DYNAMIC SIGNATURE VERIFICATION WITH TEMPLATE PROTECTION USING HELPER DATA

Manuel R. Freire, Julian Fierrez and Javier Ortega-Garcia

Biometric Recognition Group - ATVS,
Escuela Politecnica Superior, Universidad Autonoma de Madrid
C/ Francisco Tomas y Valiente 11, E-28049 Madrid, Spain
{m.freire, julian.fierrez, javier.ortega}@uam.es

ABSTRACT

A biometric template protection system for dynamic signature verification is presented. The approach uses auxiliary (helper) data that allows the matching with secure templates but do not provide information to a potential attacker. The performance of the proposed system is evaluated using the MCYT signature database comprising 330 users, with 25 genuine signatures and 25 skilled forgeries per user. The results show similar performance compared to the baseline unprotected system. However, the security of the proposed system against attacks to the template database is significantly higher.

Index Terms— Template protection, signature verification, helper data.

1. INTRODUCTION

With the growth in scale of biometric systems deployed in the last years, privacy issues related to the protection of the biometric data have emerged as a crucial challenge for the widespread use of biometric solutions. The protection of biometric patterns requires actions that should enhance data resilience against attacks while allowing the matching to be performed efficiently.

In biometric systems, a template with the features extracted from the registration of the user (enrollment) is usually stored in a database. Unprotected templates can reveal partial or complete information regarding the registered biometric, therefore becoming a threat to users' privacy and the security of the system [1].

In this context, a number of research works have proposed biometric template protection systems. The *fuzzy vault* scheme [2] allows the matching of two sufficiently similar patterns in a transformed domain. In this construction, a secret (typically, a random binary key) is encoded using an unordered set of points A, resulting in an indivisible vault V. The original secret can only be reconstructed if another set B is presented and overlaps substantially with A. The fuzzyness of this construction fits well with the intra-variability of biometrics. Its template protection capacity comes from the fact that the only value that is stored in the enrollment database is a one-way-transformed version of the biometric pattern.

When working with protected templates, the matching typically takes place in the transformed domain. However, the matching of transformed patterns is much more difficult than in the non-protected feature space. To overcome this problem, the use of helper data has been proposed [3, 4], where some auxiliary data is stored along with the protected template in order to make the matching of protected templates easier. The helper data should not reveal useful information to a potential attacker. Practical helper data systems have been proposed for several biometric traits, such as 3D face [5].

In the current contribution we propose a template protection system based on helper data for dynamic signature verification. Within biometrics, handwritten signature has interesting applications in authentication and identity management, due to its widespread social and legal acceptance [6, 7]. We use dynamic (or on-line) signature verification, which is characterized by the availability of information of the signature realization, such as the position trajectory and the pressure signal over time.

This paper is structured as follows. In Sect. 2 we explain the framework for template protection using helper data. In Sect. 3, the application to dynamic signature templates is showed. The experimental framework and results are presented in Sect. 4. Finally, our conclusions are drawn in Sect. 5.

2. TEMPLATE PROTECTION BASED ON HELPER DATA

The proposed template protection system is based on the helper data system (HDS) proposed in [3, 5]. The HDS allows the matching of an enrolled pattern X with a test pattern Y, with the help of some auxiliary data (helper data). The helper data is extracted from the enrolled pattern and should not reveal in-

This work has been supported by Spanish Ministry of Education and Science (TEC2006-13141-C03-03) and BioSecure NoE (IST-2002-507634). M. R. F. is supported by a FPI Fellowship from Comunidad de Madrid. J. F. is supported by a Marie Curie Fellowship from European Commission.

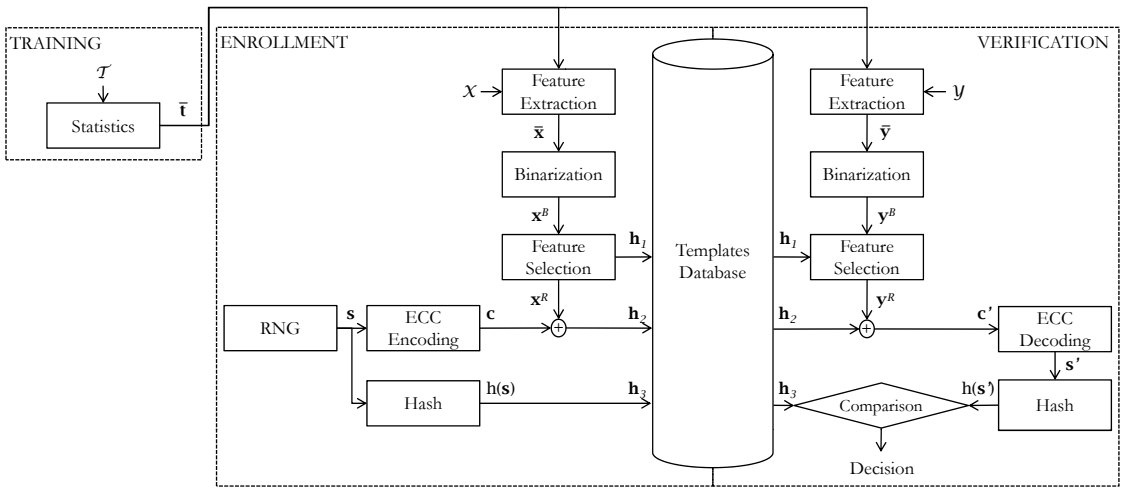

Fig. 1. Architecture of the helper data system, adapted from [5].

formation to a potential attacker. When the pattern Y emerges from the same source than X (e.g., a genuine biometric verification claim), Y can be seen as a noisy version of X. The HDS allows the correction of a given number of errors in the test pattern, by using a binary BCH error correcting code [8].

In Fig. 1 we show the architecture of the helper data system, which is divided into three phases: training, enrollment, and verification.

2.1. Training

In the training stage, a global model is built with statistics of training data from real users. Given a set of training real-valued feature vectors $\mathcal{T} = \{\mathbf{t}_i\}$, $i = 1, \ldots, N_T$, of dimension M, the mean feature vector $\overline{\mathbf{t}}$ is computed as $\overline{\mathbf{t}} = (1/N_T) \sum_{i=1}^{N_T} \mathbf{t}_i$.

2.2. Enrollment

The enrollment stage consists of a series of steps that produce a protected template with its associated helper data. First, the N_E enrollment feature vectors in the set $\mathcal{X} = \{\mathbf{x}_i\}$ are averaged extracting the mean feature vector $\overline{\mathbf{x}}$. Then each component of $\overline{\mathbf{x}}$ is binarized using each corresponding component of $\overline{\mathbf{t}}$ as the threshold, reaching the binary vector $\mathbf{x}^B$. Then, a selection of the most reliable features for the given enrollment is performed (details in Sect. 3.2), producing a reliable binary feature vector $\mathbf{x}^R$, of dimension $L \leq M$. The index of the selected features are stored as the helper feature $\mathbf{h}_1$.

On the other hand, a random binary string $\mathbf{s}$ is generated and encoded using an Error Correcting Code (ECC) such as the binary BCH code [8]. The encoded string $\mathbf{c}$ is XOR-ed with $\mathbf{x}^R$, producing the helper feature $\mathbf{h}_2$.

Finally, the random string $\mathbf{s}$ is hashed using a cryptographic hashing function, such as SHA-256, and the result is stored as the helper feature $\mathbf{h}_3 = h(\mathbf{s})$.

2.3. Verification

In the verification stage, N_V feature vectors from the verification set $\mathcal{Y} = \{\mathbf{y}_i\}$ are averaged into a mean feature vector $\overline{\mathbf{y}}$ (note that the usual case in biometric verification will be $N_V = 1$). A binarized feature vector is produced as in the enrollment stage, and then it is reduced using the selection stored in $\mathbf{h}_1$, producing $\mathbf{y}^R$. This reliable binary feature vector is then XOR-ed with $\mathbf{h}_2$, and the result $\mathbf{c}'$ is decoded using the ECC. The decoded string $\mathbf{s}'$ is transformed using the cryptographic one-way transform used in the enrollment. The comparison between this hashed value $h(\mathbf{s}')$ and $\mathbf{h}_3 = h(\mathbf{s})$ determines the final accept/reject decision of the system.

3. SECURE DYNAMIC SIGNATURE TEMPLATES

We present a system for template protection in dynamic signature verification based on the helper data system presented in Sect. 2, using global feature vectors based on statistical information of the signature.

3.1. Feature extraction

We use an on-line signature representation based on global features [7]. In particular, a 100-dimensional global feature vector is extracted from each on-line signature [9], including features based on timing information, number of strokes, geometry, pen trajectory, pressure over time, etc.

3.2. Feature selection

At the enrollment stage, the reliability measure proposed in [5] is used in order to select the most reliable features. For each

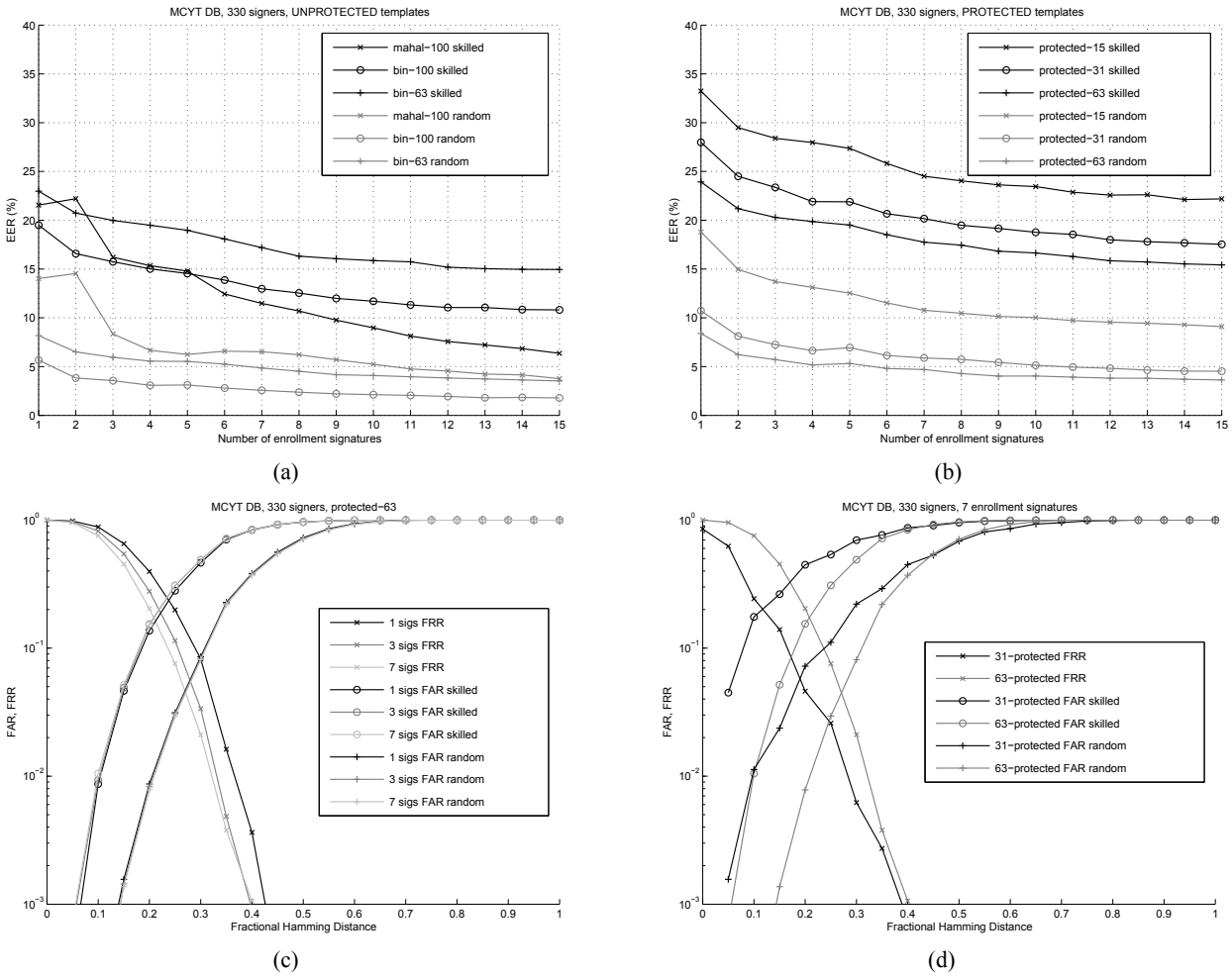

Fig. 2. Performance results of the template protection system for dynamic signature verification.

feature, the reliability is computed as:

$$r_j = \frac{|\bar{t}_j - \bar{x}_j|}{\sigma_j}$$

for all the M available features (100 in our case), where $\bar{t}_j$ is the training value of the j-th feature, and $\bar{x}_j$ and σ_j are the mean and the standard deviation of the j-th feature of the signatures presented at enrollment, respectively.

4. EXPERIMENTS

4.1. Database and experimental protocol

The MCYT on-line signature corpus is used for the experiments [10]. This database contains 330 users with 25 genuine signatures and 25 skilled forgeries per user, captured in four acquisition sites. Forgers were asked to imitate after observing the static image of the signatures to imitate, tried to copy them at least 10 times, and then they wrote the forgeries naturally without breaks or slowdowns.

For the experiments presented here, we have followed a 3-fold cross-validation strategy. The database has been divided into two sets: a *training* set, formed by one third of the users (users 1, 4, 7, ...; 2, 5, 8, ...; and 3, 6, 9, ... for the three iterations of the cross-validation procedure, respectively), and an *evaluation* set, with the remaining ones. For each user in the evaluation, the enrollment has been conducted using $1 \leq N \leq 15$ genuine signatures. False Rejection (FRR) and False Acceptance Rates (FAR) for 1-signature verification were obtained with the last 10 genuine signatures (FRR), the 25 skilled forgeries (FAR skilled forgeries) and one genuine signature from the rest of the users avoiding symmetric matches (FAR random forgeries). Equal Error Rates (EER) are also reported in some experiments [7].

Although the final accept/reject decision of the system is based on the comparison between $h(\mathbf{s})$ and $h(\mathbf{s}')$, we generate the matching scores for our experiments using the fractional Hamming distance (FHD) between $\mathbf{x}^R$ and $\mathbf{y}^R$, which highly simplifies the experiments. This can been done because $h(\mathbf{s}) = h(\mathbf{s}')$ iff $\mathbf{s} = \mathbf{s}'$ iff $\mathrm{FHD}(\mathbf{c}, \mathbf{c}') \leq n$, where n is

Table 1. Verification performance of different protected configurations for 7 enrollment signatures. Note that the number of security bits is equal to the length of the random string $\mathbf{s}$.

Configuration	EER-skilled (%)	EER-random (%)	Security (bits)	FRR (%)	FAR-skilled (%)	FAR-random (%)
protected-15	24.52	10.77	7	18.95	40.04	11.73
			11	51.68	19.68	5.22
protected-31	20.16	5.89	11	10.73	34.85	4.22
			16	36.05	15.49	1.05
protected-63	17.75	4.70	16	24.42	11.61	0.44
			18	33.83	7.93	0.23

the number of correctable bits by the ECC, and:

$$\text{FHD}(\mathbf{c}, \mathbf{c}') = \text{FHD}(\mathbf{x}^R \oplus \mathbf{h}_2, \mathbf{y}^R \oplus \mathbf{h}_2) = \text{FHD}(\mathbf{x}^R, \mathbf{y}^R)$$

In this way, we are able to estimate the performance of the complete signature verification system using protected templates without actually simulating the following modules: random string generation, ECC encoding/decoding, and hashing.

4.2. Results

Performance results are presented in Fig. 2. In Fig. 2 a) we first present the EER of different unprotected configurations, one based on the Mahalanobis distance between the real-valued feature vectors and the other two with a binarization of the features with the matching scores based in the Hamming distance, for all the available 100 features and for the best 63 according to the individual ranking in [9]. We observe that the system based on Mahalanobis distance shows the best performance when considering skilled forgeries with more than 5 enrollment signatures. With the other configurations, the binarization achieves better results in terms of the EER.

The use of binary BCH codes imposes a length for the encoded random string of $2^n - 1$, which must be smaller than the feature vector size (100 bits). Therefore, the protected configurations selected for the experiments have been 15, 31 and 63 bits, which are compared in Fig. 2 b). The results show that protected-63 is the best operating point in terms of the EER, both for random and skilled forgeries. We also observe that the protected template performs similarly than the unprotected system in the bin-63 configuration. In Fig. 2 c), FRR, FAR-skilled and FAR-random curves are displayed for 1, 3 and 7 enrollment signatures for the protected-63 configuration. Finally, Fig. 2 d) shows the error curves for different configurations with 7 enrollment signatures.

Although the EER is a good measure to compare different configurations, the template protection scheme that we used in this paper does not allow to operate at the EER point (i.e., where FAR and FRR are equal). This is because of the nature of the BCH codes introduced in the system, which restrict the maximum number of correctable bits [8]. Therefore, to measure the real performance of the template protection system we provide the FAR and FRR for fixed threshold values. A summary of the real performance of the best configurations is presented in the Table 1.

5. CONCLUSIONS

A biometric template protection system based on helper data for dynamic signature verification has been proposed. The performance has been evaluated using the MCYT signature database, comprising 16,500 signatures. The protected templates are more robust against attacks to the template database, while the results show similar performance in terms of the EER compared to baseline unprotected configurations. With the proposed system, a security of 18 bits is achieved with an FRR of 33.83% and a FAR of 7.93% and 0.23% for skilled and random forgeries, respectively.

6. REFERENCES

[1] R. Cappelli, D. Maio, A. Lumini, and D. Maltoni, "Fingerprint image reconstruction from standard templates," *IEEE Trans. PAMI*, vol. 29, no. 9, pp. 1489–1503, 2007.

[2] Ari Juels and Madhu Sudan, "A fuzzy vault scheme," *Design Codes and Cryptography*, vol. 38, no. 2, pp. 237–257, 2006.

[3] P. Tuyls, A. H. M. Akkermans, T. A. M. Kevenaar, G. J. Schrijen, A. M. Bazen, and R. N. J. Veldhuis, "Pratical biometric authentication with template protection," in *Proc. AVBPA*. 2005, vol. 3617 of *LNCS*, pp. 436–446, Springer.

[4] Umut Uludag and Anil K. Jain, "Securing fingerprint template: Fuzzy vault with helper data," in *Proc. CVPRW*. 2006, p. 163, IEEE Computer Society.

[5] E. J. C. Kelkboom, B. Gökberk, T. A. M. Kevenaar, A. H. M. Akkermans, and M. van der Veen, "3D Face: Biometric template protection for 3D face recognition," in *Proc. ICB*. August 2007, vol. 4642 of *LNCS*, pp. 566–573, Springer.

[6] Rejean Plamondon and Sargur N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. PAMI*, vol. 22, no. 1, pp. 63–84, 2000.

[7] Julian Fierrez and Javier Ortega-Garcia, *Anil K. Jain, Arun Ross and Patrick Flynn (eds.): Handbook of Biometrics*, chapter On-line signature verification, Springer, 2007.

[8] Shu Lin and Daniel J. Costello, *Error Control Coding*, Prentice-Hall, Inc., second edition, 2004.

[9] Julian Fierrez-Aguilar et al., "An on-line signature verification system based on fusion of local and global information," in *Proc. AVBPA*. 2005, vol. 3617 of *LNCS*, pp. 523–532, Springer.

[10] Javier Ortega-Garcia et al., "MCYT baseline corpus: A bimodal biometric database," *IEE Proc. Vision, Image and Signal Processing*, vol. 150, no. 6, pp. 395–401, 2003.