A Robust Authentication Methodology Using Physically Unclonable Functions in DRAM Arrays

Maryam S. Hashemian\textsuperscript{1}, Bhanu Singh\textsuperscript{1}, Francis Wolff\textsuperscript{1}, Daniel Weyer\textsuperscript{1}, Steve Clay\textsuperscript{2}, and Christos Papachristou\textsuperscript{1}

\textsuperscript{1}Dept. of EECS, Case Western Reserve University, Cleveland, OH 44106, USA
\textsuperscript{2}C.W. Consultants, Medina, OH 44212, USA

Abstract—The high availability of DRAM in either embedded or stand-alone form makes it a target for counterfeit attacks. In this paper, we propose a robust authentication methodology against counterfeiting. The authentication is performed by exploiting the intrinsic process variation in write reliability of DRAM cells. Extensive Monte Carlo simulations performed in HSPICE show that the proposed authentication methodology provides high uniqueness of 50.01\% average inter-die Hamming distance and good robustness under temporal fluctuations in supply voltage, temperature, and ageing effect over a 10-year lifetime.

I. INTRODUCTION

Over the years, dynamic random access memory (DRAM) has been used in a wide range of high-volume electronics applications. DRAM technology has a very low cost per bit, and it offers the highest density RAM, due to the simple ITIC cell structure, comprising an access transistor and a storage capacitor. DRAM is found either in stand-alone form or embedded in a system-on-a-chip (SoC). Unlike SRAM, stand-alone DRAM has a huge market in high volume applications, including servers, workstations, PCs, and game consoles.

In the past few years, the use of embedded DRAM (eDRAM) has become widespread as an alternative to embedded SRAM (eSRAM) to satisfy the high-performance and density needs in memory [1–3]. eDRAM is a DRAM embedded in the same multichip package used for mobile integrated processor graphics. Embedding DRAM onto the package adds cost because eDRAM requires additional fabrication steps compared with eSRAM. However, when it comes to bandwidth-intensive applications, eDRAM has proven to be beneficial for SoC as major area savings offset the process cost.

Over the past few years, stand-alone DRAMs have been a target for counterfeiting. Complex SoCs with multiple IP blocks are not exempt from counterfeiting either. Verification and authentication play an important role in counterfeit detection. An innovative authentication solution is the use of Physically unclonable function (PUF) [4]. PUF is a physical security primitive based on a unique challenge-response mechanism exploiting the random physical variations in the manufacturing process. PUF is easy to evaluate, but it is difficult to clone or reproduce in another chip. Random physical variations ensure that the challenge with which the PUF is queried is uniquely mapped to a response. To authenticate a device, the vendor stores for each PUF a collection of challenge-response pairs (CRPs). Later in the field, a challenge is sent to the device, and the corresponding response signature is verified against the database [5].

SRAM-based PUF has been widely investigated as an authentication mechanism [6, 7]. However, little research has been conducted to utilize DRAM for authentication purposes [8]. With the current market for stand-alone DRAM and the rapid growth of eDRAM in the future technologies with the ever-increasing demand for memory bandwidth, it is critical to investigate DRAM as an authentication and anti-counterfeiting measure.

In this work, we develop an authentication methodology based on a DRAM-PUF with the objective to provide resilience against counterfeit attacks. The advantage of the proposed methodology is based on the fact that the stand-alone DRAM or the eDRAM already present in an SoC can be used for generating device specific signatures without requiring any additional hardware. Signature generation in DRAM can be accomplished by inducing write failures. The random manufacturing process variation of the cells’ parameters across the chip cause some cells to be more vulnerable to failures. The cells’ vulnerability is then translated into a unique device-specific signature and used for authentication purposes.

In what follows, we briefly describe our contribution of this work:

• We propose a robust methodology to securely authenticate DRAM using a DRAM-based PUF. Our authentication is faster by exploiting the write duty cycle than the method in [8], while providing a much better Hamming distance. Furthermore, it provides random signatures with no dependence on initial values (seed).
• We provide extensive simulations to evaluate the effectiveness of the methodology. The results show that the generated signatures have high entropy, excellent uniqueness and very good reproducibility under temperature variation. Unlike [8], we also consider the impact of supply voltage variation and ageing effect on the reproducibility of signatures and show that these are very robust under both factors.

The remainder of the paper is organized as follows. Section 2 describes the related work. Section 3 presents the DRAM architecture followed by a description on the proposed DRAM authentication methodology. Simulation analysis and results on uniqueness and robustness are presented in Section 4. Finally, Section 5 concludes the paper.
II. RELATED WORK

In the past few years, silicon-based PUFs have become an important authentication and anti-counterfeiting measure. The concept of silicon PUF was first introduced by Devadas in 2002 [9]. Since then, several silicon PUFs have emerged. There are two main classes of silicon PUFs: (a) the PUFs exploiting delay variations and (b) the PUFs exploiting the state of memory cells [10].

Ring oscillator (RO)-based PUF [11] and Arbiter PUF [9] are the most well-studied delay-based PUFs which use the random delay variations to produce a digital signature. These PUFs are not very suitable to authenticate integrated circuits as they are power inefficient and rely on dedicated circuitry solely added for the purpose of unique signature generation and hence present large area overhead which might not be feasible for all designs.

Memory-based PUFs utilize the memory arrays already present on the hardware for unique signature generation. One example for this type of PUF is SRAM-based PUF which relies on the random start-up values of SRAM cells upon power-up [12, 13]. This PUF is currently being used for authentication purposes [6]. Another example is Memory failure-based PUF [14, 15] which exploits the random memory failures in SRAM cells to generate unique signatures.

DRAM-based PUF has been proposed in [8] by disabling the refresh signal in order to generate unique identifiers and random numbers. Due to the intrinsic process variation in DRAM cells, the dis(charge) rate of the storage capacitor is different from cell to cell, resulting in random data loss which can be extracted as a unique digital identifier. The most critical step in this approach is reading the values in a controlled time frame optimized to ensure random data loss across the memory cells.

One advantage of DRAM PUF over some power-up-based SRAM PUFs is that it does not require memory power-up values to generate unique identifiers. There is also no need to have a dedicated memory for the sole purpose of identifier generation. The identifiers can be simply generated from the regular DRAM used in the device, while it is fully functional for normal storage operations.

Despite all the advantages, there are several drawbacks to the DRAM PUF discussed above. First, the authentication method suffers from very slow speed. Based on the results in [8], the data loss is so slow that it takes 4096 seconds, i.e. more than one hour, to have only 6% of bits in a 512MByte DDR3 flipped. Secondly, the generated signature is not truly random as it is dependent on initial values (seed). Moreover, the work does not provide a uniqueness analysis on the extracted signatures from chip to chip as well as robustness analysis under supply voltage variation and ageing effect.

Understanding the impact of both technology and non-technology parameters on DRAM cells enables the design of a robust and reliable PUF to authenticate DRAM. In what follows, we discuss our development of a secure and robust DRAM authentication methodology.
driven voltage to reach the output. At the same time, the access
transistor remains open to restore the data in the DRAM cell.
There is a differential bidirectional data-bus ($D/D^*$) which is
shared by a column of sense amplifiers and is used during a
write operation. The write enable signal ($WSEN$) connects the
data-bus to the bit line pairs. In case of a write operation,
the data-bus provides a larger current to overdrive the sense
amplifier and the bit line voltage. The DRAM cell would then
be overwritten by the data values asserted to the data-bus.

B. Inserting PUF into DRAM

Large dense structures like DRAMs are particularly suscep-
tible to process variation. The existing inter-die and intra-die
variations in the device parameters, cause mismatch in the
cells’ strength. As a result, DRAM cells are not all equally
strong. Due to random nature of process variation, some cells
experience more variation and fail to operate. These failures
can be exploited to generate unique signatures which can be
used for DRAM authentication.

The failure mechanisms observed in a DRAM cell are: (a)
data retention failure, when a cell loses its value after some
amount of time; less than the amount required by DRAM
refresh, (b) access time failure caused by signal transitions
in the address decoders that are too fast or too slow, and (c)
transition failures where a cell can be written from 0 to 1, but
not from 1 to 0, or the other way around.

The idea of exploiting failures has been earlier investigated
to generate unique signatures in SRAM [14, 15]. In this paper,
the access time failure-based approach has been exploited to
generate unique signatures in DRAM. The reliability of cells
has been evaluated by inducing a write failure through precise
reduction of write duty-cycle. A delay generator, consisting
of a chain of inverters, a multiplexer, and an AND gate is
used to shorten the duty-cycle [14]. To write a value into
DRAM cell, first $WL$ is selected. Next, the sense amplifier
is temporarily forced to the desired value, causing the $BL$ to
charge or discharge the storage capacitor accordingly.

In normal scenario, the write duty-cycle is selected, such
that all DRAM cells can successfully perform the write opera-
tion under all process corners. In the proposed PUF, however,
the write duty-cycle is shortened intentionally, such that some
cells become unstable under process variation and fail the write
operation. Fig. 3 compares the read-out signature of a ran-
donally selected group of DRAM cells after a write operation
performed at 3 different write duty-cycles; (a) $Write_a$, (b) $Write_b$ (the optimized write duty-cycle ensuring
equal probability of write success and failure), and (c) $Write_c$, $<Write_b$

Fig. 3. Read-out signature of a randomly selected group of DRAM cells
after a write operation performed at 3 different write duty-cycles; (a) $Write_a$, (b) $Write_b$ (the optimized write duty-cycle ensuring
equal probability of write success and failure), and (c) $Write_c$, $<Write_b$

Finally, we can use the proposed PUF to generate a unique
CRP by following these steps: (1) choosing the address of
$n$ randomly selected cells as part of the input challenge, (2)
performing a $Write_1$ operation at the reduced write duty-cycle,
and (3) extracting the cell values as part of the $n$-bit output
response.

IV. SIMULATION ANALYSIS AND RESULTS

The proposed PUF has been implemented with a $16 \times 8$
embedded DRAM array architecture designed for the 45nm
low power Predictive Technology Model (PTM) [17]. Extensi-
tive Monte Carlo simulations were carried out using Synopsys
HSPICE for 1000 chips to generate 128-bit signatures. To
introduce process variation, we have considered 10% inter-
and 8% intra-die variation on threshold voltage ($V_{th}$).

A. Uniqueness Analysis

For Authentication purposes, PUF signatures should be as unique as possible. To quantify the uniqueness of signatures, we compute the inter-die Hamming distance ($inter-die HD$) by calculating the fractional Hamming distance between the signatures obtained from different chips, considering all pairwise comparisons. We assume $R_i$ and $R_j$ are the 128-bit signatures obtained from chip $i$ and chip $j$ respectively (where by chip we mean a packaged die). The $inter-die HD_{avg}$ \[18\] for $m$ chips is then defined as:

$$inter-die HD_{avg} = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} inter-die HD(R_i, R_j) \times 100\%$$

For high uniqueness, it is desirable to have an $inter-die HD_{avg}$ close to 50% which means that half of the cells prefer a different state. It also means that there is a low correlation between signatures which makes predicting the PUF behavior more difficult, given a collection of PUF signatures.

Fig. 5 shows the distribution of $inter-die HD$s of 1000 signatures generated by performing a $Write_1$ operation at the selected write duty-cycle. The horizontal axis represents the percentage of bits differing between two signatures, and the vertical axis represents the number of pairwise comparisons corresponding to an $HD$. As Fig. 5 shows, the $inter-die HD_{avg}$ is found to be 50.01% which is very close to the ideal 50%, showing that the proposed PUF provides very unique signatures.

It is also desirable for a PUF signature to have a fractional Hamming weight of 50% since this indicates that the PUF is unbiased, and the cells have no preference for a certain state. Fig. 6 shows the fractional Hamming weight (tendency toward 1) for each cell in a $Write_1$ operation performed at 3 different write duty-cycles. As it is shown, the middle plot performed at the selected write duty-cycle, $Write_a$, has a 50.15% average fractional Hamming weight, indicating that the PUF is perfectly unbiased. However, the top plot performed at $Write_{e}$ has a higher average fractional Hamming weight, due to a longer write duty-cycle which results in fewer write failures. Similarly, the bottom plot performed at $Write_{c}$ shows a lower average fractional Hamming weight, due to a shorter write duty-cycle which results in more write failures.

To measure the randomness of signatures, we estimate the min-entropy of DRAM PUF. Min-entropy is a conservative entropy estimation to measure the level of uncertainty associated with PUF signatures. The min-entropy should be large enough to guarantee resistance against attacks. Consider $m$ chips, $m = 1, 2, ..., 1000$, with each chip having $k$ cells, $k = 1, 2, ..., 128$. To estimate the min-entropy, we determine the fractional Hamming weight of cell $k$ denoted $HW(k)$ over all chip signatures. $HW(k)$ provides an estimate of the probability of cell $k$ to be 1. Let $p_{max}$ denote the most likely outcome of cell $k$ as follows:

$$p_{max}(k) = \max\{HW(k) ; 1 - HW(k)\}$$

The min-entropy is then defined as \[19\]:

$$H_{min} = \frac{1}{128} \sum_{k=1}^{128} - \log p_{max}(k)$$

For our methodology, the min-entropy per cell is found to be 0.97 which is very close to the ideal min-entropy of 1.

B. Robustness Analysis

Robustness measures the reproducibility of signatures under varying operating conditions, such as temperature, supply voltage, and ageing. For a PUF to be robust, the signature obtained at a non-ideal operating condition should not differ significantly from the one obtained under normal condition. To quantify the robustness of signatures, we compute the intra-die Hamming distance ($intra-die HD$) between the signatures from the same chip subject to varying operating conditions. This measure provides an indication of resilience against temporal fluctuations. To compute $intra-die HD$, a 128-bit signature, $R'$, is extracted from each chip at a non-ideal operating condition. The signature is then compared to the signature, $R$, obtained from the same chip under normal condition. Assuming $N$ samples of $R'$ are taken for each of the operating conditions, the $intra-die HD_{avg}$ \[18\] for chip $i$ is estimated as the $intra-die HD$ between $R_i$ and $R'_i$ over $N$
The total \( \text{intra-die } HD_{avg} \) is then computed by averaging all \( \text{intra-die } HD_{avg} \) values over \( m \) chips. For high robustness, it is desirable to have a total \( \text{intra-die } HD_{avg} \) close to zero.

1) Robustness Under Temperature Variation: Operational temperature has an effect on on-chip junction temperatures which in turn affects the delay of gates and wires, hence possibly causing a change in PUF response behavior. To estimate the robustness of signatures under temperature variation, we estimated the \( \text{intra-die } HD_{avg} \) among the signatures by obtaining the signature for each chip at six different temperatures under nominal supply voltage and comparing it with that obtained at room temperature. We varied the temperature from \( 20^\circ C \) to \( 50^\circ C \) in steps of \( 5^\circ C \). As Fig. 7 shows, the highest average signature variation is \( \approx 18.9 \) bits which occurs at \( 50^\circ C \) where we have the highest \( V_{th} \) shift. As temperature increases, \( V_{th} \) also increases. As a result, the selected write duty-cycle will not be long enough to fully charge the storage capacitor, resulting in a \text{Write}_1 failure (bit-flip from 1 to 0).

Fig. 8 shows the distribution of \( \text{intra-die } HD_{avg} \) of the signatures obtained from 1000 chips under temperature variation. As it is shown, most of the chips (\( \approx 94\% \)) have fewer than 18 bit-flips in their signatures, including 200 chips with not a single bit-flip, showing that the proposed PUF is stable even at varying temperatures.

2) Robustness Under Voltage Variation: As part of robustness analysis, we have also considered the impact of supply voltage on the behavior of PUF. Decreasing the supply voltage slows down the circuit and affects the reliability as different parts of the circuit suffer from a non-linear performance loss. To measure the robustness of signatures under voltage variation, we estimated the \( \text{intra-die } HD_{avg} \) among the signatures by obtaining the signature for each chip at three different voltage levels under room temperature and comparing it with that obtained at the nominal supply voltage (1.1v). We varied the supply voltage level from 80\% to 110\% of the nominal supply voltage in steps of 100mv. As Fig. 9 shows, the highest average signature variation is \( \approx 14.9 \) bits which occurs at 0.9v where the circuit performance is affected the most.

Fig. 10 shows the distribution of \( \text{intra-die } HD_{avg} \) of the signatures obtained from 1000 chips under supply voltage variation. As it is shown, most of the chips (\( \approx 90\% \)) have fewer than 25 bit-flips in their signatures, including 230 chips with not a single bit-flip, showing that the proposed PUF is stable even at varying voltage levels.

3) Robustness Under Ageing Effect: Ageing effect is an emerging reliability issue which affects the performance of a device over prolonged use. One of the most dominant ageing effects is negative-bias temperature instability (NBTI) which results in \( V_{th} \) degradation of PMOS, due to generation of trapped charges. In DRAM PUF, \( V_{th} \) degradation of PMOS transistors in sense amplifiers can affect the reproducibility of signatures. The results in [20], show the \( V_{th} \) shift of PMOS transistors under variable temperature over a 10 year lifetime of the product.

To investigate the robustness of signatures under NBTI, we updated the PMOS technology file [17] according to the expected \( V_{th} \) shift corresponding to 1, 3, 5, and 10 years lifetime of the product, prior to performing the simulations.
obtained from 1000 chips under NBTI effect. The highest average signature variation is very low due to the small number of PMOS transistors in DRAM array circuitry. The highest average signature variation is ≈1.45 bits which occurs after 10 years where we have the highest $V_{th}$ shift in PMOS.

Fig. 12 shows the distribution of $\text{intra-die } HD_{avg}$ of the signatures obtained from 1000 chips under NBTI effect. As it is shown, the majority of chips (≈98.9%) have fewer than 3 bit-flips in their signatures, including 750 chips with not a single bit-flip, showing that the proposed PUF is perfectly stable under NBTI effect.

C. Hardware Overhead Analysis
For the proposed DRAM authentication methodology, we designed a 128-cell DRAM architecture, including sense amplifiers, equalizers, and read and write circuitry using Synopsys HSPICE. Synthesis for the peripherals, such as a 4 × 16 row address decoder, 8-bit I/O with buffers, and a delay generator has been done in Synopsys Design Compiler. Since we use the DRAM that is already present in the design, the area overhead is only due to the delay generator. For our experimental DRAM which is less than a KByte, the area overhead is about 7%. However, given that the delay generator is a small fixed circuit independent of the size of DRAM, the area overhead becomes negligible for large off-the-shelf DRAMs with several Gigabytes of storage.

D. Results Comparison
The nature of our authentication methodology lends itself to be faster than the authentication method in [8]. Also using the Hamming distance metric, the number of bits differing between two random responses in our approach is 50.01% compared to 0.016% in [8]. Furthermore, the simulation results show that our methodology is more robust under temperature variation.

V. CONCLUSION AND FUTURE WORK
We have presented a robust and secure authentication methodology utilizing DRAM array. Our methodology leverages on word line duty-cycle-controlled write failures to produce unique signatures. Simulation results for 1000 chips with 10% inter- and 8% intra-die variation show that our methodology provides high uniqueness of 50.01% inter-die $HD_{avg}$. It also provides good reproducibility of 7% total intra-die $HD_{avg}$ for temperatures ranging from 20°C to 50°C and 7.4% total intra-die $HD_{avg}$ for supply voltages ranging from 0.9 to 1.2 volts. Furthermore, ageing analysis shows that under NBTI effect, signatures show only 0.92% estimated unstable bits over a 10 year lifetime. Our approach incurs very low design overhead by utilizing the DRAM array already present in the design. Future work includes validation of the simulation results, as well as exploiting other forms of memory failures in DRAM.

REFERENCES