VIA PUF-based Authenticated Key Agreement and Modelling Attack on Strong PUF

Project-II (EC47004) report submitted to

Indian Institute of Technology Kharagpur

in partial fulfilment for the award of the degree of

Bachelor of Technology

in

Electronics and Electrical Communication Engineering

by

Posina Venkata Sai Nagesh

(20EC39023)

Under the supervision of

Prof. Indrajit Chakrabarti

Department of Electronics and Electrical Communication Engineering

Indian Institute of Technology Kharagpur

Spring Semester, 2023-24

April 29, 2024

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: April 29, 2024 (Posina Venkata Sai Nagesh)
Place: Kharagpur (20EC39023)

DEPARTMENT OF ELECTRONICS AND ELECTRICAL COMMUNICATION ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

KHARAGPUR - 721302, INDIA

CERTIFICATE

This is to certify that the project report entitled "**VIA PUF-based Authenticated Key Agreement and Modelling Attack on Strong PUF**" submitted by **Posina Venkata Sai Nagesh** (Roll No. 20EC39023) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree of Bachelor of Technology in Electronics and Electrical Communication Engineering is a record of bonafide work carried out by him under my supervision and guidance during Spring Semester, 2023-24.

Date: April 29, 2024

Place: Kharagpur

Prof. Indrajit Chakrabarti
Department of Electronics and Electrical
Communication Engineering
Indian Institute of Technology Kharagpur
Kharagpur - 721302, India

Abstract

Name of the student: **Posina Venkata Sai Nagesh** Roll No: **20EC39023**

Degree for which submitted: **Bachelor of Technology**

Department: **Department of Electronics and Electrical Communication Engineering**

Thesis title: **VIA PUF-based Authenticated Key Agreement and Modelling Attack on Strong PUF**

Thesis supervisor: **Prof. Indrajit Chakrabarti**

Month and year of thesis submission: **April 29, 2024**

Physical Unclonable Functions (PUFs) provide cost-effective and secure authentication through advanced hardware technology. Many PUF-based Authenticated Key Agreement (AKA) protocols have been proposed, often relying on Error Correction Code (ECC) to address noisy responses induced by environmental variations such as temperature and voltage. The VIA PUF utilizes a distinctive method based on contact formation probability. By leveraging the stochastic nature of contact formation between metal and silicon in the chip's interconnect layer, the VIA PUF ensures reliability through intentional design choices. As we are measuring parasitic resistances of the circuit, VIA PUF is resilient to environmental variations. The PUF bit-cell was designed using the Cadence Virtuoso tool using UMC 180nm CMOS technology. This design achieves a 50.04% uniqueness value and 50.01% uniformity, closely aligining with the ideal value of 50%. Also it is proven to be reliable under various temperatures ($-40°C, 25°C$, and $125°C$) and voltages (nominal, $\pm 10\%$)

while considering all potential process corner variations, further justify the robustness of the designed 126-bit VIA PUF array with zero bit error rate (BER) after the pre-selection. Additionally, the resistive region discard ratio for 630 VIA PUF bit-cells design is noted to be 10%.

The thesis also delves into a novel type of PUF called obfuscated interconnections physically unclonable function (OIPUF), aimed at resisting modeling attacks. To validate the design of the PUF, a mathematical model of the OIPUF is computed. The extraction of challenge-response pairs (CRPs) involves implementing the OIPUF in the Xilinx Vivado tool and then running it onto an Artix 7 FPGA. Various machine learning modeling attacks, such as Support Vector Machines (SVM) and Artificial Neural Networks (ANN), are employed to evaluate the performance of the OIPUF.

Keywords - VIA PUF, authentication, key exchange, IoT devices, parasitic resistance, Error Correction codes (ECC), Cadence virtuoso, OIPUF, Field-programmable gate array (FPGA), machine learning (ML) modelling attacks, obfuscated interconnection, Xilinx Vivado, Artix 7

Acknowledgements

I wish to express my heartfelt gratitude to my thesis advisor, Prof. Indrajit Chakrabarti, whose unwavering support and guidance have been instrumental in my reasearch. Professor Indrajit's expertise, mentorship, and willingness to engage in constructive discussions have greatly enriched my academic journey. His open door policy and willingness to offer guidance and insights, while still allowing me the autonomy to explore and learn independently have been invaluable. I would also like to extend my deepest appreciation for my mentor Sivappriya Manivannan. She was with me the whole time and provided constant and effective mentorship. She helped me debug my mistakes and make steady progress. I would also like to express my heartfelt gratitude to my parents for their constant support and encouragement throughout my years of education, as well as during the rigorous process of conducting research and writing this thesis.

Contents

List of Abbreviations and Acronyms

PUF	Physically Unclonable Function
APUF	Arbiter Physically Unclonable Function
AKA	Authenticated Key Agreement
SK	Secret Key
UMC	United Microelectronics Corporation
CMOS	Complementary Metal-Oxide-Semiconductor
HW	Hamming Weight
HD	Hamming Distance
CRP	Challenge Response Pair
iPUF	Interpose Physically Unclonable function
OIPUF	Obfuscated Interconnection Physically Unclonable Function
SVM	Support Vector Machines
LR	Logistic Regression
CMA-ES	Covariance Matrix Adaption Evolution Strategy
ANN	Artificial Neural Network
OI	Obfuscated Interconnections

Chapter 1

VIA PUF: Introduction and Previous Work

1.1 Problem Statement and Research Question

The demand for affordable yet secure authentication solutions is on the rise. Ideal authentication circuits should occupy minimal space, be tamper-resistant, and exhibit low power consumption. Physically Unclonable Functions (PUFs) offer a promising solution, as they provide a primitive means of key storage without relying on volatile memory. PUFs generate a key on demand, leveraging the inherent physical characteristics of the circuit. Extracting the key from a PUF circuit without altering its physical properties is challenging.

This project specifically investigates a type of PUF known as VIA PUF or CONTACT PUF [3], which relies on contact formation probability. The contact occurs within the interconnecting layer between silicon and metal. Within a specific range of contact hole sizes, this process becomes stochastic, and the resulting randomness is utilized to generate the PUF output. The project delves into assessing the reliability and randomness of the PUF, along with the challenges encountered during its integration into device circuitry.

1.2 Objectives and Challenges

The project's goal is to design PUF circuitry and verify its performance. To achieve this, certain challenges must be addressed. First and foremost, the reliability of the PUF circuit is crucial. It should consistently produce the same output under varying environmental conditions, remaining unaffected by changes in temperature, pressure, or voltage fluctuations. Each PUF bitcell, responsible for a single-bit output, must be individually reliable. To achieve this, unreliable bitcells, referred to as resistive contacts should be discarded. This is done by thresholding PUF's output voltage.

Another key criterion for an effective PUF is its randomness. The output should be unpredictable and exhibit true randomness. The project specifically aims to design a via PUF circuit that fulfills both criteria of reliability and randomness. Determining the optimal number and range of contact hole sizes poses a challenge to ensure the randomness of the PUF output.

1.3 Previous work

1.3.1 Secure Authenticated Key Agreement (AKA)

1.3.1.1 Enrolment phase

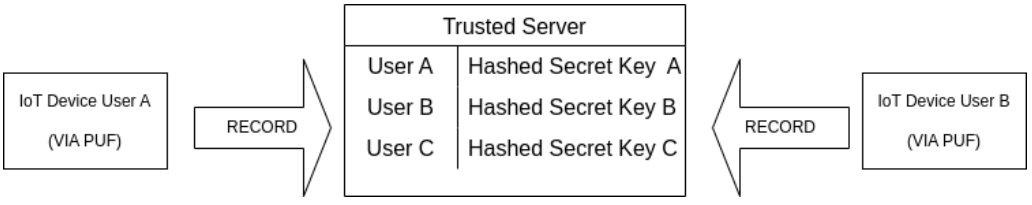

FIGURE 1.1: Enrolment phase of IoT devices with the trusted server.

The PUF enrollment is a standard, one-time (prior to deployment) and secure phase with no opportunity to "spy" for an adversary. Each user or the IoT device is registered to the server. The server stores the mapping of user to key generated by VIA PUF. Then the IoT device embedded with PUF is given to the client.

1.3.1.2 Key Exchange Protocol

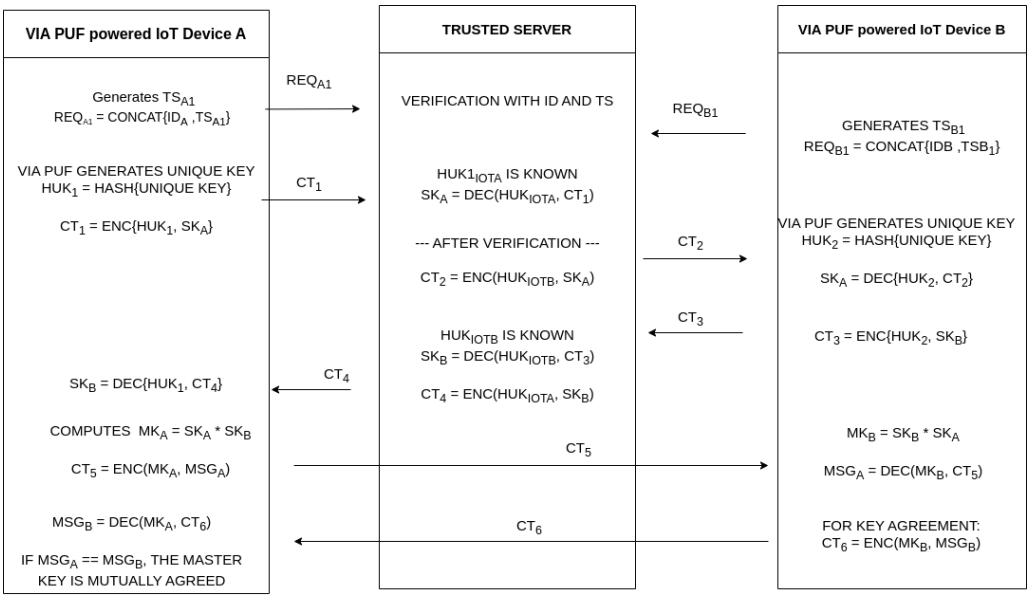

FIGURE 1.2: Authenticated Key agreement protocol (AKA) using VIAPUF.

Terminology

TS:Time-stamp , **REQ**:Request , **ID**:Unique ID of Device , **CT**:Cipher Text , **HUK**:Hashed Unique Key , **SK**:Secret Key , **MK**:Master Key , **ENC**:Encrypting operation , **DEC**:Decrypting operation , **MSG**:Message

In this secure communication protocol, IoT devices exchange secret keys via a trusted server. Upon registration, each device sends its unique ID and timestamp to the server for verification. The server facilitates key exchange by encrypting and transmitting secret keys between devices, establishing mutual agreement through message verification using derived master keys. The encryption of the secret key using the hashed key generated by the VIA PUF ensures secure transmission to the trusted server, forming the cornerstone of the protocol's security.

1.3.2 Operation and Implementation of the VIA PUF

Previously, it is discussed how the presence of active components can compromise the reliability of the PUF. A new solution is proposed to use the parasitic resistance as the entropy source for the PUF.

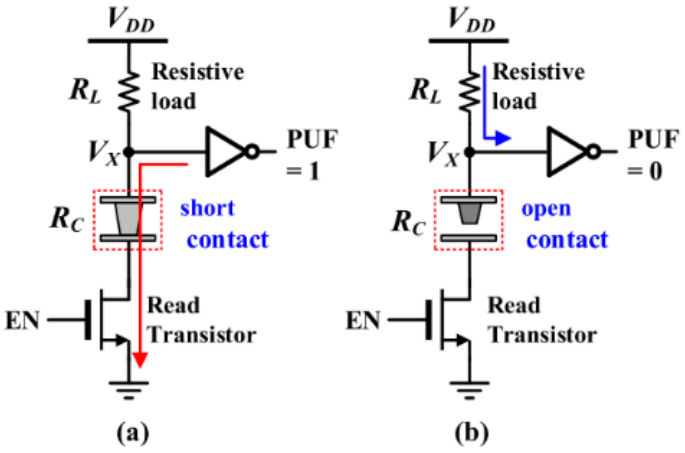

FIGURE 1.3: Basic contact PUF bitcell operation for (a) short and (b) open contacts.[3]

Figure 1.3 shows a basic contact PUF bitcell.In addition to the classification of contacts as open or short, there exists another category known as resistive contacts. These occur when the connection between the metal and silicon is only marginal. In such cases, it is possible to analyze the circuit by representing the contact as a resistor with designated value. The main idea is instead of utilizing contact holes, the parasitic resistance of the circuit can be conceptualized as contact resistance.

To observe the characteristics of the circuit (Figure 1.4) the output voltage is observed for a wide range of contact resistance values in Cadence virtuoso tool using UMC 180nm technology.

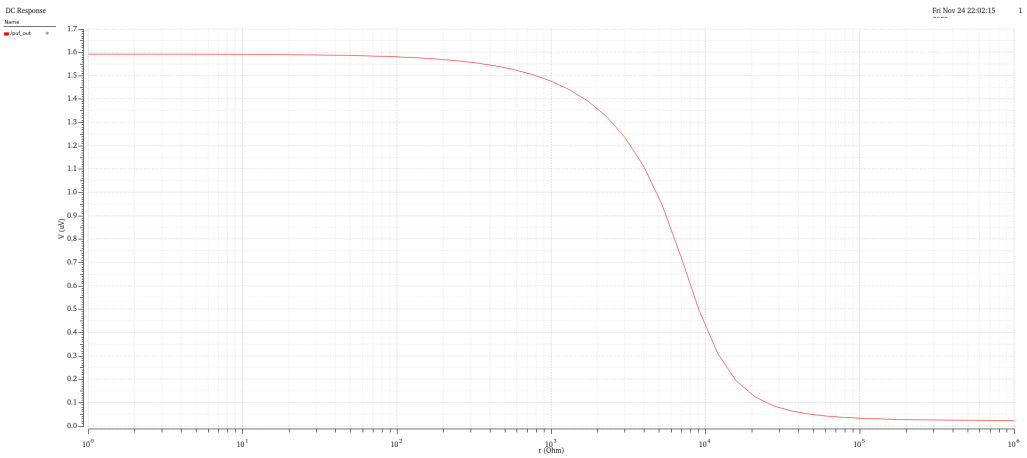

FIGURE 1.4: VIA PUF output v/s the contact resistance.

Chapter 2

VIA PUF: EXPERIMENTAL SIMULATIONS AND RESULTS

2.1 Parametric Analysis of VIA PUF

The VIA PUF is developed utilizing the Cadence Virtuoso design suite within the UMC 180 nm CMOS process technology. The typical parasitic resistances fall within the range of 10Ω. Therefore, to represent the contact resistance as parasitic resistance, it is necessary to acquire the characteristics depicted in Figure 1.4 for contact resistance within the 10Ω range, achievable by adjusting the properties of the NMOS transistors.

Once this is done, a comprehensive testing phase was conducted, involving 126 PUF bit-cell designs with different (R_L) values subjected to all possible process corners for MOSFETs (ss, tt, ff, snfp, and fnsp) and resistor (R_L) settings at min, typ, and max. The behavior of the VIA PUF bit-cell is determined by its output voltage w.r.t. contact resistance R_c value ranging from 0.3Ω to 15Ω in three R_L conditions viz., R_{LL}, R_{LM}, and R_{LU} as depicted in Figure 2.1. Given the inherently low resistance of metal contact, we opted for small resistance values in the order of a few ohms. Consequently, within the specific R_L range of 9Ω to 15.5Ω, the contact resistance (R_c) falls within the minimum resistance range. In this experiment, we exclude the VIA PUF bit-cells with resistive contacts, and this step is referred to as pre-selection.

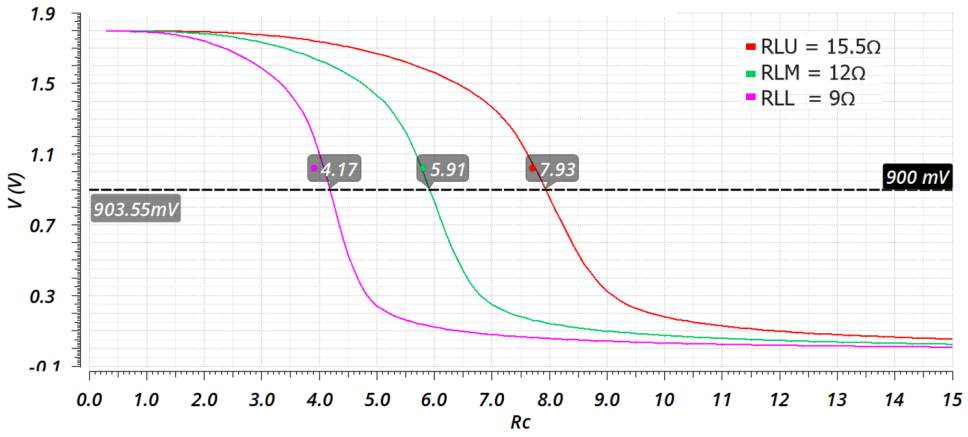

FIGURE 2.1: $V_{OUT}(V)$ vs $R_c(\Omega)$ at three different Resistive loads.

2.2 Uniformity Evaluation

Uniformity is a critical metric for evaluating a Physical Unclonable Function (PUF), indicating how evenly it produces 0s and 1s in its responses, reflecting its balanced performance. It's essential that the responses from a PUF are equally probable across all possible variations in process corners for components like the load resistor (RL) and MOS transistors. Our analysis, based on a sample of 126 bit-cells, encompassing all conceivable process corners, reveals a uniformity rate of **51.01%**. This result closely approaches the ideal value of 50%, indicating a high level of uniformity in the PUF's responses.

2.3 Randomness Evaluation

Randomness in a PUF ensures that its outputs are statistically independent and evenly distributed, making it challenging for adversaries to replicate or predict its responses. The randomness of the VIA PUF is directly related to the contact failure probability. The contact formation probability of 0.5 ensures ideal value for randomness. But this would also make the PUF unreliable to temperature and voltage fluctuations. Since the contact failure probability is dependent on the contact hole size a initial experiment is conducted to evaluate the probability of contact failure

across 630 PUF bit-cells for 100 different contact hole sizes. Figure 2.2 shows this result.

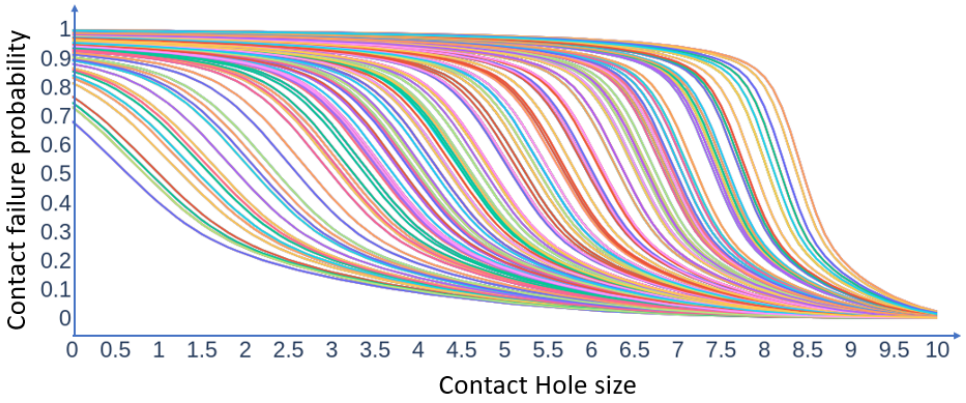

FIGURE 2.2: Contact failure probability vs Contact Hole size for 630 PUF bit-cells in all three resistive load regions.

To improve the randomness, we implement a two-step XOR operation as explained in Section 2 on the PUF output bit cells. Utilizing the existing PUF bit cells, which generate 630-bit responses for each contact hole size, totaling 100 variations, we conduct the initial 16-bit XOR operation.

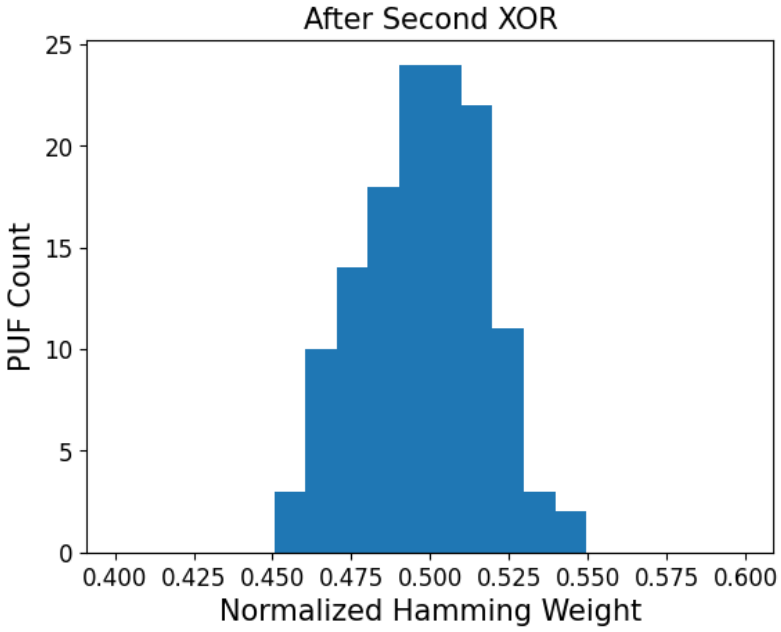

FIGURE 2.3: Randomness evaluation after 2step XOR operation.

During this operation, we sample 7,680 bits as responses. In each iteration, 12 bits are sampled and XORed to produce a single bit, repeated for 768 cycles. These resulting output bits form the response of a single PUF instance. A total of 128 such PUF instances are modeled, and Figure 2.3 presents a histogram depicting the chip count versus normalized Hamming Weight (HW). As evident, the histogram demonstrates a peak at 0.5 (normalized HW), indicating the optimal or best-case uniqueness.

2.4 Uniqueness Evaluation

The uniqueness of a PUF response gauges its individuality relative to other PUF responses, typically evaluated through the Hamming Distance (HD) metric. It is defined as follows:

$$\text{Uniqueness} = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \text{HD}_{R_i,R_j} \times \frac{1}{n} \times 100\% \tag{2.1}$$

where Ri and Rj are the n-bit responses from the ith and jth chips respectively among the k chips, and HD is the hamming distance.

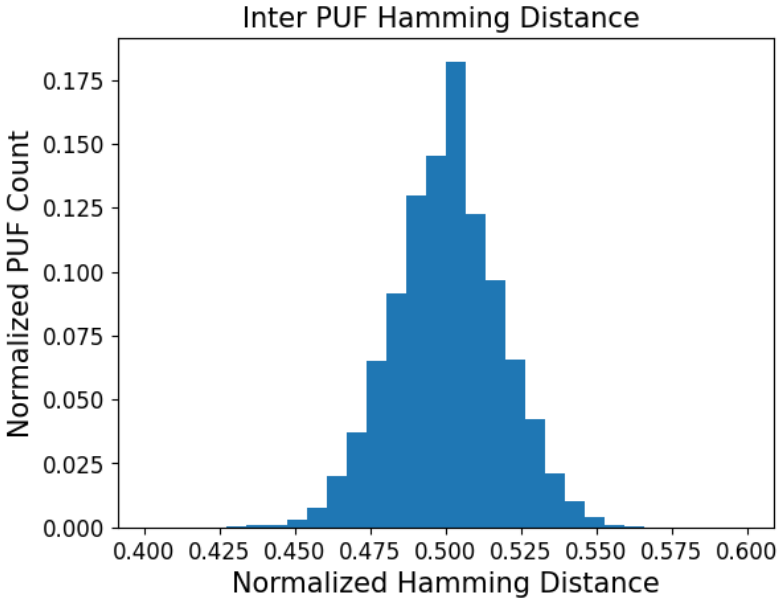

FIGURE 2.4: Uniqueness evaluation by interchip hamming distance.

TABLE 2.1: Threshold on PUF Voltage output.

Voltage range	PUF Response
$V_{out} \geq 1.28V$	Logic 1
$V_{out} \leq 0.52V$	Logic 0
$0.52V < V_{out} < 1.28V$	Resistive contacts

TABLE 2.2: Temperature Test Results for 126 VIA PUF bit-cells.

Condition	**No. of bit-cells**	**Range**	**Result**
Temperature (°C)	126	-40	Pass
		25	Pass
		125	Pass
Voltage supply (V)	126	1.62	Pass
		1.8	Pass
		1.98	Pass

In our simulation involving 128 PUF instances, we computed the HD between all possible pairs of responses, resulting in an average simulated value of 50.04%. An optimal uniqueness value of 50% indicates maximum differentiation. As depicted in Figure 2.4, the distribution of PUF bit-cells based on inter-chip HD reveals that the majority of chip pairs exhibit a 50% difference. The mean of the distribution was found to be 0.5, with a variance of 0.003, demonstrating a well-differentiated and consistent pattern among the chip responses.

2.5 Reliability Evaluation

The reliability of a PUF is measured by its capacity to produce consistent responses in varying environmental conditions.

After excluding resistive contacts with the discard ratio of 10%, bit-cell responses for 126 PUF designs were recorded under three distinct temperature conditions: -40°C, 25°C, and 125 °C for all possible process corners. It is observed that the designed PUF bit-cells produce zero BER after pre-selection.

Additionally, a similar evaluation was conducted under varying supply voltages, further confirming the reliability of the VIA PUF design. The PUF's output voltage underwent thresholding to categorize responses into logic 0, logic 1, and resistive contacts, as illustrated in Table 2.1.

The outcomes of the tests, examining the reliability of PUF bit-cells under temperature and voltage variations, are presented in Table 2.2. This information enhances our understanding of VIA PUF's suitability for practical applications in the real world, contributing valuable findings to the research.

Chapter 3

Modelling Attacks on a Strong PUF

3.1 Background

One of the fundamental characteristics of a Physically Unclonable Function (PUF) is its unclonability, meaning that it should be computationally infeasible to replicate two physical instances of the PUF with identical characteristics. PUFs are typically categorized into weak PUFs and strong PUFs. A strong PUF can be probed with an exponential number of challenges, yielding an exponential number of responses, making them suitable for authentication protocols, key generation, and storage. In contrast, weak PUFs are constrained to a limited set of challenge-response pairs. The essence of a PUF lies in its challenge-response pairs (CRPs), where a strong PUF is defined as a function that maps input challenges to output responses, with each PUF concealing a learning function responsible for this mapping. However, a drawback is that adversaries can attempt to deduce this hidden function with sufficient CRPs, employing various machine learning techniques.[10]

Proposed PUF designs aim to withstand such attacks. For instance, in the case of a basic APUF design[10], it has been demonstrated that certain machine learning

attacks can achieve accuracy rates exceeding 90%. To address this vulnerability, improvements have been made, such as incorporating XOR operations.

While a XOR-APUF may be susceptible to attacks with a small number of XOR operations, it has been shown to provide security when a sufficient number of XOR operations are utilized. In practice, the security of a PUF is often evaluated by the number of CRPs required to model the PUF with a high prediction accuracy, typically around 90%.

3.2 Problem Statement and Research Question

The security of a Physically Unclonable Function (PUF) is paramount, especially when employed for authentication purposes. An adversary armed with a sufficiently accurate software model of a PUF can potentially interfere and disrupt its operations. To combat such modeling attacks, various new PUF designs have emerged, including the interpose-PUF (iPUF) [9], a strong PUF exploiting stagewise Obfuscated Interconnections-PUF (OIPUF) [8], and a robust PUF based on the intrinsic transfer paths of a conventional digital multiplier[11].

This project focuses on investigating the security aspects of a specific type of PUF, known as OIPUF. The objective is to delve into the hardware model of OIPUF and utilize this understanding to develop improved machine learning techniques for modeling it accurately.

3.3 Classical machine learning attacks

3.3.1 Support Vector machines (SVM)

Support Vector Machines (SVMs) are a type of supervised learning model known for their effectiveness in classification and regression tasks. They operate by identifying the optimal hyperplane that best separates different classes within the feature space, maximizing the margin between the hyperplane and the nearest data points from each class, known as support vectors. SVMs are adept at handling non-linear

data through kernel functions, making them well-suited for modeling Physically Unclonable Functions (PUFs). PUF response data is often non-linearly separable, and SVMs can effectively classify and predict PUF responses even in noisy and complex environments. Additionally, SVMs offer robustness to overfitting, crucial when dealing with limited and noisy PUF response data. By maximizing the margin between classes, SVMs can generalize well to unseen data, ensuring accurate modeling of PUFs.

3.3.2 Logistic Regression (LR)

Logistic Regression (LR) offers a mathematical framework for modeling the behavior of a Physically Unclonable Function (PUF). In this context, LR aims to estimate the probability $P(r = 1 \mid \mathbf{C})$ of observing a particular PUF response $r = 1$ given the input challenges $\mathbf{C}$. This probability is expressed using the logistic function defi defined as in Eqn (3.1), where the weight vector $\mathbf{w}$ and bias term b are parameters learned during model training. During training, the LR model is optimized to maximize the likelihood of the observed PUF responses given the input challenges, adjusting its parameters iteratively. Once trained, the LR model can predict the probability of a specific response for new input challenges, facilitating classification of PUF responses based on their likelihood.

$$P(r = 1 \mid \mathbf{C}) = \frac{1}{1 + e^{-\mathbf{w}^\intercal \mathbf{C} - b}} \tag{3.1}$$

3.3.3 Reliability based attacks (CMA-ES)

[10] The main idea of the reliability-based CMA-ES attack is to extract the response multiple times for the same challenge and use the stability of the response bits to model the PUF. In traditional CMA-ES attacks, the optimization algorithm focuses solely on minimizing the difference between the predicted responses of a simulated PUF model and the actual responses observed from the target PUF. However, the reliability-based approach introduces an additional criterion that evaluates the consistency or reliability of the predicted responses.

The reliability based CMA-ES techniqueis proven to outperform all the classical machine learning algorithms for a large number of XORs.

3.3.4 Artificial Neural Networks (ANN)

ANN for modeling attacks is considered as the most powerful black-box attack on PUF. The starting layer of the neural network corresponds to the input feature vector. The feature vector plays a pivotal role in the performance of network.
An APUF can be mathematically modeled as a linear additive delay model as shown in below equation -

$$\Delta = w[0]\Psi[0] + \cdots + w[i]\Psi[i] + \cdots + w[n]\Psi[n][10] \tag{3.2}$$

where n is the number of challenge bits, w and Ψ are known as weight and parity vectors, respectively. If we transform the challenge from c to Ψ and feed it to the neural network we can achieve a prediction accuracy of up to 98.3% using only 2k Challenge-Response Pairs (CRPs) in 32 seconds. This high accuracy is also observed in the 4-XOR Arbiter PUF (APUF), where the accuracy can reach up to 98.83% with 100k CRPs. However, when attacking the APUF directly, the prediction accuracy drops significantly to only 60.7%, even with the utilization of 1 million CRPs as a training set. Moreover, the accuracy can be further reduced to approximately 50% for the 4-XOR APUF.

3.4 OIPUF Architecture

In the realm of delay-based Physical Unclonable Functions (PUFs), two prevalent architectures are the Arbiter PUF (APUF) and XOR APUF. In Equation 3.2, we examined the linear additive model of an APUF. This linear dependency arises due to the predetermined nature of all interconnections within an APUF, rather than randomness. As depicted in Figure 3.1, an APUF typically comprises only two delay lines.

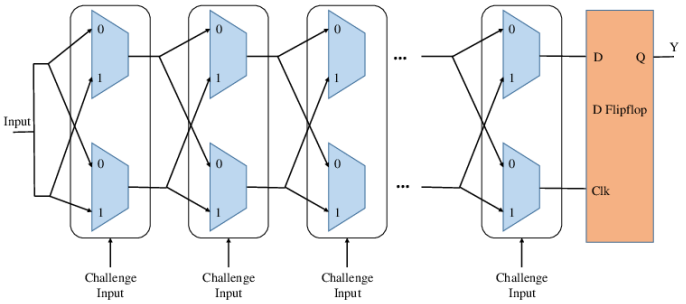

FIGURE 3.1: Architecture of Arbiter PUF.

For a k-XOR APUF it contains k independent and identical APUFs. From the knowledge of the linear additive model the response of a XOR APUF can be expressed as shown in Equation 3.3.

$$r = \theta \left(\prod_{j=0}^{k-1} \omega_j \cdot \delta_j \right) \tag{3.3}$$

Both the interconnections in the delay stage and the XOR operation introduce non-linearity, but the former can be mitigated by incorporating n+1 parity vectors. Furthermore, even in the presence of XOR operations, a 4-XOR APUF can be accurately modeled with 98.83% accuracy using just 100k Challenge-Response Pairs (CRPs).

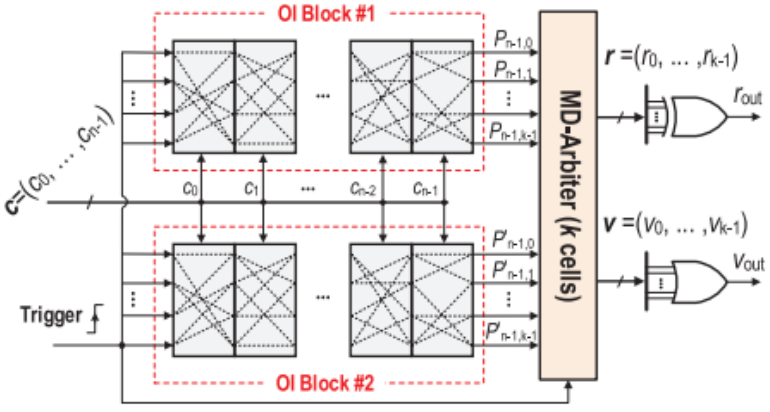

FIGURE 3.2: Proposed Architecture of (n,k)-OIPUF. [8]

Figure 3.2 presents the architecture of (n,k)-OIPUF with two identical OI blocks as entropy source. The design uses obfuscated interconnections to introduce non-linearity. Each OI block consists of n delay stages with each stage having distinct stagewise interconnections.

The obfuscated interconnections have demonstrated the ability to introduce heightened non-linearity. For instance, a (64,4)-OIPUF is purported to be resilient against all known modeling attacks, with the most effective model being the Artificial Neural Network (ANN) achieving only 56% accuracy with 100k training CRPs. This level of security surpasses that of a 4-XOR APUF with an equivalent number of PUF primitives.

Chapter 4

Methodology and Preliminary Results

4.1 Mathematical Modelling of OIPUF

We aim to formulate a mathematical equation describing the response of an OIPUF to extract pertinent features, as directly utilizing the challenge bits as a feature vector often yields suboptimal performance. The feature vector plays a crucial role in the output of the Artificial Neural Network (ANN).

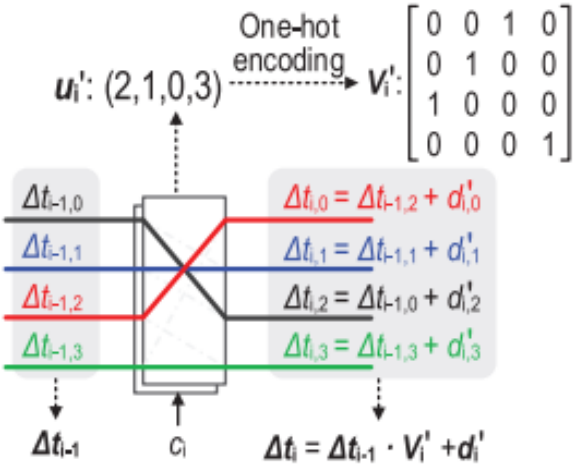

FIGURE 4.1: A single stage of OI-block. [8]

Let us consider a (n,k)-OIPUF i.e., n stages per OI-block and k multiplexers in each stage of OI-block. The time delays after ith stage are expressed as shown in Figure 4.1. $\triangle t_i$ is a $k \times 1$ vector representing the time delay after i_{th} stage, d_i is a $k \times 1$ vector representing delays introduced by i_{th} stage. Interconnections in the i_{th} stage are denoted by u_i and u'_i vectors for two OI-blocks. One hot encoding of the u_i is denoted by matrix V_i.

Let the challenge bit be C_i at i_{th} stage, then the delay after ith stage can be expressed by the Equation 4.1.

$$\triangle t_i = (1 - C_i)(\triangle t_{i-1} V_i + d_i) + C_i(\triangle t_{i-1} V'_i + d'_i) \tag{4.1}$$

$$\triangle t_i = \triangle t_{i-1} V_i + d_i - C_i(\triangle t_{i-1}(V_i - V'_i) + (d_i - d'_i)) \tag{4.2}$$

$$\text{Assume } V_i - V'_i = V_i^+ \text{ and } d_i - d'_i = d_i^+ \tag{4.3}$$

$$\triangle t_i = \triangle t_{i-1} V_i + d_i - C_i(\triangle t_{i-1} V_i^+ + d_i^+) \tag{4.4}$$

$$\text{Solving this equation recursively,} \tag{4.5}$$

$$\triangle t_1 = d_1 - C_1 d_1^+ \tag{4.6}$$

$$\triangle t_2 = (d_1 - C_1 d_1^+) V_2 + d_2 - C_2((d_1 - C_1 d_1^+) V_2^+ + d_2^+) \tag{4.7}$$

$$\triangle t_2 = \alpha_o + \alpha_1 C_1 + \alpha_2 C_2 + \alpha_3 C_1 C_2 \tag{4.8}$$

$$\text{Similarly solving for } \triangle t_3, \text{ it can be expressed as} \tag{4.9}$$

$$\triangle t_3 = (\alpha_o + \alpha_1 C_1 + \alpha_2 C_2 + \alpha_3 C_1 C_2) V_3 + d_3 - C_3(\triangle t_2 V_3^+ + d_3^+) \tag{4.10}$$

$$\triangle t_n = \alpha_o + \alpha_1 C_1 + \alpha_2 C_2 + \alpha_3 C_1 C_2 + ... + \alpha'_n C_1 C_2 ... C_n \tag{4.11}$$

$$\tag{4.12}$$

The mathematical proof presented above indicates that the response is reliant on all possible challenge bit products. Utilizing this insight, we strive to further refine our machine learning models.

4.2 Machine learning modelling attacks

Machine learning attacks were conducted using CRPs extracted from the software model of the OIPUF, leveraging insights from its mathematical representation. It was observed that the response of a (n,k)-OIPUF relies not only on the individual challenge bits $C_1, C_2, C_3, ..., C_n$ but also on their products, such as $C_1 * C_2, C_1 * C_3, ..., C_i * C_{i+1}.....C_j, ..., C_1 * C_2....C_n$. This indicates that the response is a linear combination of these challenge bits and their respective products, highlighting the complex relationship between the input challenges and the resulting PUF response.

4.2.1 Results: SVM and ANN

As previously discussed, converting challenge bits into parity bits is expected to greatly influence the model's performance. Below are experimental results conducted in python on (n,k)-OIPUF to validate this hypothesis.

Using n+1 bit parity bits as feature vector on 100k CRPs-

Table 4.1: SVM and ANN attacks on (n,k)-OIPUF with n+1 parity bits.

(n,k)	**SVM (Test Accuracy)**	**ANN (Validation Accuracy)**
(16,2)	65.76%	96.7%
(32,2)	67.14%	93.62%
(64,2)	61.99%	84.3%
(32,4)	50.79%	54.98%
(64,4)	51.32%	50.93%

The noticeable trend is a significant decline in the model's performance as k increases. This phenomenon can be attributed to the heightened complexity resulting from the increased number of XORs and obfuscated interconnections within a single stage.

Based on the mathematical model, incorporating $2^n - 1$ features, representing all possible challenge bit products of a (n, k)-OIPUF, is anticipated to enhance accuracy. However, extracting $2^n - 1$ features for n values such as 16, 32, and 64 is impractical. Therefore, an initial experiment is conducted solely to model the $(8, 3)$-OIPUF.

TABLE 4.2: Performance of modeling attacks on (8,3)-OIPUF.

Work	**Feature Set**	**SVM (Test Accuracy)**	**ANN (Validation Accuracy)**
[8]	N+1 features	42.30%	57.69%
proposed	$2^N - 1$ features	67.30%	55.77%

The SVM model demonstrates a significant performance improvement with $2^n - 1$ features, while no such enhancement is observed in the ANN. Additionally, it is noted that the ANN fails to converge, resulting in unstable validation accuracy. This instability can be attributed to the limited data availability, specifically the scarcity of sample points relative to the number of features. The same challenge arises when training the ANN for other configurations of (n,k)-OIPUF.

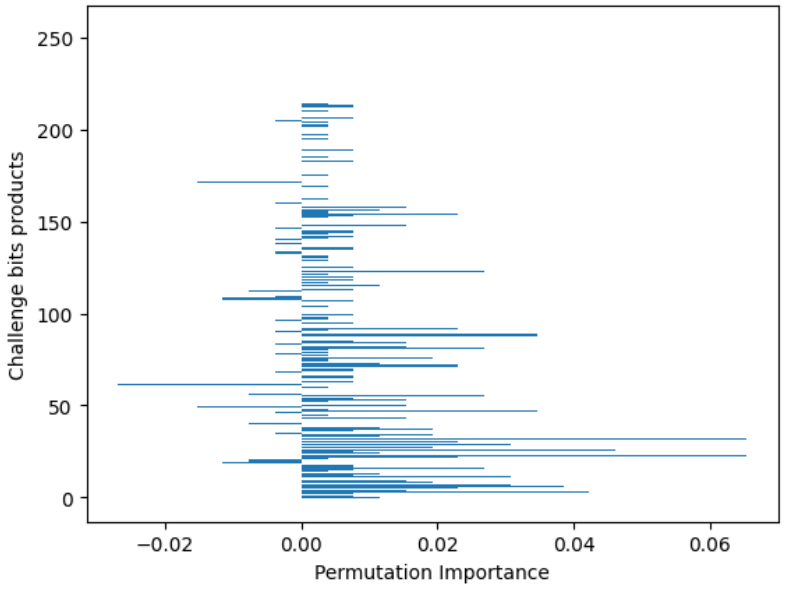

FIGURE 4.2: Feature importance of (8,3)-OIPUF.

To leverage the knowledge of $2^n - 1$ features and also to maintain the data points to the number of features ratio in ANN, we employed the feature extraction technique

only to extract the most contributing features. To do this, feature importance versus challenge bit products was plotted as shown in Figure 4.2. The feature importance was determined using the permutation importance method in SVMs.

From the above plot, it can be observed that the starting features hold more importance and contribute more to the output. These initial features include challenge bit products starting from $C_1, C_2, \ldots, C_n, C_1 \times C_2, C_1 \times C_3, \ldots$ with three products and so on. For higher values of n in a (n,k)-OIPUF, only these features can be used as the feature vector for ANN instead of employing all $2^n - 1$ features.

TABLE 4.3: Accuracy of ANN attack with most relevant features.

(n,k)	**ANN (N+1 features)**	**ANN (Relevant features)**
(16,3)	62.60%	70.75%
(32,3)	58.42%	62.51%

Chapter 5

Conclusion and Future Work

5.1 VIA PUF : Contribution of Present Work

The thesis presents a VIA PUF that capitalizes on intrinsic parasitic resistance as an entropy source, offering a superior alternative to utilizing active components prone to temperature and voltage fluctuations. Moreover, leveraging intrinsic parasitic values eliminates the necessity for additional hardware components, thereby reducing the area overhead required for the circuit. Additionally, the thesis discusses a lightweight key exchange protocol based on VIA-PUF, proven to be resilient against various attack scenarios. The VIA PUF can be used for device authentication and key generation as its high reliability and stability justify it. The reliability of 126 distinct bit-cells was assessed across three temperature conditions ($-40°$C, $25°$C, and $125°$C) and voltages (nominal, $\pm 10\%$). It achieved a 100% reliability rate with zero bit error rate (BER), obviating the need for Error Correction Code (ECC) unlike other PUFs. Additionally, the metrics used to evaluate performance, namely uniqueness and uniformity, yielded values of 50.04% and 50.01% respectively, closely approximating the nominal value of 50%.

5.2 Modeling Attacks on a Strong PUF

5.2.1 Contributions of Present Work

In the pursuit of designing PUFs resistant to modeling attacks, formidable challenges emerge due to significant advancements in machine learning. Some modeling attacks targeting specific PUF types have achieved accuracies surpassing 90%. Leveraging mathematical insights into (n,k)-OIPUF, the results showcased above aim to enhance the efficacy of classical modeling attacks. Unlike existing methods, our approach incorporates hardware-specific information, leading to notable improvements in attack resilience. These contributions highlight the potential advancements achievable in this domain.

5.2.2 Extensions of Present Work

However, there remains considerable room for improvement in this area.

- Currently, CRPs are extracted from the software model of the OIPUF, which, while feasible, is not practical. Efforts are underway to implement the OIPUF in *Xilinx Vivado* and deploy it on an *Artix 7 FPGA*. Data acquisition can then be performed directly from this FPGA, streamlining the process.
- While there has been an increase in accuracy compared to previous results, it is not significant enough to fully model the PUF. An ongoing consideration is that the obfuscated interconnections of the OIPUF, despite being random, are predetermined. An attacker could potentially obtain these interconnections through side-channel attacks. Leveraging knowledge of these interconnections, we can further extract relevant features and enhance the model.

Bibliography

[1] Pappu, Ravikanth, et al. "Physical one-way functions." Science 297.5589 (2002): 2026-2030.

[2] C. Herder, M. -D. Yu, F. Koushanfar and S. Devadas, "Physical Unclonable Functions and Applications: A Tutorial," in Proceedings of the IEEE, vol. 102, no. 8, pp. 1126-1141, Aug. 2014, doi: 10.1109/JPROC.2014.2320516.

[3] D. Jeon, J. H. Baek, Y. -D. Kim, J. Lee, D. K. Kim and B. -D. Choi, "A Physical Unclonable Function With Bit Error Rate ¡ 2.3 × 108 Based on Contact Formation Probability Without Error Correction Code," in IEEE Journal of Solid-State Circuits, vol. 55, no. 3, pp. 805-816, March 2020, doi: 10.1109/JSSC.2019.2951415.

[4] B. Park, D. Forte, M. M. Tehranipoor and N. Maghari, "A Metal-Via Resistance Based Physically Unclonable Function With Backend Incremental ADC," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 68, no. 11, pp. 4700-4709, Nov. 2021, doi: 10.1109/TCSI.2021.3105907.

[5] G. E. Suh and S. Devadas, "Physical Unclonable Functions for Device Authentication and Secret Key Generation," 2007 44th ACM/IEEE Design Automation Conference, San Diego, CA, USA, 2007, pp. 9-14.

[6] Teddy Kyung Lee, CTO, ICTK, "Via PUF Technology as a Root of Trust in IoT Supply Chain", Global Semiconductor Alliance (GSM).

[7] S. Hemavathy and V. S. K. Bhaaskaran, "Arbiter PUF—A Review of Design, Composition, and Security Aspects," in IEEE Access, vol. 11, pp. 33979-34004, 2023, doi: 10.1109/ACCESS.2023.3264016.

[8] C. Xu, L. Zhang, M. -K. Law, X. Zhao, P. -I. Mak and R. P. Martins, "Modeling-Attack-Resistant Strong PUF Exploiting Stagewise Obfuscated Interconnections With Improved Reliability," in IEEE Internet of Things Journal, vol. 10, no. 18, pp. 16300-16315, 15 Sept.15, 2023, doi: 10.1109/JIOT.2023.3267657.

[9] P. H. Nguyen, D. P. Sahoo, C. Jin, K. Mahmood, U. Rührmair, and M. van Dijk, "The Interpose PUF: Secure PUF Design against State-of-the-art Machine Learning Attacks", TCHES, vol. 2019, no. 4, pp. 243–290, Aug. 2019.

[10] Becker, Georg. (2015). The gap between promise and reality: On the insecurity of XOR arbiter PUFs. 9293. 535-555. 10.1007/978-3-662-48324-4_27.

[11] C. Xu, J. Zhang, M. -K. Law, X. Zhao, Pui-In Mak and R. P. Martins, "Transfer-Path-Based Hardware-Reuse Strong PUF Achieving Modeling Attack Resilience With200 Million Training CRPs," in IEEE Transactions on Information Forensics and Security, vol. 18, pp. 2188-2203, 2023, doi: 10.1109/TIFS.2023.3263621.