Pre-print of article that will appear at BTAS 2013.

Voice Authentication Using Short Phrases: Examining Accuracy, Security and Privacy Issues

R.C. Johnson, Terrance E. Boult
University of Colorado, Colorado Springs
Colorado Springs, CO, USA
(rjohnso9 | tboult)@uccs.edu

Walter J. Scheirer
Harvard University
Cambridge, MA, USA
wscheirer@fas.harvard.edu

Abstract

This paper examines a novel security model for voice biometrics that decomposes the overall problem into bits of "biometric identity security," bits of "knowledge security," and bits of "traditional encryption security." This is the first paper to examine balancing security gained from text-dependent and text-independent voice biometrics under this model. Our formulation allows for text-dependent voice biometrics to address both what you know and who you are. A text-independent component is added to defeat replay attacks. Further, we experimentally examine an extension of the recently introduced Vaulted Voice Verification protocol and the security tradeoffs of adding these elements. We show that by mixing text-dependent with text-independent voice verification and by expanding the challenge-response protocol, Vaulted Voice Verification can preserve privacy while addressing the problematic issues of voice as a remote/mobile biometric identifier. The resulting model supports both authentication and key release with the matching taking place client side, where a mobile device may be used. This novel security model addresses a real and crucial problem: that of security on a mobile device.

1. Introduction

Millions of mobile phones are being sold every year; a majority of those phones are now smart phones[1]. As mobile phones become more sophisticated, they are increasingly being used for all types of transactions. Many of these transactions, such as those from mobile banking applications, require user verification. The ability to establish a user's identity is a critical aspect of security engineering; thus the use of biometrics is increasingly popular as a means of verification. Much work has been done with verification using face, fingerprint and iris biometrics, but for mobile phones, voice presents a more natural choice. Issues do exist, however, when using voice as a biometric identifier. One such issue is the variability of voice. Another such issue is the trade-off between security, privacy and convenience.

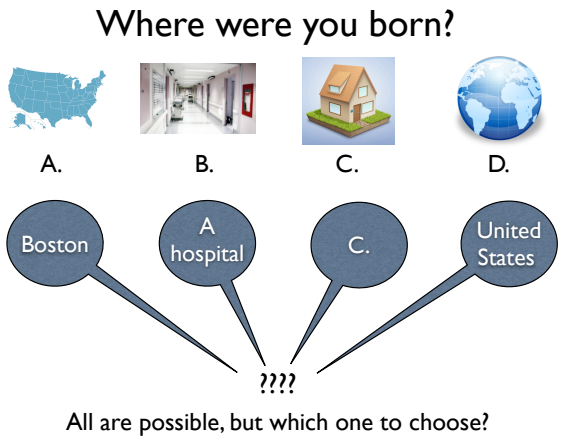

Figure 1. How would you answer this question? Would you say "C," "France," "A hospital," "Colorado," or something else? This paper looks at the potential of questions such as this one to augment the security of voice biometrics. Specifically, in this research we focus on expanding the recently introduced Vaulted Voice Verification protocol [6].

Is it possible to build a remote verification protocol that uses voice, is privacy preserving, and mixes text-dependent and text-independent modeling? This paper shows that, yes, it is. In this work we have extended the recently introduced Vaulted Voice Verification protocol [6] to increase the security of the protocol while increasing its ease of use. Vaulted Voice Verification, as is explained in Sec. 2.1, borrows techniques from the voice community, such as those described in [4, 9, 7, 8], and combines them with a vision-based technique called Vaulted Verification [20, 13].

An issue, addressed in Sec. 3, is that voice-based biometric verification protocols, such as Vaulted Voice Verification, need to get as much information as possible out of as little user interaction as possible. If it takes a user 15

[1] http://goo.gl/DHuAp

minutes to authenticate every time they use the system, they will soon stop using it. We address this concern specifically by looking at the number of bits that can be generated by a single user interaction. These "bits" represent the amount of information needed for model selection as defined by the protocol.

Another issue, also discussed in Sec. 3, is that with text-dependent voice templates/models, a system is susceptible to different voice conversion-based attacks. Such attacks are described in [10]. As shown in [1], text-independent voice templates/models are also vulnerable. Our work to extend Vaulted Voice Verification to include both text-dependent and text-independent-based models seeks to eliminate such vulnerabilities.

The main contributions of this work are:

- The inclusion of text-independent speaker models into Vaulted Voice Verification to increase the security.
- An increase of the security gain given per question.
- A security analysis of the Vaulted Voice Verification-protocol in terms of attack models.
- Analyzing the accuracy vs. security trade-off from text-independent models.

2. Relationship to Prior Work

Since we have a limited number of raw biometric features (10 fingers, two irises, one face), they cannot be directly used for verification protocols due to the threat of compromise. Typically, biometric-based verification protocols rely on templates, or models, that are created from the biometrics. There is significant interest in how to protect these templates to allow matching in an encoded domain that, at most, leaks minimal information about the original biometric features [12, 15, 19, 14, 2, 11].

One such solution is Vaulted Voice Verification, a text-dependent challenge-response-based verification protocol, which we build from in this work. Vaulted Voice Verification is a novel solution that combines work performed in the vision community [20, 13] with work performed in the voice community [4, 9, 7, 8].

The work from the vision community details a challenge-response protocol, utilizing face and iris, that mixes parts of a real image's data with chaff. To authenticate using the protocol, another image must be supplied that matches the original image close enough so that the new image can be used to distinguish the real data from the chaff.

The work from the voice community details using voice recordings to create feature vectors that can be stored securely for later use during matching. Many of these techniques utilize either Gaussian Mixture Models (GMMs) or Hidden Markov Models (HMMs) to create models from the voice recordings. Operations are then performed on these models to create the authentication systems.

2.1. An Extension to Vaulted Voice Verification

In more detail, Vaulted Voice Verification, as described in [6], is a challenge-response protocol that uses a mixture of GMMs created from a user's audio recordings and chaff GMMs created from recordings of other users or by modifying existing GMMs. The system presents the user with a series of phrases to repeat, and subsequently generates models from the responses. For each model generated, a chaff model is also generated. Multiple methods exist to generate chaff models and deciding how exactly to generate chaff is often more art than science. Examples of chaff generation include perturbing models generated from the response of the user and using real models from the same user but from a different response. The idea is that an attacker would not be able to distinguish between the real and chaff model, *i.e.* only the voice that created the real model could be used to distinguish the real model from the chaff.

As with other verification systems, Vaulted Voice Verification includes both an enrollment and a verification process. We will introduce elements of the original Vaulted Voice Verification below for the purpose of providing the necessary background to our novel extension.

RSA encryption is used to generate the server and the user key pairs in both the original and our extended version of the protocol for both enrollment and verification. Except where noted, when the user transmits data to the server, that data is first encrypted with the public key of the server so only the server can decrypt it using its matching private key. Likewise, when the server transmits data to the device of the user, that data is first encrypted with the public key of the user so only the user can decrypt it using their matching private key.

2.1.1 Enrollment

The enrollment process for our extended version of Vaulted Voice Verification is illustrated in Fig. 2. The steps described here refer to the numbered arrows in the figure. Differences in the original versus our extended version will be pointed out where appropriate for both the enrollment and verification process.

In step 1, the user enters their information into their device. Then, in step 2, the device generates keys for the user. These steps follow the steps of the original protocol.

For steps 3 and 4 in the original protocol, the device interacts with the user, asking the user to repeat series of phrases to which the user responds. In our extended protocol, we use both text-dependent and text-independent prompts for the user. For the text-dependent mode, the device prompts the user with short phrases or small passages

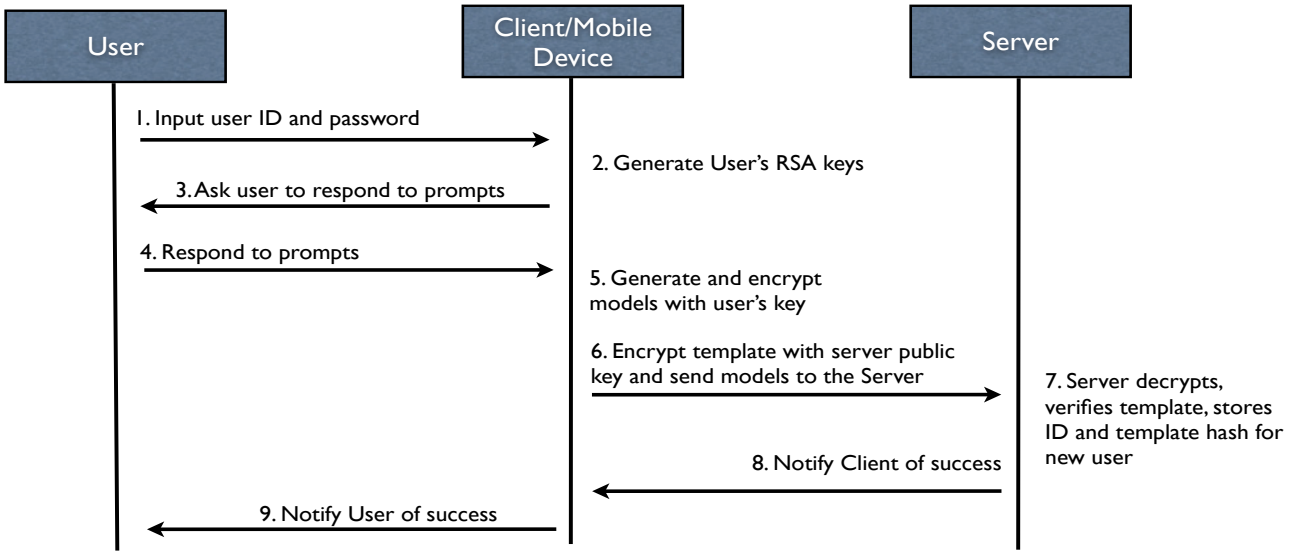

Figure 2. Vaulted Voice Verification: Enrollment Process.

to repeat. For the text-independent mode, the device shows images that need a short and non-scripted description. For each of these prompts, a real model and a chaff model are created so that the chaff model is similar to the real.

In step 5, the device then encrypts the models using the public key of the user. The encryption of the models occurs in the same manner in the original Vaulted Voice Verification protocol. The public key of the user is used here so that only the user can decrypt the models with their private key.

In steps 6 and 7 in the original protocol, the encrypted models are sent to the server for storage until verification and subsequently are removed from the client device. In our extended protocol, once the server receives the models, it creates hashes for later verification, and the models are then deleted from the server. As a result of this, the encrypted models are able to remain on the client device until verification.

Lastly, a notification is sent back to the user in steps 8 and 9. In both the original and the extended protocols the user receives a notification of success/failure of enrollment.

2.1.2 Verification

The verification process for our extended Vaulted Voice Verification is illustrated in Fig. 3. The steps described here refer to the numbered arrows in the figure. Again, differences between our extension and the original will be pointed out as appropriate.

In steps 1 and 2 the user inputs their information into the device, which sends a verification request to the server. In the original protocol for step 2, the device sent the ID and a request for verification. In our extended protocol the device also sends the encrypted template, because it is no longer stored on the server.

In step 3 of the original protocol, the server retrieves the models associated with that user and scrambles them according to some i.i.d. binary challenge string. In our extended protocol, the server first verifies the template by hashing it and comparing the hash against the previously stored value from enrollment. This allows the server to be sure it is the same data it received during enrollment without the need to store the data itself.

Steps 4 and 5 show the server sending the shuffled models to the client, which decrypts them using the information provided by the user in step 1. In steps 6 and 7 of the original protocol, the user interacts with the device, repeating the phrases as prompted. Our extension expands this interaction from only phrases to also include passages of text and images that must be described by the user.

In step 8, the device generates a new model from each response for each phrase from the user. These new models are used to select between each real and chaff model for the prompted phrases. As the device selects between the two to unscramble the models, it builds a response string. When finished building the string, in step 9, the device sends the response string to the server for verification. The device then compares the response string to the original challenge string. The remaining steps illustrate the server responding with an accept or decline decision based on the response string matching the challenge string.

While Vaulted Voice Verification provides a novel combination of techniques from two communities, the use of only text-dependent models makes it susceptible to the attacks mentioned in Sec. 1. Also, the use of binary models

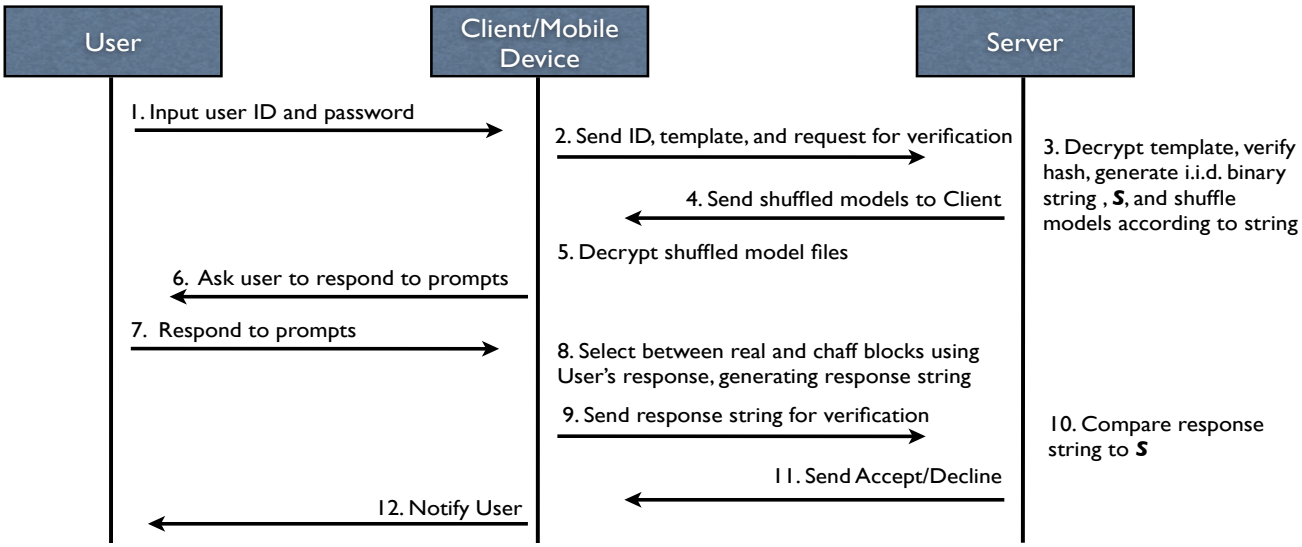

Figure 3. Vaulted Voice Verification: Verification Process.

leads to either too many questions for the user to answer in a practical interactive system, or a system that is easily compromised due to the small number of bits of security that is provided.

3. Improving Security and Usability

In this work we extend Vaulted Voice Verification beyond text-dependent modeling and single bit questions and answers. We look at mixing text-dependent modeling, where the models are based on specific word snippets, with text-independent modeling, where models are generated using larger word groupings. Also, we enhance the amount of information gathered from each question in the challenge response by extending from binary to multiple choice.

3.1. Text-dependent and Text-independent Models

Our work of mixing text-dependent and text-independent modeling is similar to that of [18] and [3], where they focused on small phrase groupings. In [18], they look at words based on frequency of occurrence in the speech samples. In [3], they look at certain keywords that are likely to occur with great frequency in conversation. In our work, we take short phrases based on images and compile them to create a more text-independent model. Our models are not completely text-independent because we're creating them based on a small number of phrases from an individual. However, our models are also not text-constrained because they are based on open-ended responses from the user. We therefore consider our models to be short phrase-based text-independent models, based on the number of available phrases.

In our earlier work, [6], we had the user respond by repeating certain phrases. With that system, the phrases are known both during the time of enrollment as well as during verification. With this, the system is limited to the number of predetermined phrases it starts with and the challenges are simple in their complexity, as well as in the security they provide. The increase in security resulting from the addition of the text-independent models as described in this section will be explored in Sec. 4.

Figure 4. Vaulted Voice Verification is extended by using open-ended challenges in the form of images instead of phrases the user must repeat. For example, how would you describe this image?

In this work, we look to extend the protocol from presenting the user with phrases to speak, to showing the user images and asking them for short descriptions of the images. Examples of this can be seen in Fig. 1 and Fig. 4. In Fig. 4 we have an example of an image the user could be presented with. The idea is that different people will use different words to describe the same image. When we

are using text-dependent models, the model implicitly incorporates the word and the voice model in the answer, thus improving the overall security. By doing this, we are able to create models that have greater degrees of freedom.

Multiple advantages exist to the proposed extensions of the protocol. The possibilities for the response to each challenge are as vast as the lexicon allows. This implicitly combines what you know with who you are. Since the models exist only on your phone, *i.e.* something you have, it's now a full three-factor authentication.

The protocol incorporates text-independent models to mitigate the threat of replay-attacks. By replay-attack, we mean an attacker recording audio responses and replaying them at a later time during an attempted attack. To do this, the server first generates a phrase, the user reads it, and then the server both generates a text-independent speech model as well as performs speech recognition to verify proper content. This way the phrase is unknown to the user until the server challenges them. This makes a replay-attack, even using a pre-recorded voice of the subject, impractical. The security not only relies on the general text-independent model, but also the spoken description. More information on the different types of attacks our research works to defend against will be detailed in Sec. 4.

3.2. Increased Information from Responses

A challenge that needs to be addressed with voice based biometrics is how to get the most information from users without it taking so long that they don't want to use the system. Initially, with the original Vaulted Voice Verification protocol, the users would answer a series of questions, with each question generating one bit of security. We have now extended the protocol so that each question is able to generate multiple bits. To achieve this, we turned each binary choice question into a multiple choice question. Our experiments used four-choice questions. This improves security by adding more bits of information to each question.

With this protocol, a portion of the security results from the server scrambling data to generate a challenge. Given this, each challenge must contain a finite number of possible choices for the server to scramble and send to the client. This differs from traditional biometrics in that this protocol does not extract data from the response and compare it to a stored value, rather it uses the response to choose between multiple possibilities as presented by the server.

In our experiments, this extension to four possible answers has doubled the amount of bits that are produced per question. For our experiments, we assume that the questions are independent of each other. We further assume that an attacker is choosing at random from the possible answers. These assumptions are for a naive attacker model. Sec. 4 will address a more sophisticated attack analysis. If the user is asked a series of five questions, the probability of getting them all correct based on the original Vaulted Voice Verification work would be 2^{-5}, with one bit per question. With the extension implemented in our experiments, two bits are generated per question. For the same given questions, there would now be 10 bits of security. With this, the probability of randomly accepting the identity would decrease to 2^{-10}.

The number of questions for the system can be chosen to balance the chance of randomly accepting an identity with the accuracy of the voice biometric. We can then balance the number of questions with the biometric error rate (*e.g.* FAR). It is important to understand that the biometric error rate provides a lower bound on identification accuracy because of the associated biometric dictionary attack [17]. Thus, the identity security offered by Vaulted Voice Verification with only five questions is already much better than the state-of-the-art biometric error rate for voice [5].

4. Security and Privacy

The security analysis for our work on Vaulted Voice Verification is straightforward, simple and similar to that of the original protocol. The security analysis for the original protocol is given in [6], but it lacks analysis based on specifically defined attack models. For our research, we looked at the security of our extended version of Vaulted Voice Verification in terms of six different attack models listed in order of likelihood of attack.

1. The knowledgable impersonator "borrows" the device: An attacker, most likely a "friend" in this case, grabs your device, knows your password and your answers.

2. Compromised transmissions: An attacker is able to capture and isolate all the data in transmission, but has not obtained any of the encryption keys.

3. Stolen device and password: An attacker is able to obtain the keys of the user and access the data, but does not have access to the server.

4. Compromised server: An attacker is able to obtain the keys for the server but cannot manipulate the server's software/network.

5. "Insider attack": An attacker is able to obtain the keys of the server and is able to access data stored there.

6. The "mission impossible" attacker: An attacker spends time to make recordings of your voice and steals your phone.

It is necessary to note that this analysis of security in terms of bits is speaking in terms of how much security is added on top of encryption. Our implementation of Vaulted Voice Verification provides P-bits of security for the encryption from the salted user password with the assumption

that the device will lock up after a set number of attempts, S-bits of security from the server encryption of the template, K-bits of "knowledge-based" security, and B-bits of "biometric-identity" security. When we refer to K-bits of knowledge-based security, we mean to say security that is gained per challenge-response question that is due to something that the authentic user knows and an attacker does not. Thus, depending on the attack model, the odds for an attacker guessing correctly mirror that of random chance. When we speak of B-bits of biometric-identity security, we mean security that is gained through the use of voice-based models that take advantage of the difference in the voices and speech patterns of different speakers.

For attack model 1, because the attacker has access to the device and the password, they can bypass the encryption. For the multiple choice questions, the attacker knows the answers. Because the attacker does not have the correct voice, this reduces to the B-bits of biometric-identity security.

For attack model 2, the attacker has obtained the data. Another way to look at this attack model is to imagine that an attacker is able to watch all transmission of data between the server and the client, but is not able to decrypt said transmissions. During both enrollment and verification, only encrypted data is being transmitted back and forth between the client and the server. Without either key, the attacker would not be able to access the data. Because the data blocks are scrambled before encryption, every time data is sent the ciphertext is different. Thus an attacker is not able to gain anything from the information they are able to collect. With this attack model, the bits of security are a total of $P + S + K + B$.

For attack model 3, similar to attack model 1, an attacker is able to compromise the keys and the data of the user. The attacker would not know the answer to the multiple choice question, reducing their chances to random guessing (again, making the assumption that the question choices are independent). The attacker would also not have the correct voice pattern for the text-dependent or text-independent matching problem. The password and user keys are compromised, but the remaining bits of security are $B + K$.

For attack model 4, the attacker has somehow obtained the encryption keys for the server and can scan the server's disks but cannot modify the operational software. Because the server has no stored data, there is not much they can do with the data found on the server. They could set up a phishing application and launch a man-in-the-middle attack, but with this they can only decrypt the template. In this attack model, there would still exist $P + K + B$ bits of security.

For the next attack model, number 5, the insider can ignore all protocols on the server, making verification inconsequential. However, to impersonate a user on any other server, there would still exist $P + K + B$ bits of security to overcome. Because the model files are encrypted, the attacker gains no information about the raw biometric and therefore cannot identify or impersonate the user anywhere else.

Attack 6 is a classic movie plot. A dedicated attacker is able to obtain multiple voice samples of the user. An outside attacker, without any of the keys, will have gained nothing from doing this. The user would still be protected by $P + S$ bits of security. If one were to believe this movie plot threat has a higher probability of occurring for this individual, they could specify more text-independent questions so that no replay attack can be used. However, this attack seems to occur only in the movies. The most likely scenario for this attack is to have malicious code on the phone, which is why we include at least one text-independent question.

Reviewing these attacks, it becomes clear that the most likely and most invasive attack is a snooping "friend". With this, the remainder of our analysis in Sec. 5 will focus on analysing the B-bit of biometric security.

5. Experimental Results

The dataset selected for our experiments is the same dataset used in the original Vaulted Voice Verification experiments: the MIT mobile device speaker verification corpus [16]. The speech corpus is composed of 48 speakers; 22 female and 26 male. Short phrases, names and ice cream flavors, were recorded in 20 minute sessions. The data was created using two enrollment sessions and one impostor session. As defined by the dataset, each person has a dedicated impostor.

For these experiments, the data was separated to yield a gallery and probe sets that contain separate audio samples. At random, 60% of the enrollment data was designated as gallery and the other 40% as probe, ensuring an appropriate balance between sessions. The impostor data was used in its entirety. In the tests, the data are further separated on a per-phrase basis; this way, speech-dependent models could be created.

Our tests are conducted by testing every person against all other people in the gallery. This method of expanding the dataset is referred to as "all vs. all" testing in the original work. In early work, we performed same sex and mixed set testing; same sex testing yielded higher accuracy. However, the individual same sex sets are small, and the GMMs easily discriminate the data within them. We felt that since mixed set testing increases the data available for testing, as well as the potential for collisions, it was a more realistic case to report. The comparison, or scoring, of the models is done using a z-score, as in the original work. For comparison, tests were first performed on the original text-dependent binary versions of the baseline and the Vaulted Voice Verification algorithms.

Fig. 5 shows four different Linear-Scale DET (LDET)

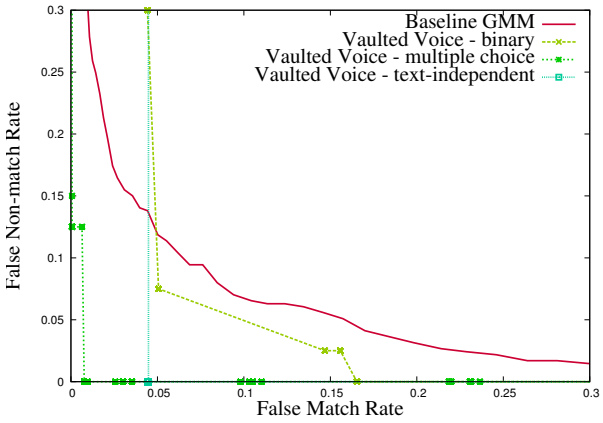

Figure 5. Linear-Scale DET (LDET) plots for three Vaulted Voice Verification variants and a baseline. The baseline GMM plot represents matching without the privacy and security that Vaulted Voice Verification provides. The three variants of Vaulted Voice Verification represent privacy and security gains through binary models, multiple choice models, and text-independent models. We do not present these results in a more tradition DET plot because scores of 0 cannot be shown on a log-based plot.

curve plots. The figure contains a baseline and highlights three different variants of Vaulted Voice Verification with varying question types and text-dependent and text-independent challenges. An operational system would likely fuse multiple choice text-dependent and text-independent models with the number of each type dependent on the desired security model. The three plots will be examined and discussed in terms of the knowledge-based security (K-bits) and biometric identity security (B-bits).

The plot labeled "Baseline GMM" is generated by scoring each binary challenge-response pair and applying a threshold across the results. By binary, we mean that for each question, two possible answers exist from which to choose. The curve represents a baseline result for a single question model, *i.e.* a single general threshold is applied for each question. It has an approximate equal error rate of 8% and represents B-bits of security.

The binary Vaulted Voice Verification plot is generated by applying the Vaulted Voice Verification protocol to binary challenge-response pairs. In Vaulted Voice Verification there is an inherent pairwise thresholding that takes place. This pairwise thresholding allows Vaulted Voice Verification to account for variation from phrase to phrase. So, for each phrase, a different threshold is applied based on which model is the closest. As shown in the figure, Vaulted Voice Verification outperforms the baseline for binary challenge-response pairs. This curve shows an approximate equal error rate of 6% and also represents B-bits of security.

The next curve we will discuss is the one generated from multiple choice questions. For this experiment, there are four possible answers for each question. Multiple choice questions expand the security of the overall system by adding K-bits of knowledge-based security. This Vaulted Voice Verification curve improved from the binary case, with an approximate equal error rate of 1%, because with multiple choice questions there are $B + K$ bits of security. This means that on top of the biometric security (something you are) additional security is provided by knowledge (something that you know).

The final curve that we will discuss is the text-independent curve. This curve not only represents $B + K$ bits of security, but also includes security added by repeating a random passage verbatim. The curve is generated by comparing models that are created from all phrases in the gallery for each user against models that are created from all the probe phrases for each given user. As such, for each user there is one real model and one impostor model, as in the case of the binary text-dependent models. For Vaulted Voice Verification, there is a single bit for each challenge response pair, resulting in a 4% equal error rate. The overall system security is improved because the security for this comes from balancing the biometric bits of security with the ability to accurately speak the correct passage to generate the model. This greatly reduces the possibility of a movie-style replay attack.

Many LDET curves are presented as false accept versus true accept rate. We instead chose to present our results as match rate versus non-match rate because acceptance is a function of the overall system – here we are analyzing only the biometric matching component. One can estimate the impact of the other layers by rescaling the false match rate by 2^{-N} where the un-compromised layers provide N-bits of security. For example, in attack model 4, the system would still have $P + K$ other bits of security. So, if P=10 and K=10, the system's false accept rate scales the false match rate by a factor of 1/1,000,000.

6. Conclusion

With this work, we have extended the Vaulted Voice Verification protocol from one that generates a single bit per challenge response pair to one that is capable of generating multiple bits. We have also extended the Vaulted Voice Verification protocol to mix text-dependent and text-independent modeling. We have shown that these extensions of the Vaulted Voice Verification protocol improve the EER from 6% to 1% when generating multiple bits per challenge response pair. We have also shown that by adding short phrase-based text-independent modeling, the EER still sees improvement, going from 6% to 4%.

We also performed a security analysis of the extended protocol. We presented this analysis in the form of the different possible attack models that exist for this protocol. We demonstrated that with the new extensions, the protocol is improved.

This paper has presented a decomposition of the security of the system into $P + B + S + K$ bits and introduced six different attack models. Attack models 1, 3 and 6 are all based on compromising the device of the user and the security is effectively limited by $B + K$. Models 2, 4 and 5 are based on factors outside the control of the user, *i.e.* data transmission and the server where security is dominated by $S + P$. Given this, a system administrator has the ability to modify the security of the overall system by defining the values set for P, S, B and K. Depending on the likelihood of attacks 2, 4 and 5, the system designer can balance convenience and cost by setting S and P; for example, by adjusting how the keys are generated and limiting passwords. The system designer can give the users a range of options for B and K that balance the perceived security versus the desired usability of the system for the user.

With these extensions, Vaulted Voice Verification is now better suited for real world application: for example, mobile access to secure information, such as calling in to check on banking or credit card information. Vaulted Voice Verification is able to ensure that the person who attempts to access the account by phone is the person they claim to be. These new extensions allow this to happen without the need to ask the user an excessive number of questions for the purpose of obtaining a reasonable amount of security bits.

References

[1] F. Alegre, R. Vipperla, N. Evans, and B. Fauve. On the vulnerability of automatic speaker recognition to spoofing attacks with artificial signals. *20th European Signal Processing Conference (EUSIPCO 2012)*, 2012.

[2] K. N. Anil K. Jain and A. Nagar. Biometric template security. *EURASIP Journal on Advances in Signal Processing*, Special Issue on Biometrics, 2008.

[3] K. Boakye and B. Peskin. Text-constrained speaker recognition on a text-independent task. In *ODYSSEY04-The Speaker and Language Recognition Workshop*, 2004.

[4] G. Z. F. Zheng and Z. Song. Comparison of different implementations of mfcc. *Journal of Computer Science andTechnology*, 16:582–589, 2001.

[5] K. Inthavisas and D. Lopresti. Secure speech biometric templates for user authentication. *The institution of Engineering and Technology Biometrics*, February 2012.

[6] R. Johnson, W. J. Scheirer, and T. E. Boult. Secure voice based authentication for mobile devices: Vaulted voice verification. *Proceedings of SPIE, Biometric and Surveillance Tech. for Human and Activity Identification*, 8712, May 2013.

[7] B.-H. Juang and L. R. Rabiner. *Automatic speech recognition A brief history of the technology development*. Elsevier Encyclopedia of Language and Linguistics, second edition, 2005.

[8] M. Just and D. Aspinall. Personal choice and challenge questions: A security and usability assessment. *Proceedings of the 5th Symposium on Usable Privacy and Security, 8*, 2009.

[9] H. H. K. LEE and M. Hwang. The sphinx speech recognition system. *Acoustics, Speech, and signal Processing (ICASSP-89)*, 1989.

[10] T. Kinnunen, Z. Wu, K. Lee, F. Sedlak, E. Chng, and H. Li. Vulnerability of speaker verification systems against voice conversion spoofing attacks: The case of telephone speech. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 4401–4404. IEEE, 2012.

[11] L. Lai, S. Ho, and H. Poor. Privacy-security tradeoffs in reusable biometric security systems. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 1722–1725. IEEE, 2010.

[12] D. Lee and K. N. Plataniotis. A novel eye region based privacy protection scheme. In *ICASSP*, pages 1845–1848, 2012.

[13] T. E. B. Michael J. Wilber, Walter J. Scheirer. Privv: Privae remote iris-authentication with vaulted verification. *Computer Vision and Pattern Recognition (CVPR)*, 2012.

[14] A. Nagar, S. Rane, and A. Vetro. Privacy and security of features extracted from minutiae aggregates. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 1826–1829. IEEE, 2010.

[15] M. A. Pathak and B. Raj. Privacy-preserving speaker verification as password matching. In *ICASSP*, pages 1849–1852, 2012.

[16] A. P. Ram H. Woo and T. J. Hazen. The mit mobile device speaker verification corpus: Data collection and preliminary experiments. *IEEE Odyssey - The Speaker and Language Recognition Workshop*, 2006.

[17] W. Scheirer, B. Bishop, and T. Boult. Beyond pki: The biocryptographic key infrastructure. In *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on*, pages 1–6. IEEE, 2010.

[18] D. Sturim, D. Reynolds, R. Dunn, and T. Quatieri. Speaker verification using text-constrained gaussian mixture models. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 1, pages I–677. IEEE, 2002.

[19] Y. Wang and K. N. Plataniotis. An analysis of random projection for changeable and privacy-preserving biometric verification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 40(5):1280–1293, 2010.

[20] M. J. Wilber and T. E. Boult. Secure remote matching with privacy: Scrambled support vector vaulted verification. *Workshops on the Applications of Computer Vision (WACV)*, 2012.