Securing Electronic Medical Records Using Biometric Authentication

Stephen Krawczyk and Anil K. Jain

Michigan State University, East Lansing MI 48823, USA
{krawcz10,jain}@cse.msu.edu

Abstract. Ensuring the security of medical records is becoming an increasingly important problem as modern technology is integrated into existing medical services. As a consequence of the adoption of electronic medical records in the health care sector, it is becoming more and more common for a health professional to edit and view a patient's record using a tablet PC. In order to protect the patient's privacy, as required by governmental regulations in the United States, a secure authentication system to access patient records must be used. Biometric-based access is capable of providing the necessary security. On-line signature and voice modalities seem to be the most convenient for the users in such authentication systems because a tablet PC comes equipped with the associated sensors/hardware. This paper analyzes the performance of combining the use of on-line signature and voice biometrics in order to perform robust user authentication. Signatures are verified using the dynamic programming technique of string matching. Voice is verified using a commercial, off the shelf, software development kit. In order to improve the authentication performance, we combine information from both on-line signature and voice biometrics. After suitable normalization of scores, fusion is performed at the matching score level. A prototype bimodal authentication system for accessing medical records has been designed and evaluated on a small truly multimodal database of 50 users, resulting in an average equal error rate (EER) of 0.86%.

1 Introduction

An increased need for a reliable authentication scheme has emerged in the health care industry as a result of the movement toward electronic medical records and the recently approved governmental regulations in the United States. Every year, billions of patients in the United States visit doctor's offices, clinics, Health Maintenance Organizations (HMO), hospitals, and other health care providers [2]. Each of these visits either generates a new medical record or adds to an existing one, necessitating the retrieval of a particular record. The procedure by which these records are stored and retrieved is undergoing a change toward a system that will better utilize modern technology. Security risks involved with this new system of archiving and retrieving patient records has brought about the onset of several government regulations pertaining to the protection and privacy of medical records which in turn has increased the need for a reliable user authentication scheme in this domain.

T. Kanade, A. Jain, and N.K. Ratha (Eds.): AVBPA 2005, LNCS 3546, pp. 1110–1119, 2005.

1.1 Electronic Medical Records

A medical record can span hundreds of pages consisting of text, graphs, and images. It contains information on treatments received, medical history, lifestyle details, family medical history, medications prescribed, and numerous other items pertinent to an individual's health. In the interests of the integrity of the health care industry and good patient care, it is recommended that these records should be retained for as long as possible. For these factors alone, it is obvious that the move toward electronic data capture will greatly assist in the storage and management of patient records. Although this change is long overdue, the health care industry has only recently begun to convert their paper records to electronic form using electronic medical record (EMR) systems [3, 14].

1.2 Federal Regulations

The automation of health care information management has created increasing governmental and societal concerns about the security of computerized health care data. While the health care industry has incorporated electronic medical records, data repositories, networking, Internet access, and other new technologies into its various process, the corresponding security measures have not been enhanced. Many weaknesses have been identified in existing health care security measures from past operations [6]. The Health Insurance Portability and Accountability Act (HIPAA), which set the standards to ensure the security and integrity of patient information that is maintained or transmitted electronically, took effect in April 2003 [5]. Patients are assured, under HIPAA regulations, that their medical records will be used only by individuals directly involved in their medical treatments, payment of their bills, and health care operations. Any other individual or organization wishing to access a patient's medical record would require specific authorization by that patient. These regulations also attempt to ensure that when the medical records are properly disclosed, only the minimum amount of information necessary shall be released.

1.3 Tablet PC

Since it is convenient for a health care professional to have a patient's record readily available when prescribing or administering treatment, many health care facilities have adopted the use of tablet PCs as access devices to retrieve and edit a patient's record. The tablet PCs are easy to use and are able to access a patient's data through wireless access points. The widespread deployment of these wireless access points in hospitals and other facilities presents new security problems where only authorized users of the tablet PC are permitted to view the requested medical records.

1.4 Biometric Authentication

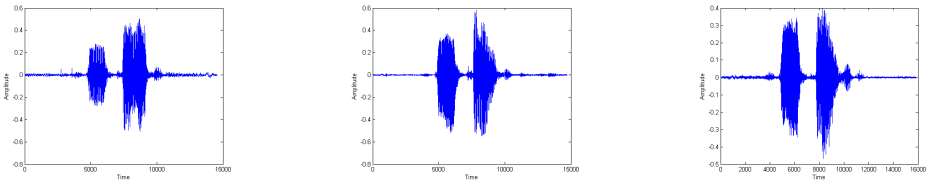
It is widely recognized that biometric authentication offers a number of advantages over traditional token-based (e.g. ID cards) or knowledge-based (e.g.

passwords) systems [12]. Several companies have realized these security benefits and have integrated biometrics into their EMR systems that use modalities such as the fingerprint and iris [1, 3, 14]. Additionally, multimodal biometric systems can overcome many of the limitations of a unimodal biometric system and will be the focus of this work [10]. In order to meet the guidelines of the HIPAA regulations, both health professionals and patients must be given access to medical records. Taking into account the requirements of both these groups (health professionals and patients), our biometric authentication system uses the voice and signature modalities. These modalities are unobtrusive and emulate the current, already well accepted system whereby a patient authenticates herself when seeking treatment or visiting a doctor's office for consultation. A typical scenario consists of a patient telling his or her name to a receptionist and then signing a release form. In addition, health professionals are already beginning to use tablet PCs to access patient records which are equipped with a stylus/pen and an internal microphone. Using the voice and signature modalities, our biometric authentication system can be seamlessly integrated into a tablet PC without any extra hardware.

2 Voice and Signature Verification

2.1 Voice Verification

In our authentication system, both voice identification and verification are utilized. The difference between voice identification and verification is that voice identification involves identifying a speaker out of a group of templates (1 to N matching) whereas verification deals with verifying whether an utterance matches with a specific user's template (1 to 1 matching). A user template is visually depicted in figure 1, and, as shown, can contain high intra-class variance. The voice biometric is used for authentication in such companies as Banco Bradesco, the largest bank in Brazil, the United Kingdom government's Intensive Supervision and Surveillance Program for fighting crime, and other major financial institutions for access to personal accounts and information [4]. In this work, both voice identification and verification are performed using the Nuance Verifier SDK [11]. The Nuance recognition and verification engines use Hidden Markov Models (HMM) to provide a mapping from sampled speech to phonetic units. Continuous-density HMMs are utilized, where the relationship between acoustic frames and states is modeled using a mixture of Gaussians [13]. These HMMs are set up in a hierarchical fashion so that after sampled speech is mapped to phonetic units, the resulting phonetic sequence is then mapped to the corresponding word sequence. The probability from the last Markov chain in the sequence is used as the verification score. Verification is text-independent while identification is text-dependent. Our system uses the same utterance for both identification and verification and accordingly the same phrase used in enrollment must also be used for verification. A minimum of two utterances is needed to train the Markov model. Each voiceprint will usually require 20KB of memory. Typical accuracy figures of the verifier are reported as being 99% or higher.



(a) (b) (c)

Fig. 1. Voice intra-class variability. (a), (b), and (c) are three waveforms (amplitude vs. time) from a single user who spoke his first and last name three different times

2.2 On-Line Signature Verification

Handwritten signatures are frequently used to authenticate financial transactions or the contents of a document, where the verification of these signatures is usually done by human visual inspection. Much work has been done in the effort to automate the process of signature verification because of its long standing acceptance in many applications. The main disadvantage of using this biometric is its inherent high intra-class variability, as shown in figure 2. The signature verification algorithm used in this work is a modified version of the algorithm reported in [7] and the details are described in [15]. The input to the algorithm is both the dynamic (temporal) and spatial information of the writing. Features such as the change in x and y coordinates between subsequent points in the signature and the pen pressure are extracted to form a feature vector at each point. An input signature is compared with an enrolled signature by using dynamic time warping (DTW), to find an alignment between the points in the two signatures such that the sum of the differences between each pair of aligned points is minimal. The resulting difference value is used as the verification score. A training set of signatures is used to both calculate user-dependent statistics and to compare against an input signature. After performing user-normalization and dimension reduction techniques, the resulting score is combined with a global feature system score to produce a final distance value. This global feature system extracts twenty global features and performs matching using the Mahalanobis distance. The size of the templates for each user is on average 30KB. The accuracy of the algorithm has an EER of 14.25% on skilled forgeries and 0.57% on random forgeries using the first 40 users from the SVC database [16].

(a) (b) (c)

Fig. 2. Signature intra-class variability. (a), (b), and (c) are three signatures from a single user

3 Biometric Fusion

Common problems that may occur in unimodal biometric systems are noise in sensed data, intra-class variations, inherent distinctiveness, and spoof attacks. Many of these limitations imposed by unimodal biometric systems can be either overcome or reduced by using multiple biometric traits. Multimodal systems have demonstrated higher accuracy due to the fact that they use multiple biometric modalities and combine independent evidence to make a more informed decision. If any of the limitations mentioned above is present in one extracted biometric trait, there will be other traits available to the system to use in its decision. Accordingly, it is necessary to determine a method by which the individual modalities are combined. There are three possible levels at which fusion can be performed; feature level, matching score level, and decision level. We are unable to perform fusion at the feature level because of the use of a commercial voice biometric system. Also, the matching scores provide much more information than the output decisions and, consequently, we will perform fusion at the matching score level. After having computed the signature and voice matching scores and before attempting to combine the scores, a normalization technique has to be applied. The signature score is a distance measure in the range $[0, \infty)$, where 0 indicates a perfect match and any non-zero value represents the degree of difference between the two signatures. The Nuance speech SDK produces a score as a similarity measure in the range $(-\infty, \infty)$, where a negative value represents a small similarity between the two voiceprints and a positive value represents a large similarity. The transformation $T_v = e^{-x_v}$ is used to convert the voice score to a distance measure, where x_v is the raw matching score and T_v is the normalized score. After this transformation, both the modalities have a similar range of $[0, \infty)$.

The problem of combining the scores from the voice and signature modalities for a given test sample T with scores (T_v, T_s) can be considered as a two-class classification problem. The sample T can fall into either the impostor (w_i) or genuine (w_g) class. A Bayesian approach would assign T to w_i if

$$P(w_i|T_v, T_s) > P(w_g|T_v, T_s) \tag{1}$$

and w_g otherwise. In the above equation, T_v and T_s are the normalized voice and signature scores, respectively, and $P(w|T_v, T_s)$ denotes the posteriori probability of class w given the voice and signature scores. The strategy used in our system is the simple sum rule described in Jain and Ross [9]. This rule assumes statistical independence among the two modalities and also assumes that the posteriori probabilities computed by the individual classifiers do not deviate much from the prior probabilities [8]. The weighted sum rule assigns a test sample $T = (T_v, T_s)$ to w_i if

$$W_v P(w_i|T_v) + W_s P(w_i|T_s) > W_v P(w_g|T_v) + W_s P(w_g|T_s) \tag{2}$$

and w_g otherwise. In equation (2), W_v and W_s are the weights assigned to the voice and signature scores, respectively. Figure 3 shows the genuine and impostor

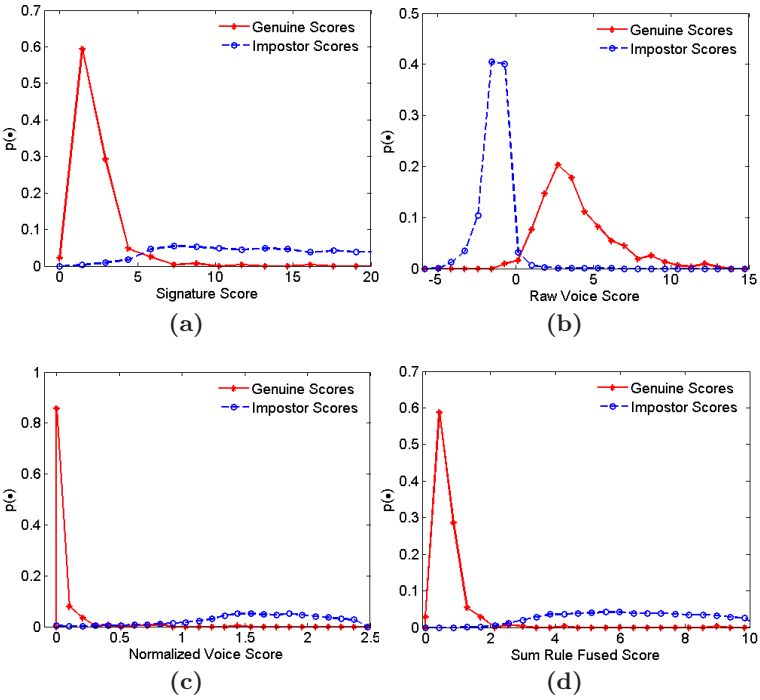


Fig. 3. Distribution of genuine and impostor scores from one trial of cross validation; (a) Signature (distance score), (b) Raw voice (similarity score), (c) Normalized voice (distance score), (d) After sum rule fusion (distance score)

distributions of the signature, raw voice, normalized voice, and fused matching scores using equal weights.

4 Results

4.1 Database

The data used for the evaluation of the authentication system was gathered from 50 individuals, each contributing 10 voiceprints and 10 signatures. The data was collected in a single session from students in various laboratories on our campus with significant ambient noise. Each individual was asked to speak his or her full name and provide a genuine signature. A Toshiba Protege tablet PC was used to perform the data collection for both the voice and signature using the stylus for the signature and the internal microphone for the voice.

4.2 Performance

The database was divided into training and testing sets by using three randomly selected voice and signature samples as the training set and the remaining seven samples as the testing set. The training voice and signature samples were used

for enrollment for each user, creating user templates for each modality. The testing samples are then used to generate authentic scores for each user. Random impostors for a user are generated by using the signature and voice samples from all the other users. The corresponding receiver operator characteristic (ROC) curves are shown in figure 4. After performing ten-fold cross validation, the average equal error rate of voice alone is 1.60% versus 3.62% for signature alone. The variance of the equal error rates of the individual voice and signature systems is 0.05 and 0.31, respectively. The combination of the two modalities using the weighted sum rule (with equal weights) has an average equal error rate of 0.86% and a variance of 0.01.

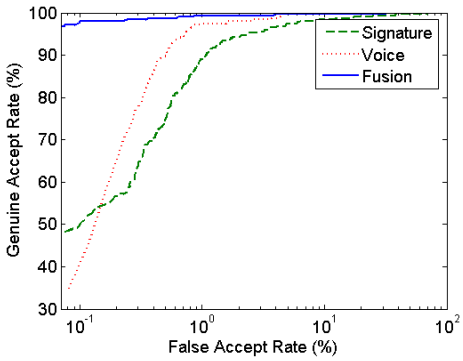


Fig. 4. ROC curves showing the results of the unimodal and multimodal systems from one trial of cross validation. The verical axis is the genuine accept rate and the horizontal axis is the false accept rate, drawn on a logarithmic scale

Figures 5, 6, and 7 show some specific examples of incorporating multiple modalities into the final decision. Figure 5 displays an example of an error in the signature verification algorithm being corrected by fusion. Here, the template and query signatures are very similar and, therefore have a low matching score. However, because the voice verification algorithm found the two voiceprints to be dissimilar, the multimodal system was able to classify the query correctly as an impostor. Figure 6 displays a situation where the query voiceprint contained a significant amount of noise and was incorrectly matched with the template. On the other hand, the signature verification algorithm found the user to be an impostor and this was able to help the system classify the query correctly. Finally, figure 7 displays an example of an error that was unable to be resolved by the mulitmodal system. Both voiceprints are greatly influenced by noise and the verification provides a misleadingly low distance score. The signatures also seem to follow the same pattern and the verification process found them to be similar. Both modalities gave wrong results and, consequently, the fusion system was unable to correctly classify the query as an impostor.

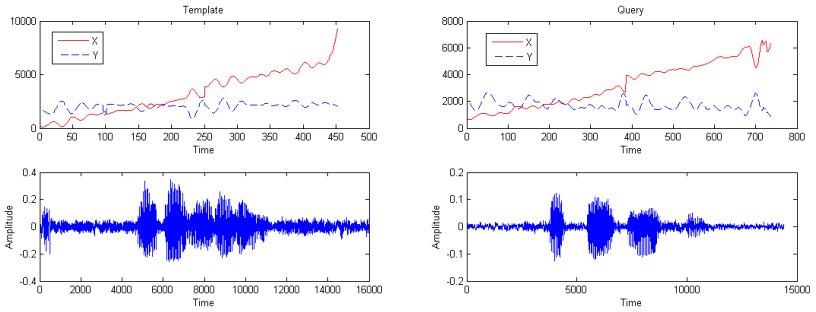


Fig. 5. Signature error resolved by fusion. The graphs show the x and y signals of signature and amplitude of voice samples plotted against time of two different users (a) and (b). The signature signals are the upper plot while the voice waveforms are depicted below. The signature score between (a) and (b) is 0.77, indicating a genuine signature. The normalized voice score between (a) and (b) is 1.5, indicating an impostor voice sample. Fusing the scores together shows the user to be an impostor

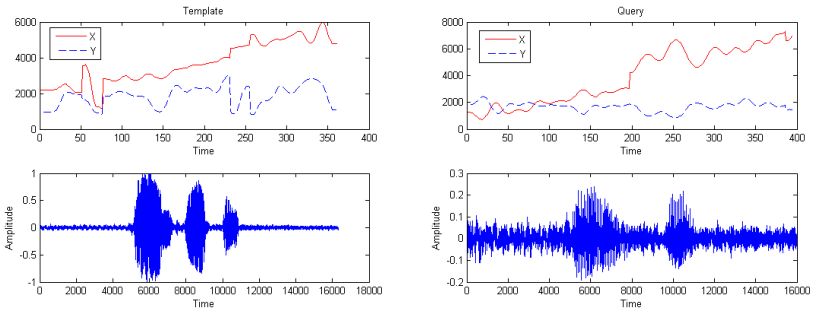
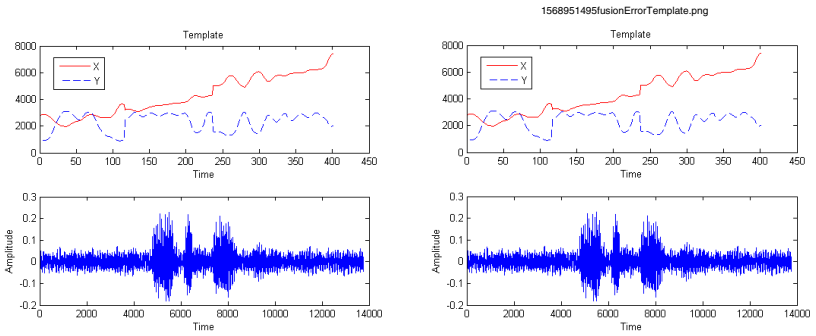


Fig. 6. Voice error resolved by fusion. The graphs show the x and y signals of signature and amplitude of voice samples plotted against time of two different users (a) and (b). The signature signals are the upper plot while the voice waveforms are depicted below. The signature score between (a) and (b) is 1.853, indicating an impostor signature. The normalized voice score between (a) and (b) is 0.02, indicating a genuine voice. Fusing the scores together shows the user to be an impostor

5 Conclusions

We have designed and implemented an authentication system based on the fusion of voice and signature data. This system was motivated by the health care industry and is designed to interact well with both patients and health care professionals. The authentication system will help medical facilities comply with the HIPAA regulations regarding protection and privacy of medical records and accountability issues. The HIPAA regulations require all patient data access to be logged. This is done in order to provide accountability (audit trail); anyone who accesses the patient records is held responsible for what they see and do.



(a) **(b)**

Fig. 7. Unresolved error after fusion. The graphs show the x and y signals of signature and amplitude of voice samples plotted against time of two different users (a) and (b). The signature signals are the upper plot while the voice waveforms are depicted below. The signature score between (a) and (b) is 0.979, indicating a genuine signature. The normalized voice score between (a) and (b) is 0.06, indicating a genuine voice. Fusing the scores together shows the user to be genuine

Accordingly, this system gives a much higher confidence in the access logs because it is very likely that the individual who logged into the system is the same as the enrolled user. To combine the voice and signature modalities, we used fusion at the matching score level and, in particular, used the weighted sum rule. Using both the modalities gives higher accuracy than either individual modality and also makes spoofing of the system a much more difficult task. Thresholds can be adjusted in this system in order to achieve the desired security in this application domain.

Acknowledgments

This work was supported by the MSU CyberSecurity Initiative. We would like to acknowledge the help of Dr. Michael Zaroukian for providing us excellent guidance during the course of this project.

References

1. A^4 Health Systems. A^4 Health Systems Electronic Medical Record Solutions. http://www.a4healthsystems.com/.
2. George J. Annas. *The Rights of Patients*. Southern Illinois University Press, Carbondale, Illinois, 2004.
3. BCBSRI. Blue Cross Blue Shield of Rhode Island. https://www.bcbsri.com.
4. Business Wire. http://www.businesswire.com
5. D'Arcy Guerin Gue. The HIPAA Security Rule (NPRM): Overview. http://www.hipaadvisory.com/regs/securityoverview.htm.
6. HHS. Protecting the Privacy of Patients' Health Information. http://www.hhs.gov/news/facts/privacy.html.

7. A. K. Jain, Friederike D. Griess, and Scott D. Connell. On-line Signature Verification. *Pattern Recognition*, 35(12):2963–2972, December 2002.
8. A. Jain, K. Nandakumar, A. Ross. Score Normalization in Multimodal Biometric Systems. To appear in *Pattern Recognition*, 2005.
9. A. K. Jain and A. Ross. Information Fusion in Biometrics. *Pattern Recognition Letters*, 24(13):2115–2125, September 2003.
10. A. K. Jain and A. Ross. Multibiometric Systems. *Communications of the ACM*, 47(1):34–40, January 2004. Special Issue on Multimodal Interfaces.
11. Nuance. Nuance Corporation http://www.nuance.com.
12. S. Prabhakar, S. Pankanti, and A.K. Jain. Biometric Recognition: Security & Privacy Concerns. *IEEE Security & Privacy Magazine*, 1(2):33–42, March-April 2003.
13. Lawrence R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of IEEE*, vol. 77, No. 2, 257-286, Feb. 1989
14. University of South Alabama Health System. http://www.southalabama.edu/usahealthsystem/.
15. S. Krawczyk. User Authentication using On-line Signature and Speech. MS Thesis, Michigan State University, Dept. of Computer Science and Engineering (May 2005).
16. D. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, and Gerhard Rigoll. SVC2004: First International Signature Verification Competition. *Proceedings of the International Conference on Biometric Authentication (ICBA)* Hong Kong, 15-17 July 2004.