Detection Technique for Hardware Trojans Using Machine Learning in Frequency Domain

*Takato Iwase, Yusuke Nozaki, Masaya Yoshikawa
Dept. of Information Engineering
Meijo University
1-501 Shiogamaguchi Tenpaku Nagoya Aichi Japan
{153430005, 143430019}@ccalumni.meijo-u.ac.jp,
masay@meijo-u.ac.jp

Takeshi Kumaki
Dept. of Electronic and Computer Engineering
Ritsumeikan University
1-1-1 Nogihigashi Kusatsu Shiga Japan
kumaki@fc.ritsumei.ac.jp

***Abstract*—Recently, the threat of hardware Trojan has been highlighted. A hardware Trojan is a hardware virus. When predetermined conditions are satisfied, that malicious virus performs subversive activities, such as a system shutdown and the leaking of important information, without the circuit users even being aware of that activity. Therefore, it is important to detect the consumer electronic devices with hardware Trojans from a viewpoint of security. This study proposes a new detection technique for hardware Trojan. The proposed method introduces machine learning for the detection. Experiments using actual devices prove the validity of the proposed method.**

Keywords—Hardware Trojan, Detection technique, Machine learning, Security module;

I. INTRODUCTION

Large-scale integrated circuits (LSIs) are mounted on almost all consumer electronic devices, including mobile phones and personal computers (PCs). Anew LSI is released every few months. Therefore, the development cycle of an LSI must be shortened. To realize the short development cycle, intellectual propertied belonging to other companies are partially used.

In general, a company does not perform all the processes of developing an LSI circuit from designing to manufacturing, but different companies are in charge of different processes. It means that many unspecified persons are involved in the development of an LSI circuit. Subsequently, the risk of a hardware Trojan has been pointed out [1] [2].

A hardware Trojan is defined as a circuit that is covertly incorporated into an LSI circuit when it is designed, manufactured, or shipped by malicious third parties or traitorous engineers. This circuit possesses an unintended function of the developer of the LSI circuit. When predetermined conditions specified by an attacker are satisfied, a hidden function is activated and extensive damage occurs. Therefore, it is important the detection of hardware Trojans. This study proposes a new detection method for hardware Trojans. Experiments using actual devices prove the validity of the proposed method.

II. PROPOSED METHOD

The proposed method utilizes the difference of power consumption between with Trojan and without Trojan. Fig.1 shows examples of the power consumption waveforms of with Trojan and without Trojan.

In the power consumption waveform data, the sample point necessary for the detection is shifted in the voltage value axis direction as shown in Fig.1. Therefore, the proposed method converts the power consumption waveform data from time domain into frequency domain using discrete Fourier transform (DFT) [3]. Fig.2 shows the converted results of power consumption waveforms.

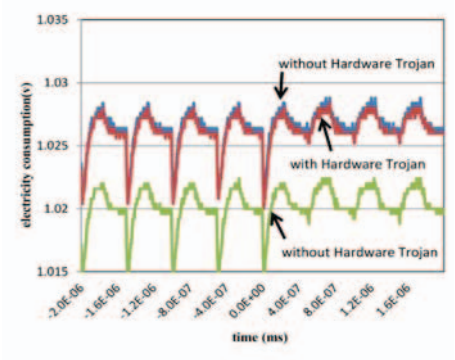

Fig. 1. Comparison of power consumption between with hardware Trojan and without

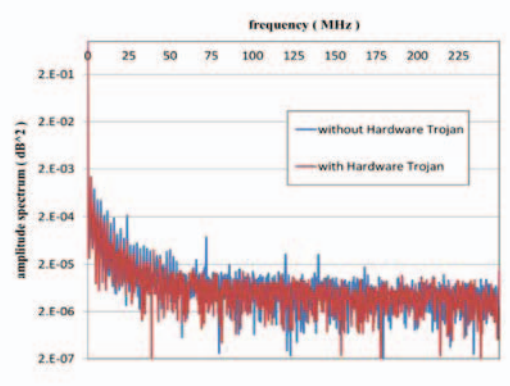

Fig. 2. Comparison of converted power consumption waveforms

978-1-4799-8751-1/15/$31.00 ©2015 IEEE

Regarding the detection, the proposed method introduces a classification technique using machine learning by support vector machine (SVM) [4].

In the learning phase, converted power consumption waveform data is used as training data. Fig.3 shows the proposed detection method based on SVM.

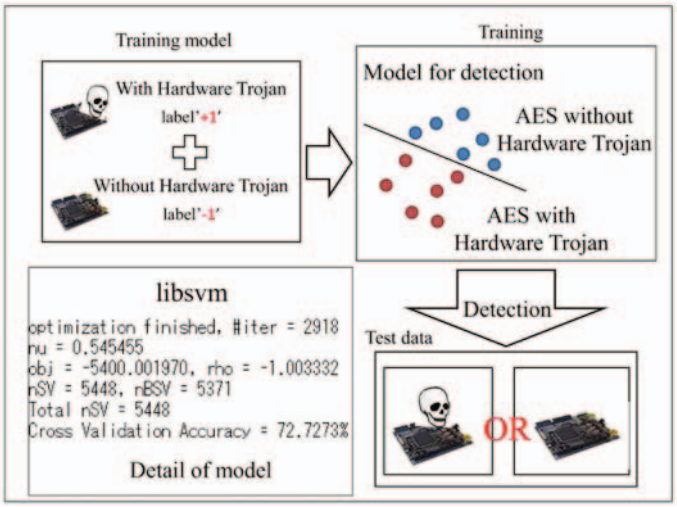

Fig. 3. Proposed method

III. EXPERIMENTS

A. Target hardware Trojans

In many studies that have been previously reported on hardware Trojans. Trojans are incorporated into cryptographic circuits. The reason for this is that because cryptographic circuits handle various types of data, including personal information and trade secrets, attackers often target cryptographic circuits. In this study, 12 kinds of hardware Trojans are developed as detection targets. All hardware Trojans are incorporated into AES which is encryption standard. Table1 shows the developed hardware Trojans.

TABLE I. PROPOSED HARDWARE TROJANS

Triggers	Events
Input round	Skip encryption
Input particular plain text	Twice encryption
Input round and input particular plain text	Rewriting register
	Outputting plain text

B. Experimental Results

In order to evaluate the proposed detection method, experiments were performed. In the evaluation experiments, AES, into which the hardware Trojan was incorporated, was described using Verilog-HDL, and an oscilloscope was used to measure power consumption. The experiments introduce 12 kinds of Trojans.Fig.4 shows the evaluation system.

Table2 shows the results obtained by the evaluation system. In Table2,"✓" indicates the case of which the proposed method detects the Trojan. The proposed method can detect all kinds of hardware Trojans when they were incorporated. It can also detect a normal circuit without hardware Trojan.

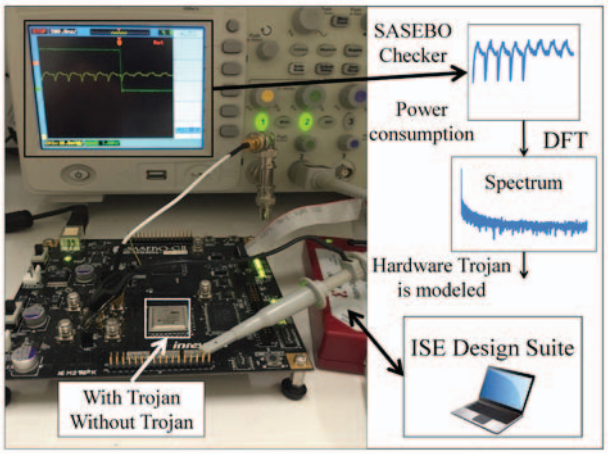

Fig. 4. Evaluation system

TABLE II. DETECTION RATE OF HARDWARE TROJANS

# Trojan	#1	#2	#3	#4	#5	#6
Detection	✓	✓	✓	✓	✓	✓

#7	#8	#9	#10	#11	#12	#without
✓	✓	✓	✓	✓	✓	✓

IV. CONCLUSION

This study proposed a new detection technique for hardware Trojan which is based on SVM. The proposed method utilized power consumption waveform data which are converted from time domain into frequency domain in classification phase on SVM. Experiments using FPGA, proved the validity of the proposed method.

Future works include other application circuits.

REFERENCES

[1] M.Yoshikawa, R.Satoh, T.Kumaki, "Hardware Trojan for Security LSI", Proc. of IEEE International Conference on Consumer Electronics, pp.31-32, 2013.

[2] N.Jacob, D.Merli, J.Heyszl, G.Sigl "Hardware Trojans: current challenges and approaches", IET Journals & Magazines Computers & Digital Techniques, pp.264-273, 2014.

[3] C.He, B.Hou, L.Wang, Y.En, S.Xie, "A failure physics model for hardware Trojan detection based on frequency spectrum analysis", IEEE International Reliability Physics Symposium, PR.1.1-1.4, 2015.

[4] B.M. Sherin, M.H.Supriya, "Selection and parameter optimization of SVM kernel function for underwater target classification", Proc. of IEEE Conference Publications, pp.1-5, 2015.