Perniciousness analysis of IC hardware Trojan based on characteristic function deployment

Jiaoying Huang, Daming Yang and Yongkang Wan

School of Reliability and Systems Engineering, Beihang University, Beijing, China

1 E-mail: dada19930921@126.com

Abstract. New hardware security threats are emerging, meanwhile detection and recognition technologies are immature for unknown types. This paper introduces a relation matrix to analysis perniciousness of IC hardware Trojan. The relation matrix is based on characteristic deployment includes concealment of trigger and harmfulness of payload. Results of this paper can be used to measure the perniciousness of different types of hardware Trojan, also provide a way to predict new hardware threats, in this way help to defend unknown hardware attacks and improve detection efficiency.

1. Introduction

In 2007, IBM Watson research center and Worcester Polytechnic Institute co-published a paper where the concept of hardware Trojan circuit is put forward for the first time [1]. Within the next ten years, experts and scholars at home and abroad have never interrupted the study of hardware Trojans [2]. Recently, more and more kinds of new hardware Trojans have been discovered [3]. The perniciousness of hardware Trojans to integrated circuits has been paid more and more attention, and the corresponding detection methods have been studied [4]. Michigan university has studied an analog hardware Trojan triggered by the cumulative charge of capacitors, A2 [5]. The researchers implanted the hardware Trojan in the Open-RISC 1200 processor. A chip entity with A2 hardware Trojan is manufactured to verify the implanted attack and can access to system permissions by attack of hardware Trojan [6]. Sanchita Mal-Sarkar predicted and analyzed the possible hardware Trojans on FPGA and he pointed out the security threats of FPGA in the process of design, manufacture and application [7]. Swarup Bhunia summarized the existing hardware Trojan and detection technology [8]. He believes that the future research needs to focus on three points: new attack modes, IC security metrics and new protection methods.

At present, the research of hardware Trojan is mainly aimed at some kinds of chips to study specific hardware Trojans. It is not comprehensive enough that perniciousness analysis of hardware Trojan affect the integrated circuit. There is no suitable analysis and detection methods for unknown hardware Trojan on chips. John.Shield at the thirteenth Australian information security conference pointed out that the current researches which focus on hardware Trojan are too narrow, and in fact, the hardware Trojan has become a systemic threat [9].

In order to provide a new idea for the safety measurement and protection of integrated circuits, this paper proposes a hardware Trojan perniciousness relationship matrix based on structural characteristic decomposition. The matrix reflects the perniciousness levels of different hardware Trojans on the integrated circuit. The matrix is formed by comprehensive analysis of the concealment and combination of hardware Trojan trigger structure and the perniciousness of payload. The results of the
matrix can measure perniciuosness severity of different hardware Trojans. It provides a preliminary prediction of potential new hardware Trojans.

2. Classification and characteristics of hardware Trojan
There are dozens of hardware Trojan that have been discovered and researched [10]. Hardware Trojans are classified according to the carrier objects, attack characteristics and effect [11]. The common classification methods are according to the hardware Trojan attack mechanism, including changing the circuit function, revealing the confidential information inside the circuit, assisting the control system and physical destroy, etc [12].

There are few actual cases of hardware Trojan attacks. At the same time, new hardware Trojans emerge in an endless stream, and the classification method based on statistics has limitations. Referring to Swarup Bhunia's analysis about the structure and function of hardware Trojan, the hardware Trojan is divided into two parts on structure and function: trigger and payload, which are classified by combination of trigger and payload, as shown in figure 1.

![Figure 1. Trojan taxonomy based on trigger and payload mechanisms.](image)

Trigger and payload are two independent parts, and can be designed separately. When hardware Trojan is triggered, there is only exchange information of feedback between the two parts. Because of this independence, it is possible that random combination of different triggers and payloads. It also confirms the current development trend of hardware Trojan: new hardware Trojans emerge in an endless stream, and hard to guard against. Various types of hardware Trojans can be derived from a new type of trigger or payload.

3. Perniciousness relation matrix of hardware Trojans
According to the features of trigger and payload, perniciousness of the hardware Trojan can be determined by two factors: concealment of hardware Trojan dodging detection and harmfulness to equipment and system [13].

The concealment is mainly decided by the trigger, and the harmfulness is mainly determined by payload. According to this, the perniciousness relationship matrix of hardware Trojans is established. Based on ‘House of quality’ method of quality function deploy (QFD), ‘House of hardware Trojan perniciousness’ is built. It can convert the concealment of trigger and harmfulness of payload into matrix of hardware Trojan perniciousness influence.

‘Roof’ is the possibility of mixing various triggering. The ‘left wall’ is the severity degree which is the damage degree of integrated circuit and the influence on the system and the whole machine under the attack of hardware Trojan. ‘Room’ represents the perniciousness level of hardware Trojans. The structure of the perniciousness relation matrix is shown in figure 2.
Figure 2. Perniciousness relationship matrix of hardware Trojan.

4. Concealment of hardware Trojan

4.1. The definition about concealment of hardware Trojan

The concealment of the hardware Trojan is defined as the difficulty of detection and recognition for hardware Trojans. The hardware Trojans which is harder to be recognized has the better concealment [14]. Based on the analysis of the hardware Trojan's functional characteristics and the characteristics of the existing detection methods, it is determined that the concealment of hardware Trojans depends mainly on the volume and trigger rate of the trigger.

The existing hardware Trojan detection techniques are mainly the following three kinds:

- **Physical detection**: By opening and taking photographs of chips to compare and identify hardware Trojan in the layout of the chip.
- **Logical test**: By producing a specific test vector to quickly activate hardware Trojan.
- **Side channel analysis**: By analyzing the side channel signal of the circuit to recognize the hardware Trojan.

Table 1 shows characteristics, advantages and disadvantages of the three existing detection techniques.

| Payload                        | Advantage                                      | Disadvantage                                |
|-------------------------------|-----------------------------------------------|---------------------------------------------|
| 1. Power & Delay              | High credibility;                             | Devastating and heavy workload;            |
| 2. Information Leakage        | Unaffected by technological noise.            | Hard to detect small Trojans.              |
| 3. Circuit Nodes              | Anti-noise;                                   | Hard to generate test vector;              |
| 4. Circuit Output             | Obvious effects for small Trojan.             | Hard to detect complex Trojans.            |
| 5. Denial-of-Service          | Obvious effects for large Trojan;             | Easy to be disturbed by process noise;      |
| 6. Physical Destruction       | Test vector is simple.                        | Hard to detect a small Trojan.             |

By summarizing and analyzing the existing method, the key factor of hardware Trojan detection is the size and activation rate of the hardware Trojan.
The result of the small size of the trigger is that the physical detection is difficult to identify hardware Trojan and the side channel analysis is also difficult because the small signals are easily concealed by process noises. The trigger has a low trigger rate, and it causes that hardware Trojan is difficult to be activated in the logic test, so it is difficult to generate the appropriate test vector.

According to the above analysis, it can be concluded that the concealment of the hardware Trojan is defined by the maturity of the detection method, and is determined by the characteristics of the trigger.

4.2. Analysis of hardware Trojan’s concealment

In this paper, three kinds of hardware Trojan are analyzed. The concealment of the hardware Trojan is defined by the orthogonal relationship between triggering rate and volume of the trigger. These three kinds of hardware Trojan are analog malicious hardware (A2), 4 input combinational logical hardware Trojan, and 16 bit asynchronous counting trigger hardware Trojan.

Hardware Trojan of 16 bit asynchronous counting trigger takes up a lot of resources. But compared to the 4 input logic trigger architecture with the same digital trigger structure, it has a much lower trigger rate. Therefore, it is easy to activate and detect the 4 input combinational logical hardware Trojan by logical testing. It is very difficult to improve its activation rate of 16 bit asynchronous counting trigger hardware Trojan, so it has a better concealment.

But for A2, it has a special setting of the trap capacitor which makes the capacitor start to charge when the trigger is in a special working state, and its characteristics of the gradual release of the charge makes the hardware Trojan more difficult to be recognized. At the same time, under the situation of meeting the requirements of process, the volume of the hardware Trojan reaches $e^{-5}$ level, which is difficult to be recognized by either physical detection or side channel signal analysis. Hence, A2 has a great concealment.

Concealment of the three triggering structures is reflected by the orthogonal relationship between volume and trigger rate. The results are as shown in figure 3.

![Figure 3. Result of the concealment analysis of the hardware Trojan.](image)

The concealment of hardware Trojan trigger which is shown in the figure is qualitatively analyzed. The results reflect the comparison of concealment levels of different kinds of hardware Trojans and it is not a specific value. The results explain the reason why it is difficult for hardware Trojan detection and makes the features of hardware Trojans more specific.
5. Perniciousness of hardware Trojan

5.1. Definition of harmfulness of hardware Trojan
When the hardware Trojan is triggered, it acted on the chip by payload to complete the attacker's function. The design of the trigger is more flexible and can be designed followed the needs of the attacker. The integrated circuit is used to be the bottom and the most core part of the equipment and system. If hardware Trojan attacks on integrated circuits, it will destroy the entire system.

As for some key applications, such as space satellites, rockets, missiles, aircraft and nuclear facilities, it may cause disastrous consequences. Referring to the perniciousness analysis in GJB/Z 1391-2006 ‘the guide for analysis of failure mode, influence and harmfulness’, the perniciousness of hardware Trojan is analyzed. Using the severity of the ESR level, which refers to the impact of the fault, the danger of hardware Trojan is defined. The perniciousness of hardware Trojan is measured by harmful severity of hardware Trojan which leads to equipment and system malfunction. Table 2 shows the severity criteria for perniciousness.

| ESR Level | Severity Grade | Severity of Failure |
|-----------|----------------|----------------------|
| 1         |                | Not enough to cause staff injury; Mild damage to the product; |
| 2         | Mild           | Mild property loss; Mild environmental damage; Lead to unplanned maintenance or repair. |
| 3         |                | Cause staff moderate injury; Medium damage to the product; |
| 5         | Medium         | Task delay or demotion; Medium property loss; Medium environmental damage. |
| 6         |                | Cause staff injury; Serious damage of products; |
| 7         | Fatal          | Failure of task; Serious damage of property; Serious environmental damage. |
| 8         |                | Cause death of people; Damage to a product (such as a plane, a tank, etc.); |
| 9         | Catastrophic   | Major property loss; Major environmental damage. |

5.2. Analysis of hardware Trojan’s harmfulness
The perniciousness of the hardware Trojan is determined by the trigger which direct effects hardware Trojan. Hence, the harmfulness of different kinds of hardware Trojan is measured based on the trigger rate of the hardware Trojan. This paper forms a hardware Trojan perniciousness matrix, which
determines the trigger mechanism, attack mode and harm degree of each hardware Trojan. It provides a theoretical basis for the determination of the key points of the integrated circuit hardware Trojan.

The perniciousness of the hardware Trojan is determined by the concealment of trigger structure and the harmfulness of payload. The symbols in the matrix, such as ‘△’, ‘○’, ‘◎’, are used to represent the perniciousness level from 1 to 9.

**Level 1:** the corresponding hardware of this intersection has a small probability or weak impact.

**Level 3:** the corresponding hardware of this intersection has a certain probability or a certain degree of harm.

**Level 9:** the corresponding hardware of this intersection has a large probability or a great perniciousness.

![Figure 4. The result of perniciousness relationship matrix analysis.](image)

Based on the existing cases, by analysis of the perniciousness relationship matrix in figure 4, it can be concluded that the combination of digital analog trigger structure has a wide variety, great concealment and is difficult to be identified. A2 belongs to a kind of trigger structure which is triggered by a special digital node and the accumulative charge of capacitance. This type of hardware Trojan can attack at the output of the chip circuit, easily acquire system permissions, change the output of the circuit or destroy the circuit directly, which has great potentially perniciousness.

### 6. Conclusions

The purpose of this paper is to qualitatively and quantitatively analyze the perniciousness of different kinds of hardware Trojan and predict potential threats of new hardware Trojan. By defining the concealment and harmfulness of the hardware Trojan, the paper built the matrix of the perniciousness relationship of the hardware and Trojan, and obtains the following conclusions:

The difficulty of detecting hardware Trojan depends on two main factors, the volume and the activation rate of hardware Trojan. The existing detection methods have limitations for the hardware Trojan with small volume and low activation rate.

New hardware Trojan in the future cannot be ignored. Through the research of hardware Trojan trigger and effective load respectively, it can be concluded that every new type of trigger or payload can derive a variety of new hardware Trojan.

The perniciousness matrix of the hardware Trojan reflects the potential threat of different hardware Trojan. The digital and analog combined hardware Trojan is difficult to detect and has great
perniciousness. In the future, the researches on the detection method for this type of hardware Trojan must be paid more attention.

Acknowledgements
This work was supported by Pre-research projects in microelectronics of the thirteenth Five-Year Plan under Grant No. 31513050304.

References
[1] Agrawal D, Baktir S, Karakoyunlu D, et al. 2007 Trojan Detection using IC Fingerprinting[C] IEEE Symposium on Security and Privacy. IEEE Computer Society 293-310
[2] Tehranipoor M, Koushanfar F 2010 A survey of Hardware Trojan Taxonomy and Detection IEEE Design Test Comput 27(1) 10-25
[3] Rad R, Plusquellic J and Tehranipoor M. 2010 A Sensitivity Analysis of Power Signal Methods for Detecting Hardware Trojans Under Real Process and Environmental Conditions IEEE Transactions on Very Large Scale Integration Systems 18(12) 1735-1744
[4] John S and Bradley H 2015 Hardware Trojans – A Systemic Threat [C] Proceedings of the 13th Australasian Information Security Conference (AISC 2015) 45-51
[5] Yang K, Hicks M and Dong Q, et al. 2016 A2: Analog Malicious Hardware[C] Security and Privacy. IEEE 18-37
[6] Jin Y, Kupp N, and Makris Y 2009 Experiences in Hardware Trojan design and implementation[C] IEEE International Workshop on Hardware-Oriented Security and Trust 50-57
[7] Malsarkar S, Karam R, and Narasimhan S, et al. 2017 Design and Validation for FPGA Trust under Hardware Trojan Attacks IEEE Transactions on Multi-Scale Computing Systems 2(3) 186-198
[8] Chakraborty R S, Narasimhan S, and Bhunia S 2009 Hardware Trojan: Threats and emerging solutions[C] High Level Design Validation and Test Workshop 166-171
[9] John S, Bradley H, and Mark B, et al 2015 Hardware Trojans – A Systemic Threat[C] Proceedings of the 13th Australasian Information Security Conference 45-51
[10] Lei L I, Shang Z, and Feng J, et al. 2013 Design and Insertion of Hardware Trojan Based on Finite State Machine Acta Scientiarum Naturalium Universitatis Pekinensis 49
[11] Rajendran J 2010 Towards a Comprehensive and Systematic Classification of Hardware Trojans[C] IEEE International Symposium on Circuits and Systems 1871-1874
[12] Bazzazi, A. M.T., Manzuri Shalmani and A.M.A. Hemmattyar 2016 Trojan Counteraction in Hardware: A Survey and New Taxonomy Indian Journal of Science and Technology
[13] Krieg C, Rathmair M, and Schupfer F 2014 A Process for the Detection of Design-Level Hardware Trojans Using Verification Methods[C]// IEEE Intl Conf on High PERFORMANCE Computing and Communications 729-734
[14] Wang X., M. Tehranipoor, and J. Plusquellic 2008 Detecting malicious inclusions in secure hardware: Challenges and solutions. in Hardware-Oriented Security and Trust,, HOST 2008. IEEE International Workshop