On the security of XOR-method in biometric authentication systems

On the Security of the XOR-Method in Biometric Authentication Systems

Tanya Ignatenko
Eindhoven University of Technology
Elec. Eng. Dept., Sign. Proc. Group
Eindhoven, The Netherlands

Frans Willems
Eindhoven University of Technology
Elec. Eng. Dept., Sign. Proc. Group
Eindhoven, The Netherlands

Abstract

A biometric authentication system can be partitioned into a layer that extracts common randomness out of a pair of related biometric sequences [Maurer, 1993] and a layer that masks a key sequence with this common randomness. We will analyze the performance of such a layered system first, and will show that an alternative method, the XOR-technique, is not always secure.

1 Introduction

Nowadays with the introduction of biometric technologies in daily life the importance of secure storage and communication of data in biometric systems increased. In this paper we concentrate on the security of XOR-based biometric authentication systems.

Biometric authentication is the process of establishing the identity of an individual using measurements of his/her biological characteristics. The attempts to create secure authentication scheme led to so called XOR-schemes [1]. In this work the biometric data is assumed to be an independent and identically distributed (i.i.d.) sequence, however, in practice, this is rarely a realistic assumption. Therefore, in this paper we analyse the impact of using non i.i.d. biometric sequences in XOR-schemes. Moreover, we consider methods to build secure authentication schemes by looking at authentication schemes as the ones partitioned into a layer that extracts common randomness out of a pair of related biometric sequences [2] and a layer that masks a key sequence with this common randomness. We show that only an XOR-scheme built using above principles will lead to an authentication scheme which is always secure.

2 The common randomness extraction layer

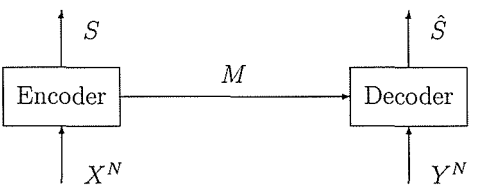

Figure 1: Randomness extraction.

Consider a system in which an encoder and a decoder have to extract the same random information S from enrollment and authentication biometric measurements X^N and Y^N, respectively. To ensure reliable communication the encoder sends helper information M to the decoder (see Figure 1). In such a system, the communication channel is assumed to be public. Using a random binning argument [3], we show that a system can be constructed such that roughly no information about the extracted randomness S is given to an eavesdropper that observes M, while ensuring that the decoder's version $\hat{S}$ of S is equal to S with probability roughly one. The amount of common randomness that can be extracted in this way is given by the following theorem.

Theorem 1 *For the common randomness extraction scheme, processing i.i.d. sequences, for each $\delta > 0$, for all N large enough, there exists a sequence of codes satisfying*

$$\begin{aligned} \Pr\{\hat{S} \neq S\} &\leq \delta, \\ H(S)/N &\geq I(X;Y) - \delta, \\ I(S;M)/N &\leq \delta. \end{aligned}$$

Conversely, there exists no secure ($I(S;M)/N \approx 0$) and reliable ($\Pr\{\hat{S} \neq S\} \approx 0$) scheme if $H(S)/N > I(X;Y)$.

Achievability: Fix an $\varepsilon > 0$. Now $\mathcal{A}_\varepsilon(X)$ and $\mathcal{A}_\varepsilon(X,Y)$ are the sets of typical and jointly typical sequences as defined in Cover and Thomas [4], based on the joint distribution of the XY-source.

We prove the achievability with a random binning argument. We assign to each sequence x^N a helper-label $m \in \{1, 2, \cdots, 2^{NR_h}\}$ with probability $\Pr\{M(x^N) = m\} = 2^{-NR_h}$. We also assign to each sequence x^N a randomness-label $s \in \{1, 2, \cdots, 2^{NR_s}\}$ with probability $\Pr\{S(x^N) = s\} = 2^{-NR_s}$.

The helper label $m(x^N)$ is sent to the decoder by the encoder. The decoder after having observed y^N looks for a unique sequence x^N with label m such that

$$(x^N, y^N) \in \mathcal{A}_\varepsilon(X,Y). \tag{1}$$

First problem is now to determine the decoder error probability averaged over the random binning:

$$\begin{aligned} \overline{P_{d,\varepsilon}} &\leq \Pr\{(X^N,Y^N) \notin \mathcal{A}_\varepsilon \cup \left(\cup_{x^N \neq X^N : (x^N,Y^N) \in \mathcal{A}_\varepsilon} M(x^N) = M(X^N)\right)\} \\ &\leq \Pr\{(X^N,Y^N) \notin \mathcal{A}_\varepsilon\} + \sum_{x^N \neq X^N : (x^N,Y^N) \in \mathcal{A}_\varepsilon} \Pr\{M(x^N) = M(X^N)\} \\ &\leq \Pr\{(X^N,Y^N) \notin \mathcal{A}_\varepsilon\} + |\{x^N : (x^N,Y^N) \in \mathcal{A}_\varepsilon\}| \cdot 2^{-NR_h} \\ &\leq \varepsilon + 2^{N(H(X|Y)+2\varepsilon)} \cdot 2^{-NR_h} \\ &\leq 2\varepsilon, \end{aligned} \tag{2}$$

for N large enough if $R_h = H(X|Y) + 3\varepsilon$.

The encoder wants x^N to be reconstructible from both the helper-label m and the randomness-label s. It looks for the unique sequence x^N with labels m and s such that

$$x^N \in \mathcal{A}_\varepsilon(X). \tag{3}$$

Now the encoder error probability averaged over the ensemble of random binnings satisfies

$$\begin{aligned} \overline{P_{e,\varepsilon}} &\leq \Pr\{X^N \notin \mathcal{A}_\varepsilon \cup \left(\cup_{x^N \neq X^N : x^N \in \mathcal{A}_\varepsilon} M(x^N) = M(X^N) \wedge S(x^N) = S(X^N)\right)\} \\ &\leq \Pr\{X^N \notin \mathcal{A}_\varepsilon\} + \sum_{x^N \neq X^N : x^N \in \mathcal{A}_\varepsilon} \Pr\{M(x^N) = M(X^N), S(x^N) = S(X^N)\} \\ &\leq \Pr\{X^N \notin \mathcal{A}_\varepsilon\} + |\{x^N : x^N \in \mathcal{A}_\varepsilon\}| \cdot 2^{-N(R_h+R_s)} \\ &\leq \varepsilon + 2^{N(H(X)+\varepsilon)} \cdot 2^{-N(R_h+R_s)} \\ &\leq 2\varepsilon, \end{aligned} \tag{4}$$

for N large enough if $R_s = I(X;Y) - \varepsilon$ (and $R_h = H(X|Y) + 3\varepsilon$).

Since $\overline{P_{d,\varepsilon}} + \overline{P_{e,\varepsilon}} \leq 4\varepsilon$ for N large enough, this implies that for N large enough there exist two random binnings such that $P_{d,\varepsilon} + P_{e,\varepsilon} \leq 4\varepsilon$. Now we focus on these codes for the rest of the proof.

Note that $H(M) \leq \log_2 2^{NR_h} = NR_h = N(H(X|Y) + 3\varepsilon)$ and that $H(S) \leq \log_2 2^{NR_s} = NR_s = N(I(X;Y) - \varepsilon)$. Now, using Fano's inequality in the last step where $\widehat{x_e^N}$ is the encoder's estimate of x^N, we find that

$$\begin{aligned} H(X^N) &= H(X^N, S, M) \\ &\leq H(S) + H(M) + H(X^N|S,M) \\ &= H(S) + H(M) + H(X^N|S,M,\widehat{X_e^N}) \\ &\leq H(S) + H(M) + 1 + NP_{e,\varepsilon} \log_2 |\mathcal{X}|. \end{aligned} \tag{5}$$

Hence the entropy of the common randomness

$$\begin{aligned} H(S) &\geq H(X^N) - H(M) - 1 - NP_{e,\varepsilon} \log_2 |\mathcal{X}| \\ &\geq NH(X) - N(H(X|Y) + 3\varepsilon) - 1 - NP_{e,\varepsilon} \log_2 |\mathcal{X}| \\ &\geq N(I(X;Y) - 3\varepsilon - 4\varepsilon \log_2 |\mathcal{X}|) - 1. \end{aligned} \tag{6}$$

From this we may conclude that for all N large enough $H(S)/N \geq I(X;Y) - \delta$ for properly chosen ε. Next we study the mutual information

$$\begin{aligned} I(S;M) &= H(S) + H(M) - H(S,M) \\ &= H(S) + H(M) - H(S,M,X^N) + H(X^N|S,M) \\ &= H(S) + H(M) - H(X^N) + H(X^N|S,M,\widehat{X_e^N}) \\ &\leq H(S) + H(M) - NH(X) + 1 + NP_{e,\varepsilon} \log_2 |\mathcal{X}| \\ &\leq N(2\varepsilon + 4\varepsilon \log_2 |\mathcal{X}|) + 1. \end{aligned} \tag{7}$$

Now we may conclude that $I(M;S)/N \leq \delta$ for all N large enough for properly chosen ε. Note that also $P_{d,\varepsilon} \leq \delta$ can be achieved in this way and that $\Pr\{\hat{S} \neq S\} \leq P_{d,\varepsilon}$. ⊠

Remark: When the source is jointly stationary ergodic a similar proof can be formulated. The difference is that now the typical sets are defined as in Cover [5].

Converse: We denote by $\widehat{S}$ the decoder's estimate of the common randomness S generated by the encoder. Then Fano's inequality yields

$$\begin{aligned} H(S|Y^N, M) &\leq H(S|Y^N, M, \widehat{S}) \\ &\leq H(S|\widehat{S}) \\ &\leq 1 + P_{s,\varepsilon} \log_2 |\mathcal{S}|, \end{aligned} \tag{8}$$

where $P_{s,\varepsilon} = \Pr\{\widehat{S} \neq S\}$ and S assumes values from $\mathcal{S}$. Note that $|\mathcal{S}| \leq |\mathcal{X}|^N$ since the encoder is deterministic. Next we consider

$$\begin{aligned} H(S) &= I(S;Y^N,M) + H(S|Y^N,M) \\ &\leq I(S;Y^N,M) + P_{s,\varepsilon}\log_2|\mathcal{S}| + 1 \\ &\leq I(S;M) + I(S;Y^N|M) + NP_{s,\varepsilon}\log_2|\mathcal{X}| + 1 \\ &\leq I(S;M) + H(Y^N) - H(Y^N|M,S,X^N) + NP_{s,\varepsilon}\log_2|\mathcal{X}| + 1 \\ &= I(S;M) + H(Y^N) - H(Y^N|X^N) + NP_{s,\varepsilon}\log_2|\mathcal{X}| + 1 \\ &= I(S;M) + NI(X;Y) + NP_{s,\varepsilon}\log_2|\mathcal{X}| + 1. \end{aligned} \tag{9}$$

For all large enough N and for $P_{s,\varepsilon} \downarrow 0$ we obtain the desired upper bound on the entropy of the common randomness per source symbol pair. ⊠

Remark: The converse result can also be reformulated for biometric sequences that are generated by a jointly stationary ergodic source.

3 The masking layer

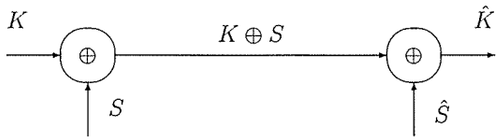

Figure 2: Masking layer.

In the previous section we described a procedure for secure randomness extraction from biometric data. However, to avoid cross-matching in the databases, to cancel compromised keys, a possibility to use different keys associated with the same biometric data X^N for authentication is required. That is why, the masking layer as in [6] (see Figure 3), which is based on the one-time pad principle, is introduced. In the masking layer, a uniform binary secret key K is generated for a biometric sequence X^N. The encoder transmits extra side information, which is the secret key K added modulo 2 to the common randomness S, to the decoder. Note that here we assume that the common randomness is a binary sequence. To perform the authentication procedure, the decoder adds modulo 2 the estimated common randomness $\widehat{S}$ extracted from Y^N to the obtained side information and uses the resulting secret key $\widehat{K}$ for authorization purposes.

Theorem 2 *If we use a masking procedure, based on a uniform binary key sequence, the system preserves its property of being secure, i.e.*

$$I(K; M, K \oplus S)/N \approx 0$$

if $H(K)/N \approx I(X;Y)$.

Proof: The proof corresponds to a masking layer added to the randomness extraction layer described in the achievability proof. Therefore the length of the key sequence is

NR_s and $H(K) = NR_s = N(I(X;Y) - \varepsilon)$. Note that $I(K; M, K \oplus S) = H(M, K \oplus S) - H(M, K \oplus S|K)$. We first consider

$$\begin{aligned} H(M, K \oplus S) &= H(M) + H(K \oplus S|M) \\ &\leq H(M) + H(K) \\ &\leq H(M) + NR_s. \end{aligned} \tag{10}$$

For the second term we find

$$\begin{aligned} H(M, K \oplus S|K) &= H(M|K) + H(K \oplus S|M, K) \\ &= H(M) + H(S|M, K) \\ &= H(M) + H(S|M) \\ &= H(M) + H(S) - I(S;M). \end{aligned} \tag{11}$$

Now we combine the two terms and obtain

$$\begin{aligned} I(K; M, K \oplus S) &\leq NR_s - H(S) + I(S;M) \\ &\leq N(I(X;Y) - \varepsilon) \\ &\quad -N(I(X;Y) - 3\varepsilon - 4\varepsilon\log_2|\mathcal{X}|) + 1 \\ &\quad +N(2\varepsilon + 4\varepsilon\log_2|\mathcal{X}|) + 1 \\ &= N(4\varepsilon + 8\varepsilon\log_2|\mathcal{X}|) + 2. \end{aligned} \tag{12}$$

Dividing both parts of the derived inequality by N finalizes the proof. ⊠

4 The XOR-scheme

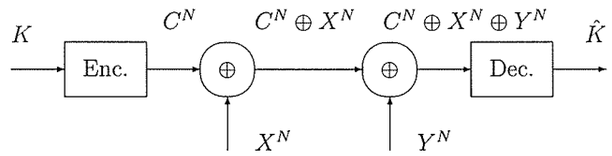

Figure 3: The XOR-scheme.

Now we consider a system (proposed in [1]) as in Figure 3. At the encoding side, a binary secret key K for biometric data X^N is uniformly chosen. This secret key is encoded into a binary codeword C^N, from the selected error-correcting code. The offset Z^N, defined as $Z^N = C^N \oplus X^N$, is released to the decoder for the authentication.

In the authentication phase, the offset Z^N is added modulo 2 to the biometric sequence Y^N, observed by the decoder $\hat{C}^N = Z^N \oplus Y^N = C^N \oplus X^N \oplus C^N$. The closest codeword in the corresponding error-correcting code is determined and this codeword is decoded to the secret key $\widehat{K}$. If $\widehat{K} = K$, the authentication decision is positive.

In the described XOR-scheme a binary error-correcting code of rate R_c is used and we assume that the code is one-to one. Since the secret key sequence K is encoded into a binary codeword C^N, this implies that $H(K) = H(C^N) = NR_c$, where R_c is the rate of the code.

Assume that the biometric sequence X^N is a stationary binary sequence with entropy

$$H_\infty(X) = \lim_{N\to\infty} H(X_1, X_2, \cdots, X_N)/N = \lim_{N\to\infty} H(X_N|X_1^{N-1}). \quad (13)$$

The binary entropy function $h(\cdot)$ is defined as $h(p) = -p\log_2(p) - (1-p)\log_2(1-p)$ for $0 \le p \le 1$. For $0 \le \alpha \le 1$ we define the inverse of the binary entropy function $h^{-1}(\alpha) = q$ if $0 \le q \le 1/2$ and $h(q) = \alpha$. Moreover for $0 \le p_1, p_2 \le 1$ let $p_1 * p_2 = p_1(1-p_2) + (1-p_1)p_2$.

Theorem 3 *For the random binary independent sequences X^N and C^N, if X^N is stationary, the following statement holds:*

$$\frac{1}{N}H(Z_1^N) \ge h[h^{-1}(H_\infty(X)) * h^{-1}(R_c)],$$

where $Z_1^N = Z_1, \cdots, Z_N = X_1 \oplus C_1, \cdots, X_N \oplus C_N$. This is an adapted version of the binary analog to the entropy-power inequality (Shamai and Wyner [7]).

Proof: For $n = 1, 2, \cdots, N$, from Shamai and Wyner ([7], last but one equation), we obtain

$$\begin{aligned} H(Z_n|Z_1^{n-1}) &\ge h[h^{-1}(H(X_n|X_1^{n-1})) * h^{-1}(H(C_n|C_1^{n-1}))] \\ &\ge h[h^{-1}(H_\infty(X)) * h^{-1}(H(C_n|C_1^{n-1}))], \end{aligned} \quad (14)$$

since $H_\infty(X) \le H(X_n|X_1^{n-1})$, and where e.g. $X^{n-1} = X_1, \cdots, X_{n-1}$. Next we use the $\cup$-convexity of $h(\beta * h^{-1}(u))$ in u (in the second inequality) and find that:

$$\begin{aligned} \frac{1}{N}H(Z_1^N) &= \frac{1}{N}\sum_{n=1}^{N} H(Z_n|Z_1^{n-1}) \\ &\ge \frac{1}{N}\sum_{n=1}^{N} h[h^{-1}(H_\infty(X)) * h^{-1}(H(C_n|C_1^{n-1}))] \\ &\ge h[h^{-1}(H_\infty(X)) * h^{-1}(\frac{1}{N}\sum_{n=1}^{N} H(C_n|C_1^{n-1}))] \\ &= h[h^{-1}(H_\infty(X)) * h^{-1}(R_c)]. \end{aligned} \quad (15)$$

This finalizes the proof. ⊠.

Based on this result we would like to analyse the XOR-scheme for biometric binary stationary sequences X^N that do not have full entropy. The side information Z^N is publicly communicated to the decoder, and we are interested in the amount of information that can be obtained by an eavesdropper from Z^N about the secret key K. Therefore, to characterize information leakage, we would like to evaluate the mutual information $I(K; Z^N)/N$. The mutual information can now be rewritten as

$$\begin{aligned} I(K; Z^N) &= H(Z^N) - H(Z^N|K) \\ &= H(C^N \oplus X^N) - H(C^N \oplus X^N|K) \\ &= H(X^N \oplus C^N) - H(X^N). \end{aligned} \quad (16)$$

where the last equality holds since C^N is determined by K and X^N and K are independent.

Theorem 4 *Information leakage is unavoidable, i.e. $I(K; C^N \oplus X^N)/N > 0$ for $H_\infty(X) < 1$ and for N asymptotically large.*

Proof: From our version of the binary analog to the entropy-power inequality

$$\begin{aligned} \lim_{N\to\infty}\frac{1}{N}I(K; Z^N) &= \lim_{N\to\infty}\frac{1}{N}H(C^N \oplus X^N) - \lim_{N\to\infty}\frac{1}{N}H(X^N) \\ &\ge h[h^{-1}(H_\infty(X)) * h^{-1}(R_c)] - H_\infty(X). \end{aligned} \quad (17)$$

Inspection shows that equality can only occur if $H_\infty(X) = 1$. Thus, we conclude that a security preserving XOR-scheme is only established if X^N is independent and uniformly distributed. ⊠.

Another, even better, lower bound follows if we use a simple binary linear code where the first NR_c information symbols are followed by $N - NR_c$ parity symbols, i.e. $H(C_n|C_1^{n-1}) = 1$ for $n \le NR_c$ and $H(C_n|C_1^{n-1}) = 0$ for $n > NR_c$, where we also assume that NR_c is integer. Therefore, from (14),

$$\begin{aligned} \frac{1}{N}H(Z_1^N) &= \frac{1}{N}\sum_{n=1}^{N} H(Z_n|Z_1^{n-1}) \\ &\ge \frac{1}{N}\sum_{n=1}^{NR_c} h[h^{-1}(H_\infty(X)) * h^{-1}(1)] + \frac{1}{N}\sum_{n=NR_c+1}^{N} h[h^{-1}(H_\infty(X)) * h^{-1}(0)] \\ &= \frac{1}{N}[NR_c + (N - NR_c)H_\infty(X)] \\ &= R_c + (1 - R_c)H_\infty(X) \\ &= H_\infty(X) + R_c(1 - H_\infty(X)). \end{aligned} \quad (18)$$

Now we obtain

$$\begin{aligned} \lim_{N\to\infty}\frac{1}{N}I(K; Z^N) &\ge H_\infty(X) + R_c(1 - H_\infty(X)) - H_\infty(X) \\ &= R_c(1 - H_\infty(X)). \end{aligned} \quad (19)$$

Using the same reasoning as before, we again conclude that a security preserving XOR-scheme is only established if $H_\infty = 1$.

Example: Let us consider an example of a biometric XOR system with the following parameters: $H_\infty(X) = 0.500$ and, therefore, $h^{-1}(H_\infty(X)) = 0.110$, and $R_c = I(X;Y)/N = 0.100$ and $h^{-1}(R_c) = 0.013$. Then $h(0.110 * 0.013) = h(0.120) = 0.530$. Therefore, $I(K; Z^N)/N \ge 0.530 - 0.500 = 0.030$, which is 30% of the information about the secret key K. If we assume that the code is linear, information digits first, we obtain the lower bound $I(K; Z^N)/N \ge 0.1 \cdot 0.5 = 0.050$, which is 50% of the information about the secret key K.

From these results, it is clear that this system is insecure. Note that we considered the asymptotic case, but it will be clear that for finite values of N there can be a security problem. Using the same reasonings, it can also be shown that side information Z^N in the XOR-scheme leaks information about biometric data.

5 Conclusions

We investigated the security properties of the XOR system (some of the provided proofs overlap with those, for example, in [6], [8]). We could conclude that full security can only be obtained for biometric sequences with entropy one. It is therefore better to use a scheme based on a randomness-extraction layer followed by a masking layer. We have analyzed such schemes and they turn out to be optimal.

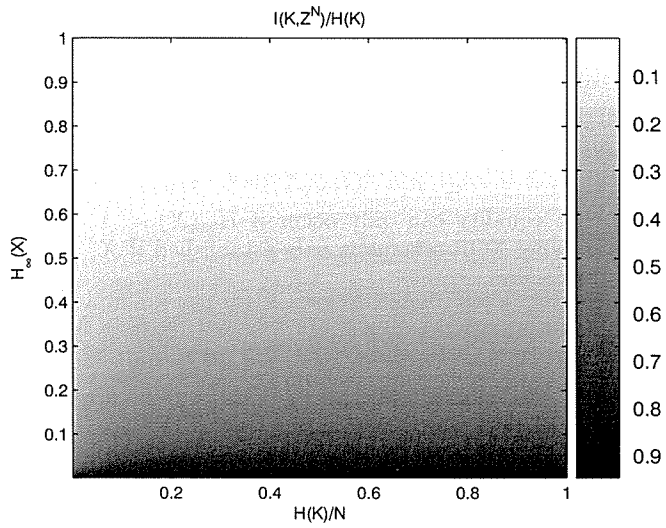

Figure 4: Lower bound plot on information leakage, $\frac{I(K,Z^N)}{H(K)}$ vs. $H_\infty(X), H(K)/N$

6 Acknowledgement

The authors thank SenterNovem and Philips Research for financial support.

References

[1] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proceedings of the 6th ACM Conference on Computer and Communications Security*, 1999, pp. 28–36.

[2] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, pp. 733–742, May 1993.

[3] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, July 1973.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: John Wiley and Sons Inc., 1991.

[5] T. Cover, "A proof of the data compression theorem of Slepain and Wolf for ergodic sources," *IEEE Transactions on Information Theory*, vol. 22, pp. 226 – 228, March 1975.

[6] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography - part I: Secret sharing," *IEEE Transactions on Information Theory*, vol. 39, pp. 1121–1132, July 1993.

[7] S. Shamai and A. Wyner, "A binary analog to the entropy-power inequality," *IEEE Transactions on Information Theory*, vol. 36, no. 6, pp. 1428–1430, November 1990.

[8] J. Goseling, T. Kalker, and P.Tuyls, "Protection of biometric templates," Koninklijke Philips Electronics NV, Tech. Rep., 2003.