# A Novel Structural Attack Against SAT-Resilient Logic Locking Techniques

**Levent Aksoy**†, Muhammad Yasin‡ and Samuel Pagliarini†

†Tallinn University of Technology, Tallinn, Estonia

‡National University of Sciences and Technology, Islamabad, Pakistan



SAFEST Workshop, Tallinn, Estonia, 19th of June 2023

# Outline

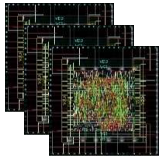
- Introduction
- Background
- Proposed Attack
- Experimental Results
- Conclusions

# Introduction

- Many security threats have been causing damages to the semiconductor industry

Piracy

Overbuilding



Modification

Reverse Engineering

- Against such threats, many efficient methods have been introduced

Digital Watermarking

Fingerprinting

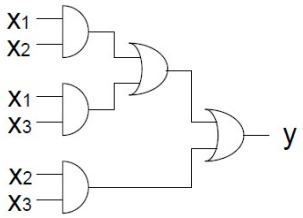
Hardware Metering

Logic Locking

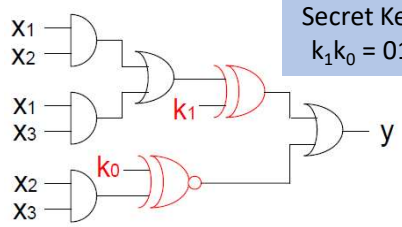
# Introduction

- Logic locking is a promising solution to a wide range of security threats
  - adds additional gates into the original design with key bits

Original Design




Locked Design




Secret Key
$k_1k_0 = 01$

# Background

## Conventional Logic Locking in the IC Design Flow

Design House

Foundry

Market

Specification → Behavioral Design → Logic Synthesis → Logic Locking → Locked Netlist → Physical Synthesis → Layout → Fabrication → Test & Packaging → Locked IC → Key Activation → Functional IC

Design House: Trusted | Foundry: Untrusted | Market: Untrusted

## Threat Models

Oracle-less (OL)

Oracle-guided (OG)

Locked Netlist →

← Locked Netlist

← Oracle

# Background

- Pre-SAT locking techniques
  - focus on output corruption and hardware complexity
    - random logic locking (RLL) [1]
- SAT-based attack [2]
  - iteratively finds distinguishing input patterns (DIPs) that eliminate wrong key(s)
- Post-SAT locking techniques
  - increase the run-time of an iteration and the number of iterations
    - SARLock [3], Anti-SAT [4], CAS-Lock [5], Gen-Anti-SAT [6], TTLock [7], SFLL-REM [8], and CAC [9]


[1] J. A. Roy, F. Koushanfar and I. L. Markov, "Ending Piracy of Integrated Circuits," in DATE, pp. 1069-1074, 2008.
[2] P. Subramanyan, S. Ray and S. Malik, "Evaluating the Security of Logic Encryption Algorithms," in HOST, pp. 137-143, 2015.
[3] M. Yasin, B. Mazumdar, J. Rajendran, and O. Sinanoglu, "SARLock: SAT Attack Resistant Logic Locking," in HOST, 2016, pp. 236–241.
[4] Y. Xie and A. Srivastava, "Anti-SAT: Mitigating SAT Attack on Logic Locking," IEEE TCAD, vol. 38, no. 2, pp. 199–207, 2019.
[5] B. Shakya, X. Xu, M. Tehranipoor and D. Forte, Domenic, "CAS-Lock: A Security-Corruptibility Trade-off Resilient Logic Locking Scheme," IACR TCHES, vol. 2020, no. 1, 175-202, 2019.
[6] J. Zhou and X. Zhang, "Generalized SAT-Attack-Resistant Logic Locking," IEEE TIFS, vol. 16, pp. 2581–2592, 2021.
[7] M. Yasin, A. Sengupta, M. T. Nabeel, M. Ashraf, J. V. Rajendran, and O. Sinanoglu,"Provably-Secure Logic Locking: From Theory To Practice," in CCS, pp. 1601-1618, 2017.
[8] A. Sengupta, M. Nabeel, N. Limaye, M. Ashraf, and O. Sinanoglu, "Truly Stripping Functionality for Logic Locking: A Fault-Based Perspective," IEEE TCAD, vol. 39, no. 12, pp. 4439–4452, 2020.
[9] K. Shamsi, T. Meade, M. Li, D. Z. Pan, and Y. Jin, "On the Approximation Resiliency of Logic Locking and IC Camouflaging Schemes," IEEE TIFS, vol. 14, no. 2, pp. 347–359, 2019.

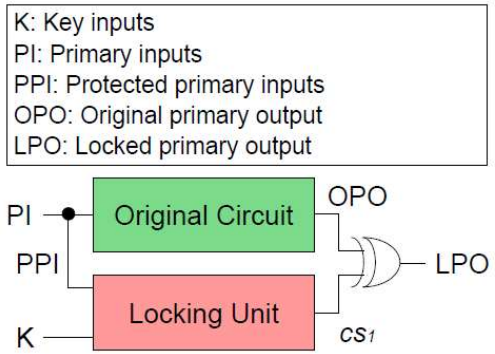
# Background

## One-point Function

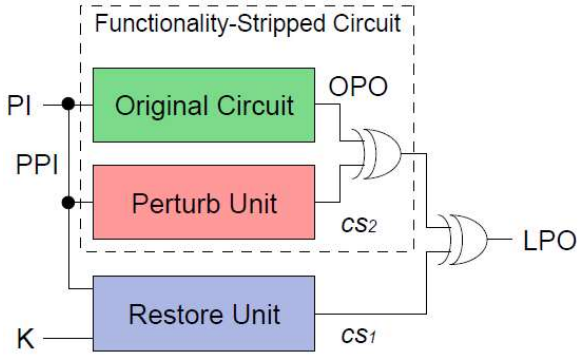
| $x_3$ | $x_2$ | $x_1$ | $K^7$ | $K^6$ | $K^5$ | $K^4$ | $K^3$ | $K^2$ | $K^1$ | $K^0$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$K^i = k_3k_2k_1 = (i)_{bin}$, $0 \le i \le 7$

## Single Flip Locking Techniques (SFLTs) [3-6]



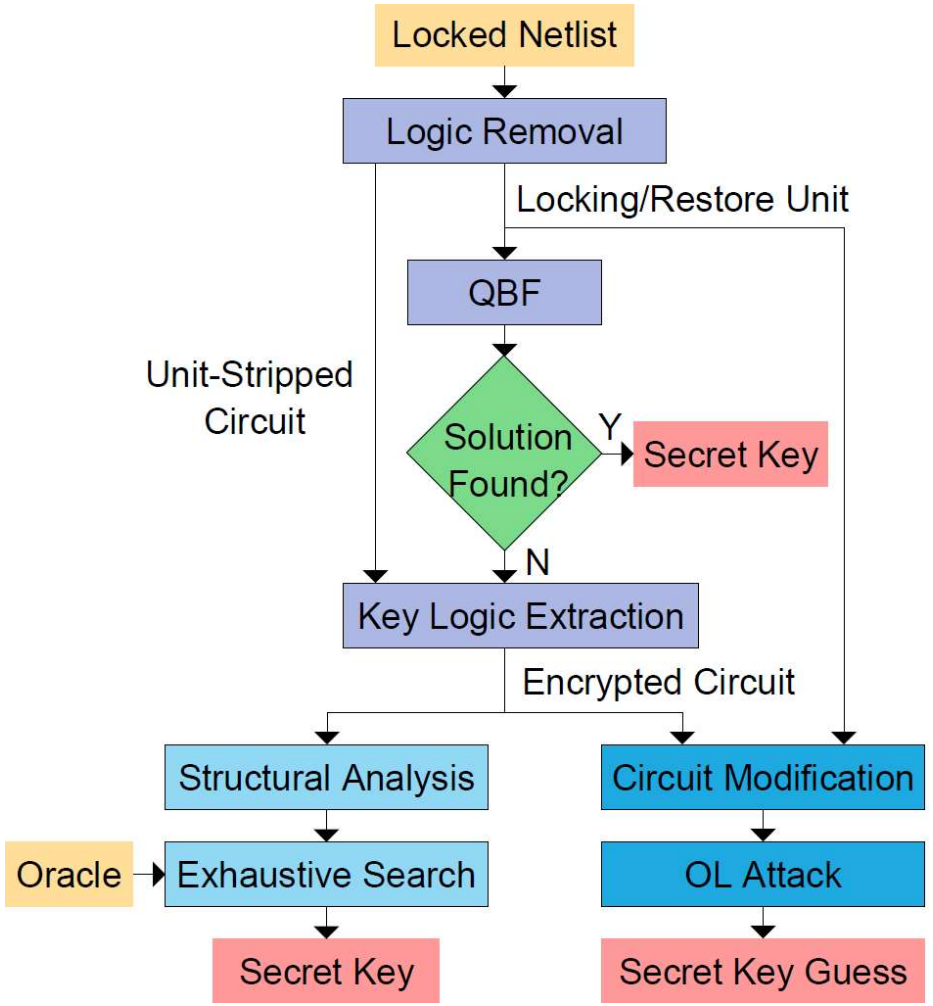

## Double Flip Locking Techniques (DFLTs) [7-9]



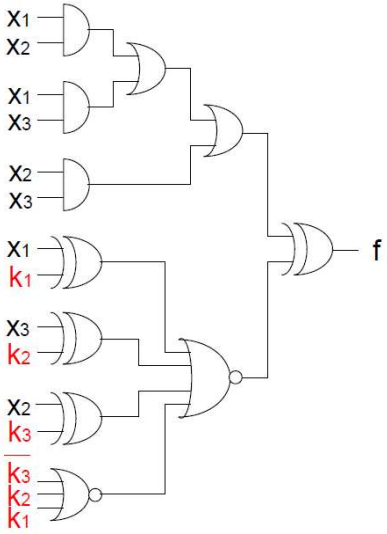
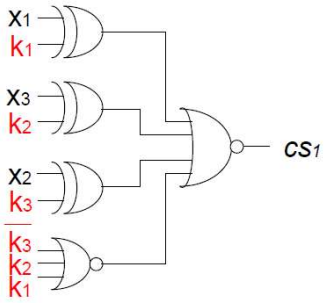
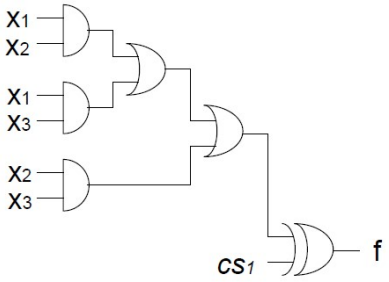
# Background

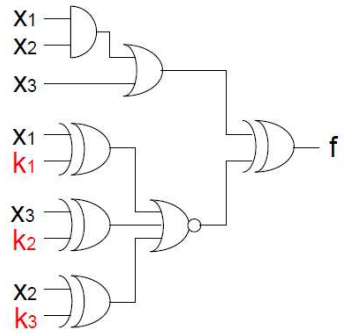
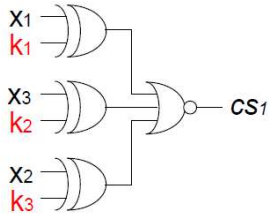
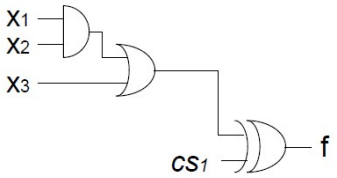
- A Boolean logic function, $\varphi: \mathrm{B}^n \rightarrow \mathrm{B}$, where $\mathrm{B} = \{0,1\}$ over $n$ variables $x_1, \ldots, x_n$
  - maps each truth assignment to 0 or 1
- The logic function $\varphi$ in the conjunctive normal form (CNF)
  - is a *conjunction* of clauses, $c_1, \ldots, c_m$, where a clause is a *disjunction* of literals $c_i = l_1 + l_2 + \ldots l_j$, $i \leq m$, $j \leq n$, where a literal is either a variable or its complement
  - $\varphi = (x_1 + x_2 + \overline{x_3})(\overline{x_1} + x_3)(\overline{x_2} + x_3)$
- The satisfiability (SAT) problem
  - is to find an assignment to the variables of a Boolean function $\varphi$ in CNF that makes $\varphi$ to be equal to 1 or prove that $\varphi$ is equal to 0
  - $\varphi$ is *satisfiable* with $x_3x_2x_1 = 111$
- The quantified Boolean formula (QBF) problem
  - is the generalization of the SAT problem including existential ($\exists$) and universal ($\forall$) quantifiers

# The Proposed Attack



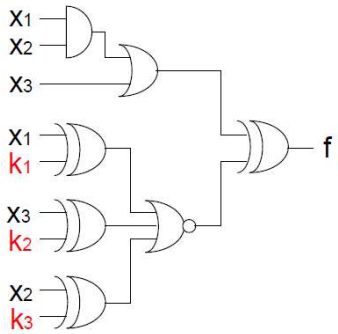
# Logic Removal and Key Logic Extraction

| Majority Circuit Locked by SARLock [3] | Locking Unit | Unit-Stripped Circuit | PPI-K Relation | QBF |
|---|---|---|---|---|
|  |  |  | $x_1 - k_1$<br>$x_2 - k_3$<br>$x_3 - k_2$ | $\exists k_1k_2k_3, \forall x_1x_2x_3, cs_1=1$<br>No Solution<br>$\exists k_1k_2k_3, \forall x_1x_2x_3, cs_1=0$<br>$k_3k_2k_1 = 100$ |

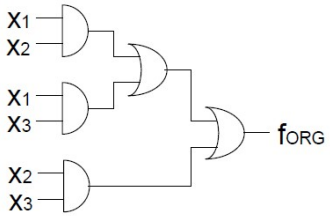
| Majority Circuit Locked by TTLock [7] | Restore Unit | Unit-Stripped Circuit | PPI-K Relation | QBF |
|---|---|---|---|---|
|  |  <br>Compares PPI and K |  | $x_1 - k_1$<br>$x_2 - k_3$<br>$x_3 - k_2$ | $\exists k_1k_2k_3, \forall x_1x_2x_3, cs_1=1$<br>No Solution<br>$\exists k_1k_2k_3, \forall x_1x_2x_3, cs_1=0$<br>No Solution |

# Structural Analysis and Exhaustive Search

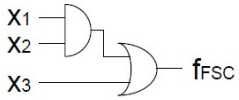
## Majority Circuit Locked by TTLock [7]




## Original Circuit

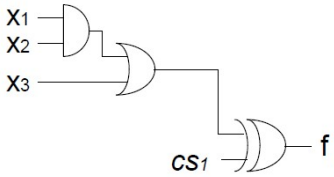



## Functionality Stripped Circuit

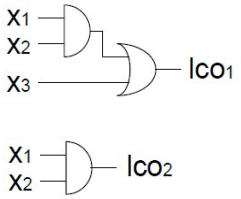



| $x_3$ | $x_2$ | $x_1$ | $f_{ORG}$ | $f_{FSC}$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 |
| 1 | 0 | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 |

## Encrypted Circuit




## Logic Cones




## Possible Set of PPI Values

| | |
|---|---|
| $lco_1(x_3,x_2,x_1) = 0$ | $lco_2(x_2,x_1) = 0$ |
| $x_3x_2x_1 = 000$ | $x_3x_2x_1 = X00$ |
| $lco_1(x_3,x_2,x_1) = 1$ | $lco_2(x_2,x_1) = 1$ |
| $x_3x_2x_1 = 100$ | $x_3x_2x_1 = X11$ |

## Exhaustive Search

$$f_{ORG}(x_3, x_2, x_1) \stackrel{?}{=} f(x_3, x_2, x_1, k_3, k_2, k_1)$$

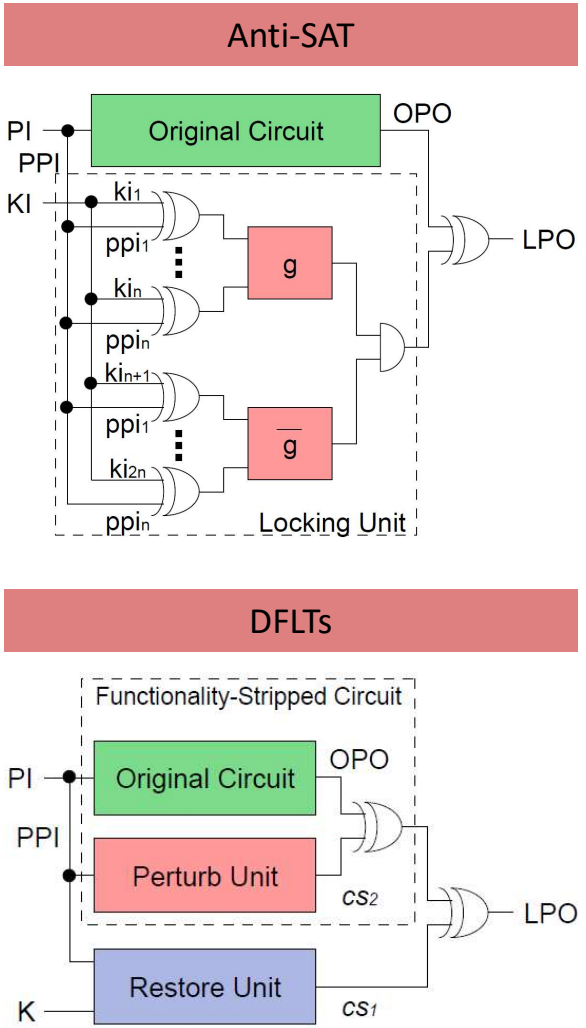
$x_1 - k_1$
$x_2 - k_3$
$x_3 - k_2$

$f_{ORG}(1,0,0) = f(1,0,0,0,1,0)$

$k_3k_2k_1 = 010$

# Circuit Modification and OL Attack

- For SFLTs, the locking unit is targeted
  - if each PPI is associated with double key inputs, e.g., Anti-SAT versions [6]
    - set PPIs to a constant logic value, e.g., 0
- For DFLTs, the encrypted circuit is targeted
  - replace the PPIs by associated key inputs
- Run SCOPE [10] on the target circuit and obtain the solution




[10] A. Alaql, M. M. Rahman, and S. Bhunia, "SCOPE: Synthesis-Based Constant Propagation Attack on Logic Locking," IEEE TVLSI, vol. 29, no. 8, pp. 1529–1542, 2021.

# Experimental Results

- Our tool KRATT was run on ISCAS'85, ITC'99, and HeLLO: CTF'22 circuits locked by
  - SFLTs including SARLock [3], Anti-SAT [4], CAS-Lock [5], and Gen-Anti-SAT [6]
  - DFLTs including TTLock [7] and CAC [9]
- It was compared to
  - OL attack SCOPE [10] and
  - OG attacks, SAT-based [2], Double DIP (DDIP) [11], and approximate SAT (AppSAT) [12]
- KRATT was developed in Perl and equipped with
  - QBF solver DepQBF [13]
    - its run-time was set to 60 seconds
  - SAT solver cryptominisat [14]
- The attacks were run on a computing server including 32 Intel Xeon processing units at 3.9 GHz with 128 GB memory

[11] Y. Shen and H. Zhou, "Double DIP: Re-Evaluating Security of Logic Encryption Algorithms," in GLSVLSI, 2017, pp. 179–184.
[12] K. Shamsi, M. Li, T. Meade, Z. Zhao, D. Z. Pan, and Y. Jin, "AppSAT: Approximately Deobfuscating Integrated Circuits," in HOST, 2017, pp. 95–100.
[13] Florian Lonsing. DepQBF Solver. [Online]. Available: https://github.com/lonsing/depqbf
[14] Mate Soos. Cryptominisat SAT Solver. [Online]. Available: https://github.com/msoos/cryptominisat

# Experimental Results

First Experiment Set – ISCAS'85 and ITC'99 Benchmarks

| Circuit | #inputs | #outputs | #gates | #key inputs |
|---|---|---|---|---|
| c2670 | 157 | 64 | 1193 | 64 |
| c5315 | 178 | 123 | 2307 | 64 |
| c6288 | 32 | 32 | 2416 | 32 |
| b14_C | 277 | 299 | 9768 | 128 |
| b15_C | 485 | 519 | 8367 | 128 |
| b20_C | 522 | 512 | 19683 | 128 |

# Experimental Results

## Results of OL Attacks on Locked Circuits

| Circuit | SFLT | | | | | | | | DFLT | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Anti-SAT | | | | SARLock | | | | CAC | | | | TTLock | | | |
| | SCOPE | | KRATT | | SCOPE | | KRATT | | SCOPE | | KRATT | | SCOPE | | KRATT | |
| | cdk/dk | CPU | cdk/dk | CPU | cdk/dk | CPU | cdk/dk | CPU | cdk/dk | CPU | cdk/dk | CPU | cdk/dk | CPU | cdk/dk | CPU |
| c2670 | 13/23 | 3.1 | 64/64 | 0.3 | 64/64 | 3.3 | 64/64 | 0.3 | 17/26 | 3.2 | 33/64 | 64.4 | 14/26 | 3.1 | 34/64 | 64.3 |
| c5315 | 13/22 | 3.9 | 64/64 | 0.6 | 64/64 | 3.9 | 64/64 | 0.5 | 12/19 | 3.9 | 33/64 | 64.6 | 16/31 | 4.0 | 34/64 | 64.5 |
| c6288 | 7/12 | 2.2 | 32/32 | 0.6 | 32/32 | 2.4 | 32/32 | 0.7 | 11/18 | 2.2 | 18/32 | 64.0 | 9/14 | 2.2 | 20/32 | 63.0 |
| b14_C | 32/55 | 15.1 | 128/128 | 4.6 | 128/128 | 15.5 | 128/128 | 10.1 | 39/71 | 15.0 | 67/128 | 74.6 | 35/59 | 14.8 | 70/128 | 74.4 |
| b15_C | 22/38 | 20.0 | 128/128 | 9.0 | 128/128 | 20.4 | 128/128 | 11.9 | 18/35 | 21.4 | 64/128 | 79.5 | 43/70 | 20.2 | 68/128 | 78.7 |
| b20_C | 24/46 | 25.8 | 128/128 | 13.6 | 128/128 | 26.2 | 128/128 | 16.9 | 30/54 | 26.1 | 58/102 | 79.3 | 24/46 | 26.1 | 68/128 | 82.3 |

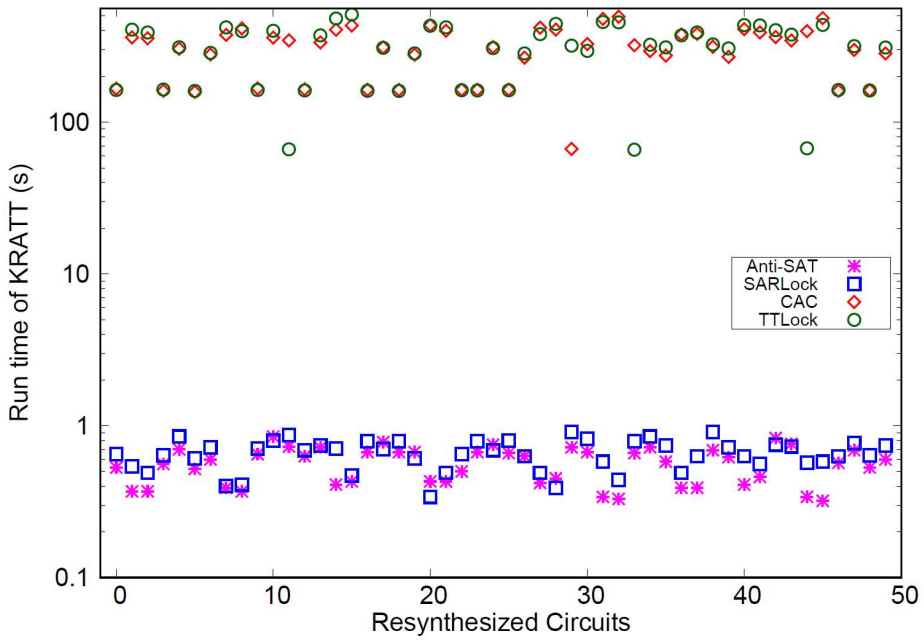
# Experimental Results

Results of OG Attacks on Locked Circuits

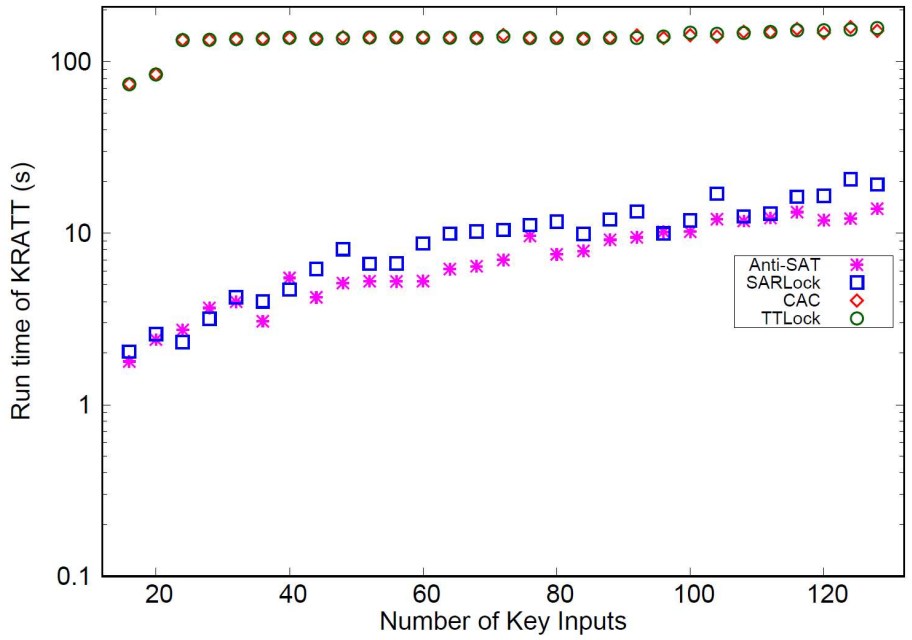
| Circuit | SFLT | | | | | | | | DFLT | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Anti-SAT | | | | SARLock | | | | CAC | | | | TTLock | | | |
| | SAT | DDIP | AppSAT | KRATT | SAT | DDIP | AppSAT | KRATT | SAT | DDIP | AppSAT | KRATT | SAT | DDIP | AppSAT | KRATT |
| c2670 | OoT | OoT | OoT | 0.3 | OoT | OoT | OoT | 0.3 | OoT | OoT | OoT | 70.7 | OoT | OoT | OoT | 70.5 |
| c5315 | OoT | OoT | OoT | 0.6 | OoT | OoT | OoT | 0.4 | OoT | OoT | OoT | 73.3 | OoT | OoT | OoT | 75.9 |
| c6288 | OoT | OoT | OoT | 0.6 | OoT | OoT | OoT | 0.7 | OoT | OoT | OoT | 163.1 | OoT | OoT | OoT | 161.2 |
| b14_C | OoT | OoT | OoT | 4.5 | OoT | OoT | OoT | 10.7 | OoT | OoT | OoT | 114.9 | OoT | OoT | OoT | 112.8 |
| b15_C | OoT | OoT | OoT | 9.1 | OoT | OoT | OoT | 11.9 | OoT | OoT | OoT | 133.3 | OoT | OoT | OoT | 131.6 |
| b20_C | OoT | OoT | OoT | 13.7 | OoT | OoT | OoT | 16.9 | OoT | OoT | OoT | 128.0 | OoT | OoT | OoT | 138.7 |

# Experimental Results

Impact of Resynthesis on Run-Time of KRATT




Impact of #Key Inputs on Run-Time of KRATT



# Experimental Results

- Second experiment set [15]
  - six ITC'99 circuits locked by SARLock, Anti-SAT, CAS-Lock, Gen-Anti-SAT, TTLock, and CAC using 128 key inputs
  - each benchmark has 10 synthesized circuits
  - a total of 360 locked circuits
- KRATT found the secret key of designs locked by SFLTs and DFLTs under the OL and OG threat model, respectively

Results of OL Attacks on Circuits Locked by Gen-Anti-SAT

| Circuit | SCOPE | | KRATT | |
|---|---|---|---|---|
| | cdk/dk | CPU | cdk/dk | CPU |
| b14_C | 9/12 | 14.3 | 127/127 | 106.3 |
| b15_C | 0/0 | 19.5 | 128/128 | 137.2 |
| b17_C | 0/0 | 51.5 | 128/128 | 533.5 |
| b20_C | 4/4 | 25.0 | 128/128 | 170.2 |
| b21_C | 0/0 | 24.8 | 128/128 | 173.4 |
| b22_C | 0/0 | 34.4 | 128/128 | 261.9 |

[15] Satwik Patnaik. Valkyrie. [Online]. Available: https://github.com/LL-Tools/Valkyrie

# Experimental Results

| Third Experiment Set – HeLLO: CTF'22 | | | | |
|---|---|---|---|---|
| **Circuit** | **#inputs** | **#outputs** | **#gates** | **#key inputs** |
| final_v1 | 767 | 755 | 17144 | 87 |
| final_v2 | 1452 | 1445 | 27440 | 47 |
| final_v3 | 522 | 1 | 93 | 29 |

# Experimental Results

| Results of OL and OG Attacks | | | | | | |
|---|---|---|---|---|---|---|
| **Circuit** | **OL Attacks** | | | | **OG Attacks** | |
| | SCOPE | | KRATT | | SAT | KRATT |
| | cdk/dk | CPU | cdk/dk | CPU | | |
| final_v1 | 0/0 | 261.9 | 73/87 | 194.6 | 1117.0 | 350.2 |
| final_v2 | 0/0 | 39.7 | 34/46 | 99.4 | OoT | 2186.5 |
| final_v3 | 0/0 | 1.9 | 25/29 | 62.6 | 20448.6 | 63.9 |

# Conclusions

- This work presented a removal and structural analysis attack KRATT against SAT-resilient logic locking techniques
  - it uses a quantified Boolean formulation to find the secret key of SFLTs
  - it uses a structural analysis and exhaustive search method to find the secret key of DFLTs
  - it can successfully handle the locked circuits under the OL threat model
- In future work, we plan to
  - extend its capabilities to break other logic locking techniques
    - multi flip and compound
  - propose a defense mechanism that thwarts structural attacks

# Questions

**THANKS for YOUR ATTENTION**

Contact: Levent Aksoy
E-mail: levent.aksoy@taltech.ee