A Novel Clustering Technique Using Backscattering Side-Channel for Counterfeit IC Detection

Luong N. Nguyen, Baki Berkay Yilmaz, Chia-Lin Cheng, Milos Prvulovic, and Alenka Zajić

School of Electrical and Computer Engineering
College of Computing
Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

Over the past few years, globalization of the semiconductor supply chain has led companies to outsource many steps of their integrated circuit (IC) production cycle. While outsourcing helps companies tremendously reduce cost and time-to-market, it arises questions on the trust level of IC, and introduces several security vulnerabilities. One of the most serious problems is the counterfeiting of ICs because it does not only negatively impact innovation and economic growth of IC industry, but also imposes serious threat and risk on systems that incorporate these counterfeit ICs. This paper proposes a novel clustering method to detect counterfeit ICs using backscattering side-channel. Backscattering side-channel is a recently introduced side-channel that has been proved to outperform other side-channels in detecting hardware changes in digital circuit. Our algorithm utilizes the backscattering side-channel to cluster counterfeit ICs from normal ICs. We test our technique on a set of ten boards over six different counterfeit IC designs. The results show that our technique can tolerate manufacturing variations among different hardware instances to report all counterfeit IC measurements with 100% accuracy and 0% false positives.

Keywords: Counterfeit IC, Hardware Security, Reliability, Cloned IC, Backscattering, Clustering

1. INTRODUCTION

A counterfeit IC is an illicit copy of a legitimate chip with discrepancy on the performance, characteristics or material, but sold or used as an authorized one. Counterfeiting of ICs has become a major challenge to the semiconductor industry because of deficiencies in the existing test techniques and lack of effective protection mechanisms. Over the past few decades, as globalization of the semiconductor supply chain has led companies to outsource many steps of their integrated circuit (IC) production cycle, the problem has been getting worse as the incidences of counterfeit ICs has increased exponentially. In 2015, it was reported that the illicit production of counterfeit ICs cost IC companies $100 billion and the cost has steadily increased every year. This has imposed a significant threat to the IC industry because it does not only negatively impact innovation and economic growth of IC industry, but also imposes serious threat and risk on systems that incorporate these counterfeit ICs. Counterfeit ICs have appeared in almost all industrial sectors, including sensitive fields such as finance, government infrastructure and military systems. As a result, demands for effective counterfeit IC detection or avoidance techniques has been tremendously escalated.

Over the past few years, a plethora of papers have been published on the topics of preventing counterfeit ICs and these work can be divided into detection techniques or avoidance mechanisms. Avoidance mechanisms rely on techniques such as adding additional signature circuitry to watermark the ICs, using multi-phase fabrication by dividing the layout of the chip and get them fabricated at different foundries, to prevent the chips from being counterfeited. However, the main drawback of avoidance techniques is the significant overhead added to the cost of ICs, which prevents them to be widely use in industry.

On the other hand, counterfeit IC detection techniques focus on distinguishing counterfeit ICs from authentic ones and do not need additional circuitry of changing the layout of ICs. These techniques can be clustered into either physical tests and electrical tests. The physical tests rely on examining the physical and chemical/material properties of the IC’s package, leads and die in order to detect procedural, mechanical, and environmental
counterfeit ICs. These techniques include incoming inspection methods such as simple external visual inspection (EVI), X-ray imaging, exterior inspection tests such as resurfacing, microscopy scanning, material analysis such as X-Ray Fluorescence (XRF), Fourier transform infrared spec. (FTIR), ion chromatography. In general, physical tests can be used to detect all type of counterfeit ICs. However, they are often destructive, time-consuming, and expensive.

Electrical tests consist of parameter tests, function tests, curve tracing, built-in tests and structural tests. Unlike physical tests, electrical tests are non-destructive, relatively fast and cheap. However, electrical tests relies on determining if functionality of IC is correct to be able to detect counterfeit ICs, which mean it does not work for counterfeit ICs that have the same functionality but different layout as authentic ones. In addition, a sheer number of electrical test techniques require additional circuitry added to the design, which significantly increases the cost of the ICs.

Motivated by the above-mentioned drawbacks of previous detection techniques, this paper proposes a novel non-destructive and fast technique using the backscattering side channel for detection of counterfeit ICs with same functionality but different layout with authentic ones. We choose to use backscattering side-channel, a new physical side-channel that has been proved to be a perfect fit, and outperforms other side-channels for hardware variation detections.

The rest of the paper is organized as follows. In Section 2, we provide a background including the overview of counterfeit IC and the backscattering side channel for this work. Section 3 explains our technique for counterfeit IC detection, while Section 4 describes our experiment setup, benchmark implementation and testing scheme formulation. Section 5 presents the results of our technique. Finally, Section 6 concludes the paper.

2. BACKGROUND

2.1 Counterfeit ICs

A counterfeit IC is an illicit copy of a legitimate chip with discrepancy on the performance, characteristics or material, but sold or used as an authorized one. Counterfeit ICs can be divided into three groups remarked/recycled ICs, out-of-spec/defective ICs, and cloned ICs. The first group includes aged ICs sold as new, ICs remarked with forged information. The second group includes out of specification ICs, rejected ICs during manufacturing tests but sold as normal ones, and tampered ICs which are infected by hardware Trojans. The last group includes overproduced ICs, and unauthorized production of an IC by illegally obtaining the design of the IC either at RTL level, netlist level or layout level.

2.2 Backscattering Side Channel

While EM and power side-channels have been used extensively, the use of the backscattering side-channel is still relatively new. EM and power side-channels are a consequence of changes in the current inside an electronic circuit. In contrast, the backscattering side-channel is a consequence of impedance changes in switching circuits, which is caused by the transistors’ two-state impedances reflecting a modulated signal.

For example, Figs. 1 (a) and (b) show a 2-input CMOS NAND gate and CMOS NOR gate respectively. When the input voltages of the NAND circuit are low, the NMOS transistors are off and the PMOS transistors are on. A direct path exists between \( V_{out} \) and \( V_{DD} \), resulting in a high output state. On the other hand, high input voltages result in a low output state. As shown in Fig. 1 (c), there exists a finite resistance \( R_1 \) between the output and \( V_{DD} \) and \( R_0 \) between the output and the ground, respectively. The values of \( R_0 \) and \( R_1 \) are different. In other words, switching between the NAND logic’s high output state (\( R_1 \)) and low output state (\( R_0 \)) creates impedance variation, which is analogous to the impedance switching in RFID tags. Similarly, the same reasoning can be applied to CMOS NOR and NOT gate circuits. Typically, every digital circuit is built by NAND (or NOR) and NOT gates so it can be generalized to any digital circuit. When a continuous-wave signal is transmitted toward a set of gates, the backscattered signal can be expected to change as the gates’ states change, thus creating an impedance-based side-channel. Similar to the traditional EM side-channel, the backscattering side-channel has a high bandwidth; however, unlike the EM side-channel, the strength of the backscattered signal can be increased when needed, its frequency can be shifted to avoid noise, interference, and poor signal propagation conditions, and it can be more accurately focused on a specific part of the chip.
3. A NOVEL APPROACH FOR COUNTERFEIT DETECTION

3.1 Using Backscattering Side Channel for Counterfeit IC Detection

Nguyen et al. have shown that HTs can be detected by analyzing impedance changes within sub-clock samples, where the changes caused by HTs happen and can be observed on the clock signal. Therefore, if we can capture the backscattered signal of sub-clock samples where the changes caused by HT can be observed, we can detect the presence of HTs. However, the problem with the time-domain signal is that they are often very noisy, therefore, difficult to extract and synchronize measurements to get samples where changes caused by HTs happen.

In contrast, the changes caused by HTs occurring abruptly at some point in the clock cycle can be observed in frequency domain by performing short Time Fourier transformation (STFT) on time-domain signal and observe which frequency components of time domain signal are affected when dormant HT is present. The signals in the frequency domain are much easier to measure, and the noise power is very small because of focusing a single frequency bin at a time. As a result, instead of measuring the time domain signal, we measure multiple harmonics of the clock in the frequency domain to observe changes in sub-clock samples for HT detection.

The change caused by HTs will be reflected in backscattered signals at the circuit’s clock harmonics: $f_{\text{carrier}} \pm f_c$, $f_{\text{carrier}} \pm 2f_c$, etc. The first clock harmonic at $f_{\text{carrier}} \pm f_c$ follows the overall RCS change during a cycle, while the remaining harmonics are affected by the rapidity of change (rise/fall times), and timing of the impedance changes within the clock cycle. For each circuit, we measure the amplitude of the first $N$ harmonics of the clock from its backscattering side-channel signals to form a vector, which characterizes the circuit’s overall amount, timing, and duration of impedance-change activity during a clock cycle. If there is a hardware Trojan in the circuit, this vector will be different from the ones recorded from an HT-free same circuit. As a result, we can represent each circuit by a vector of $N$ points, which are the amplitudes of the first $N$ harmonics of the clock from its backscattering side-channel signals: $h = [h_1, h_2, ..., h_{N-1}, h_N]$, where $h_j$ is the amplitude of the $j^{th}$ harmonic of the clock. These vectors will be used as inputs for our clustering algorithm.

If changes caused by HTs in the time-domain signal become briefer in duration, the changes among clock harmonics become smaller in magnitude and shift to higher harmonics which, compared to lower harmonics, tend to be affected more by noise. This is one of the reasons why the backscattering side-channel works better for HT detection than other traditional analog side-channels such as EM and Power side-channels. The backscattering side-channel is a consequence of the impedance changes in switching digital switching circuits, which is caused by the transistors’ two-state impedances reflecting a modulated signal. For each gate that switches, the impedance change persists for the rest of the cycle. On the other hand, the EM and power side-channels are consequences of the variation of the current flow in a circuit. As a gate switches, the current will be charged or discharged quickly, which means a current burst occurs for a very short period of time.
3.2 One-Class-Classification to Detect Counterfeits

In this section, we introduce our one-class-classification technique that accurately detect whether the layout is original. The approach is based on supervised learning techniques which contain two phases: Training and testing. In the training phase, the goal is to obtain an accurate model to claim whether a given layout is counterfeit. Then, the testing phase reveals how robust the model is to label layouts.

To achieve our goal, we first collect the harmonic magnitudes for the first $N$ harmonics as described in Section 3.1. An example of the original harmonic magnitudes are given in Figure 2. We observe that the harmonics are generally dense around the mean for each harmonic yet magnitudes vary significantly among different harmonics. Therefore, our model first considers each harmonic independently, and then combines the results of harmonics to deduce the originality of the layout. In that respect, we calculate the mean magnitude of each harmonic as

$$\hat{h}[k] = \frac{1}{M} \sum_{m=1}^{M} h_m[k]$$  \hspace{1cm} (1)$$

where $M$ is the number of training measurements, and $h_m$ is a row vector containing the harmonic values of the $m^{th}$ measurement which can be written as

$$h_m = [h_m^0 \ h_m^1 \ \cdots \ h_m^N]$$  \hspace{1cm} (2)$$

and $h_m[k] = h_m^k$. To proceed further, let assume $M_D$ be the distance of the most deviated harmonic values from the mean among all training measurements such that

$$M_D = \max \{ \abs{H - 1\hat{h}^T} \}$$  \hspace{1cm} (3)$$

where $\max\{\bullet\}$ returns a row vector which contains the maximum values at each column of its argument, $\abs{\bullet}$ returns the magnitude of its argument, $1 \in \mathbb{R}^M$ is a column vector with full of ones, and $H$ is a matrix such that each row represents a measurement, and each column contains the corresponding harmonic value.

A common method in one-class-classification is density estimation which works better when large number of measurements is present.\textsuperscript{16} \ However, having such a large sample space is difficult and takes really long time to collect signals. Therefore, by mimicking the well-known “$3\sigma$” rule, we define the confidence interval for the $m^{th}$ harmonic as

$$[\hat{h}[m] - 3 \cdot M_D[m], \ \hat{h}[m] + 3 \cdot M_D[m]]$$  \hspace{1cm} (4)$$

An example of the intervals and the average harmonic signal is given in Figure 3. One of the main observations is that the spread around each harmonic varies, therefore, considering each harmonic independently gives better insight about the layout. For an even better illustration of the harmonics, we normalize the data as follows:

$$H = (H - 1\hat{h}) ./ (1M_D)$$  \hspace{1cm} (5)$$

Figure 2. Harmonic magnitudes of the original circuit in the training phase.
where “/” is the pairwise division operation. After this normalization of the data, the confidence interval or decision boundaries are fixed to between -1 and 1. After the normalization, the training data and the boundaries are given in Figure 4. Finally, the layout is called

- *Original* if all measured harmonics are within the corresponding confidence intervals,
- *Counterfeit* if there is at least one harmonic values which violates its confidence interval.

### 4. Benchmark Implementation and Experiment Setup

#### 4.0.1 Counterfeit IC Benchmark Implementation

Specifically, for our experiments, we implement two different kinds of counterfeit IC: 1) Counterfeit ICs with the same functionality as the original but different physical implementation (position) of the circuit, and 2) Counterfeit ICs with the same functionality and position as the original but different physical layout (routing and placement) of the circuit.

- Counterfeit ICs with Different Layout: We have implemented several counterfeit IC examples by re-compiling and letting the EDA tool to change the placement and routing of the circuit. We have four different test subject designs: Original layout AES IC, 1st layout AES counterfeit IC, 2nd layout AES counterfeit IC, 3rd layout AES counterfeit IC.
Counterfeit ICs with Changed Position: We have implemented several counterfeit IC examples by moving the placement of the AES circuit from its original placement. We have four different test subject designs: original position AES IC, 1st position AES counterfeit IC, 2nd position AES counterfeit IC, and 3rd position AES counterfeit IC.

4.0.2 Experiment Setup

The experimental setup to evaluate the performance of the proposed algorithm is shown in Fig. 5. The setup includes a transmitter Aaronia E1 electric-field near-field probe\(^{17}\) connected to an Agilent MXG N5183A signal generator,\(^{18}\) and a receiver Aaronia H2 magnetic field near-field probe\(^{17}\) connected to an Agilent MXA N9020A spectrum analyzer.\(^{19}\) The devices-under-test (DuT) are Altera DE0 Cyclone V FPGA boards.\(^{20}\) An angle ruler is used as a positioner so that different DE0-CV boards can be tested using approximately the same position of probes. A laptop is used to control the devices and automate the measurements. A 3 GHz continuous sinusoid signal is generated by the signal generator, and backscattered signals are recorded by the spectrum analyzer.

5. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we provide experimental results when layout or placement position of the circuit changes. We will first perform an experiment when the layout of the circuit changes while keeping the functionality same. The results are given in Figure 6. We need to remind here that the harmonic values and the boundaries are normalized based on the formula given in (5). For the testing phase, in this equation, \(H\) corresponds a matrix where each row is a test measurement, \(h\) is the mean harmonic values obtained from training measurements, and \(M_D\) is the maximum deviation row vector for each harmonic which is also obtained in the training phase. Figure 6(a) demonstrates the behaviour of the harmonics when the layout is original. In this figure, we plot the average training signal, corresponding confidence interval, and the harmonic values of the circuit. We observe that all considered harmonics are within the confidence interval, hence, all measurements are labeled as original. Then, we experiment with counterfeit layouts that have the same functionality but different layout. Figures 6(b), 6(c) and 6(d) illustrate the harmonic values for different counterfeits. We observe that all of the measurements have at least one component that violates the confidence intervals. Therefore, our methodology achieve 100\% accuracy to detecting the fake circuits.

Another experiment is performed by changing the placement location of the circuit while keeping its functionality and layout same. Similar to the previous measurements, Figure 7(a) demonstrates the results of the
Figure 6. Normalized harmonic magnitudes of the circuit with the same functionality and different layouts.

test signals of the original circuit, and the rest illustrates the fake circuits. All of the harmonics of the original circuit measurements lie in the confidence intervals for each harmonic. Therefore, all of the original circuits are accurately labeled. If we consider the counterfeit circuits, we observe that there is at least one harmonic that is not confined in the confidence interval. Hence, all the fake circuits are accurately labeled as counterfeit. Therefore, we achieve 100% accuracy for this experiment as well.

In summary, both experiments reveal that our methodology is a very powerful and robust to identify the counterfeit circuits.

6. CONCLUSION

Over the past few years, counterfeiting of ICs has become a major challenge to the semiconductor industry because of globalization of the semiconductor supply chain, deficiencies in the existing test techniques and lack of effective protection mechanisms. This paper proposes a novel non-destructive and fast technique using the backscattering side-channel for detection of counterfeit ICs with same functionality but different layout with authentic ones. Our algorithm utilizes the backscattering side-channel to cluster counterfeit ICs from normal ICs. We choose to use backscattering side-channel, a new physical side-channel because it has been proved to be a perfect fit, and outperforms other side-channels for hardware variation detections. We test our technique on a set of ten boards over six different counterfeit IC designs. The experimental results show that our technique can tolerate manufacturing variations among different hardware instances to report all counterfeit IC measurements with 100% accuracy and 0% false positives.

REFERENCES

Figure 7. Normalized harmonic magnitudes of the circuit with the same functionality and layout, but different placement positions.


