Techniques for Improved Reliability in Memristive Crossbar PUF Circuits


©2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The online home for this paper may be found at: http://web.eecs.utk.edu/~grose4/

Citation Information (BibTex):

@INPROCEEDINGS{ISVLSI:Uddin2016,
    title="Techniques for Improved Reliability in Memristive Crossbar PUF Circuits",
    booktitle="{IEEE} Computer Society Annual Symposium on {VLSI} ({ISVLSI})",
    month="July",
    year="2016",
    pages="212--217",
    address="Pittsburgh, PA"}
Techniques for Improved Reliability in Memristive Crossbar PUF Circuits

Mesbah Uddin, Md. Badruddoja Majumder, and Garrett S. Rose
Department of Electrical Engineering and Computer Science
University of Tennessee, Knoxville
Knoxville, Tennessee 37996 USA
Email: {muddin6, mmajumde, garose}@utk.edu

Karsten Beckmann, Harika Manem, Zahiruddin Alamgir, and Nathaniel C. Cady
Colleges of Nanoscale Science & Engineering
SUNY Polytechnic Institute
Albany, New York 12203 USA
Email: {kbeckmann, hmanem, zalamgir, ncady}@sunypoly.edu

Abstract—At the same time, as technology scaling progresses further into the nanometer region, emerging nanoelectronic technologies such as memristors become viable options. Several examples of nanoelectronic memristor-based PUF circuits have been proposed in the last few years. In this paper, we analyze the behavior of crossbar memristive PUF circuits under different environmental conditions such as varying temperature, supply rail voltage fluctuations and aging. We also present an approach that improves the reliability of these circuits, taking environmental variations into consideration. The advantages and challenges associated with these PUFs are also discussed in detail. Specifically, we show results for security metrics including reliability, uniqueness and uniformity. These security performance results are presented alongside estimates for power, area and delay showing the advantages of using nanoelectronic PUFs from the perspective of efficient resource utilization.

I. INTRODUCTION

Physical unclonable functions (PUFs) have emerged as solutions to a variety of security concerns, including integrated circuit (IC) piracy, counterfeiting, and secret key storage. PUFs can be used to provide challenge-response authentication. A challenge $C$ presented to the PUF leads to a unique response $R$ as the output. PUFs exploit uncontrollable physical differences between chips such as gate delays, leakage, start-up characteristics, threshold voltages and internal resistance. These differences depend on intrinsic process variations and can be thought of as “fingerprints” for a particular chip that defines a function $f_{chip}(C) = R$.

One of the most well known PUFs is the arbiter PUF (APUF), which utilizes variations in the propagation delay of a circuit [1]. In an APUF, two-input/two-output switches are chained together to create a circuit with two separate paths for the signals to propagate through. Process variations dictate that a single input signal will propagate through one path faster than the other. An arbiter (e.g., flip-flop) is then used to translate the difference in path delays into unique signature information.

Another example is the ring-oscillator PUF (RO-PUF) [2]. A CMOS ring oscillator is a long chain of inverters used to create a particular frequency. Due to inherent process variations, two identical ring-oscillators will exhibit different frequencies. This is because there will be slightly different delays through the inverters used in each ring-oscillator [3]. A ring-oscillator PUF then compares the frequencies and converts their difference into a unique signature.

As technology scales further into the nanometer regime, the need has arisen for PUFs implemented at the nano-scale. Memristor based PUFs have been proposed as nano-scale security primitives. A 1-bit PUF based on the write-time of a memristor was first proposed by Rose et al. [4]. An experimental realization of this design was later presented by Mazady et al. [5]. Kavehei et al. presented a combination of RRAMs (memristors) and RO-PUFs named mrPUF [6]. Other works that focus on memristive PUF and memristor characteristics are presented in [7], [8], [9] and [10]. In a previous work, Rose et al. presented a PUF constructed from a dense crossbar array of memristors, the memristive crossbar PUF (XbarPUF) [11]. Because the XbarPUF depends on the relative write-time of pairs of memristive circuits, it does not require exact knowledge or control of the minimum write-time of the device. In this work, we use an XORing technique on pairs of responses from the basic crossbar architecture to improve the reliability of the circuit. We also use an improved memristor model derived from statistical data obtained experimentally from HfO$_2$ memristors. In this paper, we analyze the performance of the XORed XbarPUF circuit, and measure the reliability using the HfO$_2$ memristor model.

We first present some background on memristors, including the HfO$_2$ memristor considered for this work, and memristive PUFs in section II. We then present and describe the memristive crossbar-based PUF design and model in section III and evaluate its performance using Monte Carlo simulations in section IV. Some analyses and discussion are provided in section V on the reliability and other aspects of these memristive PUFs. Finally, future prospects and concluding remarks are provided in sections VI and VII, respectively.

II. BACKGROUND

A. Device Fabrication

The memristive devices considered for this work were designed and fabricated in-house at the SUNY Polytechnic Institute.
Institute on a 300mm wafer platform and integrated with the IBM 65nm 10LPe process technology. It should be noted that in-house facilities allow for an area-efficient and seamless flow of front-end CMOS and back-end memristive and metalization processes. A custom, cost-effective build embeds the memristor devices between metal 1 (M1) and metal 2 (M2) layers. This is achieved by replacing the standard copper with a tungsten metalization layer and introducing an additional via (W-V1) below the memristor device. The composition of the M1 and V1 layers was altered to allow for the use of front-end-of-the-line (FEOL) tools to deposit the HfO$_2$ layer with a highly precise atomic layer deposition (ALD) technique. This film serves as the active switching layer (metal-oxide layer) for the memristor device considered here. An example experimentally observed I-V characteristic of 250 switching cycles of our HfO$_2$ memristor is shown in Fig. 1.

**B. Memristor Device Characteristics and Modeling**

The model used for the simulation work in this paper has been adapted from a model first developed and presented by McDonald et al. in [12]. While the original model could be used for unipolar, nonpolar and bipolar behavior, this work is mainly focused on bipolar behavior. Model behavior based on the experimental device is shown in Fig. 2 where the positive and negative thresholds are 0.7 V and -1.0 V, respectively. Our model increments or decrements the memristance at each time step of the simulation depending on the voltage applied across the memristor. If the memristance is already either HRS or LRS, the memristor will not switch and will remain at that resistance level. During a SET (from HRS to LRS) operation the memristance is updated by:

$$M(t_{i+1}) = M(t_i) - \frac{\Delta r \Delta t |V(t_{i+1})|}{t_{swp} V_{tp}},$$  \hspace{1cm} (1)

while for the RESET (LRS to HRS) operation the memristance is updated by:

$$M(t_{i+1}) = M(t_i) + \frac{\Delta r \Delta t |V(t_{i+1})|}{t_{swp} V_{tn}},$$  \hspace{1cm} (2)

where $M$ is the memristance, $\Delta r$ is the absolute difference between the LRS and HRS values, $\Delta t$ is the simulation time step size, $V(t)$ is the applied voltage bias, $t_{swp}$ ($t_{swn}$) is the time to effect a resistance change under a bias greater than either of the thresholds, and $V_{tp}$ ($V_{tn}$) is the threshold voltage under positive (negative) polarity.

Since memristors are nanoscale devices, they show relatively large amounts of process variation contributing to inter-die variations. This translates to a larger variance in the device parameters, for example as examined by Pouyan et al. [13]. The HRS and LRS are the parameters most affected by process variation, with the variance of the HRS being somewhat greater. To account for the worst possible variance, we chose 20% and 10% for $\sigma_{HRS}$ and $\sigma_{LRS}$, respectively, in the memristor model used for simulating the XbarPUF. It is to be noted that these assumptions only consider the mean of the fabricated memristors and the added variance follows the inter-die variation. We use 10% cycle-to-cycle variance for the HRS and switching voltages and 5% variance for the remaining four parameters: $t_{swp}$, $t_{swn}$, $V_{tp}$ and $V_{tn}$. This variance is based on measurements of the experimental memristor device.

On every run of the simulation, values for all parameters are evaluated using their respective mean and variance. Fig. 2 shows the simulated I-V characteristic curve for our model. We use $HRS = 15K$, $LRS = 4K$, $t_{swp} = 10ns$, $t_{swn} = 1u$, $V_{tp} = 750mV$ and $V_{tn} = -1V$ for this simulation, all values derived from the fabricated HfO$_2$ memristor.

Another important consideration related to PUF performance is the aging of the memristors. In the worst case aging scenario, the distributions of the HRS and LRS will trend towards each other, i.e. the HRS decreases with time and the LRS increases with time [13]. We incorporate this worst case aging factor into our model using a linear aging rate for both
HRS and LRS, and account for the fact that HRS aging is faster than LRS aging. This is illustrated in Fig. 3.

To evaluate the reliability of any device, readings in different environmental conditions are needed. Since temperature has a particularly significant affect on memristor device behavior, we incorporate temperature dependence in our model. The problem is, however, temperature data for memristors are not widely available from the literature. Nevertheless, we have learned that with increasing temperature, the mean of HRS is decreased but the distribution is increased. The mean value of LRS is increased with increasing temperature and the threshold voltage values drop with increasing temperature. Taking all these factors into account, along with some partial data generated in the lab, educated guesses were made for the temperature coefficients for each of these parameters and assumed a linear dependence in the operating range:

\[ R_{new} = R_{old} \left[ 1 + \left( \text{Coeff}_{\text{temp}} \times \Delta \text{temperature} \right) \right]. \]  
(3)

C. Related Memristor based PUF Works

There are two fundamental requirements for building a PUF: random and uncontrollable variation. The variations must be random, thereby drastically reducing the probability that a unique signature will be repeated. These random variations must also be uncontrollable, otherwise an adversary could clone the device (i.e. mount an impersonation attack). Memristors inherently have both fundamental requirements. More specifically, process variability translates to variations in the minimum write-time of a memristor when using the device as a memory cell [4]. For example, a memristor being SET from HRS to LRS will only exhibit a logic ‘1’ output if the SET time (i.e. write time to SET the memristor) is greater than some minimum \( t_{\text{wr, min}} \). If, however, the SET time is chosen to be at or near the nominal \( t_{\text{wr, min}} \), then process variations will dictate that the output is nearly as likely to be a logic ‘0’ as it is a logic ‘1’. This probabilistic status for the output voltage is undesirable for conventional memory systems but can be leveraged in the implementation of PUF circuits.

The write-time based memristive PUF proposed by Rose et al. [4] is an early work that leverages the process variation of minimum SET time for a memristor to generate unique responses of a PUF over different implementations. In this PUF circuit, at first a RESET pulse is applied to the memristors for a long enough time to ensure that every memristor is initialized to the HRS. After the RESET operation, memristors are driven with a SET pulses where the pulse width is kept around the minimum time to set the memristors. Finally, memristors are read by applying a read pulse and each value is compared with another to produce the corresponding response. Since the minimum set time of a memristor is dependent on the process, this PUF gives unpredictable and unique responses for the same challenge on different processes. Rose et al. presented a more complex PUF design based on memristive crossbar array where the challenge inputs were directly incorporated into the PUF’s main source of entropy i.e write-time. Such direct dependency of the memristor’s write operation on the challenge input ensures much more uncontrollable randomness in its response bits when fabricated on different chips. This design approach improves the uniqueness of the PUF responses nearly to the ideal value.

In practice, many memristive materials, including the HfO\(_2\) devices considered here, require a forming step to initialize the devices. An elevated voltage is applied across the device to cause the first SET, after which the device can cycle between the HRS and LRS at significantly lower voltages. Prior to this step, the device operates as a regular resistor. The difference in behavior is easy to detect and thus is a prime candidate for initial tamper detection. In our designs, the memristors are only formed during device provisioning where the PUF challenge response pairs must be stored in a secure environment.

D. PUF Performance Metrics

In order to measure the performance of the XbarPUF, it is necessary to compare it to other PUF implementations using the same performance metrics. Maiti et al. described different metrics in [14] to measure the performance of a PUF. We have chosen three from [14] to evaluate the performance of the XbarPUF: uniqueness, uniformity and reliability. These metrics will measure the PUF’s performance across multiple device dimensions: inter-die space, intra-die space, and time. Uniqueness measures a PUF’s ability to produce an ID in terms of its challenge response pairs which is distinct for every chip implementation. To distinguish every chip from one another with a PUF circuit, its uniqueness should be close to 50%. It is also necessary that the IDs produced by the PUF in a chip are not too close to one another. Uniformity, another performance metric for a PUF, measures the ability to produce distinct responses across a challenge space. If just one bit of a challenge set is flipped, almost half of the response bits should also flip. Technically, uniformity estimates the ratio of 0’s and 1’s across the entire response space of a PUF. Good uniformity means a PUF will be resistant to an attack in which an attacker already has the responses of a particular chip. It
should be ideally 50%. Lastly, reliability (eqn. 4) is used to determine a PUF’s ability to perform consistently over time. If a PUF is unable to produce the same response to the same set of challenge bits consistently over time, then the PUF would not be very useful. Since in previous work, it has been shown that PUF’s uniqueness and uniformity are very close to ideal, our main focus in this paper is to evaluate the reliability of the our XORRed XbarPUF while also maintaining the fact that the uniqueness and uniformity remain excellent.

\[
\text{Reli}(\%) = 100 \times \frac{2}{N_{\text{cycles}}} \sum_{i=1}^{N_{\text{cycles}}-1} \sum_{t=1}^{N_{\text{cycles}}} r_t \oplus r_{it};
\]

(4)

for each response bit and for each chip.

While the aforementioned metrics measure a PUF’s ability to function as a security primitive, they fail to measure the PUF’s physical performance. In order to operate as a nano-security primitive, a device needs to be not only secure, but also lightweight and efficient. Performing an area analysis will determine whether an implementation is realizable on the nanoscale, and measure its ability to grow given more substantial cryptographic needs. Power analysis must be taken into account, as a device that does not operate efficiently on power would not be able to be practically utilized.

III. MEMRISTIVE CROSSBAR PUF DESIGN WITH XORING TECHNIQUE

The crossbar PUF considered in this work is illustrated in Fig. 4. There are 8 rows and 8 columns in this PUF circuit. Each two adjacent rows of memristors are driven with the inverted and non-inverted form of a single challenge input, respectively. For a high logic value of challenge input, memristors driven with the true challenge bit are set to the low resistance state and those driven with the inverted bit remain in the high resistance state. The inverse scenario occurs when a low challenge input is applied. This arrangement ensures half of the memristors in the XbarPUF are held at the HRS while the other half are driven toward the LRS for any given challenge input. One memristor from every row is connected with a single column. Identical load resistance is connected at the end of each column to measure the voltage resulting from different resistive states of the memristors of that particular column. The voltages across the loads of two adjacent columns are fed into an arbiter circuit that determines which column reaches first an effective resistance less than the load. In this way, the resistance of two columns are incrementally reduced such that the resistance values race one another until one “wins”, as determined by the arbiter. Finally, the outputs of pairs of arbiters are XORed to give the response bits. This modified crossbar PUF generates two response bits from 4 challenge bits and in terms of challenge-response combination we call it a 4 × 2 XbarPUF.

The logic voltage level found from the output of a memristive circuit is dependent on the write pulse duration which is a process dependent parameter. The challenge bits are applied directly as the write pulse for all memristors in this crossbar arrangement. Further, the output of every arbiter is dependent on all possible delays created by \(2^N\) \((N = 4\) in the example studied here) different combinations of memristive states from the columns. On top of that, every two outputs of the arbiter circuit are combined with an XOR operation. These features altogether make the response of this PUF a more complex function of its challenge inputs and internal process variations relative to previously proposed implementations such as [11].

IV. SIMULATION RESULTS

The performance of the XbarPUF based on simulated memristor parameters, was shown in a previous work [11]. Here we present our results using the updated model based on new memristor parameters obtained from experimental data. An XbarPUF circuit with 4 challenge bits and 2 response bits was modeled and simulated. Design parameters used for the memristors in our circuit are the high and low resistance states (HRS and LRS), the positive and negative threshold switching times \((t_{\text{sup}}\) and \(t_{\text{snr}}\), and the positive and negative threshold voltages \((V_{\text{tp}}\) and \(V_{\text{rn}}\)). These parameters were modeled such that they varied in time as observed in experimental data.

Circuit-level design parameters considered in the model of the XbarPUF circuit are the read and write voltages. These voltages are determined based on the given threshold voltage of the device. It is desirable that the write voltage be as low as possible in order to reduce the power consumption, but not so low that the memristors fail to get enough switching voltage across their terminals. To balance these two opposing factors, the read and write voltage for our simulations are chosen to be 0.65V and 1.3V, respectively. To assess the uniqueness, uniformity, and power of the XbarPUF circuit, 100 Monte Carlo simulation runs were performed using Cadence Spectre over several unique challenges. To measure reliability, four unique random challenges were applied across ten different chips for more than 100 cycles per chip.

Tables I and II show the results of these simulations. All device-to-device and cycle-to-cycle variations are listed in section II-B. The results show strong uniqueness and uniformity with values close to the ideal. The reliability with the XORing technique was found to be much better than the previous XbarPUF circuit without XORing [11]. We also present the reliability result with time varying temperature within a chip. The temperature is varied in small steps from 17°C to 67°C, both in ascending and descending order. The XORed XbarPUF shows a very strong reliability even under variable temperature throughout the life cycle of a chip.

V. ANALYSES AND DISCUSSION

The uniqueness of this XbarPUF circuit using the new memristor model is similar to the performance seen by Rose et al. using a model based on theoretical values [11]. However, the memristor model parameters used in our work were derived from experimental data for HfO\(_2\) memristors. Using these values, we regenerate the uniqueness and uniformity for the XbarPUF of [11]. No changes in uniqueness and uniformity are found after XORing. These values are very close to ideal
as can be seen in Table I. One notable difference between the two circuits is the increase in power consumption. This is a result of the increased write voltage, reduced LRS and additional circuitry for implementing the XORing. Changing the temperature brings about a change in the distribution of HRS and LRS and therefore a small change in power consumption is noticeable, as can be seen in Table II. In our model, increasing the temperature causes a decrease in HRS and an increase in LRS. But the change of HRS is steeper than LRS which decrease the overall effective resistance of a XbarPUF column and results in a higher current and therefore, higher power consumption.

Reliability for a memristive XbarPUF is presented for the first time in this paper and is shown in Table II. We generate the reliability for the xbarPUF both with XORing and without XORing. It is evident that the reliability of this XbarPUF with XORing is excellent and better compared to the reliability of other PUF implementations including APUF, SRAM-PUF or RO-PUF [15]–[17]. Herder et al. suggested in [18] that the reliability would be improved for APUF by XORing the output bits. Now it has been implemented and proven to be true for the XbarPUF as well. We have also considered the temperature effect in our circuit, showing reliability decreases by a small amount with increasing temperature as expected. Thus the reliability of the XORed XbarPUF is actually very close to its desired value.

Memristors are essentially nanowires, they can be placed in the cross-points of interconnects. Therefore, the crossbar representation actually requires a very little or effectively zero area. Since the XORed circuit generates one response bit from previous two, its area is double in size compared to the XbarPUF with no XORing. The delay of this circuit is the minimum clock cycle time that can be used. This minimum clock cycle time is constraint by the longest time needed for a memristor to switch to either high or low resistance state. In our model the slower switching time is $1\mu s$, so the delay of our XbarPUF circuit would be at least $1\mu s$ plus the time needed to apply a challenge and then get a valid response which is found to be 25.2ns from simulation.
VI. Future Work

We stated in our discussion of the memristor model that, currently, the data for HfO2 memristor behaviors with changing temperature are hardly available. If we are able to gather enough data, we can build a more realistic memristor model about varying environmental conditions and generate more realistic reliability results. The main concern of this circuit for now is actually the power consumption. Since we are now generating one response bit with the same crossbar that would generate two response bits previously [11], the power consumption is also increased. Including advanced power management technique like clock gating or voltage scaling would help in decreasing the power. We also haven’t considered the robustness of our XbarPUF against machine learning or other modeling attacks. Though XORing increases the robustness by a magnitude of factors, incorporating additional randomness like mixing of columns and XORing the inputs would increase its robustness ever more. These are areas we are working on now and hope to achieve good results in a short time.

VII. Conclusion

The uniqueness, uniformity and reliability of the XbarPUF are better than or similar to other existing PUFs [15]–[17]. One very important advantage of memristive PUFs is that they require a very small amount of area. Memristors are essentially crosspoints between nanowires. Thus, they can be placed between two metal layers which leads to a very dense implementation and greater complexity can be achieved in minimal space. However, the power consumption of this circuit is not as low as we hoped for. That said, the LRS can be increased, which would result in a considerable decrease in power consumption. Memristors are also non-volatile, meaning power gating techniques can be implemented in memristive PUFs, leading to further reductions in power.

The reliability of the HfO2 memristive PUFs is excellent, though not exactly ideal. The scalability of the crossbar design is limited as the parallel resistance decreases fast when more and more rows are added. However, these shortcomings are largely due to the fact that memristors are an emerging technology. The data for our experiment was generated for newly fabricated memristors. As the device technology matures, more will be learned about memristor behavior and manufacturers will be able to control their behavior better. Thus, it is possible to improve the reliability even further by incorporating newer, more stable memristors and also employing some error-correction schemes.

Acknowledgment

The authors would like to thank Lok Kwong-Yan, Bryant Wysocki, Nathan McDonald and Jillian Hallak of the Air Force Research Laboratory for interesting discussions on this topic.

References


[17] D. Sahoo, D. Mukhopadhyay, and R. Chakraborty, “Design of low area-overhead ring oscillator PUF with large challenge space,” in Int. Conf. on Reconfigurable Computing and FPGAs (ReConFig), December 2013, pp. 1–6.