ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJIREEICE

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

SITES

Smart And Innovative Technologies In Engineering And Sciences

Gyan Ganga College of Technology

Vol. 4, Special Issue 4, November 2016

Hardware Trojan Classification scheme Theory and Performance Analysis of Detection Techniques

Akanksha Dixit

Asst Professor, Dept of Electronics and Communication, Gyan Ganga College of Technology, RGPV, Jabalpur, India

Abstract: Devices produced for security sensitive fields like military, banking or government applications often make use of integrated circuits. In these areas reliability and trust are of special importance and thus an in-depth evaluation of manufactured devices with respect to security vulnerabilities is an essential task. Most of these institutions have departments that can take care of the chip-design but they rely on third party companies to actually produce the semiconductor chips in their factories. This is a very common scenario these days, where a trend towards separation of chip-design and fabrication is clearly observable. This is due to the high costs to build and maintain state- of-the-art semiconductor factories, which only pays off when the factories are used to capacity. With outsourced chip fabrication, are security sensitive chip designs still trustworthy? Every employee of the manufacturing chain from design to package assembly might maliciously modify the hardware! Therefore there is an urgent need to employ detection techniques for so called "Trojan hardware".

Keywords: Hardware Trojan; partial reconfiguration; hardware security; Trojan triggering.

I. INTRODUCTION

Today's business is global and for this reason outsourcing tasks is a common method to increase companies revenues. A company can only be competitive if their products are more advanced (i.e. higher complexity) or cheaper (due to cost pressure) than comparable products from its competitors. That is why Systems-On-Chip (SoA) or other embedded hardware devices are produced abroad. But, outsourcing poses a serious threat, especially for government agencies. Typically threatened sectors are the military, finance, power or the political sector. For example, critical applications are access control systems or ATMs that rely on embedded hardware.

Therefore, techniques to detect hardware Trojans are in focus of IT- Security research.. If a Trojan is activated the functionality can be changed, the device can be destroyed or disabled, it can leak confidential information or tear down the security and safety. Trojans are stealthy, that means the precondition for activation is a very rare event.

II. CLASSIFICATION OF TROJANS

This section describes and illustrates what classes of Trojans exist. In literature malicious hardware implantations are called hardware Trojan horse (HTH), malicious circuit or malicious logic. A Trojan is completely characterized by its physical representation and its behavior. So, its characterization can be divided into three parts:

1. Physical representation
2. Activation phase (trigger)
3. Action phase (propagate payload)

Physical Characteristics: From the perspective of a malicious circuit designer there are several physical characteristics to plan (figure 1). One of this physical Trojan characteristics is the type. The type of a Trojan can be either functional or parametric. A Trojan is functional if the adversary adds or deletes any transistors or gates to the original chip design. The other kind of Trojan, the parametric Trojan, modifies the original circuitry, e.g. thinning of wires, weakening of Flip-Flops or transistors, subjecting the chip to radiation, or using Focused Ion-Beams (FIB) to reduce the reliability of a chip.

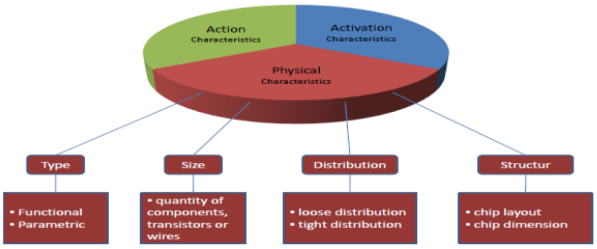


Figure 1: Classification of Trojans: Physical Aspects

Then, this kind of Trojan is called "parametric Trojan". Furthermore, an malicious designer has to define the size that is the next category. Furthermore, an malicious designer has to define the size that is the next category.

The size of a Trojan is its physical extension or the number of components it is made of. Trojan can consist of only few components, so the area is small where the malicious logic occupies the layout of the chip. In contrast this is called tight distribution.

Activation Characteristics: Figure 2 illustrates the activation characteristics. The typical Trojan is condition-based: It is triggered by sensors, internal logic states, a particular input pattern or an internal counter value. Condition-based Trojans are detectable with power traces to some degree when inactive. That is due to the leakage currents generated by the trigger or counter circuit activating the Trojan. Hardware Trojans can be triggered in different ways. A Trojan can be internally-activated, that means it monitors one or more signals inside the IC. The malicious circuitry could wait for a countdown logic an attacker added to the chip, so that the Trojan awakes after a specific timespan.

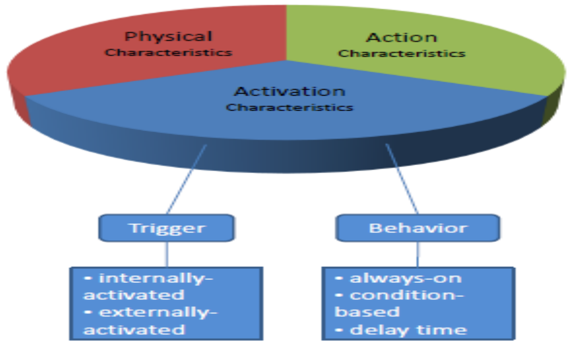


Figure 2: Classification of Trojans: Activation

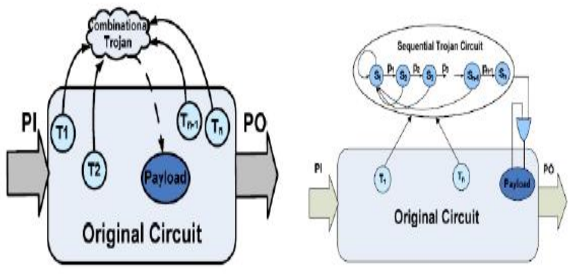


Figure 3: Illustration of a combinational and sequential Trojan (Source: [7])

The opposite is externally-activated. There can be malicious logic inside a chip that uses an antenna or other sensors the adversary can reach from outside the chip. For example a Trojan could be inside the control system of a cruising missile.

Action Characteristics: The effect of a Trojan can be seen from figure 4. It modifies the chip's function or changes the chip's parametric properties (e.g. provokes a process delay). Confidential information can also be transmitted to the adversary (transmit key information). In [8] the authors introduce the terms implicit Trojan and explicit Trojan. These terms are used to distinguish Trojans that induce small signal path delays or distinct signal path delays. The authors present a technique to measure signal path delays (see 3.4.4). Another important characteristic of these Trojans is that the chip consumes more power, so implicit Trojans are well detectable by detection methods that measure the power consumption like 3.4.1, 3.4.2 or 3.4.3. The reason for this high power consumption can be, for example, the activation of a radio transmitter or the destruction of the chip by extreme heat. The contrary term of implicit Trojan is explicit Trojan.

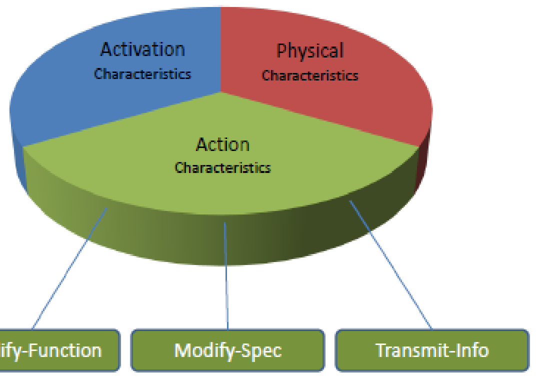


Figure 4: Classification of Trojans: Action

III. TECHNIQUES TO DETECT TROJANS

In this section we will discuss several detection techniques in some detail. In general, there are some challenges a detection technique has to master. An advanced detection technique should deal with the typical variation in the manufacturing process

- handle large and complex micro architectures
- detect small malicious Trojans
- detect any type of Trojans (classification)
- be non-destructive
- be scalable
- authenticate chips in a short time minimize the costs

IV. FUNCTIONAL TESTING

ATPG-based Trojan Detection
ATPG is acronym for automatic test pattern generation. This method is useful to detect parametric Trojans (malicious alterations that modify the circuitry): The netlist, i.e., the connectivity of an electronic design, is the

same with and without the parametric Trojan. This detection method stimulates the input ports of a chip and monitors the output to detect manufacturing faults. If the logic values of the output do not match the genuine pattern, then a defect or a Trojan could be found. A test pattern (input vector) is a digital stimulus. This stimulus is applied to the input pins of the chip, then the digital output is inspected (digital measurement). Typically, a number of input patterns that cause faulty circuit behavior are used. This set of input vectors is derived from a fault model that is a mathematical description of the fault behavior of the circuitry. It can be infeasible, depending on the complexity of the chip and the pin layout, to search for the activation vector of a condition-based functional Trojan. The complexity of this search can be calculated with the common methods (combinatory) to estimate the effort to search a secret key.

Built-In-Self-Test Techniques

A Built-In-Self-Test (BIST) is an additional functionality of a chip. The chip consists of components that provide the key functionality of the device. Furthermore, the chip design defines additional circuitry to monitor signals or detect defects. On the one hand these techniques are used to detect manufacturing errors, but on the other hand a malicious logic could be detected via manufacturing tests. BIST functionality can be successful to detect malicious logic. In the following two sections two different BIST techniques are depicted, but there exist more concepts.

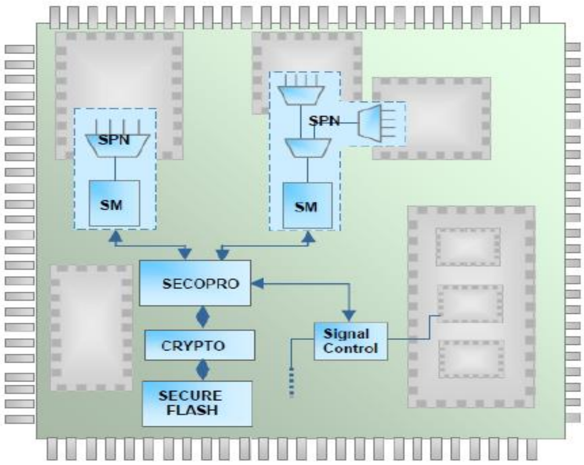
Three examples are: 1.Design for Debug: preferably used to detect manufacturing faults 2. Linear Feedback Shift Register Seeds: for more information. 3. The Design Tag System: advanced concept that is analog to DEFENSE logic

DEFENSE Logic

Another Built-In-Self-Test technique is called Design-For-Enabling-Security Chip with Additional DEFENSE Logic [13] logic (DEFENSE logic). The idea is to implement a robust inspection system in every chip analog to the immune systems in biological organisms This security logic consists of sensors, a central control unit and a signal control unit.This concept seems very expensive and complex. The design of the chip Contra takes more time because of the additional DEFENSE logic. The insertion of the DEFENSE logic is done at the design phase (at RTL (Register Transfer Language)), that the advisory is able to analyze the security system in detail. But there are some advantages: This built-in system is useful to monitor Pro the functional logic in use (on-line monitoring). It is reconfigurable because of the Security and Control Processor and the ash memory. Furthermore, it is non-destructive. Multiple and different sensors could detect functional or parametric Trojans simultaneously.

IC Fingerprinting methodology

This section deals with an effective, well-established detection methodology: power signals are used as a side channel. The advantage of this method is, that this operation detects Trojans that cannot be detected by functional analysis. This is a rough Sequence overview:



1. Generate genuine fingerprints from intensely tested chips, that is
(a) Select a few ICs out of one family of chips at random
(b) Analyze their functionality (e.g. stimulate I/O ports with input and measure the output)
(c) Measure and analyze their side channel signal
(d) Use destructive reverse engineering (see 3.1) to validate the originality of the ICs (in this step, the chips will be destroyed)
(e) Create (i.e. calculate) the fingerprint from the measured values
2. Authenticate the other chips of this family by comparing the genuine fingerprint and the measured fingerprint.

Trojan Detection with Power Traces Fingerprints are deduced from functional tests or side channels, but how is a Trojan detected in the measured values? The following method describes how a genuine fingerprint is generated from a set of values. A power trace is a discrete-time signal that consists of continuous values of power consumption. The power trace $r(t; I;C;M)$ is measured, while the chip I is processing the calculation C under influence of the power measurement M. This is the power trace of a genuine IC:

$r_G(t; I;C;M) = p(t;C) + n_p(t; I;C) + n_m(t;M)$

and this equation defines the power trace of a Trojan IC:

$r_T (t; I;C;M) = p(t;C) + n_p(t; I;C) + n_m(t;M) + \tau(t; I;C)$

I Integrated Circuit I

C actual Calculation on I during the measurement (typically, the same calculation during the analysis)

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJIREEICE

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

SITES

Smart And Innovative Technologies In Engineering And Sciences

Gyan Ganga College of Technology

Vol. 4, Special Issue 4, November 2016

M Power measurement on I
t Time points (discrete-time time-stamps)
r(t; I;C;M) Power trace
p(t;C) Mean power consumption
n_p(t; I;C) Process noise
n_m(t;M) Measurement noise
τ(t; I;C) Extra power leakage

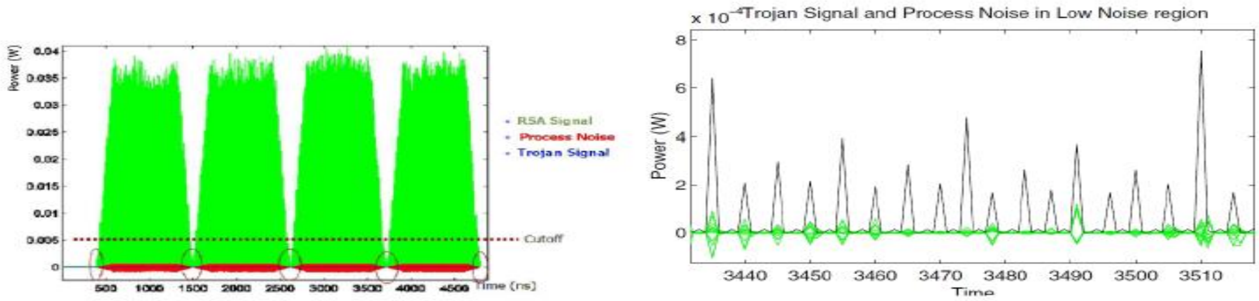


Figure 7: Left Figure: Illustration of low Process Regions, Right Figure: Process Noise Signal and Trojan IC Signal [10]

Table 1: Table of Characteristics of the Detection Methods.

Detection Technique	**resolution**	**Trojan Types**	**Comple -xity**	**Destruc -tive**	**Scal-ability**	**Costs**	**phase of IC life cycle**	**Features**
ATPG-based Trojan Detection	unknown	parametric Trojans	mid	no	high	low	test-time	
DEFENSE logic	sensor dependent	functional Trojans	high	no	high	high	run-time	reconfigurable multiple sensors, signal control unit
IC Fingerprinting	Down to 0.1% - 0.01% of total circuitry	functional Trojans, parametric Trojans	mid	no, except few IC's	mid	mid	Test-time	
Region-based Partitioning	Down to 1% of total gate count	functional Trojans	mid	no, except few IC's	mid	mid	Test-time	localisation
Path Delay Fingerprinting	Down to about1% of the chip	functional Trojans, parametric Trojans	mid	no, except few IC's	mid	mid	Test-time	

Advanced Detection with Reduced Frequency The detection results can be better if the frequency of the tested chip is reduced. Hence, the frequency must have an influence to the measured values. One common side channel is the power consumption of an IC. The total power consumption P of a chip is:

$$P = \left(\frac{1}{2}.\mathrm{C}.\mathrm{Vdd}^2 + \mathrm{Qse}.\mathrm{Vdd}\right).\mathrm{f}.\mathrm{N} + \mathrm{Ileak}.\mathrm{Vdd}$$

C; VDD and Qse technology dependent parameters
f Clock frequency
N Switching activity
I_{leak} Leakage current

Region-based Partition and Excitation Technique This approach is clearly represented in [2]. The method stimulates specific parts of the integrated circuitry to analyze the power consumption more precise. This

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJIREEICE
International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
SITES
Smart And Innovative Technologies In Engineering And Sciences
Gyan Ganga College of Technology
Vol. 4, Special Issue 4, November 2016

technique was first used for detection of manufacturing defects of sequential circuits [6]. In the context of this technique a region is an important term. It is a connected set of gates. The goal is to "cluster the flip-flops into groups that are most likely to be associated with a Trojan" [2]. This detection method uses the power consumption P of the chip as the side channel. Because the power consumption depends on frequency, the chip is clocked at low operating frequency during this analysis. The power consumption P is simplified compared to the definition in 3.4.1, because the accuracy is sufficient in this case.

$$P = C.Vdd^2.f$$

C, Vdd Technology dependent parameters
f Clock frequency

The Power Profile is a set of values .Each value represents the vector. The measured value of the power consumption is an indicator for the switching activity in a particular region of the chip.

Every detection method consists of several steps.The region-based partition and executes as follows:Sequence
1. select regions for analysis
2. generate input vectors
3. measure the power profile

Path Delay Fingerprinting

Often, the side channel signal power consumption can be too vague. In this section, a more elaborate detection technique is explained. The paper [8] describes a method to analyze the netlist of a chip, whereas [16] depicts a more practical method to measure path delays in integrated circuits.

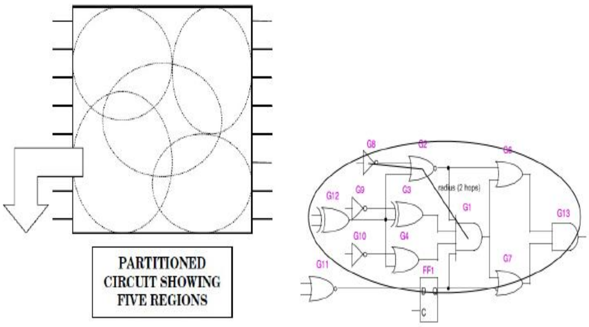


Figure 10: Illustration of the Term Regions and one Region with Radius 2[2]

The measurement of timing information seems to be slightly better than the measurement of power consumption. On the one hand a path delay fingerprint of an IC can be the set of path delay values that are measured on the pins of the chip. But, another method uses timing information measured between two registers (register-to-register path delay) to detect malicious alterations. The former and more theoretical method uses the netlist of the chip to analyse the timing behavior.

V. CONCLUSION

The analysis of ICs is expensive and there can be adversaries that apply to all Alternatives characteristics of the attacker. But then the potential of threat can be lower. In this case other techniques can be used. Obfuscation and camouflage can be useful to reduce the furor. These techniques hide the interesting chip to some extend. Further investigations should include detection methods that detect ad- Perspective aditional or malicious electronic devices like additional ICs or transistors on a complex printed circuit board (PCB). An additional wire on a PCB that bypasses an encryption chip when a particular condition occurs also is a hardware Trojan that is easy to hide. New experiments that comply with double-blind and randomized test operations could prove the practical benefit of the detection techniques.In this paper the reader was introduced to a classification scheme Theory and many terms concerning Trojan detection. The most detection methods use one side channel signal to find malicious circuitry. Some problems remain even if successful detection techniques exist. A practical detection method should be low-cost, scalable, fast-detecting and Praxis effective. In general it is more difficult to find unknown/malicious circuits than to check for known patterns or activation vectors. Several techniques exist, but some are just concepts or simulations, so tests with real fabricated ICs should be conducted.

REFERENCES

[1] Mainak Banga and Michael S. Hsiao: A Region Based Approach for the Identification of Hardware Trojans, Bradley Department of Electrical and Computer Engineering, Virginia Tech., Host'08, 2008.

[2] A. L. DSouza and M. Hsiao: Error diagnosis of sequential circuits using region-based model, Proceedings of the IEEE VLSI Design Conference, January, 2001, pp. 103-108.

[3] C. Fagot, O. Gascuel, P. Girard and C. Landrault: On Calculating Efficient LFSR Seeds for Built-In Self Test, Proc. Of European Test Workshop, 1999, pp 7-14.

[4] Rajat Subhra Chakraborty, Somnath Paul and Swarup Bhunia: On-Demand Transparency for Improving Hardware Trojan Detectability, Department of Electrical Engineering and Computer Science, CaseWestern Reserve University, Cleveland, OH, USA

[5] Yier Jin and Yiorgos Makris: Hardware Trojan Detection Using Path Delay Fingerprint, Department of Electrical Engineering Yale University, New Haven

[6] Reza Rad, Mohammad Tehranipoor and Jim Plusquellic: Sensitivity Analysis to Hardware Trojans using Power Supply Transient Signals, 1st IEEE International Workshop on Hardware-Oriented Security and Trust (HOST'08), 2008.

[7] Dakshi Agrawal, Selcuk Baktir, Deniz Karakoyunlu, Pankaj Rohatgi and Berk Sunar: Trojan Detection using IC Fingerprinting, IBM T.J. Watson Research Center, Yorktown Heights, Electrical & Computer Engineering Worcester Polytechnic Institute, Worcester, Massachusetts, Nov 10, 2006

[8] Xiaoxiao Wang, Mohammad Tehranipoor and Jim Plusquellic: Detecting Malicious Inclusions in Secure Hardware, Challenges and Solutions, 1st IEEE International Workshop on Hardware-Oriented Security and Trust (HOST'08), 2008

[9] Miron Abramovici and Paul Bradley: Integrated Circuit Security – New Threats and Solutions

[10] Tom Kean, David McLaren and Carol Marsh: Verifying the Authenticity of Chip Designs with the DesignTag System

[11] S. Mitra, and K.S. Kim: X-Compact - An Eficient Response Compaction Technique, IEEE Tran. on CAD, vol. 23, no. 3, pp. 421432, Mar. 2004.

[12] Jie Li and John Lach: At-Speed Delay Characterization for IC Authentication and Trojan Horse Detection, 1st IEEE International Workshop on Hardware-Oriented Security and Trust (HOST'08), 2008

[13] Sun Tzu: The Art of War, 6th century BC, China

[14] J. M. Soden, R. E. Anderson, C. L. Henderson: IC Failure Analysis: Magic, Mystery, and Science. In: IEEE Design & Test of Computers, Vol. 14, pp. 5969, 1997.

[15] M. Tehranipoor, H. Salmani, X. Zhang, X. Wang, R. Karri, J. Rajendran, K. Rosenfeld, ―Trustworthy hardware: Trojan detection and design-for-trust challenges,Computer (2011)

[16] M. Tarek Ibn Ziad, A. Al-Anwar, Y. Alkabani, M. W. El-Kharashi, H. Bedour, ―E-voting attacks and countermeasures, In Proceedings of the 10th International Symposium on Frontiers of Information Systems and Network Aplications (FINA 2014), held in conjunction with the 28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014), Victoria, BC, Canada, 2014, pp. 269–274.

[17] OpenCores. Online: http://opencores.org/ (accessed April 2016)