A Study of the Spatial Auto-Correlation of Memory-Based Physical Unclonable Functions

Tolga Arul
Technical University of Darmstadt & University of Passau
arul@seceng.informatik.tu-darmstadt.de
tolga.arul@uni-passau.de
ORCID: 0000-0002-2078-3976

Nikolaos Athanasios Anagnostopoulos
Technical University of Darmstadt & University of Passau
anagnostopoulos@seceng.informatik.tu-darmstadt.de
nikolaos.anagnostopoulos@uni-passau.de
ORCID: 0000-0003-0243-8594

Sergej Reißig
Technical University of Darmstadt
sergej.reissig@secuinfra.com

Stefan Katzenbeisser
University of Passau
stefan.katzenbeisser@uni-passau.de

***Abstract*—In this work, we examine the spatial auto-correlation exhibited in the responses of memory-based Physical Unclonable Functions (PUFs). In particular, we examine the responses of an SRAM PUF, a DRAM decay-based PUF, and a disturbance-based Flash PUF. For the evaluation, we use three different metrics that have already been employed in the relevant literature for measuring the spatial correlation of other PUF responses. Our results prove that the examined PUF responses exhibit little, if any, spatial auto-correlation. Thus, these PUFs can be considered as security mechanisms of high entropy, which can be utilised to enhance the security of the Internet of Things (IoT).**

***Index Terms*—correlation, spatial auto-correlation, Physical Unclonable Function (PUF), Static Random Access Memory (SRAM), Dynamic Random Access Memory (DRAM), Flash memory**

I. Introduction

In recent years, Physical Unclonable Functions (PUFs) have been proposed as a security mechanism for cost-efficient electronic devices, such as those utilised in the Internet of Things (IoT). PUFs are most often hardware modules that act as hardware-entangled functions, i.e. for each input, which is referred to as a *challenge*, they provide a specific output, which is referred to as a *response*. The pair of a challenge and its corresponding response is known as a *Challenge-Response Pair* (CRP) and the number of CRPs available for a PUF provides a first indication of its overall entropy, assuming that CRPs are independent of each other, i.e. uncorrelated. Accordingly, PUFs with only a single or a few CRPs are characterised as *weak*, while PUFs with such a large number of CRPs that their characterisation within a limited time frame is difficult, are known as *strong* [1]. Additionally, PUF responses are, most often, robust, unique per PUF instance and of high entropy, which allows their use in cryptography.

Common applications of PUFs include random number generation, identification, authentication, attestation, secure boot and secure key generation [2]–[5]. A number of established metrics have been used, in the relevant literature, to ensure the security of PUFs. In particular, the Hamming weight and the (binary) Shannon entropy of PUF responses reveal the randomness of PUF instances, whereas intra-device Hamming distance assesses their robustness and inter-device Hamming distance measures their uniqueness.

However, recent works on attacks against strong PUFs [6], [7] seem to indicate that the responses of strong PUFs exhibit a certain degree of correlation, which makes them susceptible to modelling and machine learning attacks. It has also been noted that correlation in the responses of weak PUFs can significantly affect their unpredictability [8]. For this reason, a number of recent works have investigated the spatial correlation of PUF responses of such PUFs as the Ring Oscillator (RO) PUF and the SRAM PUF, both internally and at the wafer level [8]–[11]. Other recent works have examined the correlation of the refresh and error characteristics of DRAMs [12], [13], thus providing some, albeit limited, insights into the spatial correlation of the DRAM decay-based PUF and the Row Hammer PUF, respectively.

Our work validates and extends previous works regarding the spatial correlation of memory-based PUFs, such as SRAM PUFs [8]–[11], which, while being considered as weak PUFs, have already been commercialised [14]. In particular, we study the spatial auto-correlation of some of the most well-known memory-based PUFs, i.e. the SRAM, the DRAM decay-based and the Flash PUF. In this way, we contribute to a more comprehensive overview of spatial correlation in PUFs and their general aptitude to serve as security mechanisms.

The remainder of this paper is organised as follows. Section II briefly introduces background information regarding the utilised metrics and the examined devices. In Section III, we present the results of the spatial auto-correlation measurements for the examined SRAM, DRAM decay-based, and Flash disturbance-based PUFs. Finally, Section IV provides an outlook on future work and concludes this work.

This work has been partially funded by the German Research Foundation – Deutsche ForschungsGemeinschaft (DFG), under Project: "PUFMem: Intrinsic Physical Unclonable Functions from Emerging Non-Volatile Memories" (AR 1387/1-1 — KA 2308/4-1) of the Priority Program "Nano Security: From Nano-Electronics to Secure Systems" (SPP 2253).

II. Background

In this work, we evaluate the spatial auto-correlation of memory-based PUFs, by measuring the similarity of the values of their memory cell values to those of the spatial neighbourhood of these cells in *individual* measurements. In contrast, spatial correlation or cross-correlation examines the similarity of memory cell values to their spatial neighbourhood in *different* measurements. In order to determine the spatial auto-correlation of PUF responses, we apply three metrics established in the recent literature [8]–[11], namely Moran's I [15], Geary's C [16] and the Join Count statistic [17]. The purpose of this assessment is to ascertain that PUF responses are spatially independent, which contributes to their entropy and thus to the suitability of each PUF instance to act as an adequate security mechanism. We also briefly discuss the different devices used for the implementation of the examined PUFs, namely an SRAM, a DRAM and a Flash memory found in different IoT devices.

A. Spatial Auto-Correlation Metrics

The calculation of the Moran's I and Geary's C metrics is based on a weight matrix, which is shown in Equation (1). This $n \times n$ matrix defines the impact of each memory cell i to each other cell j, as each matrix element $w_{i,j}$ is based on the distance between i and j. The larger the distance between two cells i and j is, the lower the relevant weight denoted by $w_{i,j}$. The diagonal of the matrix is populated with zeros, as by definition a cell is not adjacent to itself. Although different methods can be used to calculate the distance between two cells, in our work, we utilise the well-known concept of the Euclidean distance for this purpose.

$$\{w_{i,j}\} = \begin{bmatrix} 0 & w_{1,2} & w_{1,3} & \dots & w_{1,n} \\ w_{2,1} & 0 & w_{2,3} & \dots & w_{2,n} \\ w_{3,1} & w_{3,2} & 0 & \ddots & w_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & w_{n,3} & \dots & 0 \end{bmatrix} \tag{1}$$

1) Moran's I: Moran's I can be calculated using Equation (2), iterating over all existing cell values. Its values range between -1 and 1, with a value of -1 indicating perfect negative correlation and a value of 1 indicating perfect positive correlation. Finally, a value of 0 indicates the absence of any correlation. In Equation (2), n and $w_{i,j}$ correspond to either dimension of Equation (1) and to its elements, respectively, while x_i and x_j refer to the values of cells i and j, respectively. Finally, $\overline{x}$ is the mean value of the memory cells.

$$I = \frac{n}{\sum_i \sum_j \ w_{i,j}} \frac{\sum_i \sum_j \ w_{i,j}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i \ (x_i - \overline{x})^2} \tag{2}$$

2) Geary's C: Geary's C is conceptually similar to Moran's I, but is a metric that is more sensitive to local spatial auto-correlation. Geary's C can be calculated using Equation (3), and its values range between 0 and 2, with a value of 2 indicating perfect negative correlation and a value of 0 indicating perfect positive correlation. Finally, a value of 1 indicates the absence of any correlation. In Equation (3), n and $w_{i,j}$ correspond to either dimension of Equation (1) and to its elements, respectively, while x_i and x_j refer to the values of cells i and j, respectively. Finally, $\overline{x}$ is the mean value of the memory cells.

$$C = \frac{n-1}{2\sum_i \sum_j \ w_{i,j}} \frac{\sum_i \sum_j \ w_{i,j}(x_i - x_j)^2}{\sum_i \ (x_i - \overline{x})^2} \tag{3}$$

TABLE I
Overview of the Examined Memory-Based PUFs

Device	Memory Model	Size	PUF Type
TI Stellaris	LX4F120H5QR MCU (on-die)	2^{15}B	SRAM startup
PandaBoard ES	Micron EDB8164B3PF-8D-F	2^{30}B	DRAM decay
STM 32F429I	Samsung K9F1G08U0E	2^{30}B	Flash disturbance

3) Join Count: This metric is significantly different from the other two, since its value is the sum of neighbouring memory cells that contain a different logical value from any other cell. The cells are designated as *black* or *white*, based on whether their logical value is '1' or '0', respectively. The value of this metric can be calculated using Equation (4), where $w_{i,j}$ corresponds to the respective element of Equation (1), and x_i and x_j refer to the values of cells i and j, respectively. Based on the layout of the memory, the value for this metric that corresponds to the lack of correlation can be easily calculated, with values lower than this "expected" value indicating positive correlation and values higher than the "expected" value indicating negative correlation. Therefore, this "expected" value lies essentially in the middle of the range of values that are potentially possible for this metric.

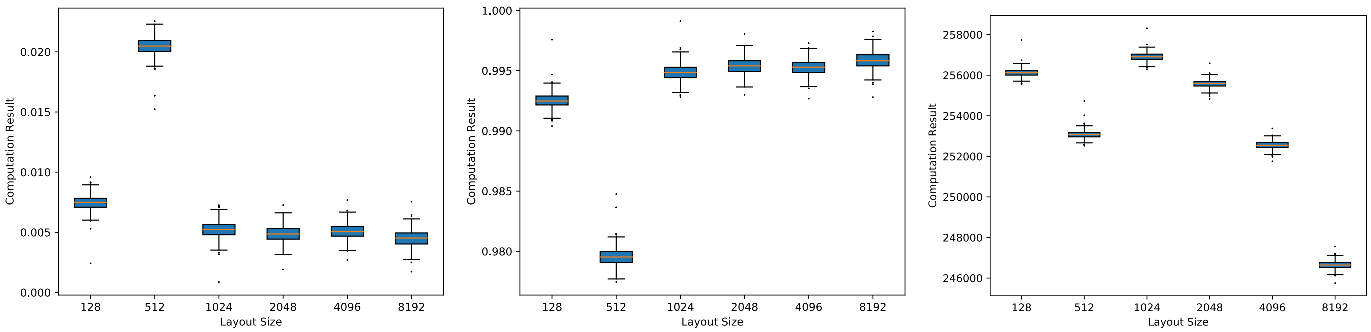
$$J_{BW} = \frac{1}{2} \sum_i \sum_j w_{i,j}(x_i - x_j)^2 \tag{4}$$

B. Memory-Based Physical Unclonable Functions Examined

In order to provide a comprehensive study of spatial auto-correlation in memory-based PUFs, we examined the most well-known such PUFs. In particular, we measured the spatial auto-correlation in the responses of an SRAM PUF, a DRAM decay-based PUF and a Flash disturbance-based PUF. An overview of the examined PUFs is provided in Table I.

1) SRAM PUF: The response of an SRAM PUF is the concatenation of the uninitialised values of the SRAM at start-up and, thus, it can only be obtained after a reboot. We implemented this PUF on the on-die SRAM of the LX4F120H5QR microcontroller, which forms an inherent component of the Texas Instruments Stellaris evaluation boards.

2) DRAM Decay-Based PUF: As the name reveals, the response of a DRAM decay-based PUF is based on the natural decay of the values stored on the DRAM cells, when these cells are not being refreshed. In order to obtain a reliable response, certain ranges of cells are set to a specific logical value and then allowed to decay, due to a lack of being refreshed. The DRAM of a PandaBoard ES Rev. B3 development board was utilised. The TI OMAP 4460 SoC processor and the Elpida (Micron) EDB8164B3PF-8D-F DRAM of this

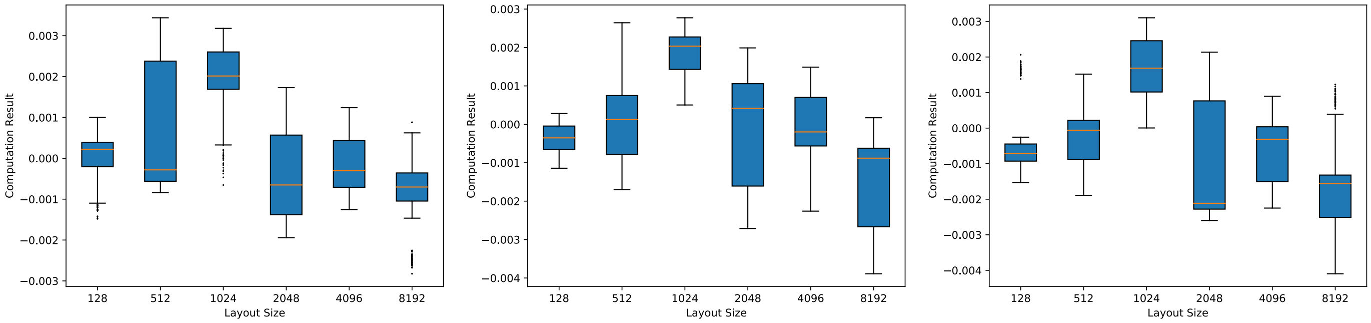


(a) Moran's I

(b) Geary's C

(c) Join Count

Fig. 1. Results of the spatial auto-correlation metrics for the examined SRAM PUF.



(a) Decay Time Duration: 120s

(b) Decay Time Duration: 240s

(c) Decay Time Duration: 360s

Fig. 2. Moran's I results for the examined DRAM PUF, for different decay time durations and the same temperature.

board are contained in the same chip, as a package-on-package scheme is utilised. As the DRAM decay PUF responses are highly dependent on both the ambient temperature and the decay time duration, responses were obtained at 40°C, 60°C and 80°C for the same decay time duration, and at the same ambient temperature for decay time durations of $120s$, $240s$ and $360s$, in order to test whether such responses can truly provide reliable security.

3) Flash Disturbance-Based PUF: The response of a Flash disturbance-based PUF is based on the disturbances that the constant erasure of memory pages causes to the values stored in other nearby pages of the Flash memory. A Samsung K9F1G08U0E NAND Flash memory chip, being controlled by an STM 32F429I Discovery board, was utilised for the implementation of this PUF.

III. Spatial Auto-Correlation of Memory-Based PUF Responses

In this section, we present the results of our study regarding the spatial auto-correlation of the aforementioned memory-based PUFs. As our results show, memory-based PUF responses seem to exhibit little to no auto-correlation. In particular, as the memory layout can have a defining role on the spatial auto-correlation of the examined memory-based PUFs, we have examined 6 different potential memory arrangements, with words of a length ranging from 128 to 8192 bytes.

A. Spatial Auto-Correlation of SRAM PUF Responses

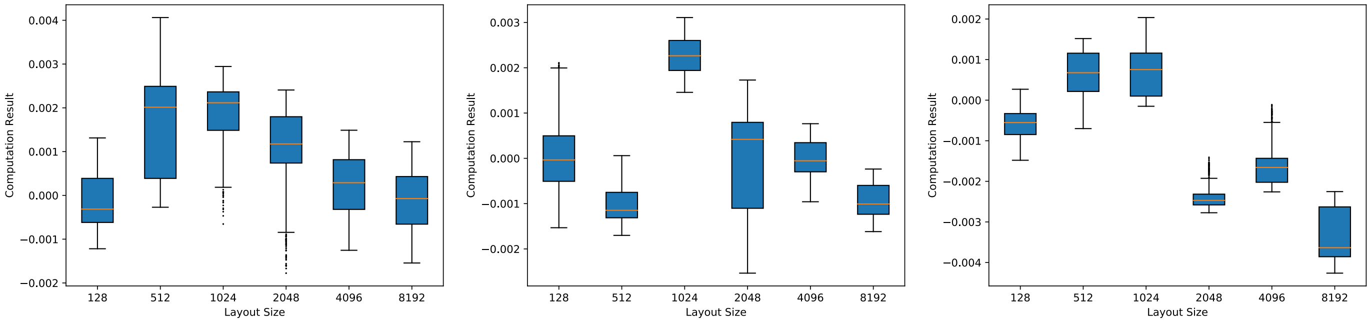
As it can be seen in Figure 1, the examined SRAM PUF exhibits only marginal positive auto-correlation, especially for a potential word length of 512 bytes. We note that, while the results for Moran's I and Geary's C are similar, the results for the Join Count exhibit a higher sensitivity to layout organization. In all three cases, however, the results are extremely close to values indicating extremely low to no auto-correlation.

B. Spatial Auto-Correlation of DRAM Decay PUF Responses

We note that, for the DRAM decay-based PUF, we do not consider the contained values in each memory cell for computing the introduced metrics but rather the positions of the cells in which a bit flip occurs due to decay for the calculation of the results. We choose this approach, since, otherwise, almost all acquired values would be identical and completely auto-correlated and, consequently, no valid conclusion would be drawn regarding the quality of the PUF response. We note that our results for Geary's C and the Join Count metrics confirm the results for Moran's I and, therefore, present only our results for Moran's I due to the brevity of this work. In particular, in Figure 2, we observe that spatial auto-correlation values are independent of the decay time duration, and that all the values for Moran's I are extremely close to the ideal value of 0, indicating little to no auto-correlation. In Figure 3, we similarly observe that also ambient temperature variations do not seem to have an effect on the auto-correlation of DRAM decay PUFs, as the values for Moran's I are extremely close to the ideal value of 0, indicating little to no auto-correlation.

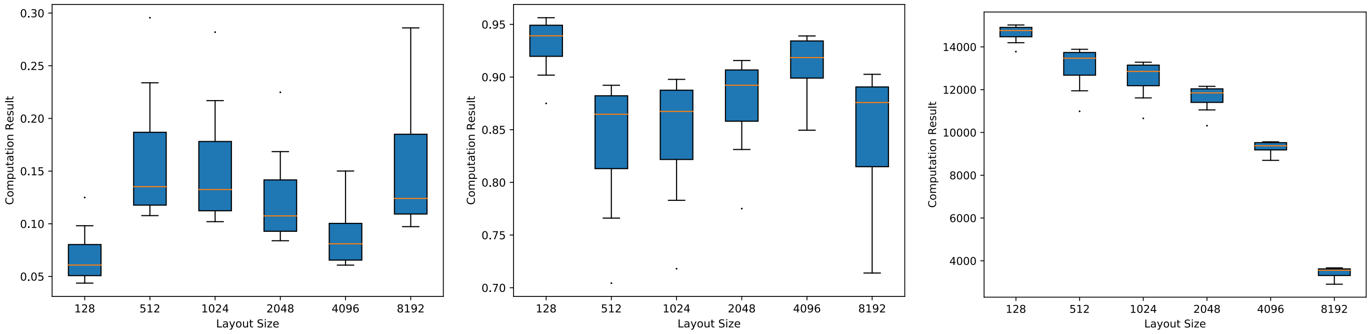
C. Spatial Auto-Correlation of Flash PUF Responses

As Figure 4 demonstrates, the examined Flash disturbance-based PUF exhibits only a low degree of positive auto-correlation, especially for a potential word length of 128



(a) Ambient Temperature: 40°C (b) Ambient Temperature: 60°C (c) Ambient Temperature: 80°C

Fig. 3. Moran's I metric result for the investigated DRAM PUF and different temperatures



(a) Moran's I (b) Geary's C (c) Join Count

Fig. 4. Results of the spatial auto-correlation metrics for the examined Flash PUF.

bytes, based on the results for Moran's I and Geary's C. However, the results for the Join Count provide are rather contradictory, calling, thus, for further investigation. In all three cases, however, the results are extremely close to values indicating low to no auto-correlation.

IV. Conclusion and Future Work

In this work, we have evaluated the spatial auto-correlation of 3 types of memory-based PUFs, namely, SRAM startup, DRAM decay and Flash disturbance PUFs by employing 3 different metrics, namely Moran's I, Geary's C and Join Count statistics. Our results indicate that the examined PUFs exhibit little to no spatial auto-correlation, with the Flash PUF requiring further investigation. Therefore, our results indicate that memory-based PUFs remain a reliable security mechanism for the IoT. Future work should examine the spatial auto-correlation of other memory-based PUFs implementations, such as the Row Hammer PUF.

References

[1] C. Herder, M. D. Yu, F. Koushanfar, and S. Devadas, "Physical Unclonable Functions and Applications: A Tutorial," *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1126–1141, Aug 2014.

[2] H. Handschuh, "Hardware-Anchored Security Based on SRAM PUFs, Part 2," *IEEE Security Privacy*, vol. 10, no. 4, pp. 80–81, July 2012.

[3] S. Schulz, A. Schaller, F. Kohnhäuser, and S. Katzenbeisser, "Boot Attestation: Secure Remote Reporting with Off-The-Shelf IoT Sensors," in *Computer Security – ESORICS 2017*. Springer, 2017, pp. 437–455.

[4] F. Kohnhäuser, A. Schaller, and S. Katzenbeisser, "PUF-Based Software Protection for Low-End Embedded Devices," in *Trust and Trustworthy Computing*. Springer, 2015, pp. 3–21.

[5] N. A. Anagnostopoulos, S. Ahmad, T. Arul, D. Steinmetzer, M. Hollick, and S. Katzenbeisser, "Low-Cost Security for Next-Generation IoT Networks," *ACM Transactions on Internet Technology*, 2020.

[6] F. Ganji, S. Tajik, F. Fäßler, and J.-P. Seifert, "Having No Mathematical Model May Not Secure PUFs," *Journal of Cryptographic Engineering*, vol. 7, pp. 113–128, 2017.

[7] N. Wisiol, G. T. Becker, M. Margraf, T. A. A. Soroceanu, J. Tobisch, and B. Zengin, "Breaking the Lightweight Secure PUF: Understanding the Relation of Input Transformations and Machine Learning Resistance," in *Smart Card Research and Advanced Applications*. Springer, 2020, pp. 40–54.

[8] B. Willsch, J. Hauser, S. Dreiner, A. Goehlich, and H. Vogt, "Statistical Tests to Determine Spatial Correlations in the Response Behavior of PUF," in *12th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME 2016)*, June 2016, pp. 1–4.

[9] F. Wilde, B. Gammel, and M. Pehl, "Spatial Correlations in Physical Unclonable Functions," in *6th Conference on Trustworthy Manufacturing and Utilization of Secure Devices (TRUDEVICE 2016)*, 2016.

[10] Z. Liao and Y. Guan, "The Cell Dependency Analysis on Learning SRAM Power-Up States," in *2018 Asian Hardware Oriented Security and Trust Symposium (AsianHOST)*, 2018, pp. 37–43.

[11] F. Wilde, B. M. Gammel, and M. Pehl, "Spatial Correlation Analysis on Physical Unclonable Functions," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 6, pp. 1468–1480, June 2018.

[12] E. Baseman, N. Debardeleben, S. Blanchard, J. Moore, O. Tkachenko, K. Ferreira, T. Siddiqua, and V. Sridharan, "Physics-Informed Machine Learning for DRAM Error Modeling," in *2018 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*, Oct 2018, pp. 1–6.

[13] A. Agrawal, A. Ansari, and J. Torrellas, "Mosaic: Exploiting the Spatial Locality of Process Variation to Reduce Refresh Energy in on-Chip eDRAM Modules," in *2014 IEEE 20th International Symposium on High Performance Computer Architecture (HPCA)*, Feb 2014, pp. 84–95.

[14] Intrinsic ID, "SRAM PUF Technology," 2020, accessed: 2020-07-01. [Online]. Available: https://www.intrinsic-id.com/sram-puf/

[15] P. A. P. Moran, "Notes on Continuous Stochastic Phenomena," *Biometrika*, vol. 37, no. 1/2, pp. 17–23, 1950.

[16] R. C. Geary, "The Contiguity Ratio and Statistical Mapping," *The Incorporated Statistician*, vol. 5, no. 3, pp. 115–146, 1954.

[17] A. D. Cliff and J. K. Ord, *Spatial Processes: Models & Applications*. Pion, 1981.