# Achieving secure fuzzy commitment scheme for optical PUFs

2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

# Achieving Secure Fuzzy Commitment Scheme for Optical PUFs

Tanya Ignatenko
*Electrical Engineering Department*
*Eindhoven University of Technology*
*Den Dolech 2, 5612 AZ Eindhoven, The Netherlands*
*Email: t.ignatenko@tue.nl*

Frans Willems
*Electrical Engineering Department*
*Eindhoven University of Technology*
*Den Dolech 2, 5612 AZ Eindhoven, The Netherlands*
*Email: f.m.j.willems@tue.nl*

***Abstract*—Fuzzy commitment of Juels and Wattenberg 1999 is a popular technique for designing secure systems based on noisy data. The scheme is easy to implement using standard error-correcting codes. However, secrecy of this scheme is only guaranteed when input data are generated by uniform i.i.d. sources, while typical input data (PUFs and biometrics) are not uniform. In this paper we address the problem of extracting robust independent uniformly distributed bits out of noisy data that can be used as entries to fuzzy commitment. The proposed techniques can serve as a building block of secure fuzzy commitment systems.**

***Keywords*-Security; secret-key extraction; PUFs**

## I. Introduction

A physical unclonable function (PUF) is defined as a function that maps challenges to responses and is embodied by a physical device. These functions were first introduced and studied by Pappu [1]. PUFs have important property that they generate responses from physical systems that are difficult to characterize and analyze. This property relies on a difficulty of taking a complex physical system, deriving its all essential parameters, and simulating the system to predict responses based on the derived information. Moreover, to measure physical parameters of a device that carries out PUFs, it is necessary to invade the device and therefore modify it, which results in breaking the PUF. Thus PUFs offer a certain degree of tamper evidence. Further, random nature of PUFs arises from random manufacturing variations. Thus even a manufacturer does not possess full knowledge about PUFs and is unable to produce two identical PUFs. This property of PUFs is called manufacture-resistance or unclonability.

The uniqueness of the responses and uncloneability of PUFs are very attractive properties for security applications. A device (token) with an embedded PUF becomes uniquely identifiable and uncloneable. Moreover, PUFs are a good source of (secret) common randomness between two terminals.

In this paper we concentrate on optical PUFs, see Pappu [1]. These PUFs consist of transparent material with randomly distributed light-scattering particles. Their suitability as a carrier of secret key material is derived from uniqueness and unpredictability of speckle patterns that result from multiple scattering of laser light in a disordered optical medium. However, measurements of PUFs are not exactly reproducible, since, for example, optical PUF displacements and laser angle differences introduce measurement noise.

Next we consider fuzzy commitment introduced by Juels and Wattenberg [2], which is a popular scheme for secret-key binding used to built secure biometric applications. This scheme relies on standard error-correcting codes (ECC) that are efficient and easy to implement. Since PUFs have similar properties to those of biometrics, fuzzy commitment is also often used to build secrecy systems based on PUFs, see Skoric et al. [3]. In [4], however, it was argued that fuzzy commitment schemes preserve their properties to be secure only if input data distribution is uniform. In practice, neither biometrics nor PUFs do not enjoy the required distribution. In the following we propose a number of methods to extract robust independent uniformly distributed bits out of noisy data. The proposed methods are validated on optical PUFs.

## II. Fuzzy Commitment Scheme

Let $(X^N, Y^N)$ be a pair of binary correlated PUF sequences, observed by an encoder and decoder, respectively. In the fuzzy commitment scheme, presented in Fig. 1, a secret key $k$ from alphabet $\{1, 2, \ldots, |\mathcal{K}|\}$ is chosen uniformly at random independently of PUF, hence $\Pr\{K = k\} = 1/|\mathcal{K}|$ for all $k \in \{1, 2, \ldots, |\mathcal{K}|\}$. The chosen secret key $k$ is observed at the enrollment side together with a PUF enrollment sequence $x^N$. The secret key $k$ is encoded into a binary codeword $c^N = (c_1, c_2, \ldots, c_N)$ with $c_n \in \{0, 1\}$ for $n = 1, 2, \ldots, N$. We write $c^N = e(k)$, where $e(\cdot)$ is the encoding function. Then the PUF enrollment sequence is added modulo 2 to the codeword. This results in the sequence $v^N = (v_1, v_2, \ldots, v_N)$ with $v_n \in \{0, 1\}$ for $n = 1, 2, \ldots, N$, hence $v^N = c^N \oplus x^N = e(k) \oplus x^N$. This sequence is referred to as helper data and is public. The helper data are released to the authentication side.

During authentication, a PUF authentication sequence $y^N$ is observed and added modulo 2 to the received helper data $v^N$, resulting in a binary sum $r^N = v^N \oplus y^N = e(k) \oplus (x^N \oplus y^N)$. This sum $r^N = \{r_1, r_2, \ldots, r_N\}$ with $r_n \in \{0, 1\}$ for $n = 1, 2, \ldots, N$ can be seen as the codeword $c^N$ to which a noise sequence $x^N \oplus y^N$ is added. This codeword $r^N$ is

978-0-7695-3762-7/09 $26.00 © 2009 IEEE
DOI 10.1109/IIH-MSP.2009.310

then decoded, hence the estimate $\widehat{k}$ of the secret key $k$ is determined as $\widehat{k} = d(r^N) = d(e(k) \oplus (x^N \oplus y^N))$, where $d(\cdot)$ is the decoding function.[1]

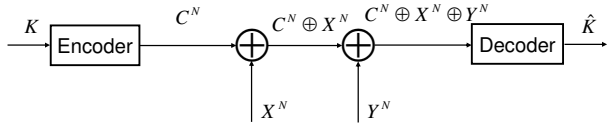



Figure 1. Fuzzy commitment scheme.

In [4] it was shown that secure fuzzy commitment is only achieved with uniform input data distribution. The properties of fuzzy commitment are summarized in the following theorem.

Assume that a PUF sequence $X^N$ is a binary stationary sequence with entropy $H_\infty(X)$. Define the binary entropy function $h(p) \triangleq -p\log(p) - (1-p)\log(1-p)$, for $0 \leq p \leq 1$, and its inverse $h^{-1}(\alpha) \triangleq q$, for $0 \leq \alpha \leq 1$, such that $0 \leq q \leq 1/2$ and $h(q) = \alpha$.

*Theorem 1 ([4]):* For fuzzy commitment that operates on binary stationary sequences $X^N$ with entropy satisfying $0 < H_\infty(X) < 1$ and uses a code with rate $0 < R_c < 1$ the secrecy leakage is lower bounded as follows.

$$\lim_{N\to\infty} \frac{1}{N} I(K; V^N) \geq h[h^{-1}(H_\infty(X)) * h^{-1}(R_c)] - H_\infty(X).$$

Thm. 1 also holds for the i.i.d. case, the proof follows from application of Mrs. Gerber's lemma (Wyner and Ziv [5]).

From the above theorem we conclude that in order to build a secure fuzzy commitment based system, a system should contain a layer that extracts uniform or very close to uniform data from enrollment sequences $X^N$.

## III. Extracting Robust Uniform Data

We concentrate on the case when PUFs are represented by binary images. Extracting independent uniformly distributed bits from PUF images requires a known model for them. We assume that PUFs are generated by Markov chains and use the CTW method [6] to find their data distribution. The CTW method is a universal sequential data compression method that finds a good, in terms of coding redundancy and complexity, coding distribution for tree sources. The method approaches entropy for one- and two-dimensional ergodic stationary sources.

### A. Selection of robust bits

Although in fuzzy commitment ECCs are used to reconstruct a secret key $K$ during the decoding phase, applying ECCs to long data sequences with high error rates is impractical. Therefore we need data that are not only uniformly random but also have lower error rates. In this section we present a number of methods to extract such data.

Let $X^N$ be enrollment, $Z_1^N, Z_2^N, \ldots, Z_L^N$ be training and $Y^N$ be authentication PUF sequences, respectively.

**Method 1.** Assume that we have access to a true PUF sequence $x^N$ and also to its noisy observations $z_1^N, z_2^N, \ldots, z_L^N$. Moreover, we assume that noise that affects data is generated by a memoryless source.

Now we define robust bits to be bits that actually occur with high probability in an authentication sequence $y^N$ after some template from a true sequence $x^N$ has been observed, see Fig. 2. Robust bits defined in this way give the smallest contribution to the codeword length, when a noisy PUF sequence is compressed given a true one. This principle is based on the assumption that the underlaying noiseless process has a structure and, given this structure, we can better compress $y^N$ if this structure is still preserved. The algorithm is summarized in the following.

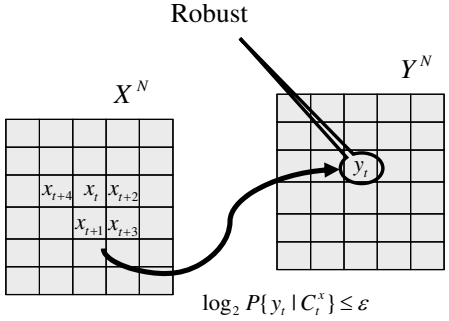



Figure 2. Robust bits, memoryless noise.

1) Using the CTW method, the distribution $p(y|x)$ is estimated as $P\{z_{L,t}|\mathcal{C}_t^x\}$, where $\mathcal{C}_t^x = \{x_{t-a}, a \in A\}$ is the context defined by a set $A$ of well-chosen integers from $\mathbb{Z}$. The context comes from an enrollment sequence $x^N$ and the bits $z_{L,t}$ come from a training sequence $z_L^N$. The context $\mathcal{C}_t^x$ is used to create nodes in the context tree and $z_{L,t}$ are used to update its counts and probabilities $(L \geq 1)$.
2) A robustified sequence $\hat{x}^N$ is constructed by analyzing the average contribution of bits $z_{l,t}$ at position $t$ to the codeword length, compressed given the context $\mathcal{C}_t^x$.
$$\hat{x}_t = \begin{cases} x_t, & \text{if } -\frac{1}{L}\sum_{l=1}^{L} \log P(z_{l,t}|\mathcal{C}_t^x) \leq \varepsilon \\ 2, & \text{otherwise} \end{cases},$$
for $t = 1, 2, \ldots, N$, where $\mathcal{C}_t^x = \{x_{t-a}, a \in A\} \subset x^N$, $L \geq 1$, and $\varepsilon$ is a threshold with its value close to 0.
3) The robustified PUF sequence is a 3-valued sequence where non-robust bits are those that take on values 0 or 1 with (almost) equal probability. Positions $t$, for which $\hat{x}_t \neq 2$, form the set of robust positions $\Pi_r$.

**Method 2.** The next method is an extension of method 1, where the assumption on the noise is modified. Unlike

[1]Note that in the original scheme a one-way function $f(\cdot)$ is applied to the secret key $k$, and then the information that is stored in a database or on a device consists of $f(k)$, helper data and/or an ID of a PUF. In this paper we omit one-way function and only concentrate on secret keys, since it will not affect our reasoning.

in the previous method, here we assume that noise is not memoryless, and robust bits are defined to be those that occur with high probability in an authentication sequence. The occurrence of a bit depends on both context from an enrollment PUF sequence and context from an authentication PUF sequence generated so far, see Fig. 3. The method reads as follows.

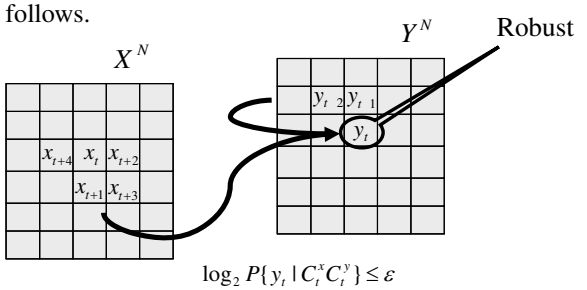



Figure 3. Robust bits, noise with memory.

1) The distribution $p(y|x)$ is estimated using the CTW method as $P(z_{L,t}|\mathcal{C}_t^x, \mathcal{C}_t^{z_L})$, where the context $\mathcal{C}_t^x = \{x_{t-a}, a \in A\}$ comes from an enrollment sequence $x^N$; and bits $z_{L,t}$ and the context $\mathcal{C}_t^{z_L} = \{z_{L,t-b}, b \in B\}$ come from a training sequence $z_L^N$, where $B$ is a well-defined set of strictly positive integers, resulting in the casual context $\mathcal{C}_t^{z_L}$.
2) A robustified sequence is constructed by analyzing the average contribution of $t$-th bit to the codeword length

$$\hat{x}_t = \begin{cases} x_t, & \text{if } -\frac{1}{L}\sum_{l=1}^{L} \log P(z_{l,t}|\mathcal{C}_t^x, \mathcal{C}_t^{z_l}) \le \varepsilon \\ 2, & \text{otherwise} \end{cases},$$

for $t = 1, 2, \ldots, N$, where $\mathcal{C}_t^x = \{x_{t-a}, a \in A\} \subset x^N$, $\mathcal{C}_t^{z_l} = \{z_{l,t-b}, b \in B\} \subset z_l^N$, $L \ge 1$ and $\varepsilon$ is a threshold with its value close to 0.
3) The set of robust positions $\Pi_r = \{t : \hat{x}_t \neq 2\}$ constitutes helper information.

### B. Marginalization

Independent uniformly distributed bits are characterized by entropy equal to 1. Therefore we are looking for bits whose probability conditional on the preceding symbols is close to 0.5. Here we propose to estimate the marginal probabilities of the robust bits and remove all bits which do not occur with probability close to 0.5. The proposed method is again based on the CTW method.

Assume that Markov chain of order $D$ governs the PUF data distribution. Then every bit $x_t$ in a PUF sequence depends on a well-chosen set of neighboring bits $\{x_{t-d}, d \in \mathcal{B}\}$. Moreover, let $\beta_t$ be a subset of $\mathcal{B}$, refering to robust bits in a given context $\{x_{t-d}, d \in \mathcal{B}\}$. We are interested in conditional probabilities $p(x_t|\{x_{t-d}, d \in \beta_t\})$ and find them as

$$p(x_t|\{x_{t-d}, d \in \beta_t\}) = \frac{p(x_t, \{x_{t-d}, d \in \beta_t\})}{p(\{x_{t-d}, d \in \beta_t\})} = \frac{\sum_{x_{t-d}: d \in \mathcal{B}\setminus\beta_t} p(x_t, \{x_{t-d}, d \in \mathcal{B}\})}{\sum_{x_t}\sum_{x_{t-d}: d \in \mathcal{B}\setminus\beta_t} p(x_t, \{x_{t-d}, d \in \mathcal{B}\})},$$

$p(x_t, \{x_{t-d}, d \in \mathcal{B}\}) = p(x_t|\{x_{t-d}, d \in \mathcal{B}\})p(\{x_{t-d}, d \in \mathcal{B}\})$, where $p(x_t|\{x_{t-d}, d \in \mathcal{B}\})$ are conditional probabilities estimated using the CTW, and $p(\{x_{t-d}, d \in \mathcal{B}\})$ are stationary probabilities found as a fraction of occurrences of a pattern $\{x_{t-d}, d \in \mathcal{B}\}$ in the enrollment PUF image $x^N$.

Computing and analyzing the bit probabilities, we only keep the robust bits satisfying $|P(x_t|\{x_{t-d}, d \in \beta_t\}) - 0.5| \le \delta$. The bits that do not satisfy this condition are processed as non-robust, and the next robust context $\{x_{t-d}, d \in \beta_t\}$ only contains robust uniformly distributed bits.

## IV. Experimental results

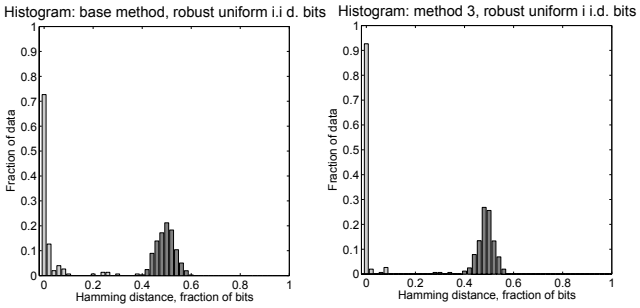
In this paper we focus on optical PUFs, see Pappu [1]. Different challenges are obtained by directing a laser beam under different angles through a PUF. Shining a laser beam through the optical medium produces speckle patterns that are picked up by a CCD camera. To obtain a compact binary representation of a speckle pattern, the measurements are preprocessed using Gabor-filtering (at $45^\circ$), thresholding and subsampling, like e.g. in Skoric et al. [3]. This results in $64 \times 64$ binary images. Note that optical PUFs are modeled as stationary and ergodic, see Skroc at al. [7], therefore the methods proposed in the previous sections are applicable.

We investigated five optical PUFs, two challenges per PUF. The PUFs were measured at different moments in time under different environmental conditions. We use 7 PUF measurements for training (robust bit selection) and 15 for testing. For robust bit selection we use the threshold $\varepsilon = 0.97$ and for uniform bit selection $\delta = 0.05$.

As a benchmark we use a commonly used reliable component method, see Skoric et al. [3] and Campisi et al. [8] extended with decimation, as proposed in Skoric et al. [7] to obtain uniform i.i.d. data. This method is referred to as base method. Due to space limitations we only present the results for the base method and method 2. We note that method 1 has comparable performance to method 2.

Comparing the results of the experiments for robust uniform bit selection, we observe that the largest number of bits is derived in method 2 (545), while the base method could detect a much smaller number of bits (301). Moreover, from the results for the base method, we observe that 90% of the robust bits are within 46 bit Hamming distance for the inter class distribution, and only 24% of the data have 0 bit Hamming distance. Similar operation points for method 2 show that 90% of the data in the inter class distribution are within 5 bit Hamming distance, and 68% of the data have 0 bit Hamming distance. In Fig. 4 we have depicted the intra and inter class distribution histograms of the extracted bits. Based on the observed performance we conclude that method 2 outperforms the base method.

The resulting robust (nearly) uniform bits can be used to design a secure fuzzy commitment scheme. Here we take BCH codes, see e.g. McEliece [9], that are used to correct




(a) The base method. (b) Proposed method.

Figure 4. Data distribution of uniformly i.i.d. robust bits. Histogram: light grey - intra class, dark grey - inter class.

multiple random error patterns. BCH codes are characterized by the triple $(n, k, t)$, where $n$ denotes the codeword length, $k$ denotes the message length and $t$ is the error correcting capability of the code. The rate of such codes is $R_c = k/n$. We select the length of a PUF sequence that masks the encoded secret to be maximum integer that satisfies BCH codeword length requirement but less or equal to the maximum number of independent uniformly distributed robust bits that can be extracted from our PUFs. This number turns out to be $n = 255$ and $n = 511$ for the base and the proposed method, respectively. The key length and error correcting capability are selected based on the data distribution.

First, we consider a system with BCH code $(255, *, *)$. The performance of such a scheme is summarized in the upper parts of Tables Ia and Ib. To compare the proposed and base methods, we look at BCH codes with the same parameters and provide the corresponding FAR, FRR and secrecy-leakage bound for them. From the obtained results we see that as the rate of the code is getting smaller, less information is leaked on the secret. However, in the latter case, the key size also reduces, making brute-force attack feasible. Method 2 clearly outperforms the base method.

In the lower parts of Tables Ia and Ib we also provide results for BCH code $(511, *, *)$. Since in the base method the number of the extracted bits is smaller then 511, in order to use BCH code (511,*,*) we append dummy bits to the PUF sequences. Note that this results in much higher secrecy leakage.

## V. Conclusions

In this paper we have proposed a number of methods to design secure fuzzy commitment systems. The methods are based on the CTW method proposed by Willems, Shtarkov, and Tjalkens [6]. We have used optical PUFs and have shown that based on the extracted keys, we can construct a reliable and secure fuzzy commitment scheme. Note that the proposed method can also be utilized for binary biometric data such as iris data or binary minutiae fingerprint data.

## VI. Acknowledgments

We would like to thank Geert-Jan Schrijen, Boris Skoric, and Pim Tuyls for useful discussions. We would also like to thank SenterNovem for funding. Project number IGC03003B.

Table I
SCHEME PERFORMANCE WITH BCH CODES

(a) The base method

| BCH triple | Rate | FRR | FAR | Secrecy leakage |
|---|---|---|---|---|
| (255,239,2) | 0.9373 | 0.2533 | 0 | 0.0066 |
| (255,155,13) | 0.6078 | 0.1200 | 0 | 0.0037 |
| (255,131,18) | 0.5137 | 0.0935 | 0 | 0.0029 |
| (255,9,63) | 0.0353 | 0.0133 | 0 | $1.07 * 10^{-4}$ |
| (511,493,2) | 0.9648 | 0.2667 | 0 | 0.3930 |
| (511,457,6) | 0.8943 | 0.1667 | 0 | 0.3575 |
| (511,421,10) | 0.8239 | 0.1400 | 0 | 0.3225 |
| (511,184,45) | 0.3601 | 0.0533 | 0 | 0.1143 |

(b) Proposed method (method 2)

| BCH triple | Rate | FRR | FAR | Secrecy leakage |
|---|---|---|---|---|
| (255,239,2) | 0.9373 | 0.0534 | 0 | 0.0066 |
| (255,155,13) | 0.6078 | 0.0400 | 0 | 0.0037 |
| (255,131,18) | 0.5137 | 0.0200 | 0 | 0.0029 |
| (255,9,63) | 0.0353 | 0.0134 | 0 | $1.07 * 10^{-4}$ |
| (511,493,2) | 0.9648 | 0.1067 | 0 | 0.0069 |
| (511,457,6) | 0.8943 | 0.0667 | 0 | 0.0062 |
| (511,421,10) | 0.8239 | 0.0574 | 0 | 0.0055 |
| (511,184,45) | 0.3601 | 0.0200 | 0 | 0.0018 |

## References

[1] R. Pappu, *Physical One-Way Functions*, Ph.D. thesis, M.I.T., 2001.

[2] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *ACM Conf. on Computer and Communications Security*, 1999, pp. 28–36.

[3] B. Skoric, P. Tuyls, and W. Ophey, "Robust key extraction from physical uncloneable functions," in *ACNS*, 2005, pp. 407–422.

[4] T. Ignatenko and F. Willems, "On privacy in secure biometric authentication systems," in *Proc. of IEEE ICASSP*, 2007.

[5] A. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications–I," *IEEE Trans. on Inf. Th.*, vol. 19, 1973.

[6] F.M.J. Willems, Y.M. Shtarkov, and T.J. Tjalkens, "The context tree weighting method: Basic properties," *IEEE Trans. on Inf. Th.*, 1995.

[7] B. Skoric, G.J. Schrijen, W. Ophey, and R. Wolters, *Security with Noisy Data: Private Biometrics, Secure Key Storage and Anti-Counterfeiting*, chapter Experimental Hardware for Coating PUFs and Optical PUFs, pp. 255–269, Springer, 2007.

[8] P.Campisi, E.Maiorana, M.G.Prats, and A. Neri, "Adaptive and distributed cryptography for signature biometrics protection," in *SPIE Conf. on Sec., Steg. and Waterm. of Mult. Cont. IX*, 2007, vol. 6505.

[9] Robert J. McEliece, *Theory of Information and Coding*, Cambridge University Press, New York, NY, USA, 2001.