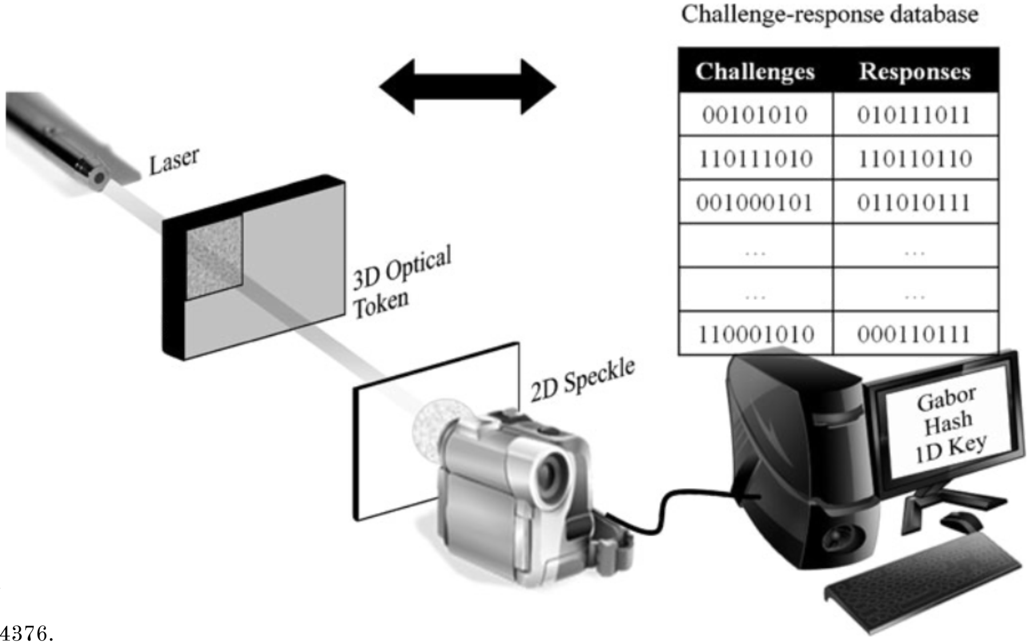
Breaking the Lightweight Secure PUF:

Understanding the Relation of Input Transformations and Machine Learning Resistance

18th Smart Card Research and Advanced Application Conference: CARDIS 2019

Nils Wisiol, Georg T. Becker, Marian Margraf, Tudor A. A. Soroceanu, Johannes Tobisch, Benjamin Zengin

Physically Unclonable Functions



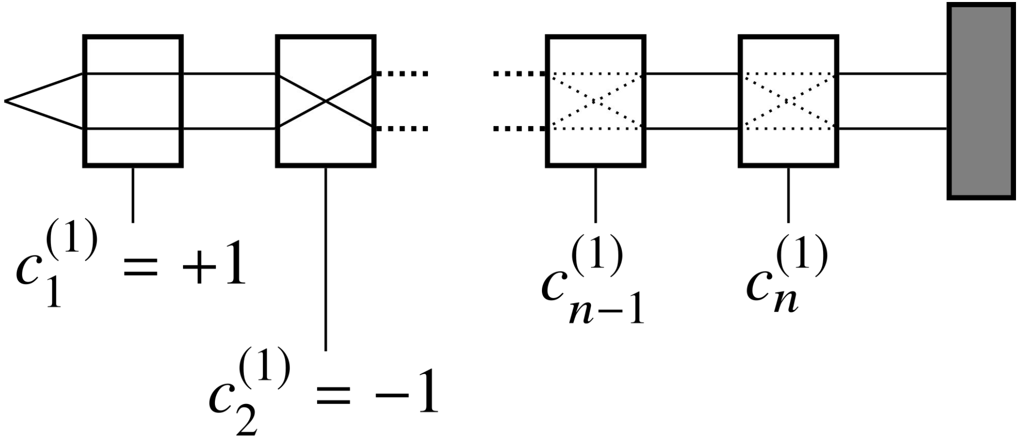
Challenge-response database

Challenges	Responses
00101010	010111011
110111010	110110110
001000101	011010111
...	...
...	...
110001010	000110111

Original research: Pappu, Ravikanth, Ben Recht, Jason Taylor, and Neil Gershenfeld. “Physical One-Way Functions.” Science 297, no. 5589 (September 20, 2002): 2026–30. https://doi.org/10.1126/science.1074376.

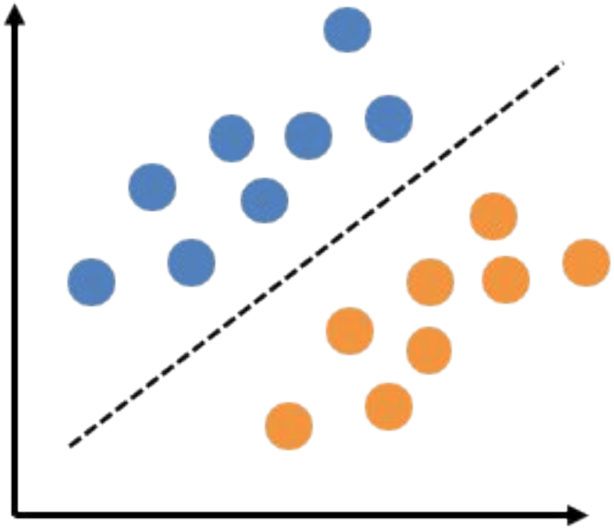
Image source: Rührmair, Ulrich, Srinivas Devadas, and Farinaz Koushanfar. “Security Based on Physical Unclonability and Disorder.” In Introduction to Hardware Security and Trust, edited by Mohammad Tehranipoor and Cliff Wang, 65–102. New York, NY: Springer New York, 2012. https://doi.org/10.1007/978-1-4419-8080-9_4.

Arbiter PUF 101

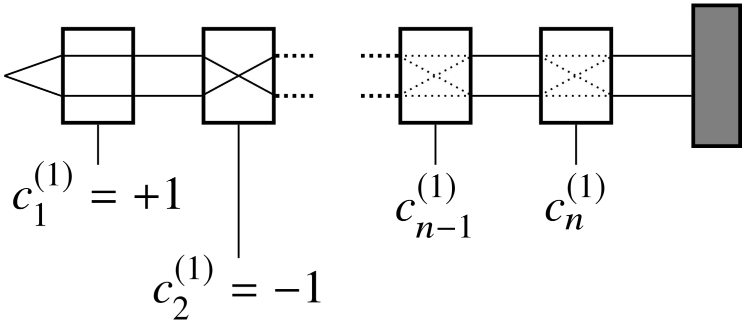


Gassend, Blaise, Dwaine Clarke, Marten van Dijk, and Srinivas Devadas. “Delay-Based Circuit Authentication and Applications.” In Proceedings of the 2003 ACM Symposium on Applied Computing, 294–301. SAC ’03. New York, NY, USA: ACM, 2003. https://doi.org/10.1145/952532.952593.

Can the behavior be modeled?



Arbiter Physical Unclonable Functions (Electric)



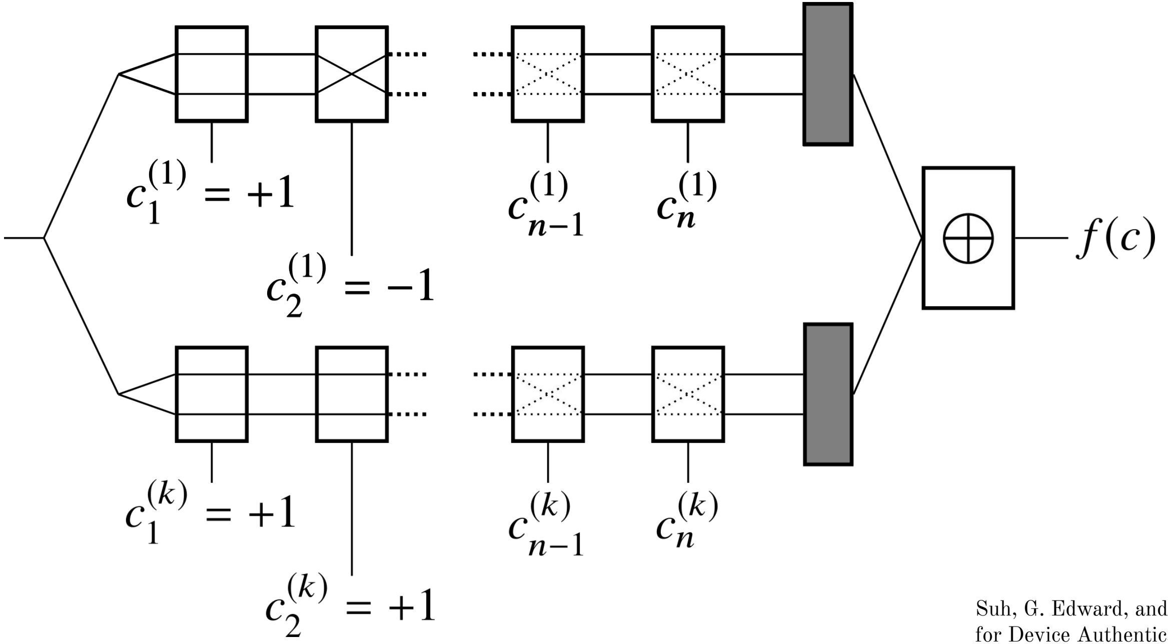
Challenge – attacker known

$$\operatorname{sgn}\langle w, x\rangle$$

Physical parameters – attacker unknown

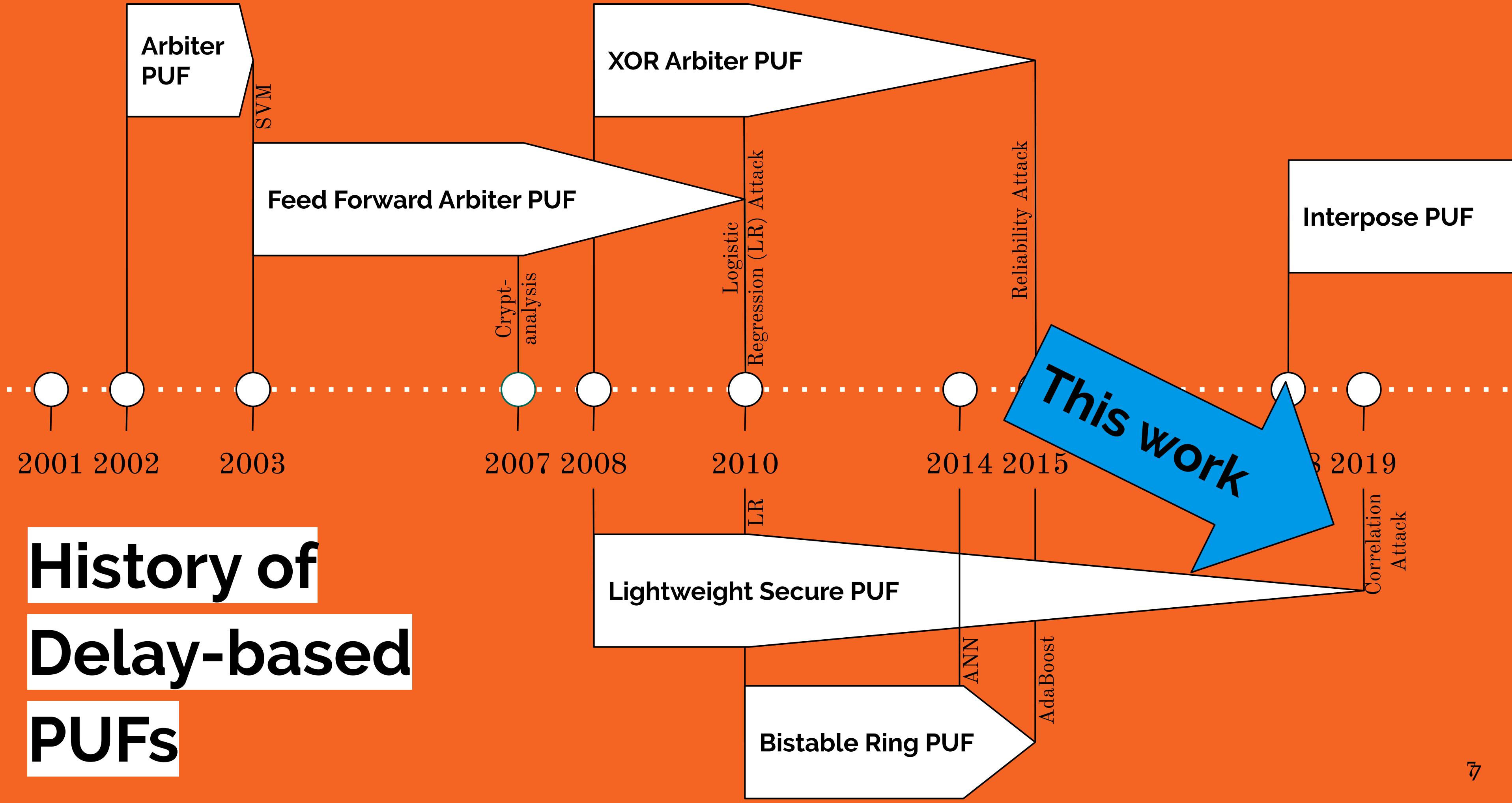
Gassend, Blaise, Dwaine Clarke, Marten van Dijk, and Srinivas Devadas. “Delay-Based Circuit Authentication and Applications.” In Proceedings of the 2003 ACM Symposium on Applied Computing, 294–301. SAC ’03. New York, NY, USA: ACM, 2003. https://doi.org/10.1145/952532.952593.

Arbiter PUF Variants: XOR Arbiter PUF



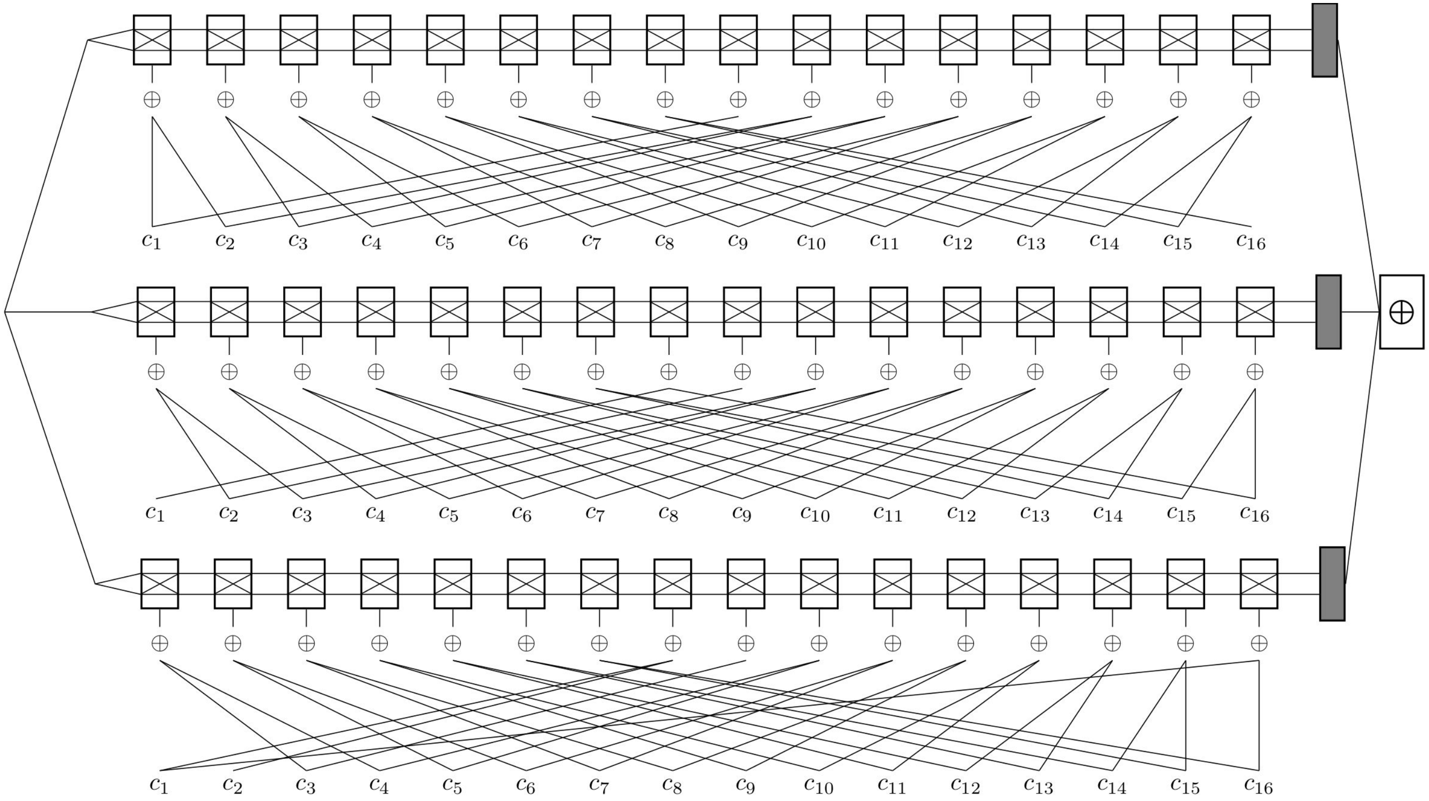
Suh, G. Edward, and Srinivas Devadas. “Physical Unclonable Functions for Device Authentication and Secret Key Generation.” In Proceedings of the 44th Annual Design Automation Conference, 9–14. DAC ’07. New York, NY, USA: ACM, 2007. https://doi.org/10.1145/1278480.1278484.

History of Delay-based PUFs



Lightweight Secure PUF

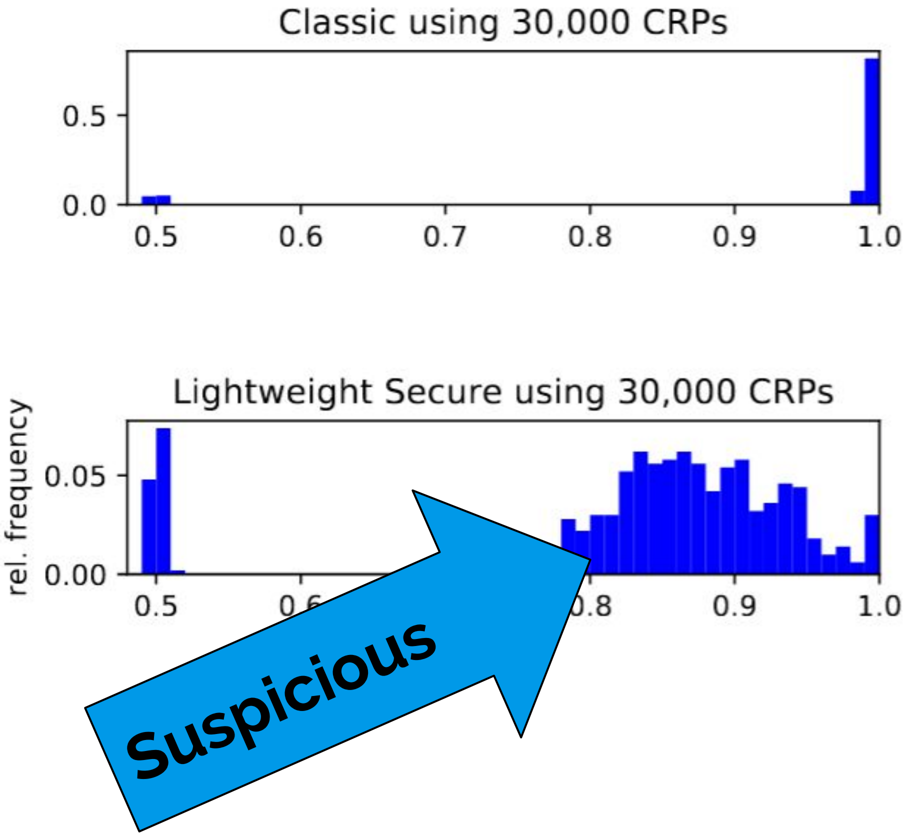
Lightweight Secure PUF


c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16}
c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16}
c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16}
⊕

Majzoobi, Mehrdad, Farinaz Koushanfar, and Miodrag Potkonjak. “Lightweight Secure PUFs.” In Proceedings of the 2008 IEEE/ACM International Conference on Computer-Aided Design, 670–673. ICCAD ’08. Piscataway, NJ, USA: IEEE Press, 2008. http://dl.acm.org/citation.cfm?id=1509456.1509603.

Correlation Attack

Logistic Regression Attack



Classic using 30,000 CRPs

Lightweight Secure using 30,000 CRPs

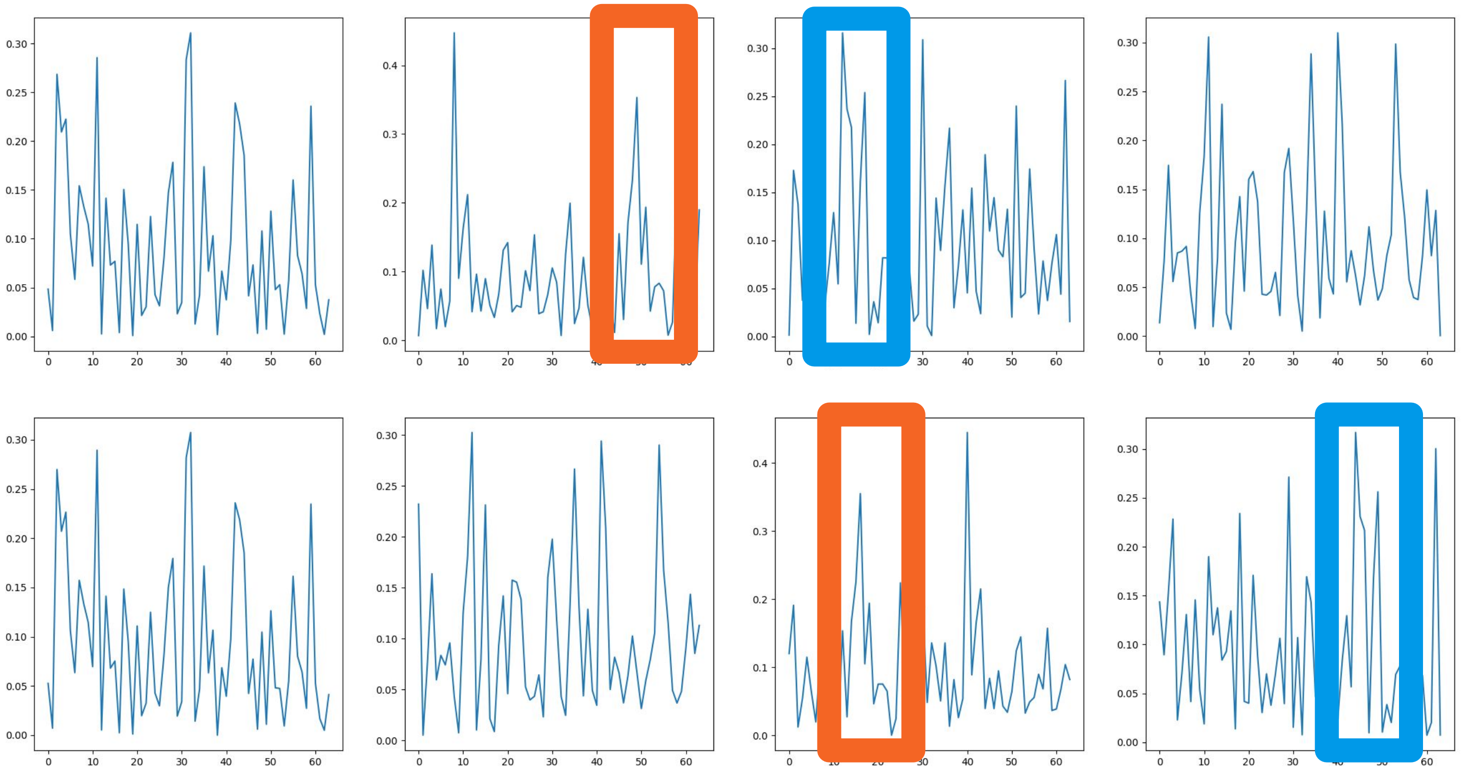
Suspicious

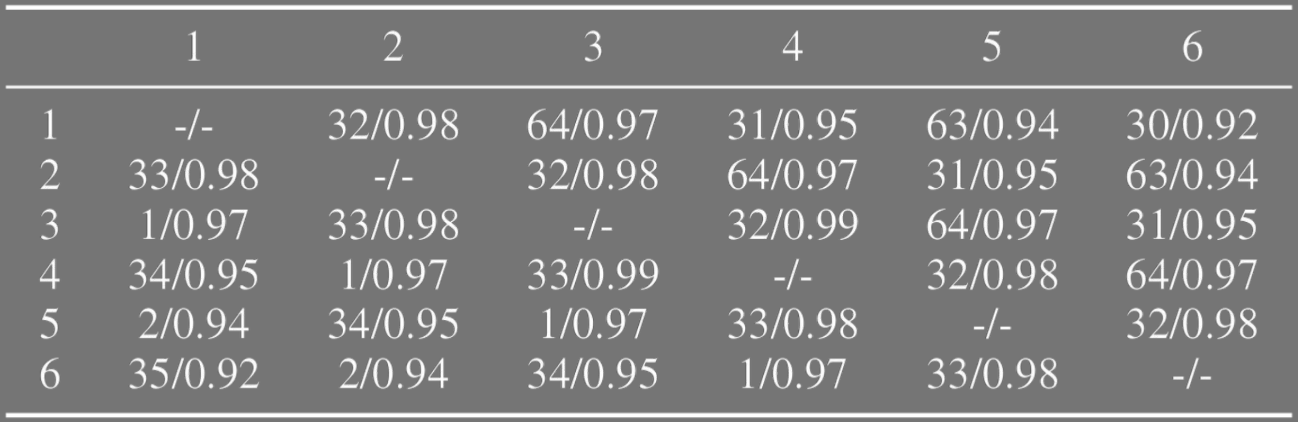
Accuracy distribution of machine learning results for 64-bit 4-XOR Arbiter PUFs and 64-bit 4-XOR Lightweight Secure PUFs.

Correlation Example (4-XOR 64-bit LW-Sec.)

Learned Weights

Simulation Weights



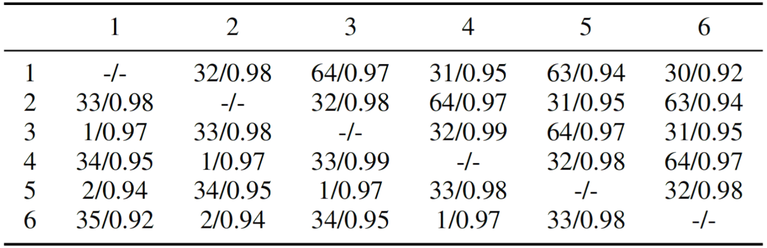


	1	2	3	4	5	6
1	-/-	32/0.98	64/0.97	31/0.95	63/0.94	30/0.92
2	33/0.98	-/-	32/0.98	64/0.97	31/0.95	63/0.94
3	1/0.97	33/0.98	-/-	32/0.99	64/0.97	31/0.95
4	34/0.95	1/0.97	33/0.99	-/-	32/0.98	64/0.97
5	2/0.94	34/0.95	1/0.97	33/0.98	-/-	32/0.98
6	35/0.92	2/0.94	34/0.95	1/0.97	33/0.98	-/-

Partial Results Reveal Information About High-Accuracy Models

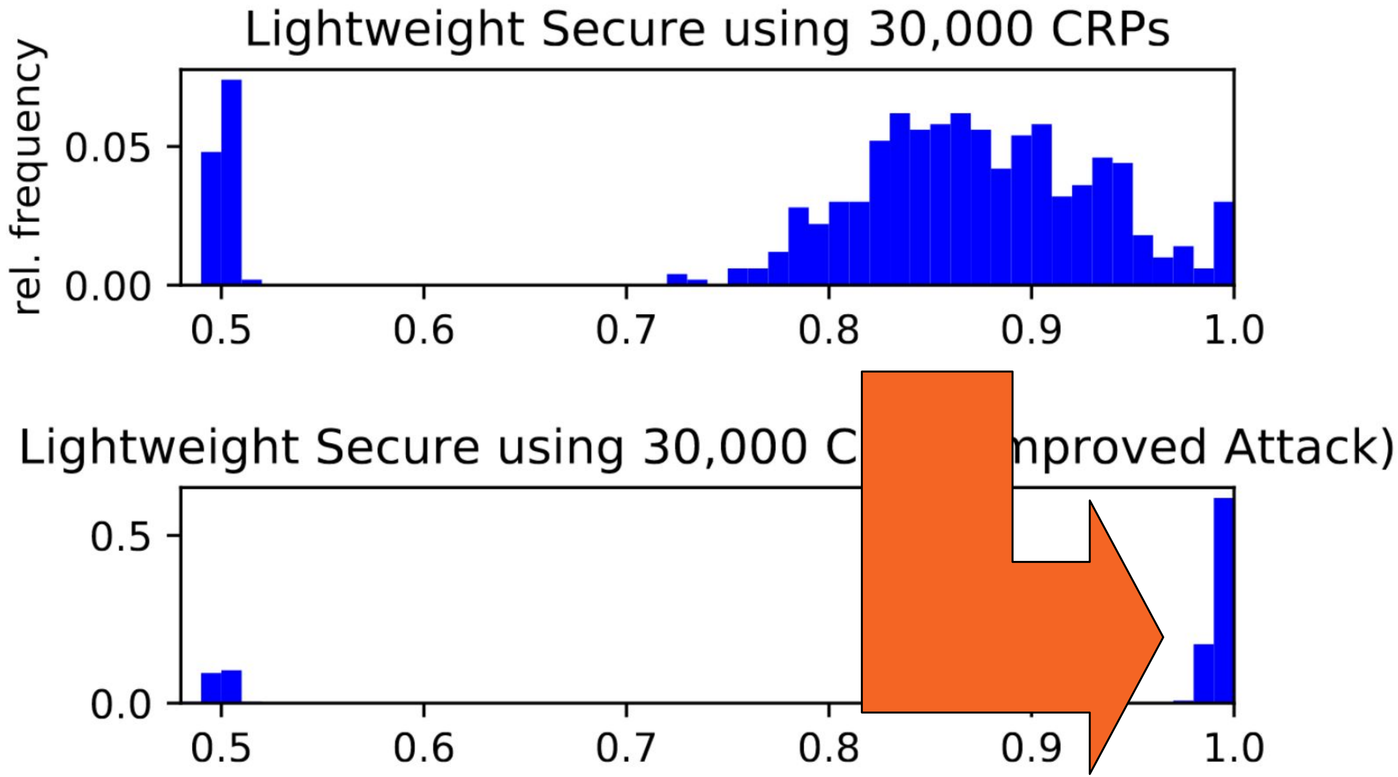
Correlation Attack

1. Train a mediocre model using the classical LR attack
2. While mediocre accuracy:
    a. Permute and switch weights
    b. Train again using LR



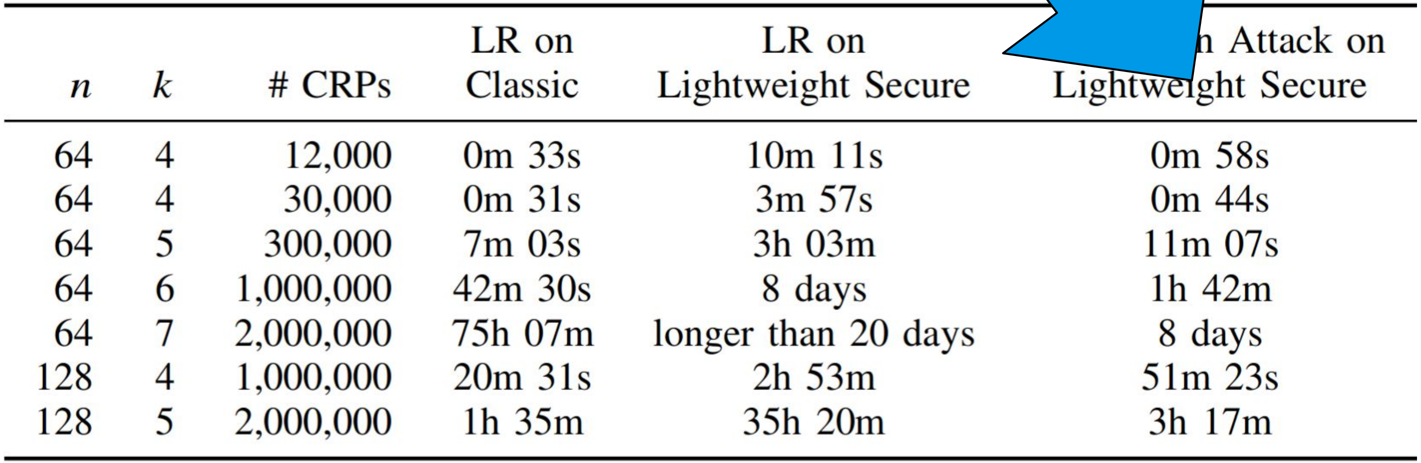
	1	2	3	4	5	6
1	-/-	32/0.98	64/0.97	31/0.95	63/0.94	30/0.92
2	33/0.98	-/-	32/0.98	64/0.97	31/0.95	63/0.94
3	1/0.97	33/0.98	-/-	32/0.99	64/0.97	31/0.95
4	34/0.95	1/0.97	33/0.99	-/-	32/0.98	64/0.97
5	2/0.94	34/0.95	1/0.97	33/0.98	-/-	32/0.98
6	35/0.92	2/0.94	34/0.95	1/0.97	33/0.98	-/-

Correlation Attack Accuracy



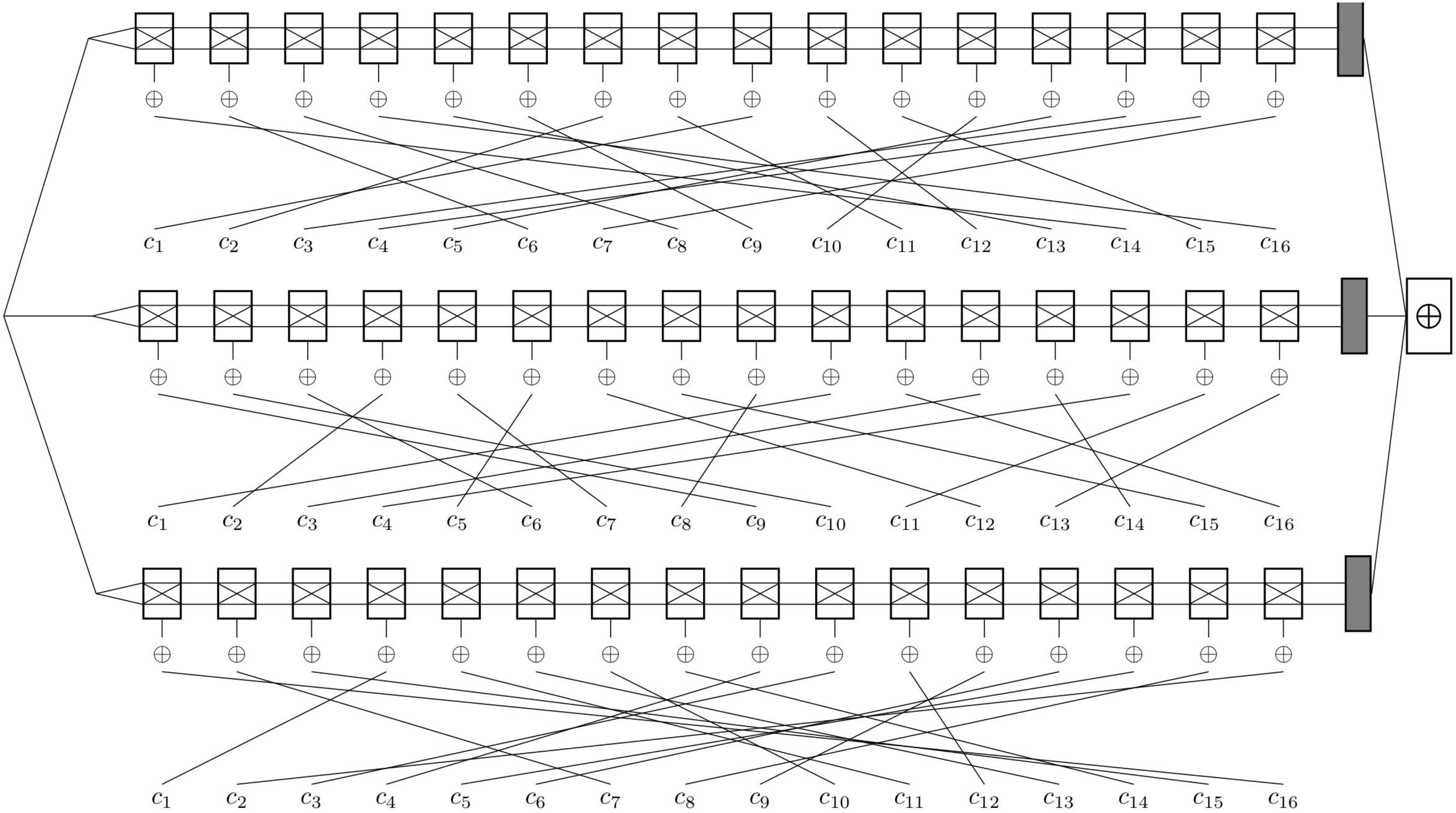
Attack Run Times

This work

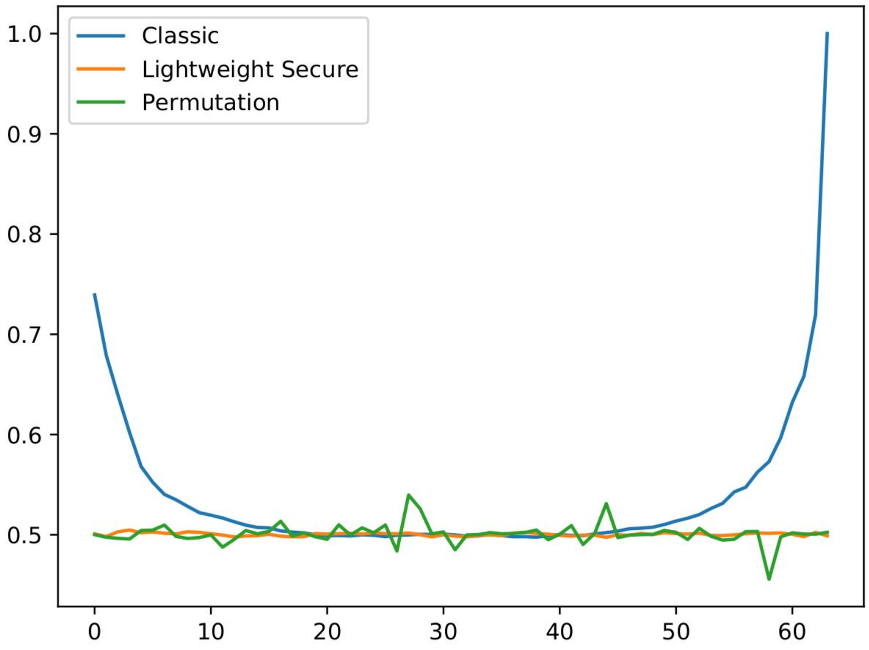


n	k	# CRPs	LR on Classic	LR on Lightweight Secure	n Attack on Lightweight Secure
64	4	12,000	0m 33s	10m 11s	0m 58s
64	4	30,000	0m 31s	3m 57s	0m 44s
64	5	300,000	7m 03s	3h 03m	11m 07s
64	6	1,000,000	42m 30s	8 days	1h 42m
64	7	2,000,000	75h 07m	longer than 20 days	8 days
128	4	1,000,000	20m 31s	2h 53m	51m 23s
128	5	2,000,000	1h 35m	35h 20m	3h 17m

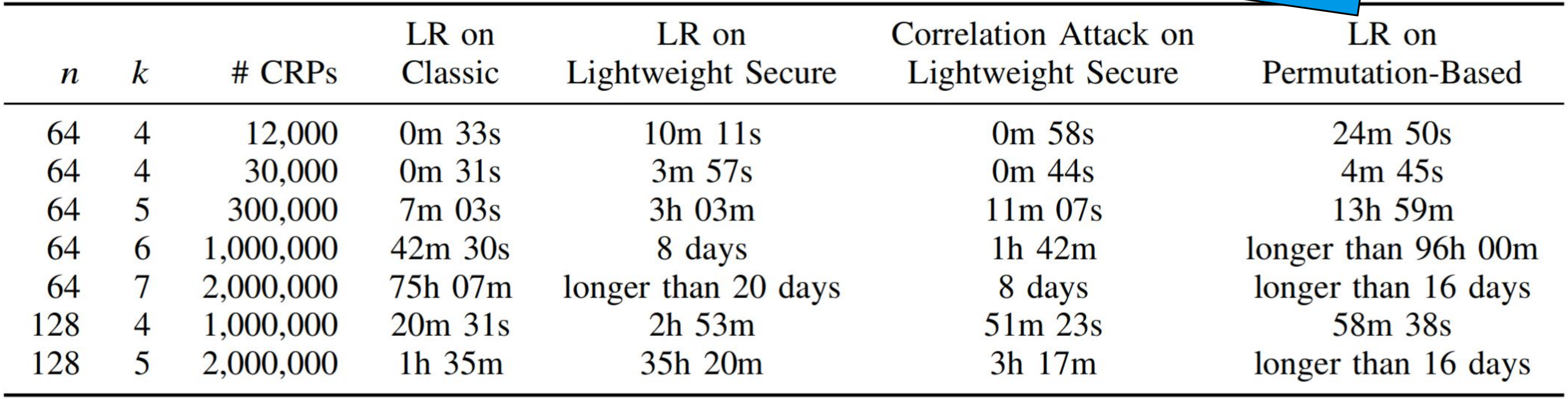
Permutation Input Transformation


c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16}
c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16}
c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16}
$\oplus$

Bit-Influence of the Permutation Input Transformation (4-XOR)



Attack Run Times



n	k	# CRPs	LR on Classic	LR on Lightweight Secure	Correlation Attack on Lightweight Secure	LR on Permutation-Based
64	4	12,000	0m 33s	10m 11s	0m 58s	24m 50s
64	4	30,000	0m 31s	3m 57s	0m 44s	4m 45s
64	5	300,000	7m 03s	3h 03m	11m 07s	13h 59m
64	6	1,000,000	42m 30s	8 days	1h 42m	longer than 96h 00m
64	7	2,000,000	75h 07m	longer than 20 days	8 days	longer than 16 days
128	4	1,000,000	20m 31s	2h 53m	51m 23s	58m 38s
128	5	2,000,000	1h 35m	35h 20m	3h 17m	longer than 16 days

Thank You!

All data and code freely available in pypuf:

`github.com/nils-wisiol/pypuf`

`nils.wisiol@fu-berlin.de`

`ia.cr/2019/799`

Breaking the Lightweight Secure PUF

Understanding the Relation of Input Transformations and Machine Learning Resistance

Nils Wisiol · {Freie, Technische} Univ Berlin
Georg T. Becker · ESMT Berlin
Marian Margraf · Freie Univ Berlin, Fraunhofer AISEC
Tudor A. A. Soroceanu · Freie Univ Berlin
Johannes Tobisch · Ruhr-Univ Bochum
Benjamin Zengin · Fraunhofer AISEC