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

Luong N. Nguyen\textsuperscript{a}, Baki Berkay Yilmaz\textsuperscript{a}, Chia-Lin Cheng\textsuperscript{a}, Milos Prvulovic\textsuperscript{b}, and Alenka Zajić\textsuperscript{a}

\textsuperscript{a}School of Electrical and Computer Engineering
\textsuperscript{b}College of Computing
\textsuperscript{1}Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

Over the past few years, globalization of the semiconductor supply chain has led companies to outsource much of the production cycle for integrated circuits (ICs). While outsourcing helps companies significantly reduce their cost and time-to-market, it also introduces concerns about the trustworthiness of an IC. One of the most serious problems is counterfeiting of ICs, which not only negatively impacts innovation and economic growth of the IC industry, but also creates serious threats and risks for systems that incorporate those counterfeit ICs. This paper proposes a novel method that uses the backscattering side-channel to cluster ICs such that counterfeits are separated from legitimate ICs. The backscattering side-channel, which has been introduced only recently, has been proven to outperform other side-channels in detecting hardware Trojan horses (HTs), i.e. ICs where additional logic gates (and connections to existing logic gates) have been added. In this work we use it to robustly separate ICs into legitimate and counterfeit ones, even when only layout or placement of the IC has changed, without any added logic or connections. We evaluate our technique on a set of ten boards over six different counterfeit IC designs, and find that our technique tolerates manufacturing variations among different hardware instances, detecting counterfeit ICs 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, typically with some difference in terms of performance, characteristics, or materials, but which is sold or used as a legitimate (authorized) IC.\textsuperscript{1} Counterfeiting of ICs has become a major challenge for the semiconductor industry, in large part because existing test techniques and protection mechanisms are not very effective in detecting counterfeit ICs. Unfortunately, over the past few decades the problem has been getting worse, because globalization of the semiconductor supply chain has led companies to outsource many steps of their integrated circuit (IC) production cycle, and incidences of counterfeit ICs have increased rapidly. In 2015, it was reported that the illicit production of counterfeit ICs has cost IC companies $100 billion,\textsuperscript{2,3} and that this cost has steadily increased. This has been, and still is, a significant threat to the IC industry, not only because it negatively impacts innovation and economic growth, but also because it represents a serious threat/risk for systems that incorporate these counterfeit ICs. In practice, counterfeit ICs have found their way into almost all industrial sectors, including ones that are highly sensitive to potential security, reliability, and other risks: cloud infrastructure, finance, government infrastructure, military systems, etc. As a result, the need for effective detection of counterfeit ICs has increased tremendously.

Over the past few years, a plethora of papers have been published on the topic of counterfeit ICs. Such work can be roughly divided into detection and avoidance techniques.\textsuperscript{4} Avoidance aims to make counterfeits easily detectable, e.g. by adding circuitry to legitimate ICs to act as a signature/watermark,\textsuperscript{1,5,6} by fabricating different parts of the chip layout in different foundries,\textsuperscript{7} etc. However, avoidance techniques significantly add to the cost of an IC, which prevents them from widespread adoption.

In contrast, techniques for counterfeit IC detection focus on distinguishing counterfeit ICs from authentic ones, usually without adding circuitry to the IC, changing its layout, etc. Detection techniques can be based on either physical tests or electrical tests.\textsuperscript{8} 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 deviations in counterfeit ICs. These techniques include external visual inspection (EVI), X-ray imaging, resurfacing, microscopy scanning, material analysis such as X-Ray Fluorescence (XRF), Fourier transform infrared spectroscopy (FTIR), ion chromatography, etc. While physical tests can, in principle, be used to detect all types of counterfeit ICs, these more reliable tests are 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 inexpensive. However, electrical tests rely on determining whether the IC’s functionality is correct, which does not detect counterfeit ICs that have the same functionality but different layout as authentic ones. In addition, reliable detection of counterfeit ICs typically necessitates use of a number of electrical test techniques, some of which require extra circuitry to be added to the design, so the total added IC cost for all these techniques can be significant.

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 the same functionality, but different layout, with authentic ones. This includes ICs that contain hardware Trojan horses, such that the functionality of the counterfeit IC is identical to that of a legitimate IC, except under very specific conditions that are highly unlikely to be encountered during electrical testing. We choose to use backscattering side-channel, a new physical side-channel that has been demonstrated to outperform other side-channels in terms of detecting hardware Trojan horses (HTs) in ICs. However, in this work we focus on detecting changes in IC layout and placement, i.e. detecting counterfeit ICs that are functionally exact equivalents of legitimate ICs, without any additional logic gates (or connections between those gates).

The rest of the paper is organized as follows. Section 2 provides background on IC counterfeiting and the backscattering side-channel. Then, Section 3 explains our new technique for counterfeit IC detection, Section 4 describes the setup for our experimental evaluation, Section 5 presents the results of that evaluation and, finally, Section 6 concludes the paper.

2. BACKGROUND

2.1 Counterfeit ICs

A counterfeit IC is an illicit copy of a legitimate chip, typically different in terms of performance, characteristics, and/or material, but which is sold and/or used as if it was a legitimate (authorized) IC. There are three major categories of counterfeit ICs: 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 to mimic more expensive (e.g. higher-rating) ICs, etc. The second group includes out of specification ICs, ICs that were rejected during manufacturing tests but sold as normal ones, and ICs that have been tampered during manufacturing (e.g. to infect them with a hardware Trojan horse). 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 electromagnetic (EM) side-channels has been used extensively for attacks and, more recently, as a way to monitor software execution, the use of the backscattering side-channel is still relatively new. The EM side-channel is a consequence of electric current inside an electronic circuit, which occurs in short bursts to charge/discharge parasitic capacitances during switching in logic gates (this also creates power, acoustic, thermal, and several other side-channels). In contrast to these, the backscattering side-channel is a consequence of the circuit’s transistors having different impedances while they are in different states, which affects the reflection coefficient of the IC as seen by an incoming EM wave (i.e. the ICs radar cross-section), thus causing the circuit-state-dependent modulation of the reflected (back-scattered) signal.

For example, Fig. 1 (a) demonstrates a 2-input CMOS NAND gate. When the input voltages of the NAND circuit are low (logical 0), the NMOS transistors are off and the PMOS transistors are on. A low-impedance path exists between \( V_{out} \) and \( V_{DD} \), as shown in Fig. 1 (b), causing the output voltage to be close to \( V_{DD} \) (logical 1). While the impedance between the output and \( V_{DD} \) is small, it is not zero, and for clarity in this example we label it as a resistance \( R_1 \). Similarly, a high input voltage result in a low output voltage, as shown in Fig.1(c),
resulting in \( R_0 \) between the output and the ground. The values of \( R_0 \) and \( R_1 \) are typically different.\(^{21}\) In other words, from the perspective of the circuit’s output and the power supply, switching between the NAND logic’s high output state (\( R_1 \)) and low output state (\( R_0 \)) creates impedance variation. Similar reasoning can be applied to CMOS NOR, NOT, and other logic gates, and a digital circuit is composed of connecting the gates’ outputs to other gates’ inputs. 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, called backscattering side-channel. Like the traditional EM side-channel, the backscattering side-channel has a high bandwidth. However, unlike the EM side-channel, the magnitude of the backscattered signal can be increased as needed (by changing the transmitted signal), 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 IC chip.

![Diagram of CMOS NAND gate and its two equivalent impedance circuits](image)

**Figure 1.** CMOS NAND gate (a) and its two equivalent impedance circuits (b).

### 3. A NOVEL APPROACH FOR COUNTERFEIT DETECTION

#### 3.1 Using Backscattering Side Channel for Counterfeit IC Detection

Nguyen et al.\(^{13}\) have shown that HTs can be detected by analyzing how impedance changes during a clock cycle, and comparing the impedance changes of the IC under test to those of a “golden” IC (i.e. an IC that is known to be free of HTs). As pointed out by Nguyen et al.,\(^{13}\) however, time-domain reception/recording of the backscattered signal for this would require extremely high bandwidth (many times the clock frequency of the IC), and the signal distortion due to radio-frequency (RF) noise, quantization noise (ADC resolution), imperfect synchronization (jitter), etc. would make it difficult to identify small changes caused by the presence of a stealthy HT. Thus, the backscattered signal was instead measured in the frequency domain, at multiple harmonics of the ICs clock frequency, which directly correspond (through Discrete Fourier Transform) to time-domain samples during the clock cycle. These frequency-domain measurements can be very accurate because they can be collected as separate frequency-bin measurements (with slower but more accurate ADCs), averaged over many clock cycles to reduce the impact of RF noise and jitter.

In these measurements, the change caused by an HT will be reflected in backscattered signals at the harmonics of the circuit’s clock frequency: \( f_{\text{carrier}} \pm f_c \), \( f_{\text{carrier}} \pm 2f_c \), etc. The first clock harmonics at \( f_{\text{carrier}} \pm f_c \) follow 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.

In the time domain, when state changes caused by HTs become briefer in duration, the corresponding frequency-domain changes at the clock harmonics become smaller in magnitude and shift to higher harmonics. Compared to lower harmonics, the higher harmonics tend to be affected more by noise, clock jitter, and other measurement impairments. One of the key reasons why the backscattering side-channel is highly suitable for HT detection (compared to traditional analog side-channels such as EM and power) is that impedance changes, once they happen, persist for the rest of the cycle, so larger changes in impedance tend to affect relatively low harmonics of the clock in the backscattered signal. In contrast, the current bursts that create EM and power signals are already very brief, and the presence of an HT tends to change the magnitude of these bursts and/or shift their timing within the cycle, so the presence of the HT tends to only affect higher harmonics of the clock.
In this work, we follow the measurement approach of Nguyen et al.\textsuperscript{13} Specifically, 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. Thus, 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: $\mathbf{h} = [h_1, h_2, ..., h_{N-1}, h_N]$, where $h_j$ is the amplitude of the $j^{\text{th}}$ harmonic of the clock. These vectors will be used as inputs for our clustering algorithm that identifies counterfeit ICs, even when they only differ from legitimate ICs in layout or placement on the chip, without any additional gates (or connections among gates) compared to the original IC. This is in contrast to Nguyen et al.,\textsuperscript{13} which focused on detection of HTs, i.e. on detection of IC where additional logic gates (and connections to existing gates) are present.

### 3.2 One-Class-Classification to Detect Counterfeits

In this section, we introduce our one-class-classification technique that accurately detects whether or not the IC’s layout is the legitimate one. The approach is based on supervised learning techniques which contain two phases: Training and testing. In the training phase, the goal is to obtain parameters of the cluster that corresponds to back-scattered signals at clock-frequency harmonics for a legitimate IC. The testing phase then determines whether measured back-scattered signals for an IC-under-test map within or outside of the legitimate-IC cluster.

To achieve our goal, we first collect magnitudes of the first $N$ harmonics as described in Section 3.1. An example of the original ICs 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

\[
\hat{h}[k] = \frac{1}{M} \sum_{m=1}^{M} h_m[k]
\]

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

\[
h_m = [h_m^0, h_m^1, \ldots, h_m^N]
\]

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

\[
M_D = \max \left\{ \text{abs} \left( H - 1\hat{h}^T \right) \right\}
\]
where max\{\bullet\} returns a row vector which contains the maximum values at each column of its argument, \(\text{abs}\{\bullet\}\) 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.

Density estimation, a common method in one-class-classification, works better with large numbers of measurements.\(^{22}\) 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

\[
[\hat{h}[m] - 3 \cdot M_D[m], \quad \hat{h}[m] + 3 \cdot M_D[m]].
\]

(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:

\[
H = (H - \hat{1}h) ./ (1M_D)
\]

(5)

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, when testing an IC, the layout is called

- \textit{Original} if all measured harmonics are within the corresponding confidence intervals,
- \textit{Counterfeit} if there is at least one harmonic value which violates its confidence interval.
4. BENCHMARK IMPLEMENTATION AND EXPERIMENT SETUP

4.0.1 Counterfeit IC Benchmark Implementation

For our experimental evaluation, 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.

![Measurement setup for counterfeit IC detection using backscattering side-channel.](image)

4.0.2 Experimental 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 connected to an Agilent MXG N5183A signal generator, and a receiver AAronia H2 magnetic field near-field probe connected to an Agilent MXA N9020A spectrum analyzer. The devices-under-test (DuT) are Altera DE0 Cyclone V FPGA boards. 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 sinusoidal 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 functionality the same. The results are given in Figure 6. Recall that the harmonic values and the boundaries are normalized based on
the equation given in (5). During testing, in this equation $H$ is a matrix, where each row is a test measurement, $\hat{h}$ 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) shows the harmonics for the original layout, along with confidence intervals obtained from training. We observe that all considered harmonics are within the confidence interval, and thus all these measurements are labeled as corresponding to the original IC (0% false positive rate).

Next, we experiment with counterfeit layouts, which have the same functionality as the original IC but different layout. Figures 6(b), 6(c) and 6(d) illustrate the harmonic values for different counterfeits of this kind. We observe that all of the measurements have at least one component that violates the confidence intervals (100% true positives, i.e. 100% accuracy in detecting counterfeits of this kind).

![Graph](image1.png)

(a) Normalized harmonic magnitudes of the original circuit.

![Graph](image2.png)

(b) Normalized harmonic magnitudes of the circuit with the same functionality and a different layout.

![Graph](image3.png)

(c) Normalized harmonic magnitudes of the circuit with the same functionality and a different layout.

![Graph](image4.png)

(d) Normalized harmonic magnitudes of the circuit with the same functionality and a different layout.

Figure 6. Normalized harmonic magnitudes of the circuit with the same functionality and different layouts.

Another experiment is performed by changing the placement location of the circuit within the chip, while keeping its functionality and layout same. Figure 7(a) shows the results of our testing for instances of the original circuit which, like in Figure 6(b), correctly labels all these instances as original. The rest of Figure 7 show the results of position-change counterfeits. For each of these counterfeits, at least one harmonic is outside of the confidence interval, causing all these counterfeits to be accurately labeled as counterfeit. Overall, for this type of counterfeit IC, our method again achieves 100% detection accuracy.

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, globalization of the semiconductor supply chain has led companies to outsource much of the production cycle for integrated circuits (ICs). While outsourcing helps companies significantly reduce their cost and time-to-market, it also introduces concerns about the trustworthiness of an IC. One of the most serious
problems is counterfeiting of ICs, which not only negatively impacts innovation and economic growth of the IC industry, but also creates serious threats and risks for systems that incorporate those counterfeit ICs. This paper proposed a novel method that uses the backscattering side-channel to cluster ICs such that counterfeits are separated from legitimate ICs. In this work, we utilized backscattering side-channels to robustly separate ICs into legitimate and counterfeit ones, even when only layout or placement of the IC has changed, without any added logic or connections. We evaluated our technique on a set of ten boards over six different counterfeit IC designs, and found that our technique tolerates manufacturing variations among different hardware instances, detecting counterfeit ICs with 100% accuracy and 0% false positives.

7. ACKNOWLEDGMENTS

This work has been supported, in part, by NSF grants 156399, 1651273 and 1740962, DARPA LADS contract FA8650-16-C-7620, and ONR grants N00014-17-1-2540 and N00014-19-1-2287. The views and findings expressed in this paper are those of the authors and do not necessarily reflect the views of NSF, DARPA, and ONR.

REFERENCES


