In Proceedings WIC Symposium on Information
Theory in the Benelux, Boekelo, May 24-25, 2012

Optimized Helper Data Scheme for Biometric Verification Under Zero Leakage Constraint

Joep de Groot Jean-Paul Linnartz

Eindhoven University of Technology
Signal Processing Systems, Faculty of Electrical Engineering
P.O. Box 513, 6500 MB Eindhoven, The Netherlands

{j.a.d.groot, j.p.linnartz}@tue.nl

Abstract

In biometric verification, special measures are needed to prevent that a dishonest verifier can steal privacy-sensitive information about the prover from the template database. We introduce an improved version of the zero leakage quantization scheme, which optimizes detection performance in terms of the false rejection ratio. Our scheme ensures zero leakage, that is, zero mutual information between auxiliary verification data and the protected enrolled secret and guarantees a uniformly distributed secret. Moreover, our solution replaces the helper data in the template database by a user specific threshold and eliminates the pre-distortion in the verification phase, which allows very rapid verification algorithms. Although the false rejection rate was only reduced by 20% for well correlated biometric features, it can be shown that this particular scheme achieves optimal detection under our requirements.

1 Introduction

The application of biometrics for authentication or identification purposes remains a challenging problem. A secret derived from the biometric features can be protected against theft by applying a cryptographic hash function. However, this protection requires the secret to be exactly reproducible, which is difficult with noisy biometrics.

To handle the variations caused by the noise, various techniques have been proposed that store user-dependent data to reproduce the enrolled secret. At the same time this data can leak information about the enrolled secret, which effectively weakens the secret. For an adversary who possesses this auxiliary data it becomes easier to guess the enrolled secret. Therefore we wish to minimize the information in the auxiliary data about the enrolled secret. In theory such leakage even can be made zero [1, 2].

Template protection, as described above, suffers from a reduced detection performance compared to continuous classifiers [3]. Especially for a practical application in which the number of biometric features is limited, one wishes to minimize the false rejection probability, since the alternative, namely applying an error correcting code that can handle a large number of errors, will reduce the number of effective bits in the secret even further.

In this work we introduce a modification of the zero leakage quantization scheme [1], which provides an improved solution in terms of false rejections, while maintaining

zero leakage. Moreover, the solution requires a less complex verification phase, which allows a very rapid verification procedure. Although our new solution does not require a specific distribution of the biometric features, we quantify the performance based on a Gaussian distribution.

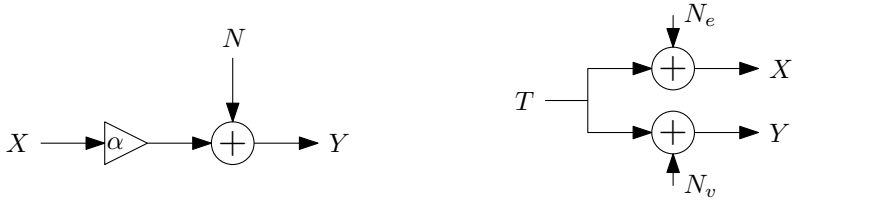
The paper first gives an overview of various biometric descriptions that exist in literature and describes under what conditions these descriptions are equivalent. Although the results are highly intuitive, we think that this gives a better understanding of the properties of the biometric features we have to deal with. The derived relations allow us to benchmark results with papers using a different biometric description.

The remainder of the paper is structured as follows. In section 3 we will derive the constraints required to obtain zero leakage. This is the basis of our optimization approach and low-complexity implementation explained in section 4. The detection results are given in section 5, followed by a discussion and conclusion on our results in sections 6 and 7 respectively.

2 Gaussian biometric models

In literature there appears to be a common generic model, but with various mathematical descriptions. Yet, some of these descriptions only cover a subset of the entire class of descriptions by the generic model. A common modality is the assumption of M i.i.d. features, at least after pre-processing. The values observed during enrollment are commonly noted as X, whereas the verification vector is represented by Y, which are both of length M.

However, the various papers use different descriptions for the relation between the enrollment and verification sample. We have identified at least four approaches which we will refer to as the 'additive channel', a common communication theoretical concept [4, Chap. 9], the 'hidden template' [5], the 'within-/between-class distribution' [3] and the 'joint Gaussian' [6].



(a) The 'additive channel' description with additive noise N and attenuation α

(b) The 'hidden template' description with two additive noise channels

Figure 1: Biometric feature relation descriptions

The 'additive channel' description, depicted in Figure 1(a), is simply a noisy channel with independent additive Gaussian noise. Since this introduces a different variance at the in- and output, i.e. between the enrollment and verification samples, an additional attenuation α can be introduced. The variables in this description are distributed as

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2) \qquad \text{and} \qquad N \sim \mathcal{N}(\mu_N, \sigma_N^2) \tag{1}$$

Table 1: Overview of biometric descriptions found in literature

Model	$\mathbb{E}X$	$\mathbb{E}Y$	$\mathbb{E}X^2$	$\mathbb{E}Y^2$	$\mathbb{E}XY$	$\mathbb{E}Y\|X$	SNR
Channel	μ_X	$\mu_X + \mu_N$	σ_X^2	$\alpha^2\sigma_X^2 + \sigma_N^2$	$\alpha\sigma_X^2 + \sigma_N^2$	αx	$\alpha^2\sigma_X^2/\sigma_N^2$
Template	$\mu_T + \mu_{N_e}$	$\mu_T + \mu_{N_v}$	$\sigma_T^2 + \sigma_{N_e}^2$	$\sigma_T^2 + \sigma_{N_v}^2$	σ_T^2	Eq (6)	$\sigma_T^2/\sigma_{N_v}^2$
W/B-cls	0	0	σ_t^2	σ_t^2	σ_b^2	$(\sigma_b^2/\sigma_t^2)x$	σ_b^2/σ_w^2
Joint G	0	0	1	1	ρ	ρx	$\rho/(1-\rho)$

The 'hidden template' description, depicted in Figure 1(b), consists of two parallel channels, both with independent additive Gaussian noise N_e and N_v for the enrollment and verification sample respectively. The common input of both channels is the biometric template T. The variables in this description are distributed as

$$T \sim \mathcal{N}(\mu_T, \sigma_T^2), \qquad N_e \sim \mathcal{N}(\mu_{N_e}, \sigma_{N_e}^2) \qquad \text{and} \qquad N_v \sim \mathcal{N}(\mu_{N_v}, \sigma_{N_v}^2), \tag{2}$$

The notion of a 'within-/between-class distribution', with variance σ_w^2 and σ_b^2 respectively, is a special case of the 'hidden template', in which

$$\begin{cases} \mu_T &= 0 \\ \mu_{N_e} = \mu_{N_v} &= 0 \end{cases} \qquad \text{and} \qquad \begin{cases} \sigma_T^2 &= \sigma_b^2 \\ \sigma_{N_e}^2 = \sigma_{N_v}^2 &= \sigma_w^2 \\ \sigma_b^2 + \sigma_w^2 &= \sigma_t^2 \end{cases} \tag{3}$$

In this work we prefer to use the 'joint Gaussian', which was introduced by [6]. This is described as a probability density function with parameter ρ, which directly relates to the Signal-to-Noise ratio (SNR)

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right) \tag{4}$$

The joint probability density function allows us to derive the conditional density function

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(y-\rho x)^2}{2(1-\rho^2)}\right) \tag{5}$$

which is required to calculate the genuine verification error. Note that this conditional distribution is symmetric around ρx.

An overview of the description parameters is shown in Table 1. The conditional expectation $\mathbb{E}Y|X$ for the 'Hidden template' description can be derived by using Bayes' rule

$$\mathbb{E}Y|X = \mathbb{E}T|X + \mu_{N_v} = \frac{(x - \mu_{N_e})\sigma_T^2 + \mu_T\sigma_{N_e}^2}{\sigma_T^2 + \sigma_{N_e}^2} + \mu_{N_v}. \tag{6}$$

Although the relations between the descriptions are straightforward, they are relevant since it enables us to compare the results obtained with different biometric descriptions.

3 Zero leakage constraint

In [1] we have introduced the pre-distortion function to achieve zero leakage. Once this function is set to the cumulative distribution function (CDF) of the biometric feature x, i.e.

$$u = g(x) = F_X(x) \qquad \text{and} \qquad v = g(y) = F_X(y), \tag{7}$$

the transformed features u and v will be uniformly distributed.

Subsequently, helper data defined in the transformed domain will be uniformly distributed. However, a uniform distribution is not required to obtain zero leakage. The actual constraint is given on the conditional helper data function, namely

$$f_W(w|s = n) = f_W(w) \qquad \forall n \in \{0, \ldots, 2^N\} \tag{8}$$

as can be seen from the definition of leakage

$$\begin{aligned}
I(W;S) &= H(S) - H(S|W) && (9)\\
&= H(S) - \int_{-q}^{q} H(S|W = w) f_W(w)\, dw && (10)\\
&= H(S) - \int_{-q}^{q} \sum_n \frac{f_W(w|S = n)}{f_W(w)} \mathrm{P}(S = n) \cdot && (11)\\
&\qquad \log_2 \left(\frac{f_W(w|S = n)}{f_W(w)} \mathrm{P}(S = n) \right) f_W(w)\, dw && (12)\\
&= H(S) - H(S) \int_{-q}^{q} f_W(w)\, dw && (13)\\
&= H(S) - H(S) = 0 && (14)
\end{aligned}$$

Therefore, if the constraint of Equation (8) is satisfied, we can apply all possible transformations to the helper data and maintain zero leakage, since a transformation will not undo it. The transformation to a uniform distribution is just a convenient step to make sure that the conditional distributions are equal.

4 Distortion adjustment

For the moment we will limit ourselves to a verification scheme that extracts just one bit per dimension, i.e. $N = 1$. However, our method can be extended to a scheme in which multiple bits are assigned per dimension, but we consider this to be beyond the scope of this paper.

In the verification phase we added the helper data w to our distorted sample $v = g(y)$ and compared it with $1/2$ to determine the bit value. However, for a single bit

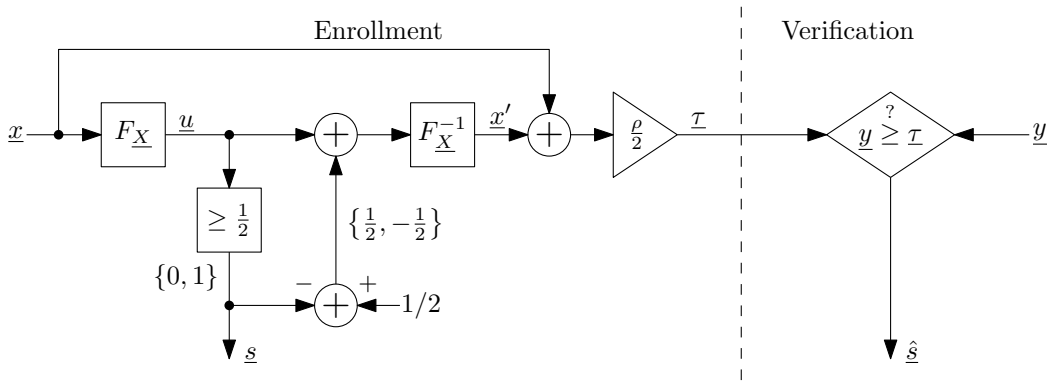


Figure 2: The improved zero leakage quantization scheme, which determines the optimal threshold during enrollment instead of additive helper data, so we store/transmit τ which is the quantization threshold during verification.

this can be done differently, as is shown for $\hat{s} = 0$

$$v + w = F_X(y) + w < \frac{1}{2} \tag{15}$$

$$F_X(y) < \frac{1}{2} - w \tag{16}$$

$$y < F_X^{-1}\left(\frac{1}{2} - w\right) = \tau \tag{17}$$

So, instead of sending w we can transfer helper data in the form of a user-specific threshold τ to the verifier. This still achieves zero leakage according to our conclusion of the preceding section. However, it requires the inverse pre-distortion function $F_x^{-1}(w)$ to exist. This is no problem in practice, but excludes discontinuous distributions.

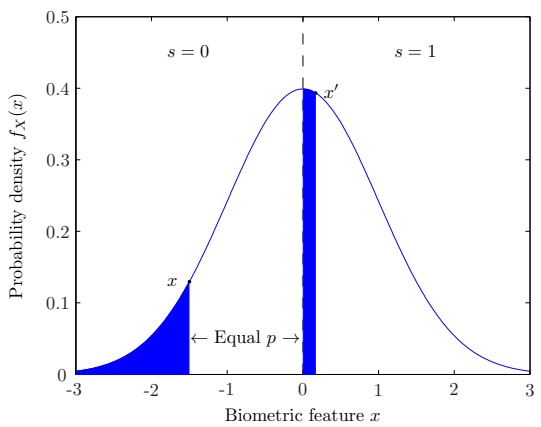
For each enrollment sample x, we introduce a sister point x' with the same helper data $w = w'$, but with opposite secret $s' = \neg s$. The relation between the two points is given by

$$F_X(x') = \begin{cases} F_X(x) + \frac{1}{2} & x < 0 \\ F_X(x) - \frac{1}{2} & x \geq 0 \end{cases} = F_x(x) - \left(s - \frac{1}{2}\right) \tag{18}$$

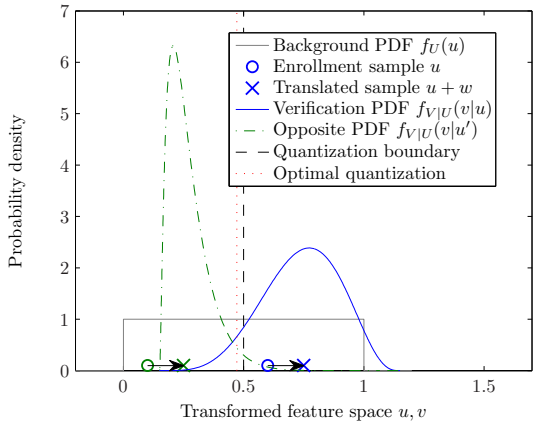
The relation between these points is graphically explained in Figure 3(a). Analogously we define $u' = u - (s - 1/2)$ the corresponding sister point in the transformed domain.

Any helper data distribution function that satisfies Equation (8) maintains zero leakage. This can be used to optimize detection performance, since the pre-distortion function $F_X(x)$ introduces a strong unequal scaling between centered and deviating verification sample distributions (equation (5)). A distribution close to zero yields a much wider distribution in the transformed domain than the distributions further away from the center.

Figure 3(b) shows that a threshold at $u = 1/2$ is not optimal for this verification distribution. In this example the large probability mass below $u = 1/2$ leads to a high false rejection ratio. Ideally we would like our helper data to shift the verification distribution away from the quantization boundary in order minimize this error. However,



(a) Given x, the sister point x' can be found in the other quantization interval with an equal probability mass from the beginning of the interval up to that point. The relation between x and x' is thus given by the CDF of the features.



(b) Example of distorted verification sample distributions with a non-optimal quantization boundary during verification. For this example $u = 0.6$ and $\rho = 0.9$, i.e. SNR ≈ 9.5 dB.

Figure 3: Background and noise distributions.

this leads to a trade-off, since the sister distribution at u' will inevitable be moved closer to the quantization boundary.

Centering the distorted distributions on the quantization interval is not optimal in terms of error probability for a genuine user. In order to improve our performance we try minimize the error probability for a genuine user. This error is defined as the probability of observing a different secret during verification, hence

$$p^g_{e,s} = \mathrm{P}(s \neq \hat{s}) = \begin{cases} \mathrm{P}\left(y \geq \tau\right) & s = 0 \\ \mathrm{P}\left(y < \tau\right) & s = 1 \end{cases} = \begin{cases} 1 - F_{y|x}(\tau|x) & s = 0 \\ F_{y|x}(\tau|x) & s = 1 \end{cases} \tag{19}$$

in which $F_{y|x}(y|\,x)$ is the conditional Cumulative Distribution Function (CDF) of our biometric feature. Moreover, we use the above defined threshold τ, so we can do the optimization in the original untransformed domain. For the optimization we solve the following equation

$$\tau = \arg\min_{\tau}(\underbrace{F_{Y|X}(\tau|\,x)}_{p^g_{e,1}} + \underbrace{1 - F_{Y|X}(\tau|\,x')}_{p^g_{e,0}}), \tag{20}$$

in which x' is the biometric sister point of x.

The biometric samples x and x' in Equation (20) can be interchanged, but this yields the same result. For this particular derivation we assumed $u \geq 1/2$, i.e. $s = 1$ and $s' = 0$. An example of this situation is depicted in Figure 3(b).

Equation (20) can be solved by taking the derivative with respect to τ and setting it equal to zero

$$\frac{d}{d\tau}\left[F_{Y|X}(\tau|\,x) + 1 - F_{Y|X}(\tau|\,x')\right] = 0, \tag{21}$$

which yields the following solution for our model described in Equation 5

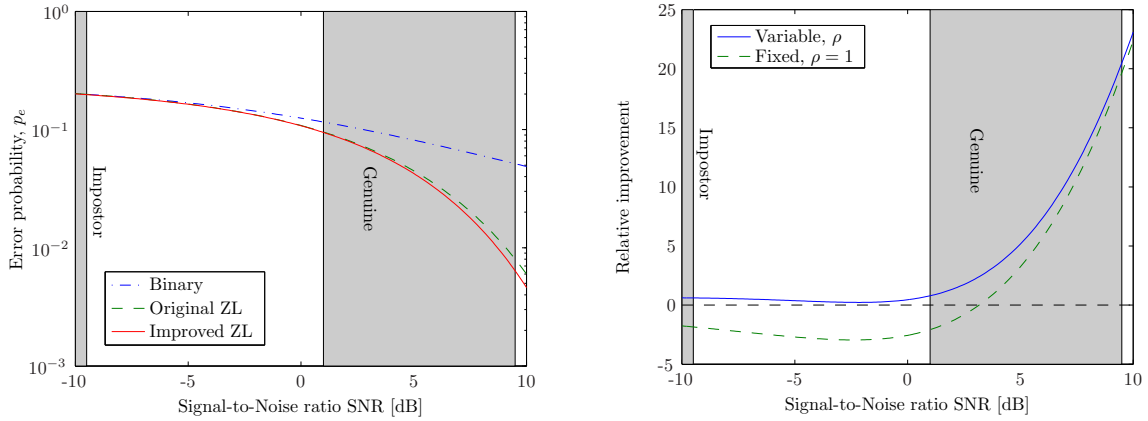
$$\tau = \rho \frac{x + x'}{2}. \tag{22}$$

A schematic representation of this verification scheme is depicted in Figure 2.

Although the math is seemingly simple, the solution gives an important insight. The auxiliary data, in our case a threshold, does not provide any information to the adversary on which side the original sample was. This can be understood by noting that x and x' can simply be swapped while obtaining the same result ($\tau = \tau'$).

5 Detection performance

The detection improvement, i.e. the reduction in genuine verification error, is for most features comparable to that of the initial zero leakage scheme [1] as can be seen from Figure 4(a). However, an relative improvement of approximately 20% is obtained for well correlated biometric features, i.e. $\rho = 0.95$ or SNR ≈ 9.7 dB, as is depicted Figure 4(b). Both figures indicate the approximate regions a genuine and impostor feature is likely to be in. This assumes the genuine user's features to have a correlation $0.75 < \rho < 0.95$, while the impostor's features have correlation $\rho = 0$.



(a) in comparison to other schemes

(b) relative w.r.t. the original scheme

Figure 4: Detection performance of the improved zero leakage scheme

The influence of parameter ρ can directly be noticed from the relative detection improvement. In case one would fix this parameter to $\rho = 1$, i.e. assuming a very well correlated biometric, the detection rate for inferior biometric features, that is features with a relatively low correlation ρ or SNR, are reduced as can be seen in Figure 4(b). This is caused by the fact that the features during verification are distributed around ρx according to our model, which requires the scaling by ρ.

6 Discussion

An important limitation shared with the original zero leakage scheme is that the system designer needs to know the distribution of the biometric features in advance. Moreover,

the improved scheme even requires the existence of a continuous inverse of this function to exist. However, this might not be a serious limitation since in general the biometric feature distribution will be continuous, which implies that the inverse CDF will also be defined everywhere on the domain $(0, 1)$.

A limitation of our optimization is that a viable solution can only be found for a single bit per dimension ($N = 1$) and if the conditional CDF of the verification features is a convex function for $F_{X|Y}(x|y) \leq 1/2$, otherwise multiple solutions exist for Equation (20). Our model, Equation (5), perfectly fulfills this requirement, but a simple counter example is a Gaussian Mixture Model (GMM) with a local minimum in its PDF. We do not yet have a general solution to handle these kind of distributions.

7 Conclusion

We have proposed an optimized solution for a biometric verification system that offers zero leakage and a uniformly distributed secret. We have shown that the optimal solution can simply be implemented as a user-dependent threshold. Our solution reduced genuine verification errors by approximately 20% for well correlated biometric features.

8 Acknowledgment

The authors would like to thank Boris Škorić, Frans Willems and Niels de Vreede for the valuable discussions during this research.

References

[1] J.A. de Groot and J.-P.M.G. Linnartz. Zero leakage quantization scheme for biometric verification. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process*, May 2011.

[2] E.A. Verbitskiy, P. Tuyls, C. Obi, B. Schoenmakers, and B. Skoric. Key extraction from general nondiscrete signals. *Information Forensics and Security, IEEE Transactions on*, 5(2):269 –279, June 2010.

[3] E.J.C. Kelkboom, J. Breebaart, and R.N.J. Veldhuis. Classification performance comparison of a continuous and binary classifier under gaussian assumption. In *Proc of the 31st Symp on Inf Theory in the Benelux*, 2010.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., second edition, 2005.

[5] J.-P.M.G. Linnartz, P. Tuyls, and B. Skoric. *A Communication-Theoretical View on Secret Extraction*, chapter 4, pages 57–77. Security with Noisy Data. Springer, 2007.

[6] F.M.J. Willems and T. Ignatenko. Quantization effects in biometric systems. In *Information Theory and Applications Workshop*, 2009.