Visible Hardware Security Techniques

Mehrdad Zaker Shahrak*
Department of Computer Science and Eng.
University of Nebraska-Lincoln

Sheng Wei †
Department of Computer Science and Eng.
University of Nebraska-Lincoln

Abstract

The most recent cyber security attacks have raised significant public concerns about hardware security against physical attacks. In particular, with increasing tendency of connecting various devices, or more specifically hardware systems to the Internet, any vulnerabilities of the hardware would become more exposable to high probability of being attacked. A general procedure of detecting such hardware physical attacks, such as hardware Trojans, is to check determined side channels of the hardware, such as power, delay, temperature and electromagnetic fields. However, it is very difficult to identify between the normal power variation and the one caused by Trojans, which may result in high rates of false positives/negatives in the detection results. To deal with the problems, we develope a big data visualization-based method to effectively identify the malicious behavior or hardware Trojans. In particular, we organize the side channel analysis of hardware and visualize it in a meaningful manner for a security analyst. In this way, it incorporates human intelligence factor into the detection process and thus addresses the issue of identifying malicious from non-malicious components in the side channels. Also, by using big data analysis, we transform the sampled power data into spatially visible forms, which enables the diagnosis of hardware Trojan locations.

Index Terms: Trojan detection, big data visualization, integrated circuit spatial correlations

1 Introduction

The most recent cyber security attacks have raised significant public concerns about hardware security against physical attacks, e.g., the July 2015 Chrysler car hijacking attack that results in the recall of 1.4 million vehicles [1][2][3]. In particular, with the ever increasing trend of connecting various devices, appliances, or other hardware systems to the public Internet, any vulnerabilities or security backdoors in the hardware would become more exposable to physical attacks, resulting in a high potential of even more severe attacks. Even though issuing a firmware update or a recall may help address some of the existing security vulnerabilities, a deeper level of hardware security concern was originated from the fact that hardware design companies often outsource the manufacturing tasks to the foundries overseas for cost reduction. Since there exist untrusted/malicious foundries who have full access to the hardware during manufacturing, there is a potential risk of security breach, which may enable future cyber security attacks.

A typical method of detecting such hardware physical attacks, such as hardware Trojans, is to monitor certain side channels of the hardware, such as power, delay, temperature and electromagnetic fields. For example, we can measure the leakage power of the target circuit and detect anomalies possibly caused by Trojans. While it eliminates the huge cost of physical hardware inspection, the challenge in side channel-based approach is that the side channels are often unstable. Therefore, in practice, it is very difficult if not impossible to distinguish between the normal power variation and that caused by Trojans, which may result in high false positives/negatives in the detection results.

This has been a long-standing problem in the hardware security community. Even though we may leverage techniques like circuit partition or variation/error modeling [4][5][6] to scope the natural variation and make it more differentiable from the malicious variation, there are often unrealistic assumptions being adopted in this process regarding the natural variation in side channels and, more importantly, it is extremely difficult to incorporate all these variations and the complex inter-correlations between them in order to draw a final conclusion.

To address the problems, we for the first time propose a big data visualization-based method to effectively identify the malicious behavior and thus hardware Trojans in the real time via side channels. Our key idea is to bring human expertise and feedback into the detection process to address the aforementioned complex problem of uncertainty. In particular, we organize and visualize the side channel traces of hardware in an efficient way and present it to a security analyst who has rich expertise in both security and hardware domains. In this way, we can achieve more robust and accurate detection results. Obviously, in this case the data to be visualized is huge in size in the sense that it could involve very frequent side channel samples (e.g., every few seconds or less), a combination of multiple data sources (e.g., sampling power, delay, and temperature at the same time), and different types of data (e.g., sampling both power consumption and application behavior like CPU usage). We address the challenge by leveraging a multi-level exploration-based visualization approach, where multiple types of events from multiple data sources are well visualized and presented to the security analysts for a comprehensive analysis. More importantly, we further look into the spatial correlations among power consumption in different regions of the circuit and use visualization-aided consistency analysis to identify abnormal behavior possibly caused by the hardware Trojans. The spatial correlation-based visual analysis could also enable us to determine the location of the Trojan, solving the challenging hardware Trojan diagnosis problem.

2 Related Work

Many research efforts have been made on Trojan detection analysis with the help of side channel analysis technique, such as design and(or) analysis of physical unclonable functions [7][8][9]. However, these researches are limited with the number of non-malicious factor and variations are involved in the measurement of side channels and, more importantly, the malicious factor is often intentionally hidden by the attackers and have extremely confined and correlated representations in the observable signals. In addition, the existing approaches have not delt with the hardware Trojan diagnosis problem which locates the Trojan, due to the huge and complex spatial correlations in the integrated circuit.

The proposed visible hardware security approach is capable of addressing both of the gaps and provide an answer for them as described below. (1) It includes human intelligence factor in the detection process which is by far the most effective way to address

[*] e-mail: mehrdad@cse.unl.edu
[†] e-mail:shengwei@unl.edu

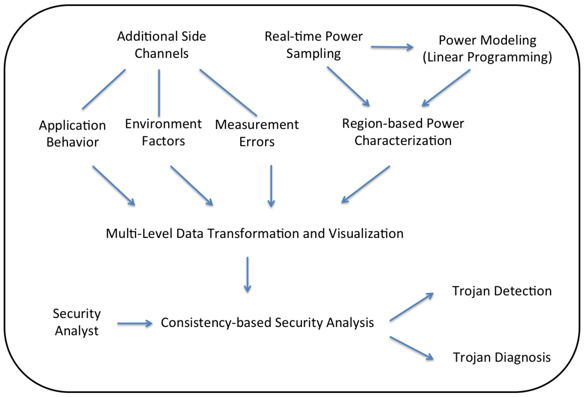
Additional Side Channels
Real-time Power Sampling
Power Modeling (Linear Programming)
Application Behavior
Environment Factors
Measurement Errors
Region-based Power Characterization
Multi-Level Data Transformation and Visualization
Security Analyst
Consistency-based Security Analysis
Trojan Detection
Trojan Diagnosis

Figure 1: Flow of visible hardware Trojan detection and diagnosis.

the issue of identifying malicious from non-malicious components in the side channels; and (2) It uses big data analysis that transforms the sampled power data into spatially visible forms, which directly simplifies the diagnosis of hardware Trojan locations.

Also, the proposed research has the potential of visualizing hardware security vulnerabilities via side channels without physically examining the hardware itself. The only requirement for measuring the power consumption of the target hardware is inexpensive and widely available power measuring equipments, together with high performance region-based power computations and data transformations for visualization. The success of the project will enable a real-time digital scan of the hardware for security screening with no intrusion to the hardware and very low costs in timing and area.

3 Visible Hardware Security Techniques

We develop a systematic approach for visible hardware Trojan detection and diagnosis, including power sampling, characterization, and visualization, as shown in Figure 1. The approach applies to the post-silicon testing phase of the integrated circuit, in order to identify potential Trojans embedded by the untrusted foundry. In particular, there are four critical steps in our approach as presented in the following subsections.

3.1 Real-time power sampling

We first conduct continuous power sampling on the target circuit at the interval of milliseconds while the system is operation. The sampled power values serve as a data stream and source to the modeling and characterization process. We conduct the power experiments based on the FPGA implementation of hardware Trojan benchmarks available on TrustHub [10], which is widely used by the hardware security community for the evaluations of hardware Trojan research.

3.2 Power modeling and region-based characterization

Given the sampled power stream, we model the total power consumption and that of the particular regions using a linear programming approach, which enables us to derive the regional power values to visualize. The linear power model is determined by the fact that the total power consumption of the chip is the sum of the power consumption in all the regions, plus additional variations such as measurement errors. Then, a linear programming-based approach can be adopted to determine the regional power values while varying the input signals to the chip and creating various forms of equations. Based on our prior work in this domain [5][8] , in this research, we specifically aim to significantly improve the performance of the modeling and characterization process, with a goal of adapting it to the high power sampling frequencies and real-time requirement in identifying Trojans.

3.3 Data transformation and multi-level visualization

We transform the data from multiple sources, including the regional power consumption, as well as additional side channels that may impact the power consumption, such as application behavior, environment factors, and measurement errors. Then, we employ the multi-level based approach to visualize the multiple sources of data and present it to the security analyst. The visualized data will show not only the abnormal variations caused by various signals but also the spatial correlations of the variations with certain regions of the chip, enabling the diagnosis of the Trojan locations if there are any.

3.4 Visualization-based correlation and consistency analysis

While viewing the multi-level visible results, the security analyst conducts consistency and correlation analysis combining the visible data and his/her domain expertise. Two types of conclusions will be drawn in the real time based on the security analysis: (1) hardware Trojan detection, which indicates whether there is any hardware Trojan embedded in the target chip; and (2) hardware Trojan diagnosis, which indicates the region where the Trojan is located, in order to provide evidence for hardware physical attacks or mask the Trojan infected region for further processing and mitigation.

4 Conclusion

In summary, by introducing big data analysis and visualization techniques into the hardware security domain, we take advantage of the human expertise in security screening and anomaly detection, which helps improving the accuracy of the detection. Based on the goal of the proposed approach, we plan to present the following technical content and results in the poster: (1) a real-time sampling, modeling, and data transformation flow to generate the power streams for visualization; (2) a multi-level, multi-phase visualization mechanism to identify the Trojan infection and its specific region; and (3) a case study or user survey to evaluate the proposed approach.

References

[1] Chrysler recalls 1.4 million hackable cars, CNN Money, July 24, 2015, http://money.cnn.com/2015/07/24/technology/chrysler-hack-recall/

[2] The Ashley Madison hack explained, August 21, 2015, http://www.wired.com/2015/08/ashley-madison-hack-everything-you-need-to-know-your-questions-explained/

[3] Target Hackers Broke in via HVAC Company, February 2014, http://krebsonsecurity.com/2014/02/target-hackers-broke-in-via-hvac-company/

[4] M. Banga, M. Hsiao, A Region Based Approach for the Identification of Hardware Trojans, IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), pp. 40-47, 2008.

[5] S. Wei, M. Potkonjak, Self-Consistency and Consistency-Based Detection and Diagnosis of Malicious Circuitry, IEEE Transactions on Very Large Scale Integration Systems (TVLSI), Vol. 22, No. 9, pp. 1845-1853, 2014.

[6] S. Wei, S. Meguerdichian, M. Potkonjak, Gate-level Characterization: Foundations and Hardware Security Applications, Design Automation Conference (DAC), pp. 222-227, 2010.

[7] J. Li, J. Lach, At-Speed Delay Characterization for IC Authentication and Trojan Horse Detection, IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), pp. 8-14, 2008.

[8] S. Wei, J. Wendt, A. Nahapetian, M. Potkonjak, Reverse Engineering and Prevention Techniques for Physical Unclonable Functions Using Side Channels, Design Automation Conference (DAC), pp. 1-6, 2014.

[9] M. Potkonjak, S. Meguerdichian, A. Nahapetian, Sheng Wei, Differential Public Physically Unclonable Functions: Architecture and Applications, Design Automation Conference (DAC), pp. 242-247, 2011.

[10] The TrustHub, https://www.trust-hub.org/