Information-Theoretic Security Analysis of Physical Uncloneable Functions

P. Tuyls, B. Škorić, S. Stallinga, A.H.M. Akkermans, W. Ophey

Philips Research Laboratories, Prof. Holstlaan 4,
5656 AA Eindhoven, The Netherlands

We propose a general theoretical framework to analyze the security of Physical Uncloneable Functions (PUFs). We apply the framework to optical PUFs. In particular we present a derivation, based on the physics governing multiple scattering processes, of the number of independent challenge-response pairs supported by a PUF. We find that the number of independent challenge-response pairs is proportional to the square of the thickness of the PUF and inversely proportional to the scattering length and the wavelength of the laser light. We compare our results to those of Pappu and show that they coincide in the case where the density of scatterers becomes very high. Finally, we discuss some attacks on PUFs, and introduce the Slow PUF as a way to thwart brute force attacks.

Keywords: Physical Uncloneable Function, entropy, speckle pattern, Challenge-Response Pair

1 Introduction

1.1 Physical Uncloneable Functions

A 'Physical Uncloneable Function' (PUF) is a function that is realized by a physical system, such that the function is easy to evaluate but the physical system is hard to characterize [1, 2]. PUFs have been proposed as a cost-effective way to produce uncloneable tokens for identification [3]. The identification information is contained in a cheap, randomly produced (i.e. consisting of many random components), highly complicated piece of material. The secret identifiers are read out by performing measurements on the physical system and performing some additional computations on the measurement results. The advantage of PUFs over electronic identifiers lies in the following facts: (1) Since PUFs consist of many random components, it is very hard to make a clone, either a physical copy or a computer model, (2) PUFs provide inherent tamper-evidence due to their sensitivity to changes in measurement conditions, (3) Data erasure is automatic if a PUF is damaged by a probe, since the output strongly depends on many random components in the PUF. Additionally one can extract cryptographic keys from a PUF. This makes PUFs attractive for Digital Rights Management (DRM) systems.

The physical system is designed such that it interacts in a complicated way with stimuli (*challenges*) and leads to unique but unpredictable *responses*. Hence, a PUF is similar to a keyed hash function. The key is the physical system consisting of many "random" components. In order to be hard to characterize, the

system should not allow efficient extraction of the relevant properties of its interacting components by measurements. Physical systems that are produced by an uncontrolled production process, i.e. one that contains some randomness, turn out to be good candidates for PUFs. Because of this randomness, it is hard to produce a physical copy of the PUF. Furthermore, if the physical function is based on many complex interactions, then mathematical modeling is also very hard. These two properties together are referred to as *Uncloneability*. From a security perspective the uniqueness of the responses and uncloneability of the PUF are very useful properties. Because of these properties, PUFs can be used as unique identifiers for smart-cards and credit cards or as a 'cheap' source for key generation (common randomness) between two parties.

At the moment there are several main candidates: optical PUFs [3, 4], silicon PUFs [2, 5], coating PUFs [6] and acoustic PUFs [6]. Silicon PUFs make use of production variation in the properties of logical gates. When these are probed at frequencies that are out of spec, a unique, unpredictable response is obtained in the form of delay times. Coating PUFs are integrated with an IC. The IC is covered with a coating, which is doped with several kinds of particles of random size and shape with a relative dielectric constant differing from the dielectric constant of the coating matrix. An array of metal sensors is laid down between the substrate and the passivation layer. A challenge corresponds to a voltage of a certain frequency and amplitude applied to the sensors at a certain point of the sensor array. The response, i.e. the capacitance value, is then turned into a key. In an acoustic PUF, one measures the response of a token to an acoustic wave. An electrical signal is transformed to a mechanical vibration through a transducer. This vibration propagates as a sound wave through the token and scatters on the randomly distributed inhomogeneities. The reflections are measured by another transducer which converts the vibration back into an electric signal. It turns out that the reflections are unique for each token.

Fig. 1. Left: Card equipped with an optical PUF. Right: Reading device.

Optical PUFs contain randomly distributed light scattering particles. A picture of an optical PUF and its reading device is shown in Fig. 1. They exploit the uniqueness of speckle patterns that result from multiple scattering of laser light in a disordered optical medium. The challenge can be e.g. the angle of incidence, focal distance or wavelength of the laser beam, a mask pattern blocking part

of the laser light, or any other change in the wave front. The response is the speckle pattern. An input-output pair is usually called a *Challenge-Response Pair* (CRP). Physical copying is difficult for two reasons: (i) The light diffusion obscures the locations of the scatterers. At this moment the best physical techniques can probe diffusive materials up to a depth of ≈10 scattering lengths [7]. (ii) Even if all scatterer locations are known, precise positioning of a large number of scatterers is very hard and expensive, and this requires a process different from the original randomized process. Modeling, on the other hand, is difficult due to the inherent complexity of multiple coherent scattering [8]. Even the 'forward' problem turns out to be hard[1].

The goal of this paper is to show how cryptographic tools based on (classical) physical functions can be modeled and rigorously analyzed in a cryptographic context. We derive an information-theoretic framework for PUFs and investigate the security level of optical PUFs. More in particular, we analyze the number of *independent* CRPs of a PUF, i.e. CRPs whose responses are not predictable using previously obtained CRPs. We derive a formula that gives the number of independent CRPs supported by an optical PUF in terms of its physical parameters. In section 2.1, we derive the model starting from the physics of multiple scattering. The security analysis, and in particular the computation of the number of independent CRPs, is presented in section 3. Finally, in section 4 we discuss brute force attacks. In particular, we introduce the 'slow PUF' as a way of thwarting these attacks.

1.2 Applications

Optical PUFs are well suited for identification, authentication and key generation [3, 6]. The goal of an identification protocol is to check whether a specific PUF is present at the reader. The goal of an authentication protocol is to ensure that received messages originate from the stated sender. For authentication it is therefore the objective to extract the same cryptographic key from the PUF as the one that is stored at the Verifier's database during enrollment, while for identification it is sufficient if the response is close to the enrolled response.

In order to use PUFs for above mentioned purposes they are embedded into objects such as smartcards, creditcards, the optics of a security camera, etc., preferably in an inseparable way, meaning that the PUF gets damaged if an attacker attempts to remove the PUF. This makes the object in which a PUF is embedded uniquely identifiable and uncloneable. Secret keys can be derived from a PUF's output [6] by means of protocols similar to those developed in the context of biometrics [10, 11].

The usage of a PUF consists of two phases: enrolment and verification. During the enrolment phase, the Verifier produces the PUF and stores an initial set of CRPs securely in his database. Then the PUF is embedded in a device and given

[1] Given the details of all the scatterers, the fastest known computation method of a speckle pattern is the transfer-matrix method [9]. It requires in the order of $N_{\mathrm{mod}}^3 d/\lambda$ operations (see section 3.2 for the definition of N_{mod}, d and λ).

to a user. The verification phase starts when the user presents his device to a terminal. The Verifier sends a randomly chosen PUF challenge from his database to the user. If the Verifier receives the correct response from the device, the device is identified. Then this CRP is removed from the database and will never be used again.

If, additionally, the device and the Verifier need to exchange secret messages, a secure authenticated channel is set up between them, using a session key based on the PUF response. We present the following protocols.

Identification Protocol:

- User: Puts his card with PUF in the reader and claims its ID.
- Verifier: Randomly chooses a challenge C from his CRP database and sends it to the User.
- Reader: Challenges the PUF with the Challenge C, measures the Response R and computes an identifier S'. S' is sent back to the Verifier.
- Verifier: Checks whether S' equals the identifier S stored in his database during enrollment. Then he removes the pair (C, S) from his database and never uses it again.

We note that the security of this protocol relies on the fact that an attacker who has seen (C_1, S_1) cannot predict the identifier S_2 corresponding to the challenge C_2, and on the fact that the PUF supports a large number of CRPs.

Authentication Protocol:

- User: Puts his card with PUF in the reader and claims its ID.
- Verifier: Randomly chooses a challenge C from his CRP database and sends it to the User, together with a random nonce m.
- Reader: Challenges the PUF with the Challenge C, measures the Response R and computes a key S'. $M_{S'}(m)$ is sent to the Verifier, where $M_{S'}(m)$ denotes a MAC on m, using the key S'.
- Verifier: Computes $M_S(m)$ with the key S stored in his database and compares it with $M_{S'}(m)$. If they are equal, then $S = S'$ with very high probability. The key S is then used to MAC and/or encrypt all further messages.

The security of this scheme depends on the fact that (when the key S is unknown) the MAC $M_S(m)$ is unpredictable given that the attacker has seen the MAC on a message $m_1 \neq m$.

1.3 Notation

We introduce the following notation. The power of the laser is denoted by P and its wavelength by λ. The thickness of the PUF is denoted by d. Scattering is assumed to be elastic[2], with mean free path[3] ℓ. We further assume diffusive

[2] Elastic scattering means that the photons do not lose energy when they are scattered.

[3] The mean free path is the average distance travelled by the photons between two scattering events.

scattering, i.e. $\lambda \ll \ell \ll d$. The illuminated area of the PUF is $A = W^2$. For simplicity the output surface area is also taken to be A. The detector needs time $\triangle t$ to record a speckle pattern. The following numerical values will be used by way of example: $W = 1$ mm, $d = 1$ mm, $\ell = 10$ μm, $\lambda = 500$ nm, $P = 1$ mW, $\triangle t = 1$ ms. Note that the total area of the PUF (A_{PUF}) can be much larger than the *illuminated area* A. We will use A_{PUF} =5cm^2. Throughout this paper we will mostly calculate properties of one specific volume Ad, and only afterwards adjust our results by a factor A_{PUF}/A. This effectively amounts to treating the PUF of area A_{PUF} as a collection of independent PUFs of area A.

2 Information Theory of PUFs

2.1 General PUF Model

A PUF can be modeled as a function mapping challenges to responses. We denote the challenge space by $\mathcal{A}$ and the response space by $\mathcal{R}$. A PUF is then a parametrized function $\pi_K : \mathcal{A} \to \mathcal{R}$ whose behaviour is determined by the physical interactions. The parameter K belongs to the parameter space $\mathcal{K}$ and is determined by a large number of random variables, namely the physical structure of the PUF. Hence, the space $\mathcal{K}$ models the space of all possible PUFs and there is a one to one correspondence between the elements of $\mathcal{K}$ and the set of PUFs. In order to express the uncertainy about the random variable K, described by the probability measure η, we use the Shannon entropy $\mathsf{H}_\eta(K)$

$$\mathsf{H}_\eta(K) = -\sum_{i=1}^{|\mathcal{K}|} \eta(K_i) \log \eta(K_i), \tag{1}$$

where $|\mathcal{K}|$ denotes the size of $\mathcal{K}$. Sometimes we will also need the conditional entropy $\mathsf{H}(K|R)$, representing the uncertainty about K given that one knows a response R. The mutual information between K and R is denoted as $\mathbf{I}(K;R)$. For the precise definitions of these notions we refer the reader to textbooks on information theory, e.g. [12]. The notation "log" denotes the logarithm with base 2.

One of the important quantities used in this paper is the size of the parameter space $\mathcal{K}$, representing the information content of a PUF. Therefore we have to make a precise definition of this quantity. To this end, we start by defining some abstract notions and later we make those notions concrete by means of an example. First, we present a brief computation that motivates the definitions that we introduce. The amount of information about the PUF that is revealed by one response is given by the mutual information $\mathbf{I}(K;R) = \mathsf{H}(K) - \mathsf{H}(K|R)$. We show that the mutual information is actually equal to $\mathsf{H}(R)$. First we observe that $\mathsf{H}(K) = \mathsf{H}(K,R)$, since given the PUF, the speckle pattern is fixed. Using the identity $\mathsf{H}(K,R) = \mathsf{H}(R) + \mathsf{H}(K|R)$ we obtain

$$\mathbf{I}(K;R) = \mathsf{H}(R). \tag{2}$$

2.2 Definitions

The information content of a PUF ($\mathsf{H}_\eta(K)$) and of its output ($\mathsf{H}(R)$) depends on the measurements that can be performed on the system. This is formalized as follows. We identify a measurement with its possible outcomes.

Definition 1 *A measurement $\mathcal{M}$ is a partition $\{R_1, \cdots, R_m\}$ of $\mathcal{K}$.*

Here R_j is the set in $\mathcal{K}$ containing all PUFs that produce outcome j upon measurement $\mathcal{M}$, and m is the number of possible outcomes. Two measurements give more (refined) information than one. The composition of two measurements is denoted as $\mathcal{M}_1 \vee \mathcal{M}_2$ and is defined as follows:

$$\mathcal{M}_1 \vee \mathcal{M}_2 = \{R_i^{(1)} \cap R_j^{(2)}\}_{i,j=1}^{m}. \tag{3}$$

$R_j^{(i)}$ is the set of all PUFs that produce outcome j upon measurement $\mathcal{M}_i$. By induction this definition extends to composition of more than two measurements.

Definition 2 *Let η denote a probability measure on $\mathcal{K}$. The information obtained about a system $K \in \mathcal{K}$ by performing measurement $\mathcal{M}$ is defined as*

$$\mathsf{h}_{\mathcal{M}}(\mathcal{K}) = -\sum_{i=1}^{m} \eta(R_i) \log \eta(R_i).$$

We note that the following monotonicity property can easily be proven

$$\mathsf{h}_{\mathcal{M}_1 \vee \mathcal{M}_2} \geq \mathsf{h}_{\mathcal{M}_1}, \tag{4}$$

which corresponds to the fact that finer measurements give more information. Due to the physics, one will often only have a finite set $\mathcal{A}$ of challenges available. This set restricts the amount of information that can be obtained.

Definition 3 [4] *Given the set $\mathcal{A}$ of possible measurements, the total amount of information that can be obtained about a system $\mathcal{K}$ is*

$$\mathsf{h}_{\mathcal{A}}(\mathcal{K}) = \sup_{\mathcal{M}_1, \ldots, \mathcal{M}_q \in \mathcal{A};\ 0<q\leq|\mathcal{A}|} \mathsf{h}_{\mathcal{M}_1 \vee \ldots \vee \mathcal{M}_q}(\mathcal{K}).$$

It follows from the monotonicity property (4) that $\mathsf{h}_{\mathcal{A}}(\mathcal{K}) \leq \mathsf{H}(K)$, i.e. the maximum amount of information that can be obtained about a system is upper bounded by the amount of uncertainty one has in the measure η. If η is given by the random measure $\eta(K_i) = 1/|\mathcal{K}|$, we find that $\mathsf{H}(K) = \log(|\mathcal{K}|)$. In the remainder of this text, we will assume that η is given by this measure.

Definitions 1 and 2 are very general and apply to many kinds of PUFs. In this framework, the couple $(\mathcal{K}, \mathcal{A})$ has to be specified for a well-defined notion of PUF security. We consider two extreme cases to illustrate the definitions. If $\mathcal{A}$ contains

[4] We note that this definition is in agreement with the theory of dynamical systems and dynamical entropy [13].

a CRP measurement that distinguishes PUFs perfectly, then the PUF supports only one independent CRP. The opposite extreme case is a set of measurements $\mathcal{A} = \{\mathcal{M}_j\}_{j=1}^{n}$ that can be represented as an extremely coarse partitioning of $\mathcal{K}$, say $|M_1^{(j)}| = |M_2^{(j)}| = |\mathcal{K}|/2$, where the combined measurements $(\mathcal{M}_1 \vee \ldots \vee \mathcal{M}_n)$ suffice to distinguish all elements of $\mathcal{K}$. In this case a minimum of $\log|\mathcal{K}|$ measurements is needed to reveal all details of the PUF. For good PUFs, all available measurements are fuzzy, revealing little about the physical structure.

2.3 Optical PUFs

We illustrate Definition 2 for optical PUFs. As the probing light has wavelength λ, it follows from the theory of electromagnetism [14] that details of size smaller than λ are difficult to resolve. It is natural to divide the volume into elements ('voxels') of volume λ^3. The number of voxels is $N_{\text{vox}} = Ad/\lambda^3$. In the example of section 1.3 we have $N_{\text{vox}} = 8 \cdot 10^9$ and a total number of $4 \cdot 10^{12}$ voxels in the whole PUF. For the sake of simplicity, we assume that light can only distinguish whether a voxel contains a scatterer or not. Hence, the information content of a voxel is at most 1 bit, and the PUF can be represented as a bit string of length N_{vox}. The entropy derived from the probability distribution η is[5] $\mathsf{H}(K) = N_{\text{vox}}$.

$\mathcal{A}$ is the full set of non-compound measurements that can be performed by means of a beam of monochromatic light. Combining all these available measurements, the maximum amount of information $\mathsf{h}_{\mathcal{A}}(\mathcal{K})$ that can be extracted from the PUF is $\mathsf{H}_\eta(K) = N_{\text{vox}}$. The couple $(\mathcal{K}, \mathcal{A})$ as defined here is used in the remainder of the text.

3 Security analysis

3.1 Security parameter C

The main goal of this paper is to estimate the number of independent CRPs. This number is denoted as C. It represents the minimal number of CRP measurements that an attacker has to perform to characterize the PUF.

Definition 4 *Measurements $\mathcal{M}_1, \ldots, \mathcal{M}_t$ are mutually independent iff*

$$\mathsf{h}_{\mathcal{M}_1 \vee \ldots \vee \mathcal{M}_t} = \mathsf{h}_{\mathcal{M}_1} + \cdots + \mathsf{h}_{\mathcal{M}_t}.$$

Note that $\mathsf{h}_{\mathcal{M}_1 \vee \ldots \vee \mathcal{M}_t} = t \cdot \mathsf{h}_{\mathcal{M}_1}$ if all measurements give the same amount of information, which by symmetry arguments is a reasonable assumption.

Independent measurements are also called *independent CRPs* since responses are implicitly incorporated in definition 4. In words, knowledge of independent

[5] It is possible to refine this model, taking into account the number of photons taking part in the measurement process. This gives rise to an extra factor proportional to the log of the number of photons. We will not discuss this refinement.

CRPs $\{\mathcal{M}_j\}_{j\neq i}$ does not give any information about the response to the i'th challenge. The number of independent CRPs is hence naturally defined as

$$C = \frac{\mathsf{h}_{\mathcal{A}}(\mathcal{K})}{\mathsf{h}_{\mathcal{M}}(\mathcal{K})} = \frac{\mathsf{h}_{\mathcal{A}}(\mathcal{K})}{\mathsf{H}(R)}, \tag{5}$$

where $\mathcal{M} \in \mathcal{A}$ and $\mathsf{H}(R)$ denotes the information content of a response. The second equality in (5) follows from (2). As we have already argued that $\mathsf{h}_{\mathcal{A}}(\mathcal{K}) = N_{\mathrm{vox}}$, the remainder of this section focusses on the computation of $\mathsf{H}(R)$.

In practice the independent challenges may turn out to be very complicated combinations of basic challenges. However, for the security analysis it is not necesary to have precise knowledge about them. The number C provides a basic security parameter which is not affected by technological and computational advances. An "adaptive chosen plaintext" attack (in the PUF context: trying to model a PUF by collecting responses to self-chosen challenges) requires at least C speckle pattern measurements, irrespective of the attacker's capabilities.

In practice many mutually *dependent* challenges may be used safely by the verifier. Even if some mutual information exists between the responses, it is computationally hard to exploit it, since that would require a characterisation of the physical function. It is not a priori clear how much mutual information between responses can be tolerated before the system becomes insecure, only that the answer depends on the capabilities of the attacker and that the 'safe' number of challenges is proportional to C. Therefore, the best available measure of the security level offered by a PUF is the parameter C, the number of challenges that can be used safely if the attacker has infinite computation power.

3.2 Speckle pattern entropy

In order to define the information content $\mathsf{H}(R)$ of a speckle pattern, we investigate the physics of multiple coherent scattering and speckle formation. Based on the physics, we turn this problem into a counting problem of the distinguishable photon states in the light leaving the PUF. First, we show that the PUF can be modeled as a strongly scattering waveguide of thickness d, cross-section $A = W^2$ and scattering length ℓ, satisfying $\lambda \ll \ell \ll d$. The waveguide allows a number of transversal modes N_{mod}. The scattering process is represented by an $N_{\mathrm{mod}} \times N_{\mathrm{mod}}$ random scattering matrix S_{ab}, specifying how much light is scattered from incoming mode b to outgoing mode a. Given a single incoming mode, the speckle pattern is fully determined by one column of the S-matrix. Hence the question is how much information is contained in one such column.

Then we calculate the speckle pattern entropy in the case where all S-matrix elements are independent. This yields an upper bound on $\mathsf{H}(R)$. In this calculation, the finiteness of the speckle pattern entropy is ultimately based on the discretisation of light in terms of photons. Finally, we take correlations between the matrix elements into account to compute a lower bound on $\mathsf{H}(R)$.

Wave guide model

First, we compute the number of incoming and outgoing modes N_{mod}. The complex amplitude of the electric field at the PUF surface can be represented as

$$E(\boldsymbol{r}) = \int_{|\boldsymbol{q}|\leq k} \frac{\mathrm{d}^2 q}{(2\pi)^2} \tilde{E}(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \; ; \; \tilde{E}(\boldsymbol{q}) = \int_{|x|,|y|\leq W/2} \mathrm{d}^2 r \; E(\boldsymbol{r}) e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}, \qquad (6)$$

where $\boldsymbol{r} = (x, y)$ denotes the position and $\boldsymbol{q} = (q_x, q_y)$ the lateral wave vector. A mode is propagating if the longitudinal (z) component of the wave, $q_z = \sqrt{k^2 - \boldsymbol{q}^2}$, is real (where $k = 2\pi/\lambda$). Hence the integration domain is a circle in $\boldsymbol{q}$-space with radius k. Note that both $E(\boldsymbol{r})$ and $\tilde{E}(\boldsymbol{q})$ are band-limited functions. Applying the Shannon-Whittaker sampling theorem [14] to the expression for $\tilde{E}(\boldsymbol{q})$ in (6), it follows that $\tilde{E}(\boldsymbol{q})$ can be characterized by discrete samples,

$$\tilde{E}(\boldsymbol{q}) = \sum_{a_x, a_y = -\infty}^{\infty} \tilde{E}(a_x \frac{2\pi}{W}, a_y \frac{2\pi}{W}) \frac{\sin(q_x W/2 - a_x \pi)}{q_x W/2 - a_x \pi} \frac{\sin(q_y W/2 - a_y \pi)}{q_y W/2 - a_y \pi}.$$

Next, we use the fact that the electric field is band-limited in q-space as well. The integers a_x, a_y have to satisfy $(a_x^2 + a_y^2)(2\pi/W)^2 \leq k^2$. The number of modes is therefore finite and is given by the number of pairs (a_x, a_y) satisfying the momentum constraint $|\mathbf{q}| \leq k$. Denoting the transverse modes as $\boldsymbol{q}_a$, we have[6]

$$\boldsymbol{q}_a = \frac{2\pi}{W}(a_x, a_y) \; ; \quad N_{\mathrm{mod}} = \#\left\{(a_x, a_y) \text{ with } |\boldsymbol{q}_a| \leq k\right\} = \frac{\pi A}{\lambda^2}. \qquad (7)$$

The integers a_x, a_y lie in the range $(-W/\lambda, W/\lambda)$. In the example of section 1.3 there are $N_{\mathrm{mod}} = 1.3 \cdot 10^7$ transversal modes. The angular distance between outgoing modes corresponds to the correlation length present in the speckle pattern as derived by [15]. The scattering process can be represented as a complex random matrix S, whose elements map incoming states to outgoing states,

$$\tilde{E}_a^{\mathrm{out}} = \sum_{b=1}^{N_{\mathrm{mod}}} S_{ab} \tilde{E}_b^{\mathrm{in}}. \qquad (8)$$

We take the distribution function of S to be symmetric in all modes. We introduce $T_{ab} = |S_{ab}|^2$, the transmission coefficient from mode b to mode a, which specifies how much light *intensity* is scattered. Given a basic challenge, consisting of a single incoming mode b, a speckle pattern corresponds to an N_{mod}-component vector $\boldsymbol{v}$, namely the b'th column of the T-matrix,

$$v_a = T_{ab}, \; b \text{ fixed}. \qquad (9)$$

Hence, the entropy of the response is given by $\mathsf{H}(\boldsymbol{v})$. Because of the mode symmetry in the distribution of S, the entropy does not depend on b. In the more

[6] If polarisation is taken into account, the number of modes doubles. In this paper we will not consider polarisation.

general case where the challenge is a linear combination of basic challenges, one can always perform a unitary transformation on the modes such that the challenge is given by a single mode in the new basis. The response is then a single column of the transformed matrix S'. Since S' has the same probability distribution as S, the entropy contained in one column of S' is the same as the entropy of one column of S. Hence, $\mathsf{H}(\boldsymbol{v})$ (9) is valid for composite challenges as well.

Weak PUFs: Upper bound on $\mathsf{H}(\boldsymbol{R})$

Here we derive an upper bound for the entropy of a speckle pattern. We start with a simplified situation, assuming the outgoing modes to be independent. This is generally not true but it gives an upper bound on $\mathsf{H}(R)$ and hence a lower bound on C. For this reason we refer to such a PUF as a *weak* PUF. It is clear that a speckle pattern cannot carry more information than the light leaving the PUF. We therefore derive an upper bound on the information content of N_{mod} light intensity states. Although the physics of multiple scattering is classical, we need the quantum description of light in terms of photons for our computation.[7] We have to count the number of *distinguishable* ways in which N_φ photons can be distributed over N_{mod} outgoing modes. To this end we estimate the number of distinguishable photon states (energy levels) N_{states} in one mode. The energy in the mode is Nh/λ, [8] where N is the number of photons in the mode. We restrict ourselves to the case of photon number statistics governed by $\langle N^2\rangle - \langle N\rangle^2 = \langle N\rangle$ without thermal noise. This Poisson relation holds for lasers and thermal light at room temperature. The more general case is treated in [14]. The energy states have a width of approximately $2\sqrt{N}$. Given the level density $1/(2\sqrt{N})$, the number of distinguishable energy levels with photon number lower than N is

$$N_{\mathrm{states}} \approx \int_0^N \frac{\mathrm{d}x}{2\sqrt{x}} = \sqrt{N}. \tag{10}$$

The energy level of the i'th mode is denoted by the integer L_i and the corresponding number of photons by $n_i \approx L_i^2$. We assume that all configurations $\{n_i\}$ have the same probability of occurring, as long as they satisfy the conservation $\sum_i n_i = N_\varphi$. From (10) we see that this translates to $\sum_i L_i^2 = N_\varphi$. Hence, the number of distinguishable configurations is given by the area of a section of an N_{mod}-dimensional sphere of radius $\sqrt{N_\varphi}$ (the section with positive L_i for all i). The area of an n-sphere is $2\pi^{n/2} r^{n-1}/\Gamma(n/2)$. Our upper bound on $\mathsf{H}(R)$ is

$$\mathsf{H}_{\mathrm{up}}(R) \approx \log\left[(\tfrac{1}{2})^{N_{\mathrm{mod}}} 2\pi^{\frac{1}{2}N_{\mathrm{mod}}} \sqrt{N_\varphi}^{N_{\mathrm{mod}}-1}/\Gamma(\tfrac{1}{2}N_{\mathrm{mod}})\right]. \tag{11}$$

Since N_{mod} is large, we can use Stirling's approximation and obtain

$$\mathsf{H}_{\mathrm{up}}(R) \approx \tfrac{1}{2}N_{\mathrm{mod}} \log\left(\tfrac{1}{2}\pi e N_\varphi/N_{\mathrm{mod}}\right). \tag{12}$$

[7] A similar situation arises in statistical mechanics, where a discretisation of classical phase space, based on quantum physics, is used to count the number of microstates.

[8] h denotes Planck's constant.

We have assumed $N_\varphi > N_{\rm mod}$, so the log in (12) is positive. The entropy increases with the number of photons, but only in a logarithmic way. Hence, errors in estimating N_φ will have a small effect on $\mathsf{H}_{\rm up}(R)$. The number of participating photons is proportional to the measurement time $\triangle t$,

$$N_\varphi = P\triangle t \cdot \lambda/(hc), \tag{13}$$

where c is the speed of light. In principle, it is possible to completely characterize the PUF by performing a single very long measurement. However, as seen from (13) and (12), substituting $\mathsf{H}_{\rm up}(R) \to \mathsf{H}(K)$, $\triangle t$ is then exponential in $\mathsf{H}(K)$. Information can be extracted from the PUF much faster, namely linearly in $\triangle t$, by doing many fast measurements. Using the example numbers of section 1.3, we have $N_\varphi = 2.5 \cdot 10^{12}$ and the upper bound is $\mathsf{H}_{\rm up}(R) < 1.2 \cdot 10^8$.

Strong PUFs: Lower bound on $\mathsf{H}(R)$

In multiple scattering PUFs, the modes at the outgoing surface are correlated. In [16] a correlation function was obtained for the elements of the T-matrix,

$$\begin{aligned}\frac{\langle \delta T_{ab}\delta T_{a'b'}\rangle}{\langle T_{ab}\rangle\,\langle T_{a'b'}\rangle} &= D_1\delta_{\triangle \boldsymbol{q}_a,\triangle \boldsymbol{q}_b}F_1(\frac{d}{2}|\triangle \boldsymbol{q}_b|) \\ &+\frac{D_2}{4gN_{\rm mod}}\left[F_2(\frac{d}{2}|\triangle \boldsymbol{q}_a|)+F_2(\frac{d}{2}|\triangle \boldsymbol{q}_b|)\right]+\frac{D_3}{(4gN_{\rm mod})^2}\end{aligned} \tag{14}$$

where $\delta T_{ab} = T_{ab} - \langle T_{ab}\rangle$, $\langle\cdot\rangle$ is the average over all scatterer configurations and

$$F_1(x) = x^2/\sinh^2 x\,; \qquad F_2(x) = 2/(x\tanh x) - 2/\sinh^2 x \tag{15}$$

with $\triangle \boldsymbol{q}_a = \boldsymbol{q}_{a'} - \boldsymbol{q}_a$, g the transmittance $N_{\rm mod}^{-1}\sum_{ab} T_{ab} \approx \ell/d$, and D_i constants of order unity. Due to the correlations, the number of degrees of freedom $N_{\rm dof}^{\rm out}$ in the speckle pattern is less than $N_{\rm mod}$. We calculate $N_{\rm dof}^{\rm out}$ following the approach of [8], but we make use of (14). We sum over the correlationsin the vector $\boldsymbol{v}$ to obtain the *effective cluster size* μ. μ represents the number of variables correlated to a given v_a (for arbitrary a). The vector $\boldsymbol{v}$ can be imagined to consist of uncorrelated clusters of size μ, where each cluster contains exactly one degree of freedom. This means that we approximate $\boldsymbol{v}$ by a vector of $N_{\rm dof}^{\rm out} = N_{\rm mod}/\mu$ independent cluster-size entries. Denoting the variance of v_a as σ_a and neglecting the D_3 term, the correlations within $\boldsymbol{v}$, obtained from (14), are given by

$$C_{aa'} = \langle \delta v_a\ \delta v_{a'}\rangle/(\sigma_a\sigma_{a'}) = \delta_{aa'} + D_2/(D_1 4gN_{\rm mod})\left[\tfrac{4}{3} + F_2(\tfrac{1}{2}d|q_a - q_{a'}|)\right]. \tag{16}$$

The D_2 term consists of the sum of a short-range (F_2) term and a long-range contribution (4/3). From (16) we obtain μ and the number of degrees of freedom,

$$\mu = \sum_a C_{aa'} \approx \frac{D_2}{3D_1}\frac{d}{\ell}\,; \quad N_{\rm dof}^{\rm out} = \frac{N_{\rm mod}}{\mu} \approx \frac{3D_1}{D_2}\frac{\pi A}{\lambda^2}\frac{\ell}{d}. \tag{17}$$

Here we have neglected the summation over the F_2-term, since $F_2(x)$ asymptotically falls off as $2/x$ for large x. We have also neglected the contribution $\sum_a \delta_{aa'} = 1$ with respect to d/ℓ.

The speckle entropy is calculated by repeating the level counting computation of the 'Weak PUFs' section, but now with modified parameters. Every output mode within a cluster of size μ emits exactly the same amount of light. Consequently, the problem of distributing N_φ photons over N_{mod} modes is equivalent to the problem of distributing N_φ/μ bunches of μ photons over N_{mod}/μ clusters. Performing the substitution $\{N_{\mathrm{mod}} \to N_{\mathrm{mod}}/\mu, N_\varphi \to N_\varphi/\mu\}$ into (12) we obtain

$$\mathsf{H}_{\mathrm{low}}(R) = \frac{N_{\mathrm{dof}}^{\mathrm{out}}}{2} \log\left(\frac{\pi e}{2}\frac{N_\varphi}{N_{\mathrm{mod}}}\right) = \frac{3\pi D_1}{2D_2}\frac{A\ell}{\lambda^2 d}\log\left(\frac{\pi e}{2}\frac{N_\varphi}{N_{\mathrm{mod}}}\right). \qquad (18)$$

Substituting into (18) relation (13) and the numbers given in section 1.3, we have $\mathsf{H}_{\mathrm{low}}(R) \approx 4 \cdot 10^6$. By assuming that several modes carry the same photon state, we have underestimated $N_{\mathrm{dof}}^{\mathrm{out}}$. Therefore, the result (18) is indeed a lower bound on $\mathsf{H}(R)$. Furthermore, we have assumed that all the information present in the outcoming light is recorded by an ideal detector, capturing all the light in a sphere surrounding the PUF. This is the optimal situation for an attacker. Hence we err on the side of safety.

3.3 The security parameter

We now use the results of section 3.2 to estimate the security parameter. We assume that we are in the regime where the PUF can be probed to such an extent that all bits can be determined by measurements. In this regime we have $C = \mathsf{H}(K)/\mathsf{H}(R)$, and after substitution of the upper bound (12) and the lower bound (18) for $\mathsf{H}(R)$ we find that C lies in the interval

$$\left(\min\left\{\frac{2}{\pi}\cdot\frac{1}{\log(\frac{\pi e}{2}\frac{N_\varphi}{N_{\mathrm{mod}}})}\cdot\frac{d}{\lambda},\ N_{\mathrm{mod}}\right\}, \min\left\{\frac{2}{3\pi}\cdot\frac{1}{\log(\frac{\pi e}{2}\frac{N_\varphi}{N_{\mathrm{mod}}})}\cdot\frac{d^2}{\lambda\ell},\ N_{\mathrm{mod}}\right\}\right) \qquad (19)$$

The $\min\{\cdots, N_{\mathrm{mod}}\}$ function reflects the fact that there are no more than N_{mod} basic challenges. The result (19) has the following properties:

- C grows with increasing d/λ, since the PUF entropy is proportional to d/λ.
- In addition, the upper bound on C grows with increasing d/ℓ. This is a measure for the number of scattering events N_{sc} taking place before a photon exits the PUF. (Assuming a random walk, $d/\ell \propto \sqrt{N_{\mathrm{sc}}}$). Hence, multiple scattering increases the cryptographic strength of a PUF.
- (19) refers to one illuminated area $A = W^2$. By shifting the laser over a distance equal to the diameter of the laser spot, one illuminates a new subvolume of the PUF with the same number of challenges. This means that the total number of independent challenges C_{tot} is given by

$$C_{\mathrm{tot}} = C \cdot A_{\mathrm{PUF}}/A. \qquad (20)$$

Using the numbers from section 1.3, (19) gives $3 \cdot 10^4 \leq C_{\mathrm{tot}} \leq 1 \cdot 10^6$. In order to achieve this many distinct challenges in practice, an angular accuracy of laser positioning is required of the order of 1 mrad. This is easily achievable.

We emphasize once more that the security parameter has been computed from an *information-theoretic* point of view. This means that an attacker who has gathered $C_{\rm tot}$ CRPs in principle knows the complete CRP behaviour. He is *in principle* able to compute the response to a new challenge (one for which he has not seen the response before). *In practice* the security might be much stronger and is based on the following assumptions: (i) the so-called *forward problem* (computing the response for a given challenge) is difficult and (ii) interpolation techniques do not allow for an accurate prediction of a response, given responses to nearby challenges. This means that one can use more than $C_{\rm tot}$ CRPs safely.

Finally, we compare our result (19) to [3, 4]. Their approach is based on the memory angle $\delta\theta \propto \lambda/d$ [16] and does not take the density of scatterers into account. Dividing the half-sphere into pieces of solid angle $\delta\theta^2$, they obtain a number of CRPs proportional to d^2/λ^2, representing the number of obtainable responses that look mutually uncorrelated. This number is larger than our upper bound for C by a factor $\propto \ell/\lambda$. The two approaches give comparable results only in the limit of extremely strong scattering, $\ell \approx \lambda$.

4 Attacks and countermeasures

We discuss the following threat. An attacker steals somebody's PUF and tries to characterize the PUF without being noticed by the owner. In particular this means that the PUF has to be returned to the owner within a short time period.

4.1 Brute force

It follows from the definition of C that *in principle* only C measurements are required to fully characterize a PUF. However, an attacker faces the problem that CRP intrapolation is difficult. Consequently, a brute force attack may be more feasible. The brute force attack is an attempt to exhaustively record the full set of CRPs. The responses are stored in a database. Let us assume that a challenge takes the form of a single incoming transverse momentum mode; it is clear that the number of possible challenges is of order $N_{\rm mod}$. The required storage space is relatively small, since it is not necessary to store complete speckle patterns, but only the keys/identifiers derived from them. The measurement duration for this attack is $N_{\rm mod}\triangle t \cdot A_{\rm PUF}/A = \pi A_{\rm PUF}\triangle t/\lambda^2$. Using the wavelength and the PUF area from the example in section 1.3, and taking $\triangle t$ of the order of 10ms, we have a total duration in the order of hundreds of days. This is too long for the attack to go unnoticed in the scenario sketched above.

We emphasize the necessity of enforcing "long" measurement times $\triangle t$.

4.2 The Slow PUF

A long integration time in the detector can be achieved by attaching a gray filter (irremovably) to the PUF. Let us denote the transmission of the combined PUF and gray filter by $\eta_{\rm PUF}$. In the detector the incoming photons are converted to

electrons with quantum efficiency η_{Q}. The actual signal of each detector cell is the number of electrons N_{e} collected in time $\triangle t$. The number of cells in the detector is denoted as N_{cells}. The generation of photo-electrons is a Poisson process,

$$\left\langle N_{\mathrm{e}}^2 \right\rangle - \left\langle N_{\mathrm{e}} \right\rangle^2 = \left\langle N_{\mathrm{e}} \right\rangle = \frac{\eta_Q N_\varphi}{N_{\mathrm{cells}}} = \frac{\eta_{\mathrm{Q}} \eta_{\mathrm{PUF}} P}{N_{\mathrm{cells}}(hc/\lambda)} \triangle t. \tag{21}$$

The signal to noise ratio SNR can at most be equal to $\left\langle N_{\mathrm{e}} \right\rangle^2 / (\left\langle N_{\mathrm{e}}^2 \right\rangle - \left\langle N_{\mathrm{e}} \right\rangle^2) = \left\langle N_{\mathrm{e}} \right\rangle$, since there may be other noise sources which are not taken into account here. Hence, (21) gives a lower bound on $\triangle t$, proportional to the SNR. According to Shannon's theorem [12], the number b of useful bits that can be extracted from any signal is limited in the following way,

$$b \leq \tfrac{1}{2} \log(1 + \mathrm{SNR}). \tag{22}$$

In our case, b represents the number of bits used for gray level representation. Combining (21) and (22) we obtain

$$\triangle t \geq \frac{(hc/\lambda) N_{\mathrm{cells}}}{\eta_{\mathrm{Q}} \eta_{\mathrm{PUF}} P} (2^{2b} - 1). \tag{23}$$

For example, taking $\eta_{\mathrm{Q}} = 0.3$, $\eta_{\mathrm{PUF}} = 0.001$, $N_{\mathrm{cells}} = 3 \cdot 10^6$ and $b = 4$ in the example of section 1.3, we get $\triangle t \geq 1\mathrm{ms}$.

This gives a fundamental physical lower bound on the integration time, which can therefore never be reduced by technological advances. Given a challenge-response setup with fixed P and η_{Q}, (23) shows that the integration time can be increased in the following ways: (i) by decreasing the transmission η_{PUF}, (ii) by increasing the number of gray levels 2^b that should be detected in order to correctly process the speckle pattern and (iii) by increasing the number of pixels that should be resolved by the detector to ensure correct processing. All three methods have their limitations.

An attacker can of course use the best detector in existence (high η_{Q}). He can also use any laser that he desires (high P). Especially the use of a high-intensity laser can dramatically shorten the integration time. This can be prevented by adding to the PUF a photo-layer which darkens (permanently or temporarily) when irradiated by light above a certain threshold intensity.

5 Conclusions

We have introduced a general theoretical framework to study the secret key capacity of PUFs. We propose to use the number C of independent CRPs as a security parameter. This parameter represents a security measure that is not affected by technological or computational advances. In practice one may use many more CRPs safely, under the assumption that correlations between CRPs are hard to exploit computationally.

For optical PUFs we have derived an analytical expression for C. In order to make brute force attacks difficult, long measurement times have to be enforced. This has been analyzed for the case of optical PUFs.

References

1. B. Gassend et al., *Controlled Physical Random Functions*, Proc. 18th Annual Computer Security Applications Conf., Dec. 2002.
2. B. Gassend et al., *Silicon Physical Unknown Functions*, Proc. 9th ACM Conf. on Computer and Communications Security, Nov. 2002.
3. R. Pappu, *Physical One-Way Functions*, Ph.D. thesis, MIT 2001.
4. R. Pappu et al., *Physical One-Way Functions*, Science Vol. 297, Sept 2002, p.2026.
5. B.L.P. Gassend, *Physical Random Functions*, Master's Thesis, MIT 2003.
6. P. Tuyls, B. Škorić, *Secret Key Generation from Classical Physics*, Proceedings of the Hardware Technology Drivers for Ambient Intelligence Symposium, Philips Research Book Series, Kluwer, 2005.
7. M. Magnor, P. Dorn and W. Rudolph, *Simulation of confocal microscopy through scattering media with and without time gating*, J.Opt.Soc.Am. B, Vol. 19, no. 11 (2001), pp 1695–1700.
8. J. F. de Boer, *Optical Fluctuations on the Transmission and Reflection of Mesoscopic Systems*, Ph D thesis, 1995, Amsterdam.
9. H. Furstenberg, *Noncommuting Random Matrices*, Trans. Am. Math. Soc. 108, 377, 1963.
10. J.P. Linnartz, P. Tuyls, *New Shielding Functions to enhance Privacy and Prevent Misuse of Biometric Templates*, Proc. 4th International Conference on Audio and Video based Biometric Person Authentication, LNCS2688, Guildford UK, June 9-11, 2003.
11. E. Verbitskiy, P. Tuyls, D. Denteneer, J.P. Linnartz, *Reliable Biometric Authentication with Privacy Protection*, Proc. of the 24th Symposium on Information Theory in the Benelux.
12. T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Wiley ans Sons, New York, 1991.
13. K. Petersen, *Ergodic Theory*, Cambridge University Press, 2000.
14. D. Gabor, *Light and Information*, in E. Wolf, Ed., Progress in Optics Vol. I, North-Holland, Amsterdam 1961.
15. J. W. Goodman, *Statistical properties of laser speckle patterns*, in Laser Speckle and Related Phenomena, 2nd ed., J. C. Dainty, Ed. New York: Springer-Verlag, 1984.
16. S. Feng, C. Kane, P.A. Lee and A.D. Stone, *Correlations and Fluctuations of Coherent Wave Transmission through Disordered Media*, Phys.Rev.Lett. Vol.61 No.7 (1988), pp 834–837.