pass/Fail Rates for Name Algorithm

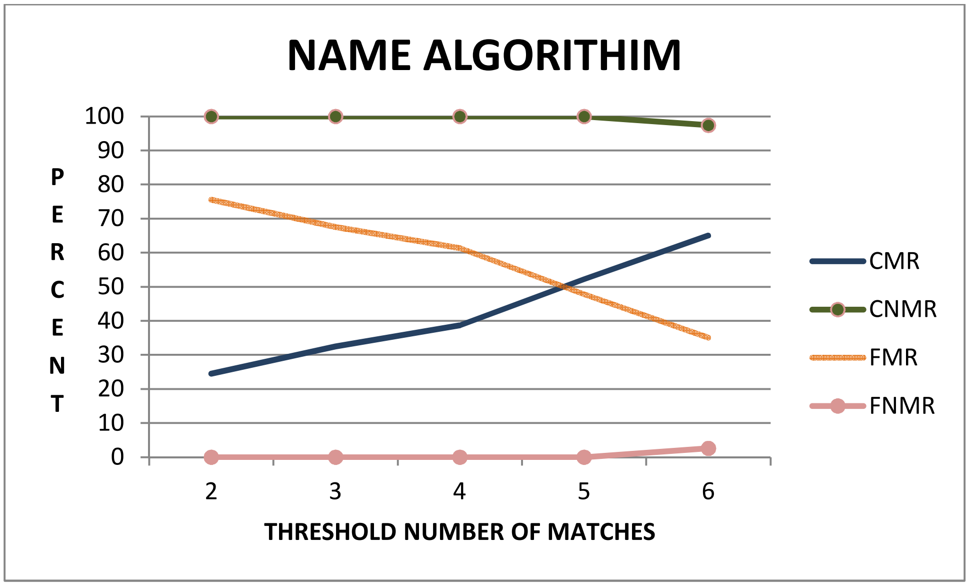


Figure 16 Graph of Pass/Fail Rates for Name Algorithm

As the threshold value increases the number of matches decrease, but the CMR increases. The name algorithm trends as expected. The threshold of 6 and 5 matches show a trend for higher correct matching and lower false matching. The trend is an example that the Name algorithm can be used at a threshold of 6 to allow a better CMR.

III.3 OTHER LANGUAGE GREETINGS PASS/FAIL RATES

The other language greetings (Spanish, Russian and Farsi) only had 7 sample files for each language. The native tongue of the speakers is the one highlighted yellow in the appendix table.

Spanish pass fail rates shown in Table 18 and Figure 17:

Threshold 6	CMR	7/16	43.7	CNMR	12/12	100.0	FMR	9/16	56.3	FNMR	0/12	0.0
Threshold 5	CMR	7/18	39.0	CNMR	9/9	100.0	FMR	11/18	61.0	FNMR	0/9	0.0
Threshold 4	CMR	7/20	35.0	CNMR	8/8	100.0	FMR	13/20	65.0	FNMR	0/8	0.0
Threshold 3	CMR	7/25	28.0	CNMR	3/3	100.0	FMR	18/25	72.0	FNMR	0/3	0.0
Threshold 2	CMR	7/27	26.0	CNMR	1/1	100.0	FMR	20/27	74.0	FNMR	0/1	0.0

Table 18 Pass/Fail Rates for Spanish Greeting Algorithm

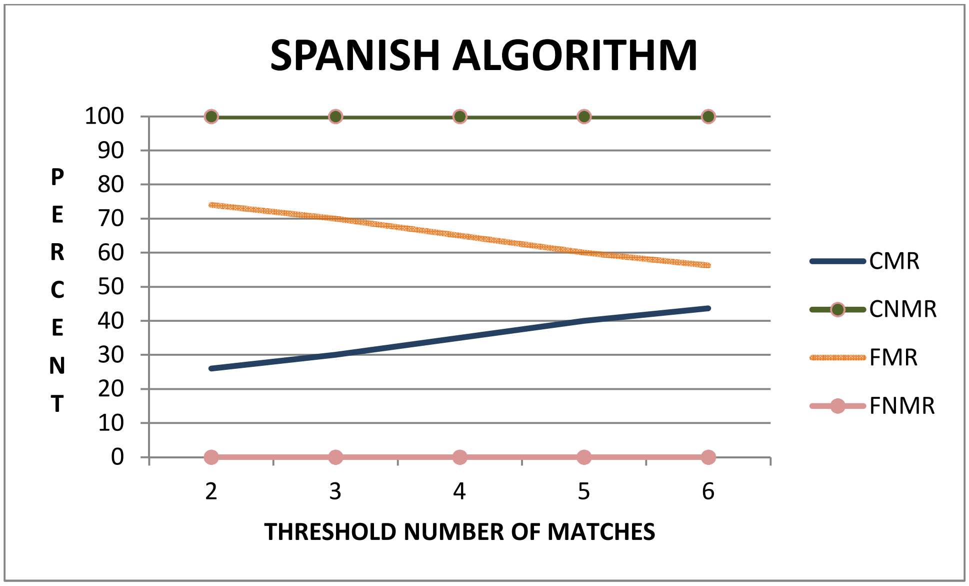


Figure 17 Graph of Pass/Fail Rates for Spanish Algorithm

For the Spanish algorithm voice authentication system if a high reward is desired then having a threshold is 2 is desired. The combined value of the matches (correct matches (CM) and false matches (FM)) is higher. If the high reward is to block imposters then the ideal threshold is 6. The combined value of the false match (correct non-matches (CNM) and false non-matches (FNM)) is higher.

Russian pass fail rates shown in Table 19 and Figure 18:

Threshold 6	CMR	7/16	43.7	CNMR	12/12	100.0	FMR	9/16	56.3	FNMR	0/12	0.0
Threshold 5	CMR	7/17	41.0	CNMR	11/11	100.0	FMR	10/17	59.0	FNMR	0/11	0.0
Threshold 4	CMR	7/23	30.0	CNMR	5/5	100.0	FMR	16/23	70.0	FNMR	0/5	0.0
Threshold 3	CMR	7/25	28.0	CNMR	3/3	100.0	FMR	18/25	72.0	FNMR	0/3	0.0
Threshold 2	CMR	7/27	26.0	CNMR	1/1	100.0	FMR	20/27	74.0	FNMR	0/1	0.0

Table 19 Pass/Fail Rates for Russian Greeting Algorithm

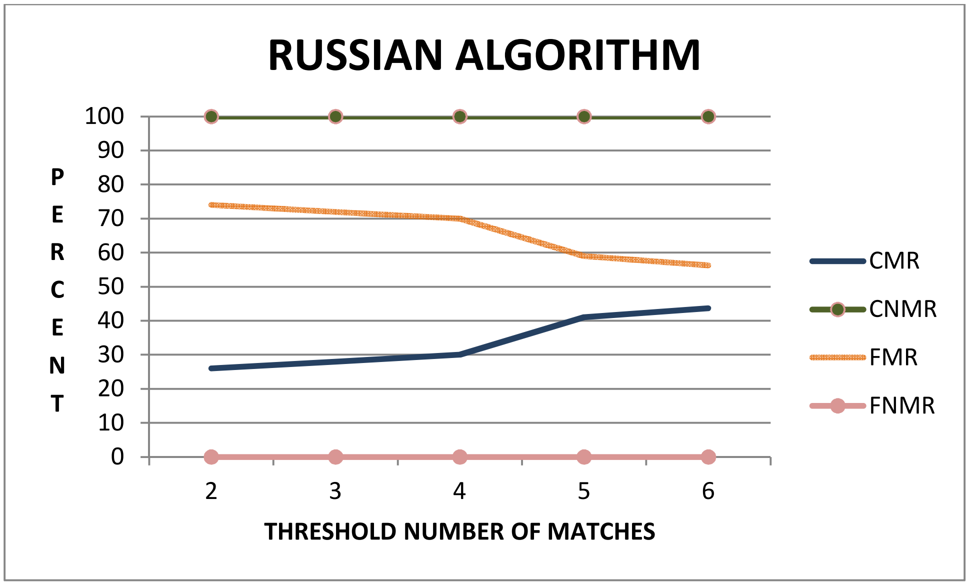


Figure 18 Graph of Pass/Fail Rates for Russian Algorithm

The way the Russian algorithm voice authentication system was implemented there cannot be a display of high reward for false non-matches.

Farsi pass fail rates shown in Table 20 and Figure 19:

Threshold 6	CMR	7/16	43.7	CNMR	12/12	100.0	FMR	9/16	56.3	FNMR	0/12	0.0
Threshold 5	CMR	7/19	37.0	CNMR	9/9	100.0	FMR	12/19	63.0	FNMR	0/9	0.0
Threshold 4	CMR	7/20	35.0	CNMR	7/7	100.0	FMR	13/20	65.0	FNMR	0/7	0.0
Threshold 3	CMR	7/22	32.0	CNMR	6/6	100.0	FMR	15/22	68.0	FNMR	0/6	0.0
Threshold 2	CMR	7/27	26.0	CNMR	1/1	100.0	FMR	20/27	74.0	FNMR	0/1	0.0

Table 20 Pass/Fail Rates for Farsi Greeting Algorithm

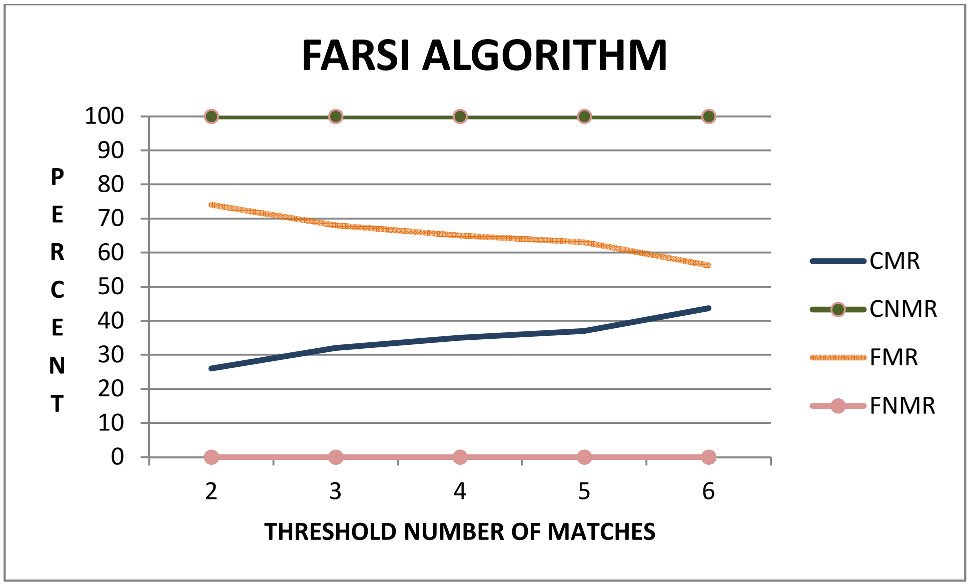


Figure 19 Graph of Pass/Fail Rates for Farsi Algorithm

For the Farsi algorithm voice authentication system there is higher reward for CMR when the threshold is 6 than the other thresholds used.

All three of the of these language trend as expected. When the threshold was decreased the match became larger, but the CMR decreased for all three languages. The numbers are very similar between the three languages. The English language files have better CMR than these do. The better CMR could be because there were more English speaking people in the study, or because there were more samples available.

III.4 TEXT INDEPENDENT GREETING PASS/FAIL RATES

Text independent greeting pass/fail comparison is between the maximum values of all English, Spanish Russian and Farsi greetings. The text independent greeting rates are shown in Table 21 and Figure 20:

Threshold 6	CMR	84/307	27.4	CNMR	288/434	89.4	FMR	223/307	72.6	FNMR	46/434	10.6
Threshold 5	CMR	101/404	25.0	CNMR	307/337	91.1	FMR	303/404	75.0	FNMR	30/337	8.9
Threshold 4	CMR	114/511	22.3	CNMR	224/230	97.4	FMR	397/511	77.7	FNMR	6/230	2.6
Threshold 3	CMR	107/594	18.0	CNMR	134/147	91.2	FMR	487/594	82.0	FNMR	13/137	8.8
Threshold 2	CMR	116/668	17.4	CNMR	66/73	90.4	FMR	552/668	82.6	FNMR	7/73	9.6

Table 21 Pass/Fail Rates for Text Independent Greetings Algorithm

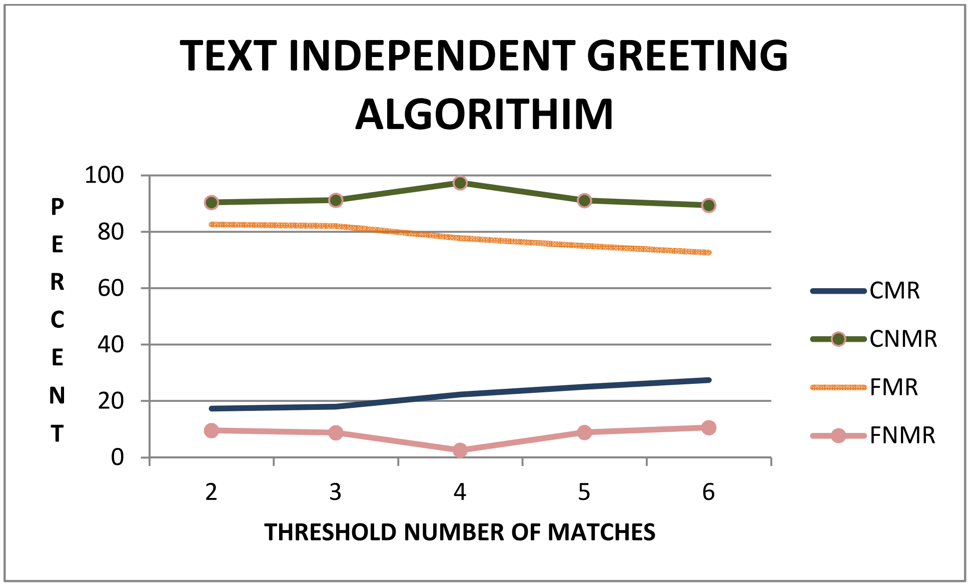


Figure 20 Graph of Pass/Fail Rates for Text Independent Greeting Algorithm

The text independent algorithm trends as expected for the CMR and FMR. The correct non-match rate has an unexpected trend. The English greeting has better CMR values but worse CNMR values. The rates may indicate that a system using a native tongue may be more secure. This observation is the opposite of what was expected.

III.5 TIME PASS/FAIL RATES

Comparing the length of time a person uses to say the English greeting was also used to evaluate the time algorithm. While the time algorithm was not expected to generate good comparisons the time algorithm does show that the method will work even when the algorithm is less than ideal. The 1 sec, .5 sec and .25 sec are the difference of time between the sound files. The time pass/fail English rates are shown in Table 22 and Figure 21:

t =< 1 SEC	CMR	48/119	40.3	CNMR	30/34	88.3	FMR	71/119	59.7	FNMR	4/34	11.7
t =< .5 SEC	CMR	29/67	43.3	CNMR	69/86	80.3	FMR	38/67	56.7	FNMR	17/86	19.7
t =< .25 SEC	CMR	26/54	48.2	CNMR	78/99	78.8	FMR	28/54	51.8	FNMR	21/99	21.2

Table 22 Pass/Fail Rates for Time Algorithm

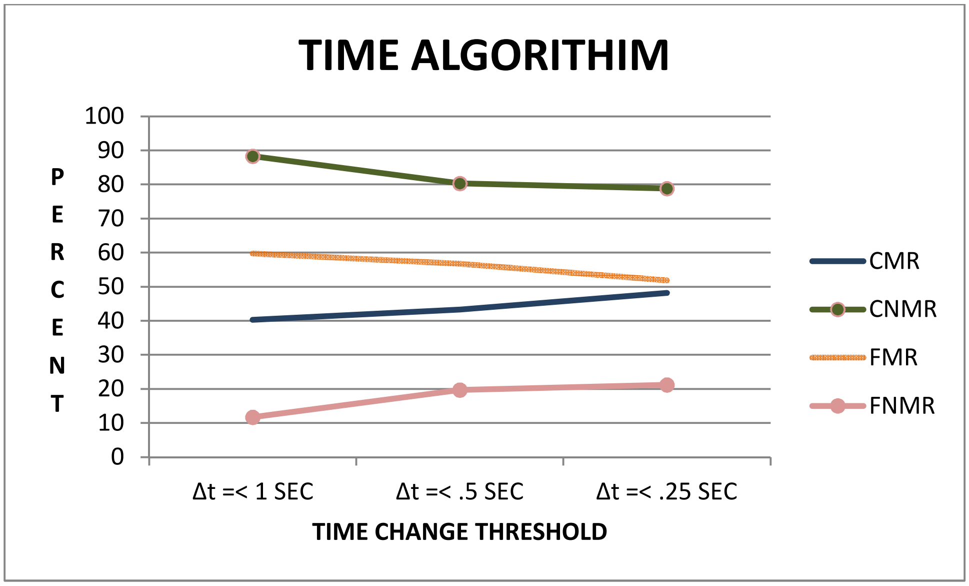


Figure 21 Graph of Pass/Fail Rates for Time Algorithm

The time trends were as expected. The change in time is from 1 second to .25 seconds to show the trend from least stringent matching to most stringent as trending in all preceding algorithms. When time increased the matching numbers went up, but the CMR decreased. As anticipated time is not the best algorithm to choose for matching. The time algorithm shows how the evaluation method can still work even for a less than ideal algorithm.

III.6 RELATIVE AMPLITUDE PASS/FAIL RATES

The relative amplitude rates are used to compare the maximum amplitude to the average amplitude of the English greeting. The threshold comparison of 1, 2 and 3 below are the rates of the relative amplitude comparison.

English relative amplitude pass/fail rates shown in Table 23 and Figure 22:

1	CMR	26/69	37.7	CNMR	63/84	75.0	FMR	43/69	62.3	FNMR	21/84	25.0
2	CMR	38/110	34.5	CNMR	34/43	79.1	FMR	72/110	65.5	FNMR	9/43	20.9
3	CMR	41/130	31.5	CNMR	17/23	73.9	FMR	89/130	68.5	FNMR	6/23	26.1

Table 23 Pass/Fail Rates for Relative Amplitude

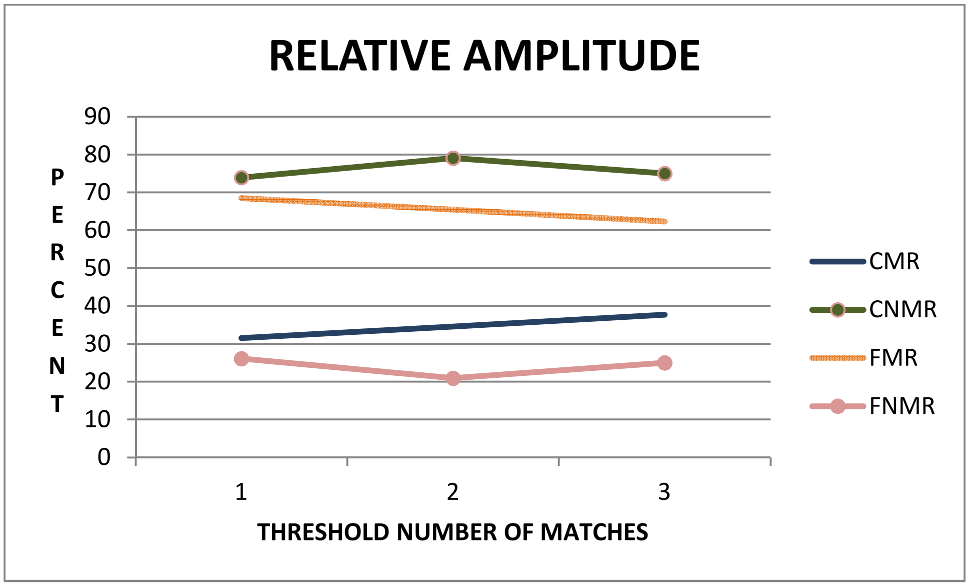


Figure 22 Graph of Pass/Fail Rates for Relative Amplitude

Relative amplitude has an anomaly with the non-match rates. The non-match rates are not straight lines as expected. When the threshold is increased the CMR went up as anticipated.

III.7 OBSERVATION

Observations of the different pass/fail rates are as follows. Observation one, the Name (text independent) files were superior for CMNR compared to the English greeting. Observation two, the all language greeting comparison (text independent) had poorer use of CMR, yet a slightly better CNMR, also the threshold of 4 was higher for CNMR that the threshold of 2 (this observation was not expected). Observation three, the English MIN decreased the matches for all five thresholds and improved the CMR of thresholds 5, 4, 3 and 2 (this observation is as expected) and decreased the CM for all five thresholds (also as expected). Observation four, the time and relative amplitude algorithms demonstrate that the method can be used for non-traditional comparisons as well. Observation five, using a language that is not a native tongue did not produce a better CMR.

III.8 APPLICATIONS

Four sample applications have been chosen to demonstrate the method developed in this thesis. These sample applications will describe the steps a company can use to either optimize their current voice authentication system or chose between two or more systems. The pass/fail rates of the English greeting maximum algorithm, figure 14, and name algorithm, figure 16, will be used to give theoretical suggestions to the applications.

The first sample application is for a convenient store. The owner has chosen to use a voice authentication system on the alarm keypad after several break-ins. The owner is considering two different voice authentication systems to be added to the new keypad lock. This method can help with the decision using the following steps. Step one is to create a database of sound files with the employees speaking a predetermined phrase at least two times. Step two is to convert the sound files to mathematical files; this can be done with Fast Fourier Transforms. Step three is to use the two systems to generate pass/fail rates using several thresholds. Step four is to evaluate between the pass/fail rates which work the best for store access. There are two things that must be considered when choosing the rate for this application. The first consideration is that ideally all employees are granted access; this must be weighed with the second consideration keeping out thieves from the store. The owner must decide how many times they want to be interrupted by an employee because they cannot enter the store versus a thief getting into the store. Using the English greeting maximum algorithm the suggested threshold is 2.

The second sample application is for an office building. The owner an office building is considering upgrading a voice authentication system on the side doors to the complex. This method can help with the decision using the following steps. Step one is to create a database of sound files with the employees speaking a predetermined phrase at least two times. Step two is to convert the sound files to mathematical files; this can be done with Fast Fourier Transforms. Step three is to use the system to generate pass/fail rates using several thresholds. Step four is to evaluate between the pass/fail rates which work the best for store access. The building owner must weigh building access with risk of a hostile or thieving person entering the complex through the side doors. Using the English greeting maximum algorithm the suggested threshold is 3.

The third sample application is for a bank. The board of directors has chosen to use a voice authentication system for their customers when using the ATM and their employees when accessing the bank's computers. The two VASs have very different goals so the steps for each will be looked at independently. This method can help with the decision for the customer system using the following steps. Step one is to create a database of sound files with the customers speaking a predetermined phrase at least two times. Step two is to convert the sound files to mathematical files; this can be done with Fast Fourier Transforms. Step three is to use the system to generate pass/fail rates using several thresholds. Step four is to evaluate between the pass/fail rates which work the best for ATM access. The bank must weigh alienating customers (especially very rich ones) by denying them the money versus allowing a thief to steal the customer's money. For employee computer access this method can help with what to set the algorithm threshold: Step one is to create a database of sound files with the employees speaking a predetermined phrase at least two times. Step two is to convert the sound files to mathematical files; this can be done with Fast Fourier Transforms. Step three is to generate pass/fail rates using several thresholds. Step four is to evaluate between the pass/fail rates which work the best for computer access. The bank should consider a higher CMR for employee computer versus the customer account access. This reason for this consideration is because while a person may leave a bank over not getting their money an employee is less likely quit because they may have to take extra steps to log onto their computer. Also, the risk of a thief access a bank computer is higher than one ATM account. Using the English greeting maximum algorithm the suggested threshold is of the customers is 4 and the suggested threshold for the employees is 6.

The fourth sample application is for a nuclear plant. The plant manager has chosen to use a voice authentication system at the gate to verify employee identity. The plant manager is considering want to verify employee identity and upon failure compare the voiceprint to known terrorist. This method can help with the decision using the following steps. Step one is to create a database of sound files with the employees speaking a predetermined phrase at least two times. Step two is to work with government agencies to create a terrorist sound file database. Due to the nature of this the second comparison will most likely be text independent. Step three is to convert the both database sound files to mathematical files; this can be done with Fast Fourier Transforms. Step four is to use the system to generate pass/fail rates using several thresholds. Step five is to evaluate between the pass/fail rates which work the best for the nuclear plant access. There is also how effective the algorithm is at identifying the simulated terrorist accessing the system. Using the English greeting maximum algorithm the suggested threshold for the employees is 6 and using the name algorithm for the terrorist check the suggested threshold is 4.

CHAPTER IV

SUMMARY

Security is necessary in our complex world. People rely on security measures to make positive identifications of individuals. To establish identity at least one of the three components is needed. What a person has, what a person knows or what a person is.

The players that are legitimate communicators are Alice, Bob, Carol and Dave. Persons that access the system and should not be there are Eve, Imelda, Mallory and Randy. The persons that facilitate communication and establish identity are Peggy, Randy, Trent and Walter. These players can be players in a biometric system.

Biometrics (what a person is) is the component used in this thesis. Biometrics are measurable characteristics of an individual that can be used for identification. Voice authentication is a human biometric. Voice authentication was chosen for personal reasons as well as how many parameters of voice authentication were met favorably.

The parameters of biometrics are cost, time, universality, distinctiveness, permanence, collectability and acceptability. Voice authentication has favorable parameters with cost, time, universality, collectability, acceptability, and availability of technology.

Current research is centered on improving algorithms. That is why this thesis focuses on selecting existing ones. This can make choosing an algorithm simpler. The method takes users and creates user sound file database. The files are converted using FFT to allow comparisons.

The algorithm(s) under test is used to generate a pass/fail rate with several thresholds. The pass/fail rates are used to select the best solution to the application.

To demonstrate this method for evaluating voice authentication algorithms, voice files were collected. A voice capturing system was used to generate 48 sound files. The sound files are seven people saying their name, speaking an English greeting, a Spanish greeting, a Russian greeting and a Farsi greeting. There are twelve extra files of Dr. John Acken and Leslie Nelson saying their names and the English greeting, or impersonating the other person. After the sound files were created, they needed to be mathematically analyzed.

These sound files were transformed from time domain to frequency domain using Fast Fourier Transforms (FFT). The transformation was so they would have numerical values. The amplitude versus frequency waves were compared to generate pass/fail rates. Different algorithms were used in this research.

The initial analysis was of a 14-peak maximum amplitude comparison algorithm. Addition algorithms were also chosen: Maximum amplitude with minimum/maximum amplitude match cancelation, time taken to speak the phrase, and ratio of maximum versus the average amplitude. These different algorithms have different matching and non-matching values. The adjustment of the threshold changes how many matches are made. When the matches and non-matches change the pass/fail rates also change.

CONCLUSIONS

Voice authentication was used as the biometric due to personal interest as well as favorable parameters. The majority of the population can speak therefore the universality of voice authentication is high. If asked to speak into a microphone most people will comply. Voice authentication is more acceptable than some of the other biometrics, such as fingerprints. Due to the plethora of sound recording devices readily available voice authentication has an ease of collectability. Voice authentication should not be used alone to secure a system.

Research has proven that using multiple components of security over just using one is superior. To secure a bank transaction having a pin number and debit card, than just a debit card for security. Also a voice authentication system can be added to an automatic teller machine to help identify a user. Since digital biometrics in security is a budding industry there is an importance in being able to pick the best algorithms.

The demonstrated method takes into consideration the application of voice authentication algorithms. The security needs of a grocery store are significantly different than the needs of a nuclear power plant. The grocery store system should allow for a low rejection rate of authentic users (CMR). In contrast the nuclear power plant system would want high rejection rate for imposters (CNMR). The method provides the mechanics to evaluate voice authentication systems.

The devised method provides the mechanics for evaluating voice authentication algorithms. The method can evaluate even weak voice authentication systems such as length of time to say a phrase. The method can be applied to both text dependent and text independent algorithms.

When compared using the demonstrated method text independent and text dependent did not show a significant difference in it rates. This can be observed with the data from the English maximum algorithm when compared to the name and all greeting algorithms. The name algorithm is comparable with the CMR and the FMR, but the CNMR is better than the English maximum rates. The all language greeting has a very high FMR rate and does not show a much change when the threshold is varied.

A methodology to evaluate voice authentication algorithms has been created. The method is essential to select the algorithm and the threshold needed as well. This selection is made by considering the system that is being secured.

REFERENCES

1 Ferrand, T. Carole: Speech Science, Pearson Education Inc., Boston, MA, 2007.

2 Perkins, William H.; Kent, Raymond D.: Functional Anatomy of Speech, Language, and Hearing, A College Hill Publication, Boston, MA, 1996.

3 Woodward, John; Orleans, Nicholas; Higgins, Peter: Identity Assurance in the Information Age, Biometrics, Osborne, Andover, KS, 2003.

4 Nanavati, Samir; Thieme, Michael; Nanavati, Raj: Biometrics, Identity Verification in a Networked World, Wiley Computer Publishing, Hoboken, NJ, 2002.

5 Robert Grover Brown: Introduction to Random Signal Analysis and Kalman Filtering, John Wiley and Sons, New York, NY, 1983.

6 Jain, A. K.: An Introduction to Biometric Recognition, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 14, NO. 1, JANUARY 2004.

7 Schneier, Bruce: Applied Cryptography, John Wiley and Sons, Inc., New York, NY, 1996.

8 Pfleeger, Charles and Shari: Security in Computing, Pearson Education, Inc., New York, NY 2003.

9 Walpole, Ronald and Myers, Raymond: Probability and Statistics for Engineers and Scientists, The Macmillan Company, New York, NY, 1972.

10 Gattaca. Dir. Andrew Niccol. Perf. Ethan Hawke, Uma Thurman, and Jude Law. DVD. Colombia Picture, 1997. Shows examples of circumventing the system. One of the easiest circumventions is the hardest to prevent.

11 Band, Stephen R. PhD.; Cappelli, Dawn M.; Moore, Andrew P.; Fischer, Lynn F. PhD.; Shaw, Eric D. PhD. and Trzeciak Randall F.: Comparing Insider IT Sabotage and Espionage: A Model-Based Analysis, Carnegie Mellon University, Pittsburgh, PA, 2006.

12 Sohoni, Sohum; Shaver, Clark D.; Acken, John M.; Mertz, Doug; Nelson, Leslie E.; Remington, Justin; Sadr, Behnaz and Sundararajan, Gopal: Evaluation Criteria for Biometric Based Identity Authentication Systems, 2009.

13 Nazar, Muhammad Noman: Speaker Identification Using Cepstral Analysis, Student Conference 2002, ISCON '02. Proceedings, IEEE Volume 1, Aug. 16-17, 2002, page(s) 139-143.

14 Denes, Peter; Pinson, Elliot: The Speech Chain, W. H. Freeman and Company, New York, NY, 1993.

15 Bracewell, Ronald N.: The Fourier Transform and its Applications, McGraw-Hill, New York, NY, 1978.

16 Acken, John M. and Nelson, Leslie E.: Identity Authentication Systems for Web Based CAD, Computer –Aided Design and Applications, 5(1-4), 2008.

17 Acken, John M. and Nelson, Leslie E.: Security Based for Testing and Security of Digital Systems for Identity Authentication, CCCT, 2008.

18 Gates, Kelly A.: Our Biometric Future, New York University Press, New York, NY, 2011.

19 Kanade, Sanjay G., Petrovska-Delacretaz, Dijana, and Dorizzi, Bernadette: Enhancing Information Security and Privacy by Combining Biometrics with Cryptography, Synthesis Lectures of Information Security, Privacy, and Trust, June 2012, Vol. 3, 2012.

BIBLIOGRAPHY

1 Wilson,Scott, http://ccrma.stanford.edu/courses/422/projects/WaveFormat/, updated 2003. Gives a byte by byte detail of the heading of a PCM .wav file.

2 http://www.shareup.com/Frequency-Analyzer-download-7919.html, 2003-2004. Is the website where the program can be downloaded. More information will follow.

3 Zhang, David: Biometric Solutions For Authentication In An E-World, Kluwer Academic Publishers, 2002.

4 Lane, David, http://davidmlane.com/hyperstat/A18652.html , updated 4/14/2006. Give detail about hypothesis and matching statistics.

GLOSSARY

Authentic Person - The person who originally made the data sample.

Biometric – A physical characteristic of an individual.

Identification Characteristics – Physical characteristics such as face dimensions, voice pattern, or fingerprints. (8)

Individual – A distinct person.

Imposter – Person trying to impersonate the authentic person.

Test Data – The data sample being taken by the person under test to determine if that person is the authentic person or an imposter.

Person under test – The test subject that is trying to gain access to the system.

Test Database – collections of samples changes into algorithm test platforms taken by the authentic person.

Algorithm – Mathematical procedure for problem solving.

False Match Rate (FMR) – An imposter is allowed access to the system.

False Non-Match Rate (FNMR) – An authentic person is denied access to the system.

Voice Authentication – System in which a data sample from a known authentic person is created, transformed into a database by an algorithm, and then compared to a test subject to determine if the test subject is the authentic person or an imposter.

Voice Characteristics – Physical pattern that a voice has such as the amount of time it take to complete a phrase, the peaks and valleys of the pattern, and the frequency of the voice.

Pass rate – The rate at which the system correctly allows access to the authentic person and denies access to the imposter.

Fail rate – The rate at which the system incorrectly denies access to the authentic person and allows access to the imposter.

Threshold – A setting of the importance of allowing the authentic person access to the system and denying the imposter.

Capture Devices – Laptop computer with Microsoft windows XP, windows media player (with .wav file capability) and a microphone.

APPENDICES

	Name	English Greeting	Farsi Greeting	Spanish Greeting	Russian Greeting
Leslie E Nelson	LEN1 LEN1b	LEN2 LEN2b- LEN2f	LEN3	LEN4	LEN5
John M Acken	JMA1 JMAb JMAasLEN	JMA2 JMA2b- JMAf	JMA3	JMA4	JMA5
Berta Kadimov	BK1	BK2	BK3	BK4	BK5
Ebb Tesfidit	ET1	ET2	ET3	ET4	ET5
Jesus Lugo	JL1	JL2	JL3	JL4	JL5
Martin D Crossland	MDC1	MDC2	MDC3	MDC4	MDC5
Navid Amiliagona	NA1	NA2	NA3	NA4	NA5

Table 24 File Format for Data Samples

DATA Descriptions

Name of file – Person how spoke, file format, file sample rate, file bits/sample, mono speaker file, .WAV audio file, size of file in bytes, how long file is in seconds

1. BK1-Berta Kadamov stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 54684 samples (b), 2.48sec
2. BK2-Berta Kadamov greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 48510 samples (b), 2.2sec
3. BK3-Berta Kadamov greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 55125 samples (b), 2.5sec
4. BK4-Berta Kadamov greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 63945 samples (b), 2.9 sec
5. BK5-Berta Kadamov greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 58212 samples (b), 2.64 sec
6. ET1- Ebb Tesfidit stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 90405 samples (b), 4.1 sec
7. ET2- Ebb Tesfidit greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 62181 samples (b), 2.82 sec
8. ET3- Ebb Tesfidit greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 58212 samples (b), 2.64 sec
9. ET4- Ebb Tesfidit greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 63504 samples (b), 2.88 sec
10. ET5- Ebb Tesfidit greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 63063 samples (b), 2.86 sec
11. JL1- Jesus Lugo stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 50274 samples (b),2.282sec

12. JL2- Jesus Lugo greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 0131 samples (b), 1.82sec
13. JL3- Jesus Lugo greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 6746 samples (b), 2.12 sec
14. JL4- Jesus Lugo greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 63504 samples (b), 2.88 sec
15. JL5- Jesus Lugo greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 64386 samples (b), 2.92 sec
16. JMA1- Dr. John Acken stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 91287 samples (b), 4.14 sec
17. JMA2- Dr. John Acken greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 57330 samples (b), 2.6 sec
18. JMA3- Dr. John Acken greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 66591 samples (b), 3.02 sec
19. JMA4- Dr. John Acken greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 71001 samples (b), 3.22 sec
20. JMA5- Dr. John Acken greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 69678 samples (b), 3.16 sec
21. LEN1- Leslie Nelson stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 63063 samples (b), 2.86 sec
22. LEN2- Leslie Nelson greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 42777 samples (b), 1.94 sec
23. LEN3- Leslie Nelson greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 43659 samples (b), 1.98 sec
24. LEN4- Leslie Nelson greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 57330 samples (b), 2.6 sec

25. LEN5- Leslie Nelson greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 51156 samples (b), 2.32 sec
26. MDC1- Dr. Martin Crossland stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 62181 samples (b), 2.82 sec
27. MDC2- Dr. Martin Crossland greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 56889 samples (b), 2.58 sec
28. MDC3- Dr. Martin Crossland, greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 47628 samples (b), 2.16 sec
29. MDC4- Dr. Martin Crossland greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 64827 samples (b), 2.94 sec
30. MDC5- Dr. Martin Crossland greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 49392 samples (b), 2.24 sec
31. NA1- Navid Amiliagona stating name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 147294 samples (b), 6.68 sec (repeated name)
32. NA2- Navid Amiliagona greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 40572 samples (b), 1.84 sec
33. NA3- Navid Amiliagona greeting in Spanish, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 64827 samples (b), 2.94 sec
34. NA4- Navid Amiliagona greeting in Russian, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 41895 samples (b), 1.9 sec
35. NA5- Navid Amiliagona greeting in Farsi, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 47628 samples (b), 2.16 sec
36. JMA1b- Dr. Acken saying name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 71,338 samples (b), 3 sec
37. JMAasLEN- Dr. Acken saying Leslie Nelson’s name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 81,530 samples (b), 3 sec

38. LEN1b- Leslie Nelson saying name, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 56,602 samples (b), 2 sec
39. JMA2b- Dr. Acken greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 121,344 samples (b), 2 sec
40. JMA2c- Dr. Acken greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 132,370 samples (b), 3 sec
41. JMA2d- Dr. Acken greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 187,500 samples (b), 4 sec
42. JMA2e- Dr. Acken greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 99,292 samples (b), 2 sec
43. JMA2f- Dr. Acken greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 121,344 samples (b), 2 sec
44. LEN2b- Leslie Nelson greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 88,266 samples (b), 2 sec
45. LEN2c- Leslie Nelson greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 55,188 samples (b), 1 sec
46. LEN2d- Leslie Nelson greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 66,214 samples (b), 1 sec
47. LEN2e- Leslie Nelson greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 88,266 samples (b), 2 sec
48. LEN2f- Leslie Nelson greeting in English, RIFF, Sample rate 22050 kb/s, 8bits/sample, mono, WAVE Audio File, 55,188 samples (b), 1 sec

VITA

Leslie Ellen Butts

Candidate for the Degree of

Master of Science

Thesis: PROPOSED METHOD FOR EVALUATING VOICE AUTHENTICATION SYSTEMS

Major Field: Electrical Engineering

Biographical:

Education:

Completed the requirements for the Master of Science/Arts in your major at Oklahoma State University, Stillwater, Oklahoma in December, 2012.

Completed the requirements for the Bachelor of Science/Arts in Electrical Engineering at Oklahoma State University, Stillwater, OK/USA in 2012.

Experience: