Chapter 3
Hardware Trojan Detection: Untrusted Manufactured Integrated Circuits

Trojan circuits are designed to avoid detection, triggering only under rare conditions. Trojans are silent most of their lifetimes and have a very small size, relative to their host circuits, and make only limited contributions to circuit characteristics. These qualities suggest that they most likely connect to nets with low controllability and/or observability [1,4].

Hardware Trojan detection depends on the Trojan's full or partial activation. In the full activation scenario, circuit functionality deviates from the genuine specifications, and the Trojan can cause catastrophic failures. In the partial activation scenario, however, the Trojan impacts the circuit power profile or its delay characteristics. Several techniques have been proposed to address these rare triggering conditions or to capture the impact of a Trojan on side-channel signals.

For full activation of a Trojan, in [2] circuit nets with low probabilities of "1" or "0" are distinguished. To avoid the Trojan's detection, an adversary may utilize a combination of low transition nets as a Trojan trigger to reduce the probability of activation during authentication. In this work, patterns are generated to make those nets more switch in order to increase the probability of Trojan activation. Many techniques have been also proposed based on Trojan partial activation by studying their impact on the delay or the power characteristics of circuit under authentication (CUA).

3.1 A Case Study for Hardware Trojan Detection in Integrated Circuits

The amount of current drawn by a Trojan can be so small that it submerges into the envelope of noise and process variation effects, where it cannot be detected by measurement equipment. However, Trojan detection probability can improve greatly when the current is measured locally from multiple power pads. The local current refers to the current drawn from a power port near a Trojan circuitry. The more

M. Tehranipoor et al., *Integrated Circuit Authentication: Hardware Trojans and Counterfeit Detection*, DOI 10.1007/978-3-319-00816-5_3,

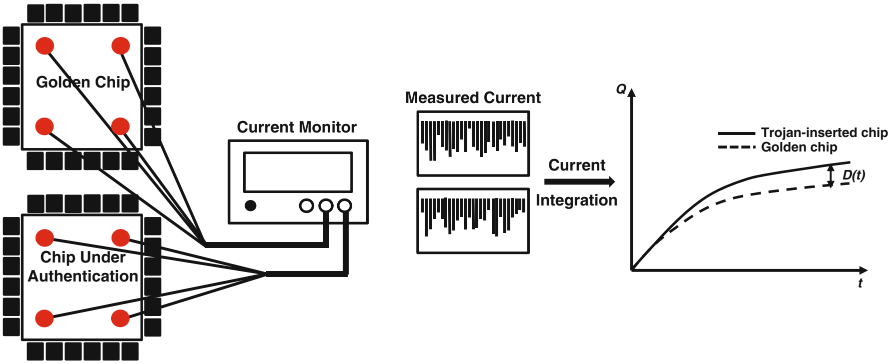

Fig. 3.1 Current integration method

instances of switching on Trojan inputs and inside Trojan circuitry, the greater the Trojan's power consumption. Since small Trojans are expected to be inserted into chips to reduce the probability of detection, the local current impact could be more significant than the global current measured from power pins.

Any partial activity in a Trojan circuit demands current. On the other hand, variations in process parameters, such as gate channel length and voltage threshold, sometimes increase or decrease the amount of circuit current consumption with different input vectors over time. Based on this fact, a current integration methodology is presented in [3] which accumulates the impact of a Trojan over the time while it is expected that the process variations impact is canceled out by integration. Figure 3.1 shows the current integration methodology for detecting hardware Trojans.

It is assumed that an adversary inserts Trojans randomly into a selected number of chips. Using an exhaustive test on a few randomly selected chips can help identify some golden chips. After identifying the golden chips, an average current waveform is formed in response to a pattern set. Next, the pattern set is applied to each CUA, and the current is measured locally via power pads or C4 bumps. Using this current integration method, the small current consumption difference between Trojan-inserted and Trojan-free circuits can be increased through the integration process. In the case of a Trojan's existence in a chip, more current difference can be measured by applying more patterns to the chip, making the Trojan detection task easier. When the current difference surpasses a pre-defined threshold, it indicates that Trojan detection and pattern application has stopped.

If $I_{Trojan-free}$ and $I_{Trojan-inserted}$ denote current drawn by Trojan-free and Trojan-inserted circuits at time t, respectively, the integrated current at time t for Trojan-free and Trojan-inserted circuits ($Q_{Trojan-free}(t)$ and $Q_{Trojan-inserted}(t)$) can be expressed by (3.1) and (3.2):

$$Q_{Trojan-free}(t) = \int_0^t I_{Trojan-free}(t).\mathrm{d}t \tag{3.1}$$

Table 3.1 Trojan characterization

Trojan	Type	Size (%)	Distribution	Structure
Counter	1-bit	0.04	Tight	No-change
	3-bit	0.10	Tight	No-change
	7-bit	0.31	Tight	No-change
	9-bit	0.42	Tight	No-change
Comparator	3-input	0.02	Loose	No-change
	5-input	0.04	Loose	No-change
	20-input	0.15	Loose	No-change

$$\begin{aligned} Q_{Trojan-inserted}(t) &= \int_0^t I_{Trojan-inserted}(t).\mathrm{d}t \\ &= \int_0^t I_{Trojan-free}(t) + I_{Trojan}(t).\mathrm{d}t \end{aligned} \tag{3.2}$$

where $I_{Trojan}(t)$ denotes the current drawn by the Trojan. Since the same pattern set is applied to both a golden chip and a CUA, the difference between $I_{Trojan-free}$ and $I_{Trojan-inserted}(t)$ comes from (I) the additional current drawn by Trojan gates and (II) changes in the circuit current due to process variations. By integrating the current along the time axis for both chips, their cumulative difference at time t, denoted by $D(t)$ in (3.3), can be increased by applying more patterns.

$$D(t) = Q_{Trojan-inserted}(t) - Q_{Trojan-free}(t) = \int_0^t I_{Trojan}(t).\mathrm{d}t \tag{3.3}$$

When $D(t)$ reaches a predefined Trojan detection threshold D_{th}, i.e. $D(t) \geq D_{th}$, then the chip is identified as a Trojan-inserted chip. It should be noted that D_{th} is determined by the Trojan detection timing budget as well as the current measurement device resolution.

The proposed current integration technique is effective for detecting both tightly- and loosely-distributed Trojans. Further, its capability does not depend on the location of a Trojan in a circuit, since current is measured locally through power pads or C4s. Since there are a large number of considerable power pads on the power distribution network, a Trojan circuitry impacts at least one power pad.

The current integration technique is used to detect Trojans inserted into the s38417 benchmark. First seven layouts of the original s38417 benchmark are generated using Synopsys physical design tools in the 180nm technology node. A 1-bit, 3-bit, 7-bit, 9-bit counter and 3-input, 5-input, 20-input comparator Trojan is inserted into each of these seven layouts (i.e. only one Trojan in each layout). Table 3.1 shows the type, size, distribution, and structure of the Trojans. The impact of variations in voltage threshold (V_{th}), channel length (L), and oxide thickness (T_{ox}) on Trojan detection is investigated as well. Table 3.2 shows the applied process variations.

Table 3.2 Process variations applied during trojan detection

Parameter	Inter-die (%)	Intra-die (%)
Threshold voltage (V_{th})	5	20
Channel length (L)	2	8
Oxide thickness (T_{ox})	1	4

Figure 3.2a shows the simulation results obtained using Synopsys Nanosim for s38417 containing the Trojan 9-bit counter. The patterns are applied with a frequency of 100 MHz. As seen in the figure, after applying 700 clock cycles (7us pattern application time), $D(t) > 1 \times 10^{-9}$, which is easily detectable using measurement devices. In fact, for such a Trojan, depending on D_{th}, a shorter application time can be even sufficient for detection. Shown in Fig. 3.2b, the results obtained for s38417 with a 7-bit counter also confirm that such a Trojan can be easily detected. In general, detecting a counter is easier than a combinational Trojan since a counter continuously receives the clock and consumes power. No process variations are considered for the results shown in Fig. 3.2, although the process variations would not be significant enough to change the detection outcome for such Trojans.

Figure 3.3a,b show the simulation results for the circuit with 3-bit and 1-bit counters, respectively, without considering process variations. The results imply that the smaller Trojans consume significantly lower power (i.e. current) which makes their detection more difficult. Note that Trojan detection depends on two important factors: (1) process variations and (2) the resolution of measurement device. It is believed that process variations are a limiting factor in Trojan detection.

Figure 3.4 shows the simulation results for the circuit containing a 3-bit counter considering the worst case process corners for both Trojan-inserted and Trojan-free circuits. The process corner used in the Trojan-free circuit increases the current within the circuit while the process corner used in the Trojan-inserted circuit reduces the total current. This is done to evaluate the efficiency of the technique in detecting the Trojan. The Trojan-inserted circuit with process variations still consumes more current compared with the Trojan-free circuit with process variations. However, detecting smaller Trojans, such as 2-bit and 1-bit counters, is not possible considering worst-case process variations. As seen in Fig. 3.3b, the current difference between Trojan-free and Trojan-inserted circuits containing 1-bit counters is negligible. The existence of process variations makes Trojan detection even more difficult.

Figure 3.5 shows simulation results for the 20-input comparator Trojan inserted in the s38417 benchmark after applying 300 random patterns. As seen in the figure, the 20-input comparator can be easily detected, even in presence of the two process corners. However, the results shown in Fig. 3.6 demonstrate the difficulty of detecting the 5-input comparator even without process variations. This is the case for the 3-input comparator Trojan as well. To further increase the probability of detection, more test patterns should be applied. The application time depends upon where the Trojan-inserted circuit's results fall outside of the results of Trojan-free

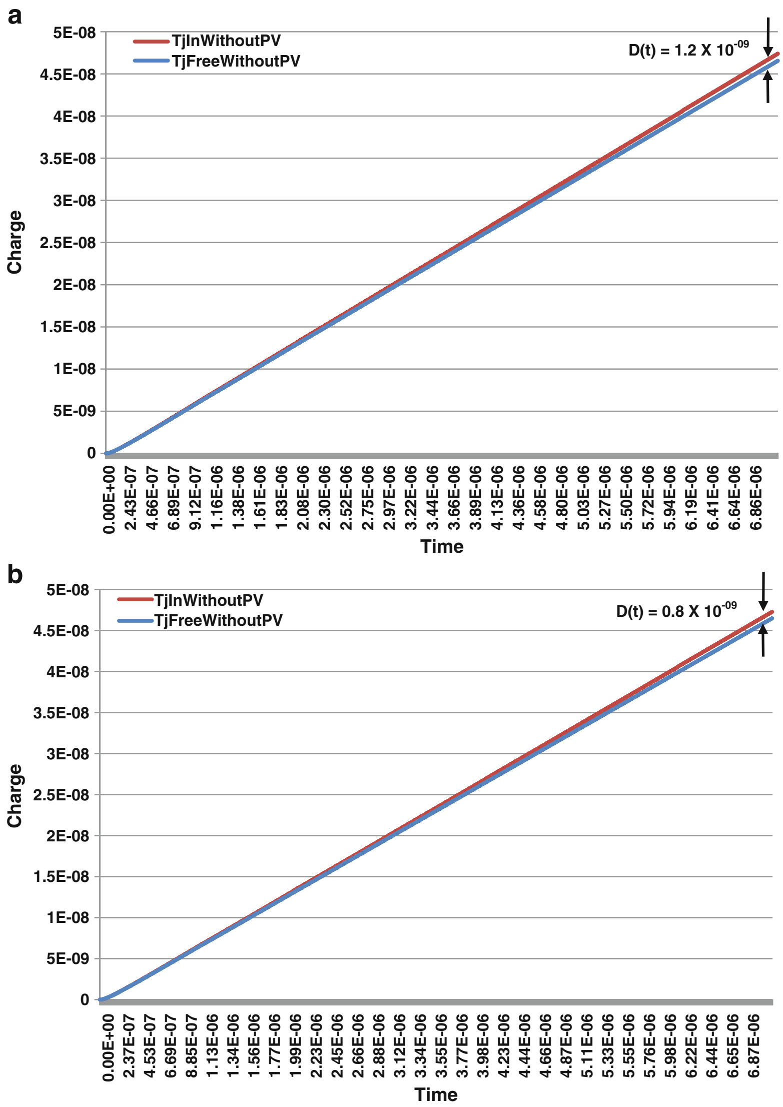

Fig. 3.2 Current integration for s38417 with 9- and 7-bit counters. (**a**) 9-bit counter. (**b**) 7-bit counter

circuit while considering process variations. The total number of patterns required to detect such small Trojans can be estimated from the results shown in Fig. 3.6 based on D_{th}.

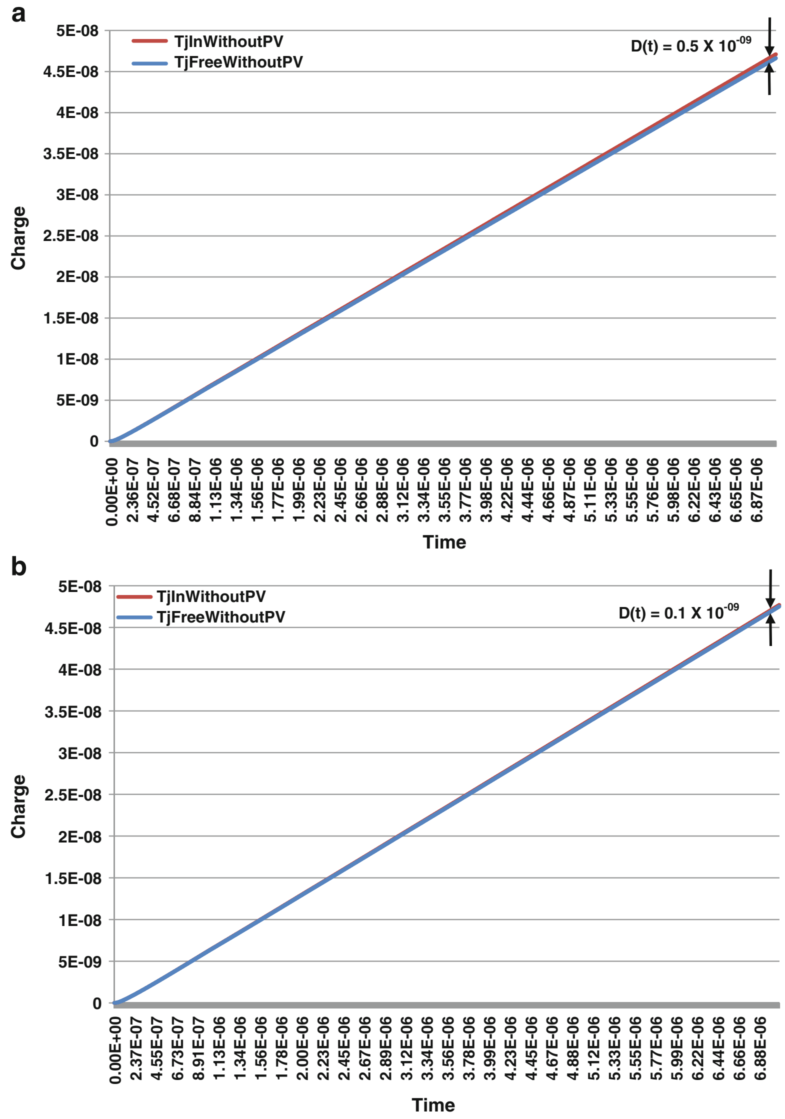

Fig. 3.3 Current integration for s38417 with 3- and 1-bit counters. (**a**) 3-bit counter. (**b**) 1-bit counter
]

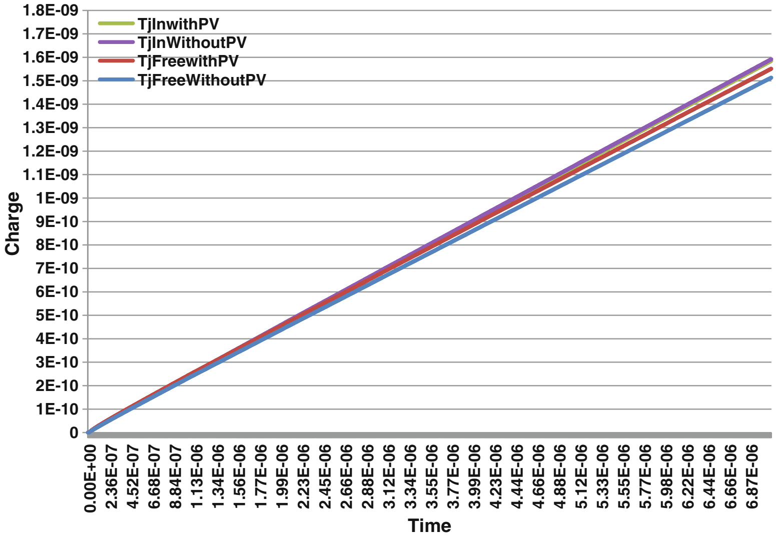

Fig. 3.4 Current integration for s38417 with 3-bit counter considering the two process corners

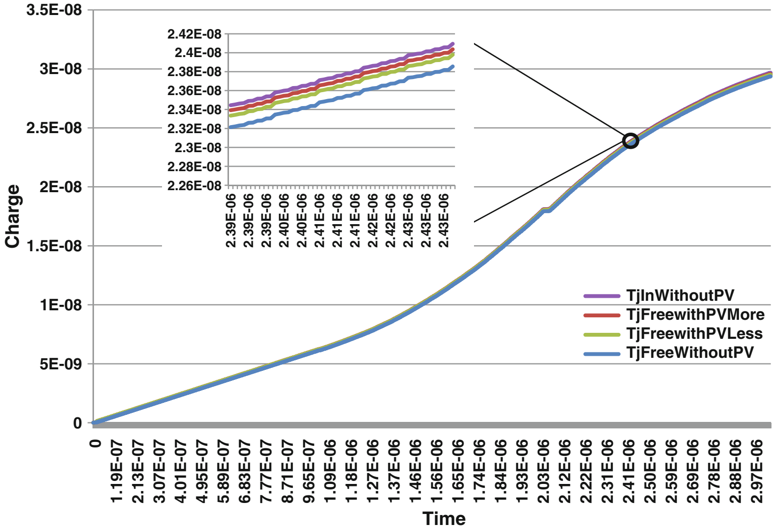

Fig. 3.5 Current integration for s38417 with 20-input comparator and two process corners

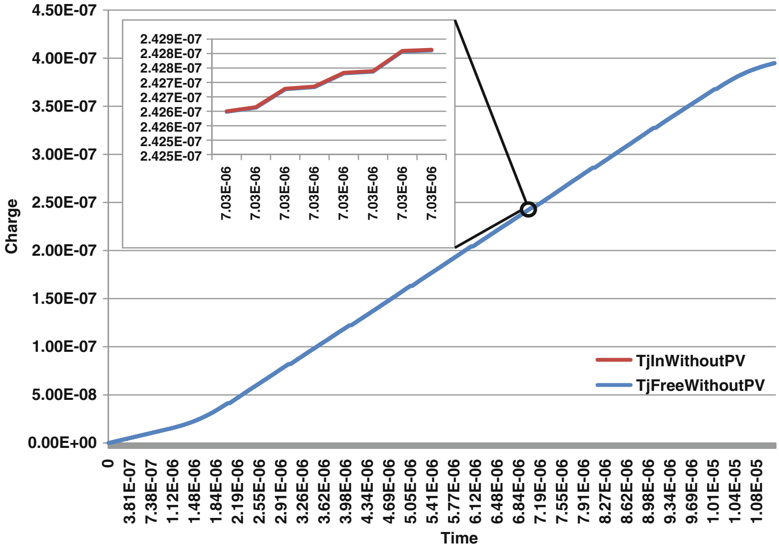

Fig. 3.6 Current integration for s38417 with 5-input comparator and no process variations

3.2 Summary

The current integration technique for Trojan detection and isolation was presented. The technique measures the current locally from the on-die power pads. Comparing the results obtained for a golden chip against a CUA, it can be seen that Trojans can be detected if the current integration results for the CUA fall outside that of the golden chip. It was shown that the technique can easily detect Trojans as small as 0.1 % within the circuit area.

References

1. M. Banga and M. S. Hsiao, "A Novel Sustained Vector Technique for the Detection of Hardware Trojans," in Proc. of the *International Conference on VLSI Design*, pp. 327–332, 2009.
2. F. Wolff, C. Papachristou, S. Bhunia and R.S. Chakraborty, "Towards Trojan-Free Trusted ICs: Problem Analysis and Detection Scheme," in Proc. of the *Design, Automation and Test in Europe(DATE '08)*, pp. 1362–1365, 2008.
3. X. Wang, H. Salmani, M. Tehranipoor and J. Plusquellic, "Hardware Trojan Detection and Isolation Using Current Integration and Localized Current Analysis," in Proc. of the *International Symposium on Fault and Defect Tolerance in VLSI Systems (DFT08)*, pp. 87–95, 2008.
4. X. Wang, M. Tehranipoor and J. Plusquellic, "Detecting Malicious Inclusions in Secure Hardware: Challenges and Solutions," in Proc. of the *IEEE International Workshop on Hardware-Oriented Security and Trust (HOST 2008)*, pp. 15–19, 2008.