Quality Assessment of RFET-based Logic Locking Protection Mechanisms using Formal Methods

Marcel Merten* Sebastian Huhn*† Rolf Drechsler*†

*University of Bremen, Germany
{mar_mer,huhn,drechsle}
@informatik.uni-bremen.de

†Cyber-Physical Systems, DFKI GmbH
28359 Bremen, Germany

***Abstract*—Nowadays, the fabrication of Integrated Circuits (ICs) is highly distributed over foundries yielding long and untrustworthy supply chains. This circumstance leads to security, privacy, and reliability concerns of the fabricated ICs. One technique to protect these ICs against malicious usage and counterfeit is logic locking, i.e., the design is encrypted and obfuscated by newly introduced key gates. The correct functional behavior of the IC has to be unlocked by applying a secret key. The emerging technology of Reconfigurable Field-Effect Transistors (RFETs) has recently been utilized to implement new polymorphic logic mechanisms to protect intellectual property. However, no appropriate assessment technique for determining the protection quality is available yet. This work proposes a novel method to assess the quality of these RFET-based logic locking structures allowing for detecting weak protection structures the designers can then improve.**

I. Introduction

Due to the globalization of *Integrated Circuits* (ICs) manufacturing, designers can benefit from access to advanced technology nodes without the large capital expenditure of operating their own semiconductor foundries. The distribution of the chips' manufacturing is one of the main security challenges. Therefore, a growing threat prevails about compromising the integrity of once trusted IC processes by unauthorized or untrusted users [1]. During the last decade, *Complementary Metal-Oxide-Semiconductor* (CMOS)-based protection mechanisms have been the dominant technology for implementing various protection measures. However, a trade-off exists between the achievable protection level and the resulting cost overhead. In particular, layout-level obfuscation using CMOS-based camouflaging causes a significant overhead with respect to the required area and the resulting power consumption [2].

Recent works like [1], [3], [4] have been focusing on achieving high protection while still preserving low overhead by utilizing reconfigurable silicon nanowire field-effect transistor-based polymorphic logic gates [1]. In [1], an algorithm is proposed that replaces gates with high impact on the original circuit's behavior by reconfigurable polymorphic logic gates. Afterward, the quality of the resulting logic locking functionality is assessed by a metric based on the Hamming distance of the outputs over certain applied stimuli. The result is considered optimal if the Hamming distance is 50% of the maximal Hamming distance. However, [1] has solely been tested on combinatorial circuits, and even a considered optimal result potentially exposes the correct behavior (in 50% of the applied stimuli) even though a wrong key is applied. Due to the simulation-based nature of the existing approaches, they cannot cover all input/key combinations.

To tackle the shortcomings of existing approaches, this work proposes a novel technique to assess the quality of introduced *Reconfigurable Field-Effect Transistor* (RFET)-based logic locking protection mechanisms by heavily orchestrating the *Boolean Satisfiability* (SAT) problem and *And-Inverter Graphs* (AIGs) as the underlying data structure. In the end, a seamless framework for the automated quality assessment of the circuit's protection has been designed, allowing to evaluate the strength of the *correct* key. In contrast to other techniques, all possible stimuli combinations are considered for its evaluation and *incorrect* keys are being exploited that either lead to functional equivalent behavior and the circuit's state in terms of silent-data corruption (if enabled in framework). First experiments are conducted on ITC'99 benchmark circuits showing the completeness of the proposed technique on sequential circuits and, hence, clearly outperforming existing works.

II. Preliminaries

Within the last decade, a lot of research work has been conducted to enhance electronic systems further while the classical CMOS technology has exceeded its physical limits. In order to realize even more complex systems, reconfigurable technologies have gained a lot of attention. This emerging technology employs polymorphic logic gates and, hence, is a promising candidate to exceed the boundaries of the current state-of-the-art circuit design.

A. Reconfigurable Field-Effect Transistors

Different approaches have been proposed to realize a device-level reconfiguration capability like RFETs. An RFET can be configured between an n-channel and p-channel behavior [2] by adding a new control gate. It is taken advantage of the reconfiguration capability of this new technology to, among others, implement new protection mechanisms like on-chip key storage by the polymorphic logic behavior [2]. Since a RFET realizes two functionalities in the same cell, it provides an effective way to realize protection mechanisms against optical reverse-engineering attacks.

A well-known approach to avoid reverse engineering, even given the entire layout, is about introducing logic locking mechanisms. Logic locking uses a secret key to encrypt the correct functional behavior of a circuit. Typically, CMOS-based approaches result in a huge overhead in the area- and power-consumption [1] by introducing XOR/XNOR key gates [5]–[7] or MUX gates [8]–[11] to obfuscate the correct functional behavior of the circuit.

NAND				NOR			
a	b	key	out	a	b	key	out
0	0	0	1	0	0	1	1
0	1	0	1	0	1	1	0
1	0	0	1	1	0	1	0
1	1	0	0	1	1	1	0

Figure 1: NAND/NOR-RFET

Polymorphic logic gates like RFETs realize multiple functionalities in the same cell, whereby the actual functionality is chosen by configuring a control signal. These polymorphic gates can replace the corresponding gates of the original circuit that directly impact on the primary outputs to insert efficient key gates. By this, the high performance overhead of CMOS-based techniques are avoided [1].

Various RFET-based cells are available that implement different functionalities like the NAND/NOR- or the XOR/XNOR-RFET. An example of the RFET is visualized in Figure 1, which can be configured as a NAND or NOR gate depending on the control signal serving as a key bit.

III. Quality Assessment Framework

This section describes the generation of the SAT-based model for the quality assessment of RFET-based protection mechanisms that have been introduced into a circuit, yielding the *Circuit-under-Assessment* (CuA).

At first, a miter circuit is generated from the CuA while considering the a-priori known *correct* key $\mathcal{K}$ yielding the SAT instance $\Phi_{\mathcal{K}}$ and a corrupting key $\widehat{\mathcal{K}}$ yielding $\Phi_{\widehat{\mathcal{K}}}$. The basic principle of this construction is given in Sub-figure 2a. The CuA is unrolled for N clock cycles since sequential elements – meaning *Flip-Flops* (FFs) – have to be considered for an exact assessment in terms of sequential circuits' unrolling [12]. Here, the value N has to be adjusted with respect to the CuA characteristics. Furthermore, 0 is assumed as the initialization value for all FFs in cycle $n = 1$. For keeping the resulting miter model small, only the relevant combinatorial logic, i.e., the transitive fan-in, for the corresponding output is calculated, which is done for all cycles n (with $1 \leq n \leq N$). The FFs are modeled as *Pseudo Primary Inputs* (PPIs) in cycle $n+1$ and are connected to the corresponding *Pseudo Primary Outputs* (PPOs) of the previous cycle n. Consequently, the introduced miter compares the unrolled $\Phi_{\mathcal{K}}$ with the unrolled $\Phi_{\widehat{\mathcal{K}}}$, i.e., considering any corrupting key $\widehat{\mathcal{K}} \neq \mathcal{K}$. More precisely, both the state – defined by the stored FFs' values – and the primary output values are being compared. Furthermore, the primary inputs are equally driven for both unrolled instances (of the CuA) and are kept constant during the unrolling. The entire model is stored as one SAT instance and processed by a state-of-the-art SAT solver. Sub-figure 2b presents the general approach about how corrupting keys are being evaluated. If a satisfiable solution is determined, a corrupting key has been detected that yields a functional equivalent behavior of the CuA, forming a potential security breach.

For a qualitative assessment of the discovered security threat, every determined corrupting key (if any) is evaluated against the number of possible stimuli leading to this breach. More precisely, the individual corrupting key is enforced $\Phi_{\widehat{\mathcal{K}}}$, and the solving process is repeated iteratively. After each iteration, the problematic stimuli are extracted for later analysis and excluded from the further search process of the SAT solver. The process ends as soon as no further stimuli could have been determined – the SAT instance gets unsatisfiable – or a user-defined limit has been exceeded.

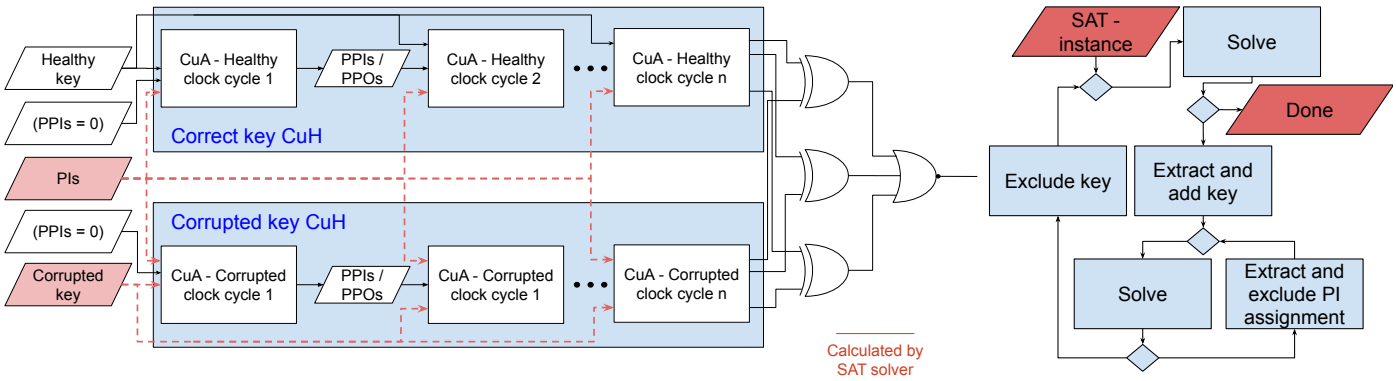
IV. Experimental Evaluation

This section describes the experimental evaluation of the proposed quality assessment framework for RFET-based logic locking protection mechanisms and discusses the obtained results, i.e., the number of determined corrupting keys with respect to the stimuli.

All experiments have been conducted on an *AMD 4750U* processor with 32 GB system memory. The proposed technique has been solely implemented in C++. For an initial evaluation, different benchmark circuits of the *ITC'99* benchmark suite are considered. For each of the considered circuits, ten of the *NOR* and *NAND* gates have been randomly replaced by RFETs, while the functional behavior is retained if the correct key is applied.

An experimental evaluation has shown that ten RFETs can be considered a sufficient number of key gates to create diverse logic locking structures with minor or mayor weaknesses, depending on the considered circuit. Consequently, each circuit holds ten control signals resulting in $2^{10} = 1,024$ possible keys. Furthermore, a maximum of 1,024 stimuli (per corrupting key $\widehat{\mathcal{K}}$) is captured – for limiting the computation time per corrupting key – if the CuA behaves functionally correct even though a corrupting key is applied.

Table I shows the absolute number of identified corrupting keys, the minimum, the average, and the maximum number of corrupting stimuli per key. For the conducted experiments, each of the specified circuits has been unrolled for five clock cycles. The results show various corrupting keys per circuit with differing numbers of PI assignments with equivalent behavior. Considering circuits like the b11, b12 and b15, every incorrect key is corrupting considerably large number of stimuli. Consequently, those weak logic locking structures are easy to detect, even using approaches with a limited amount of observed key and stimuli pairs. However, for example, the circuit b09 has one key out of 1,024 incorrect keys, that corrupt all stimuli of the circuit. Meaning this incorrect key behaves equivalent to the correct key and, hence, is a maximal security threat. Since it's the only key with any equivalent behavior, this security breach is hard to unveil. In comparison to a Hamming distance-based technique, the proposed formal approach considers all possible stimuli and key combinations and, hence, is complete. When invoking other approaches, that rely on the observation of limited stimuli, it is likely that worst-case keys $\widehat{\mathcal{K}}_w$ (with the maximum equivalent behavior per key)



(a) Resulting abstract model of the CuA

(b) SAT-based key- and stimuli extraction

Figure 2: Quality assessment technique

TABLE I: Results

circuit	$\#\{\widehat{\mathcal{K}}\}$	#stimuli		
		minimum	average	maximum
b05	3	1	1,333	2
b06	35	1	1.6	4
b07	7	2	2	2
b08	3	256	256	256
b09	1	2	2	2
b10	63	896	991	1,024
b11	1,023	127	127	128
b12	1,023	16	16	32
b13	1	512	512	512
b14	0	0	0	0
b15	1,023	1,024	1,024	1,024
b20	63	1,024	1,024	1,024
b21	31	1,024	1,024	1,024

remains undetermined. Furthermore, the keys of sequential circuits were assessed in an RFET-based logic locking approach for the first time.

V. Conclusions

This paper presented a novel method for assessing the quality of RFET-based logic locking protection systems by heavily orchestrating formal techniques and efficient data structures like SAT and AIGs. In the end, the proposed framework allows determining corrupting keys and evaluates their threat to the protection system. In contrast to other approaches, the assessment is conducted exactly considering the fully functional state space of the circuit. Future work will enhance the SAT-based model by incorporating Pseudo-Boolean Optimization techniques and investigates a compositional approach allowing for processing even larger industrial-sized designs.

VI. Acknowledgements

This work was financially supported by the German Federal Ministry of Education and Research BMBF under the framework of VE-CirroStrato and the AI initiative of the Free Hanseatic City of Bremen.

References

[1] Q. Alasad, J.-S. Yuan, and Y. Bi, "Logic locking using hybrid CMOS and emerging SiNW FETs," *Electronics*, vol. 6, no. 3, 2017.

[2] S. Rai, S. Srinivasa, P. Cadareanu, X. Yin, X. S. Hu, P.-E. Gaillardon, V. Narayanan, and A. Kumar, "Emerging reconfigurable nanotechnologies: Can they support future electronics?" in *IEEE/ACM International Conference on CAD*, 2018.

[3] Q. Alasad and J. Yuan, "Logic obfuscation against IC reverse engineering attacks using PLGs," in *IEEE International Conference on Computer Design*, 2017, pp. 341–344.

[4] Q. Alasad, J.-S. Yuan, and P. Subramanyan, "Strong logic obfuscation with low overhead against IC reverse engineering attacks," *IEEE Transaction on CAD of Integrated Circuits and Systems*, vol. 25, no. 4, 2020.

[5] J. A. Roy, F. Koushanfar, and I. L. Markov, "EPIC: Ending piracy of integrated circuits," in *Design, Automation and Test in Europe*, 2008, p. 1069–1074.

[6] J. Rajendran, Y. Pino, O. Sinanoglu, and R. Karri, "Security analysis of logic obfuscation," in *Design Automation Conference*, 2012, pp. 83–89.

[7] J. Rajendran, H. Zhang, C. Zhang, G. S. Rose, Y. Pino, O. Sinanoglu, and R. Karri, "Fault analysis-based logic encryption," *IEEE Transaction on Comp.*, vol. 64, no. 2, pp. 410–424, 2015.

[8] Q. Alasad, Y. Bi, and J.-S. Yuan, "E_2LEMI:Energy-efficient logic encryption using multiplexer insertion," *Electronics*, vol. 6, p. 16, 02 2017.

[9] J. B. Wendt and M. Potkonjak, "Hardware obfuscation using PUF-based logic," in *IEEE/ACM International Conference on CAD*, 2014, pp. 270–271.

[10] S. M. Plaza and I. L. Markov, "Solving the third-shift problem in IC piracy with test-aware logic locking," *IEEE Transaction on CAD of Integrated Circuits and Systems*, vol. 34, no. 6, pp. 961–971, 2015.

[11] Y. Lee and N. Touba, "Improving logic obfuscation via logic cone analysis," 05 2015.

[12] R. Arora and M. Hsiao, "Enhancing SAT-based bounded model checking using sequential logic implications," in *International Conference on VLSI Design*, 2004, pp. 784–787.