Dhirubhai Ambani Institute of Information and Communication Technology
Gandhinagar, Gujarat 382007

Testability Based Victim Signal Identification in Hardware Attacks

By

Nikita Shah
Rohan Mistry

Under the guidance of

Dr. Sreeja Rajendran

I. Abstract:

Malicious changes in an IC design are commonly referred to as hardware Trojans. The presence of multiple entities in the VLSI design cycle has made HT detection extremely difficult. This article describes a novel method for detecting HT at the gate level of abstraction. The algorithm detects malicious nets in the circuit by comparing the netlists of the genuine (design) and Trojan-inserted circuits.This method uses testability analysis as a metric to separate malicious nets in the compromised circuit netlist. The testability parameters, such as controllability and observability of the nets, are determined using the fabricated IC's netlist obtained through reverse engineering and the original circuit's design netlist. Based on the variation in testability metrics of a signal in the original circuit, the algorithm identifies the malicious nodes inserted into the original circuit. Using the list of malicious gates and compromised circuit netlist, it is possible to identify the Trojan nets inserted by the adversary.

II. Introduction:

In recent years, there has been an increase in research into hardware security. This is primarily due to the fact that multiple entities are involved in the VLSI design process. As a result, several stages of the VLSI design cycle have been designated as untrusted. Intellectual property (IP) vendors, design houses, and fabrication units all play important roles in the development of a secure chip. Therefore, the possibility of tampering with the design cannot be completely eliminated. This is critical because undetected hardware Trojans (HT) have a negative impact on circuit reliability and expected lifetime.Prevention of HT insertion as well as HT detection are equally important for chips employed in safety-critical applications, such as aeronautics, military, etc. The hardware security and trust community has developed various preventive mechanisms known as Design for Trust which rely on modification of the current IC design flow. However, these methods face many limitations like area overhead, the requirement of a large number of test vectors for improving test coverage, etc. Apart from these, the security of the added protection units is also a concern. Some of these techniques are based on the assumption that attackers will only use rare events for triggering the intended action and therefore are oblivious to “always on” Trojans. As a result, research continues in the quest for a full proof mechanism to thwart hardware attacks. The purpose of this article is to detect functional Trojans inserted during the fabrication process. The work in this article focuses on an offline HT detection mechanism that uses reverse engineered netlists and operates at the gate level abstraction. The method does not require Trojan activation. The design netlist of logic circuits serves as a reference.

The project evaluates both combinational and sequential testability measures in VLSI circuits. Testability assesses the ease with which faults in a combinational circuit can be detected, whereas sequential testability assesses the effectiveness of testing methods for sequential elements such as flip-flops. The project's goal was to identify critical testing points and optimize testing strategies for better fault coverage by analyzing controllability and observability metrics. The goal was to develop techniques that ensure thorough testing of both the combinational and sequential components of the circuit, thereby improving its reliability and performance.

III. Hardware Attacks:

Any attacks that result in a deviation in the functionality of a chip from the expected output can be broadly termed as hardware attacks. These attacks can result in leakage of information through an implanted backdoor, damage the chip or even result in Denial of Service (DoS). Hardware attacks can be launched at various levels of abstraction like circuit, gate and RTL. Attacks at the circuit level are initiated through changes in manufacturing process parameters like doping, channel width etc. At gate level, logic gates are added or deleted to launch attacks. Hardware attacks at RTL level can be introduced by the addition of states to the finite state machine. The motive behind hardware attacks can be economic reasons or even for intelligence purposes. Integrated circuits are the backbone of all safety critical systems. Launching hardware attacks on such sensitive systems can result in catastrophic effects. To protect digital systems from these hardware attacks, designers should have a vast knowledge about security vulnerabilities and the measures to overcome them. To build a secure system, the designers should be aware of the possible attacks to watch out for and about how to protect the system from such malign attacks. Analyzing the susceptibility of a design to malicious intrusions is the key to developing a secure system. Thorough analysis of a circuit design is essential to determine its susceptibility to Hardware Trojan insertion. Nodes with poor testability have been identified as the potential sites for HT insertionAnd therefore testability analysis is an integral step to assess the vulnerability of circuit designs.

Data Attacks		Design Attacks	Functionality Attacks
Invasive	Non-invasive	Reverse Engineering (RE)	Hardware Trojans (addition/ deletion of gates)
RE	Side Channel attacks		
Micro probing	Hardware Trojans		

Table 1. Classification of Hardware Attacks on IC intended actions

Testability Measures:

An attempt to quantify testability by Goldstein and Grason resulted in two testability measures:

- **Controllability:**
  The measure of difficulty of setting a particular logic signal to 0 or 1.
- **Observability:**
  The measure of difficulty in propagating the value at a particular signal to primary output (i.e. difficulty of observing the logic 0 or 1 at any signal).

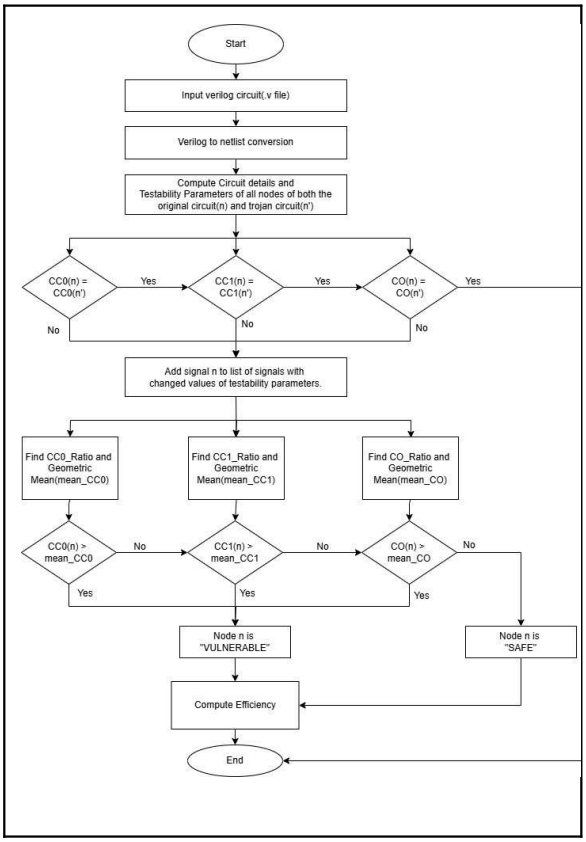
Out of several methods available to find testability measures, here we follow:

SCOAP: (Scandia Controllability Observability Analysis Program)

IV. Problem Statement:

In the realm of hardware security, identifying victim signals susceptible to attacks is paramount for ensuring robust protection against malicious exploits. The challenge lies in effectively pinpointing these vulnerable signals within complex integrated circuits and systems, where traditional testing methodologies may fall short. This necessitates the development of innovative testability-driven approaches that can accurately identify potential victim signals by leveraging hardware testing and analysis techniques. Addressing this issue not only enhances the security of hardware systems but also minimizes the risk of unauthorized data access and manipulation, thereby fortifying the overall integrity and reliability of electronic devices.

V. Project Workflow:



VI. Example:

The HT detection algorithm has been applied to the c17 ISCAS benchmark circuit. Three different combinational trojans and 2 hybrid trojans have been inserted and the algorithm is tested. The algorithm has accurately detected the malicious nets in all the cases. The combinational trojans 1–3 inserted into the c-17 benchmark the circuit consists of two gates, namely, NOR and XOR. The same trojan circuit with different input signals to the NOR gate defining the trigger conditions. The Trojan has been designed thereby to complement the output signal N23 with different trigger conditions set by the input nets to the NOR gate. The c17 benchmark circuit has been chosen for simplicity to explain.

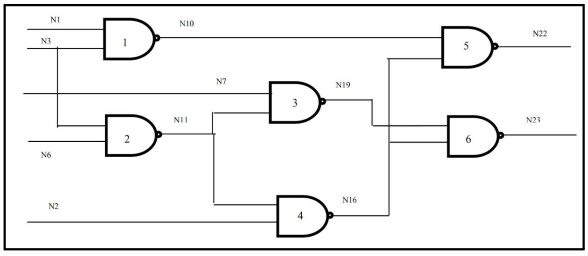


Fig 1. c17 benchmark circuit

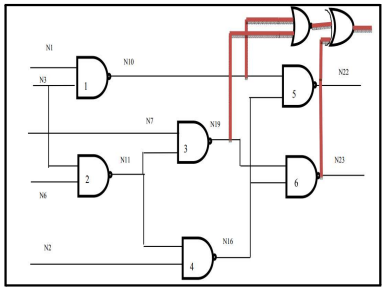


Fig 2. c17 benchmark circuit inserted with Combinational Trojan-

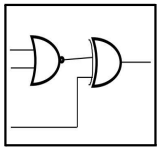


Fig 3. Hardware Trojan Circuit

Signal	Trojan Free	Trojan-1	Trojan-2	Trojan-3
N10	3/2/3	3/2/3	3/2/3	3/2/3
N11	3/2/5	3/2/5	3/2/5	3/2/5
N16	4/2/3	4/2/3	4/2/3	4/2/3
N19	4/2/3	4/2/7	4/2/7	4/2/7
N22	5/4/0	5/4/0	5/4/0	5/4/0
N23	5/5/0	5/5/4	5/5/4	5/5/4

Table 2. Comparison of Testability Parameters of c-17 benchmark circuit with and without Trojans

VII. Results and Discussion:

The testability parameters have been computed for various benchmark circuits and are listed in the table. The functionality of the benchmark circuits and the total number of logic gates in the circuit are given in the table. It gives a picture of the complexity of the logic circuit. The maximum and minimum values for the controllability parameters CC0(Combinational Controllability of 0) and CC1 (Combinational Controllability of 1) as well as the values for observability are tabulated. The maximum and minimum value of every parameter is indicated in the table so that the reader gains an understanding of how large the variation is within a particular logic circuit and also its level of complexity. Ideally it is good if the CC/CO (Combinational Observability) values are closer to the theoretical minimum which is 1/0. Though SCOAP defines the method to determine testability parameters of the circuit, it does not provide a range of values to classify the nodes in a logic circuit as safe and susceptible ones. Classification of the nodes on the basis of testability parameters will help designers to identify critical nodes/regions in the logic circuit which need to be attended to for safeguarding against malicious attacks.

	c2670	c3540	c5315	c6288
No. of gates	686	1143	2307	2416
No. of Nodes	1527	2318	4792	4864
Execution Time(ns)	2993000	4143000	16649000	2193000
CC0/CC1/CO	219/246/308	322/273/399	223/266/340	307/419/749
No. of trojan circuits analyzed	101	101	111	111

Table 3. Benchmark Circuit details (Combinational)

	s1423	s13207	s15850	s35932
No. of gates	701	4121	4447	12022
No. of Flip Flops	74	625	513	1728
No. of Nodes	1634	9975	10414	29010
Execution Time(ns)	27831000	1.18E+08	1.59E+08	2.38E+08
CC0/CC1	224/233	1772/1063	1065/785	48/44
SC0/SC1	25/26	171/104	41/34	5/4
No. of trojan circuits analyzed	92	151	151	104

Table 4. Benchmark Circuit Details (Sequential)

Circuits		CC0	CC1	CO
c2670	**c2670**	219	246	308
	c2670_T091	219	246	361
	c2670_T093	296	300	308
c3540	**c3540**	322	273	399
	c3540_T024	327	278	404
	c3540_T005	322	273	816
c5315	**c5315**	223	266	340
	c5315_T065	195	266	442
	c5315_T043	412	399	415
c6288	**c6288**	307	419	749
	c6288_T021	561	666	749
	c6288_T024	1018	1025	763

Table 5. Variability In Testability Parameters of Trust-Hub Combinational Circuits Due to Trojan Insertion

		CC0	CC1	SC0	SC1
s1423	**s1423**	224	233	25	26
	s1423_T614	1141	27858	10032	27856
	s1423_T426	779	18569	6689	18567
s13207	**s13207**	1772	1063	171	104
	s13207_T489	1825	8499	3065	8497
	s13207_T478	2208	10721	3863	10719
s15850	**s15850**	1065	785	41	34
	s15850_T476	1074	800	573	570
	s15850_T202	1078	805	543	327
s35932	**s35932**	48	44	5	4
	s35932_T439	277	6075	2185	6073
	s35932_T435	306	6597	2379	6595

Table 6. Variability In Testability Parameters of Trust-Hub Sequential Circuits Due to Trojan Insertion

VIII. Advantages:

i) Can detect the presence of any digital Trojan inserted in the foundry.

ii) Can be performed offline and therefore eliminates the need for test vectors to detect the presence of HT.

IX. Conclusion:

This report presents a novel method for the detection of HTs inserted in a chip during the fabrication phase. The proposed method employs design netlists of the original circuit as well as the reverse-engineered netlist of the fabricated chip to detect the presence of malicious elements in the original circuit. An efficient general purpose software is developed to compute the testability parameters of combinational and sequential digital circuits of any level of complexity. The computation of testability parameters of various benchmarked digital circuits has been carried out using the cpp code based testability analysis tool that has been developed. The tool is developed based on the Scandia Controllability Observability Analysis Program(SCOAP). It falls under the category of topology based testability analysis procedure. The tool eases the process of computation of controllability and observability values of various signals in the circuit. The testability parameters computed using the tools can be used to develop methods to classify the nodes of a logic circuit into safe and susceptible categories. This in turn will lead to secure digital system design for highly sensitive applications. The tool can also be used for HT detection through the process of reverse engineering. It is found that introduction of HTs results in a change in testability parameters of certain nodes in a circuit. Subsequently, the software can be used to identify the location of malicious nodes in the particular circuit.

The GitHub Repository provided contains all the code and resources related to this project.

X. References

[1] S. Rajendran and M. L. Regeena, "A Novel Algorithm for Hardware Trojan Detection Through Reverse Engineering," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 41, no. 4, pp. 1154-1166, April 2022, doi: 10.1109/TCAD.2021.3073855.
keywords: {Trojan horses;Registers;Logic gates;Hardware;Integrated circuits;Feature extraction;Security;Hardware Trojan (HT) detection;path retrace algorithm;reverse engineering (RE);testability analysis},

[2] Rajendran, Sreeja, and Mary Lourde. "Testability Analysis and its Application to Hardware Security." International Journal of Circuits and Electronics 5 (2020).