**IMPLEMENTATION OF SYMMETRIC ALGORITHM MODIFICATION SYSTEM TO RESIST POWER BASED SIDE CHANNEL ATTACKS**

Project Id: 17-044

Project Proposal Report

Pathirana K.P.A.P Lankarathne L.R.M.O Hangawaththa N.H.A.D.A

SUPERVISOR

…………………………………………………..

Mr. Kavinga Yapa Abeywardena

CO-SUPERVISOR

…………………………………………………..

Mr. Nuwan Kuruwitaarachchi

B.Sc. Special (Honors) Degree in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

March 2017

**DECLARATION**

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

|  |  |  |
| --- | --- | --- |
| Name | Student ID | Signature |
| Pathirana K.P.A.P | IT14120930 |  |
| Lankarathne L.R.M.O | IT14116216 |  |
| Hangawaththa N.H.A.D.A | IT14100376 |  |

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

.....………………………… ……………………

Signature of the supervisor: Date

**ABSTRACT**

Side Channel attacks are an easy way to launch a powerful attack against cryptographic modules by examining its physical specificities without analyzing cipher text or plain text. Sensitive information of a cryptographic module can be easily tracked by evaluating the side channel information such as power consumption, heat, and electromagnetic emissions that outputs from the cryptographic device. Side channel attacks are getting much more popular since it is easy to mount an attack in a short time with only a few hundred dollars’ worth of devices and it is time saving comparing to other methods of breaking encryption also deep knowledge about algorithms isn’t required to perform an attack. This creates a huge impact on the security of the cryptographic modules as it is an efficient technique to break cryptographic algorithms by analyzing the patterns of the side channel information. The solution proposed in this paper is an algorithm modification system which inputs randomness to the algorithm by analyzing power consumption fluctuations that outputs from the algorithm in order to mitigate side channel attack possibility. In the suggested solution, a hardware device tracks down the patterns in power consumption in an accurate way and analyze those meter readings by going through advanced machine learning techniques. System suggests appropriate countermeasures in order to break down the power consumption patterns and add randomness to the algorithm. After coming up with suitable countermeasures, system will add those modifications in to the algorithm without altering the output of the algorithm in order to resist side channel attacks.

**TABLE OF CONTENTS**

1. INTRODUCTION5
   1. Background & Literature survey5

Research Gap9

Research Problem10

1. OBJECTIVES11
   1. Main Objectives11

Specific Objectives 11

1. METHODOLOGY12
   1. Device implementation for extracting power consumption measurements 12

Data classification, analysis and module training using machine learning 14

* 1. Automate Code embedding mechanism using machine learning 14

1. DESCRIPTION OF PERSONAL AND FACILITIES15
2. BUDGET AND BUDGET JUSTIFICATION16

REFERENCE LIST16

APPENDICES

Appendices 1: List of Figures 19

Appendices 2: List of Tables 20

Appendices 3: List of Abbreviations 21

1. **INTRODUCTION**
   1. **Background & Literature survey**

As the need of the protecting and securing information from adversaries the cryptographic [1] methods were born and this goes back to thousands of years. Also the breaching of the cryptographic systems (cryptanalysis) to gain access to the contents of the encrypted massages is also there since beginning of cryptography. The modern cryptographic methods are mainly based on mathematics, computer science and electrical engineering. Up until now mathematical cryptanalysis [2] is the heavily used technique to gain information from encrypted data. But currently one of the most trending cryptanalysis method that been used is side-channel-attacks (SCA). An attacker always does not need to find the keys by complex mathematical calculations in order to break encryption because the cryptographic algorithms are not purely mathematical objects they also need to be implemented. In SCA these information also take in to the account. Cryptographic algorithms are implemented on software but also on physical devices, these devices will interact with the environment and the changes in the environment around the device can be closely monitored and gather cryptanalysis information [3] like patterns in noise, electromagnetic emissions, power consumption and these are called side-channel-information. These information can be used to break cryptographic algorithms and they are called side-channel-attacks.

As for the history the first side channel attack dates back to year 1965 by the British intelligence agency trying to brake an encryption of the cipher used by Egyptian embassy London [3] by placing microphone to get the noise of the rotter-cipher machine’s click sounds but this attempt was not a complete success due to lack of computational power at the time.

Mainly this research focus on the power consumption because among the many methods that can collect side-channel-information (noise, electromagnetic emissions, power consumption) power consumption analysis is one of the most effective and powerful methods. And it is the most frequently used method to collect side-channel-information by attacker and intruders because power consumption analysis is also considered as one of the most reliable side-channel-information sources and if the cost is considered and compared with other methods of information gathering (electromagnetic, noise etc…) power consumption is the most cost effective, the equipment needed for measure electromagnetic waves are far more expensive

(At least $500) but for measuring power consumption that can be done around $100. Another main reason is power analysis can be mounted easily and the availability of the parts that are needed for the device are high and as for the software side it is less complex when isolating necessary side-channel-information.

Considering power based side channel attack there few different kind attacks. **Simple power analysis (SPA)** attack [4] this is done by getting a basic visual representation of the power consumption of the encryption device while performing and try to pinpoint the different amounts of power used for different operations like s-boxes, permutation, exponentiations in DES rounds. With SPA sequences of operations can be detected. SPA is mostly used gather side-channel-information on symmetric encryption algorithms. In **differential power analysis (DPA)** [4, 5] is much more advance than SPA it contains not only visual presentation but also it uses statistical analysis and error correction statistical analysis methods to extract information about keys. The DPA can analyze the measurements which contains too much noise that is harder to handle by SPA. Commonly DPA is used for analyzing asymmetric encryption algorithms which has high mathematical complexity like RSA implementation by defining functions like Chinese remainder theorem (CRT). DPA attacks are much harder to prevent the SPA attacks.

For this task we use simple power analysis (SPA) for gathering side-channel-information on symmetric encryption algorithms (3DES, AES, BLOWFISH, MARS, RC 2 4 5 6 etc…). The reason to choose symmetric encryption algorithms are because those algorithms are the once that use for encrypting large bulks of data. Most companies or organization use to encrypt their most important data before store them in hard drives and for this purpose they use symmetric encryption algorithms. Google uses 128 bit- AES to encrypt user data [6]. So the focus of this research is to protect the stored encrypted data.

Side-channel-attacks (SCA) are one of the most trending type of attacks which has gotten a huge attention from the people in the cryptography field. While being easy to implement SCA are a very powerful attack type. SCA attacks are based on side-channel-information (power, EM, noise) and their targets can be ranged from small IC’s to large scale computer systems [3, 4] SCA has been there for around half a century according to the paper “side-Channel Attacks: Ten Years after Its Publication and the Impacts on Cryptographic Module Security Testing” [3] the first official SCA was conducted in 1965. It was simply about using the sound of the rotor-cipher device to break encryption. The field of side-channel-attack has come long way since then. Today there are many methods of attacks that conduct by gathering necessary side-channel-information from various ways. Some of them are [3, 7]

Timing attacks – measure the time taken for operations

Power monitoring attacks – monitor the power consumption to identify operations

EM attacks – measuring the electromagnetic radiations leaked from the device.

Acoustic attacks – monitor the sound produce during the computation take place.

Differential fault analysis – introducing faults to a computation to reveal information.

In the present there are devices that are design to gather side-channel-information and analyze them to perform an attack. Most of these devices they have the capability to analyze the side-channel-information of a cryptographic algorithms and derive a key.

**SASEBO (Side channel Attack Standard Evaluation Board)** [8] is a board that enables us to perform SCA experiments. This board has 4 types of platforms [9] SASEBO-G, SASEBO-B, SASEBO-R, SASEBO-W. SASEBO G and R both have FPGAs (field programmable gate arrays) [10] which allows user to reconfigure the functions of the circuit to implement cryptographic algorithms and these LSIs can be mounted to SASEBO-R. SASEBO-G is a much powerful than other version of the board and SASEBO-GII is the much more improved version of SASEBO-G. The board is equipped with USB interface and serial port for communication. This board has the ability to measure power consumption and EM radiation during the cryptographic operation. And according to paper “Side-Channel Attack Standard Evaluation Board SASEBO-W for Smartcard Testing” they have successfully derived a keys for a smartcard using SASEBO board.

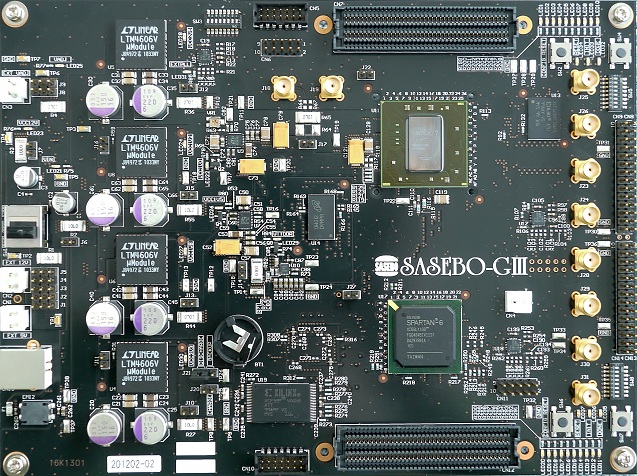


Fig. 1: SASEBO-GIII

**CHIP WHISPERER** is also another board produced by newAE Technology Inc that is in the market which can perform side channel analysis and attacks. Chip Whisperer-Lite, Chip Whisperer-Lite 2-Part Version, Chip Whisperer-Pro, Chip Whisperer-Capture Rev2 are 4 kinds of board that can select and they offer different features. This device is capable of performing side channel power analysis and glitching attacks. The board will provide several features [11] headers for mounting ADC (analog to digital converter) or DAC (digital to analog converter) boards, AVR (a microcontroller) programmer, voltage level translator, clock inputs, external phase locked loop (PLL), low noise amplifier (LNA) etc…

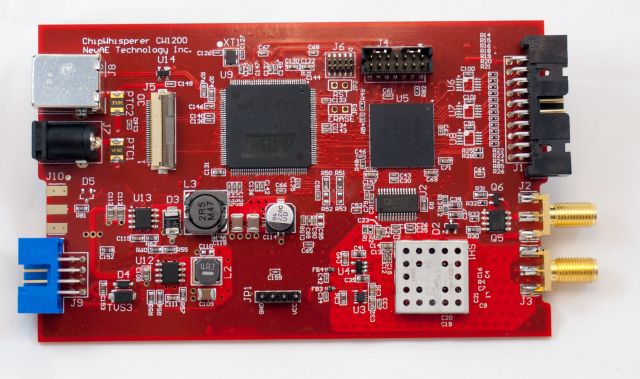


Fig. 2: Chipwishperer

**SAKURA-G FPGA** [12] design to do side-channel-attacks (SCA), Fault Injection Attacks (FIA), Physical Unclonable Function (PUF) and dynamic reconfigurations. Chip Whisperer was designed using SAKURA-GII as the base. 2 FPGAs are used as controller and the main security circuit of the board. All these board which they have in common is they used **FPGAs** [10]. The reason for using FPGAs are,it is an integrated circuit that can be configured to perform the functions we want. The different between normal ICs and FPGAs is normal IC cannot be configure as we wish because the functions are defined and the interconnection between the transistors in IC cannot be changed but in FPGAs the interconnection between transistor can be fix according to the function we defined. Another important task FPGA can do is parallel processing which enables to perform multiple transaction at the same time. So a board that uses to analyze side-channel-information should be able to handle multiple operations at the same time.

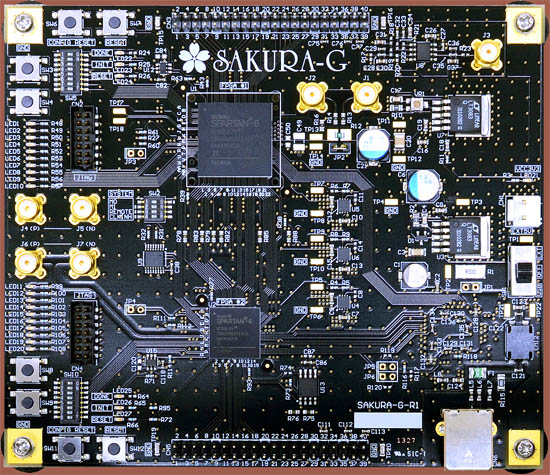


Fig. 3: SAKURA-G

Through these devices a key or several partial key can be extracted by analyzing side-channel-information of cryptographic algorithms operations. So these are the side channel techniques and technologies that has been developed up until now. Though there are system to break cryptographic algorithms there no system that can prevent those attacks by modifying the algorithms automatically.

Keywords:

Side Channel information, Cryptography, Power analysis, FPGA, DES, AES

* 1. **Research Gap**

During the literature review, we were able to find out few key areas which were not addresses by the present available products and conducted researches. According to the literature survey details gathered in previous section, most of present solutions are developed to trace down the side channel details with the help of customized board and it predicts the internal/secret key for the decryption. It demonstrates end to end side channel attack which outputs a final result that can be used to decrypt the encrypted information. Following key areas were not discussed and implemented by current products and these areas were covered in proposed solution in relation to prevent in identifying the patterns using side channel information.

At present there are methods that capable of deriving a key or a partial keys.[] For example every side channel analysis system (chip whisperer, Sakura-G, SASEBO) [9,11,12] that we come across is capable of extracting a key through power consumption. But there is no any proper method implemented to mitigate these attacks or take any preventive actions by taking vulnerabilities cause by side channel patterns (e.g. power spike patterns) of the algorithm in to consideration. So here in our solution the aim is to provide a code level modification method to the algorithm, which means changing the source code of the algorithms to resist power consumption base side channel attacks by adding a randomness to mask the patterns of power consumption of the original algorithm and preventing side channel information analysis systems to recognize those patterns.

In the proposed solution, there is an implementation of a database which consists of power consumption pattern information for publically available symmetric cryptographic algorithms. In the most of the previously defined solutions, machine learning techniques are mainly used to derive an internal key. To achieve the required solution which is suggesting from this paper, a separate module is trained to identify the vulnerable patterns of an algorithm by letting it learn from the past data of the previously analyzed algorithms which we collected before.

There is no way of knowing for a user with a little knowledge of cryptography or side channel attacks, the level of security of his algorithm. So in here with the information gathered from the previous analyzed information, a rating system is implemented which the end user can get a rough understanding about the resilient percentage for algorithm by comparing it with the previously gathered data from publically available cryptographic algorithm.

* 1. **Research Problem**

Side channel attacks are the attacks based on “side channel information” which are the information that can be retrieved from the encryption device that is neither the plain text nor the cipher text resulting from the encryption process. Side channel information are collected from heat, electromagnetic emissions, power consumption statistics etc… by these information we can determine information like the time that take for an operation and use these information to analyze an encryption algorithm to get information. At the present side channel attacks are getting much more popular since it is easy to mount an attack in a short time with only a few hundred dollars’ worth of devices and it is time saving comparing to other methods of breaking encryption also deep knowledge about algorithms isn’t required to perform an attack.

In present, publically available algorithms as well as customized algorithms are used for encryption purpose. There are no specific methods to determine whether these algorithms are vulnerable to a side channel attack and if it is what are those vulnerable points and necessary suggestions to overcome those vulnerabilities. As mentioned in the literature review, most of the current products are providing an internal key by analyzing the side channel information of a cryptographic algorithm. There is no in built mechanism to modify the algorithm by analyzing the power consumption pattern in order to mitigate side channel attack. In current products related to side channel analysis, there is no any product to recognize the relevant pattern which outputs from side channel information and find the vulnerable points in algorithm and patch it automatically without altering the algorithm in order to mitigate side channel attacks. As discussed there is no any pre-defined method in current world by adding randomness using techniques like adding noise, Obfuscation, Reduce signal size and leakage reduction.

1. **OBJECTIVES**

* 1. **Main Objective**

Securing encrypted data of a company by modifying the algorithm uses for encryption to resist power based side-channel-attacks.

When we take a company all the important data will be encrypted before storing or transmission. So the algorithms that uses for this purposes should be secure from any method of unauthorized decryption. An encryption algorithms should not only be secure in a conventional cryptanalysis methods like mathematical analysis, brute force etc… but also should consider securing the algorithm from side channel attack that exploit weakness in the cryptographic algorithm themselves. Implementing a method to provide the user to get a clear understanding about the extent of vulnerability of the algorithm to the side channel attacks and provide possible solutions and apply necessary precautions to secure the algorithm and mitigate the risks.

* 1. **Specific Objectives**

Implementing a mechanism to get the power consumption readings from the algorithm

As the first step several physical devices are needed for this information gathering (power consumption) task. We need to implement a device that can get the power consumption readings out and to get accurate power readings we have to highlight signals, isolate features, reduce noise, remove alignment errors etc... For this we use an Arduino [13] environment and while executing a sample of the encryption algorithm with a data bulk it enables us to capture necessary power consumption information.

Analyzing the gathered power consumption information of the algorithm and storing in a database

With this objective the main target is to analyze the algorithms using machine learning techniques (supervised/unsupervised) [14] by taking different criteria’s in to consideration like power spike and analyzing them we can form patterns and get information like frequency of the power spike pattern and use them to identify vulnerable point of the algorithm. Also implementing a database with power consumption information of publically available algorithms we can use these information to compare and correlate with algorithms that are testing and analyze their vulnerabilities.

Implementing necessary countermeasures by modifying the algorithm

After achieving the above objective at this stage the development of a mechanism to suggest suitable modifications and apply them to the algorithm without altering its outcome and make it resist to side channel attacks. To implement this mechanism machine learning techniques are used and the target is to prevent the algorithm by making any recognizable power consumption patterns and adding randomness by using methods like adding noise, Obfuscation, Reduce signal size, Leakage reduction etc…

1. **METHODOLOGY**

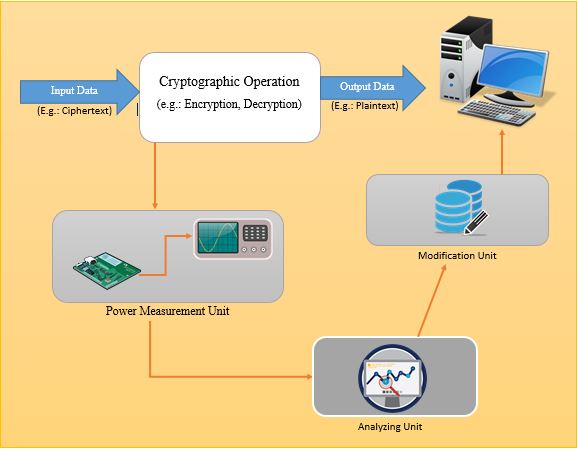


Fig. 4: High-Level Diagram

* 1. **Device implementation for extracting power consumption measurements**

Implementing a device which can trace down power consumption measurements is the initial and the most significant step that have to be carried out in an accurate way. The purpose of this implementation is to come up with a physical device that enable us to take the power consumption readings from a particular target chip while it executing in a cryptographic algorithm. As mentioned in the literature review, there are hardware solutions [9,11,12] which can get side channel information more accurately. Most of these devices are capable enough to trace down different types of measurements for different types of attacks (Fault Injection Attacks (FIA), Physical Unclonable Functions (PUF) dynamic reconfigurations) [8] which reduces the focus of identifying a side channel information. As the scope of this paper is to come up with modifications to algorithms by only examining the power consumption measurements. Above mentioned boards are consists with lots of features that are not useful for our scope and it’s not worth enough to pay such a huge amount of money in order to achieve our task. After considering the above facts, we decided to come with a board which suitable for our defined scope and its more cost effective rather going through the previously defined commercialized products.

|  |  |  |
| --- | --- | --- |
| Circuit Name | Price | Reference Links |
| **ChipWhisperer**  ChipWhisperer-Lite (CW1173) Basic Board  ChipWhisperer-Lite (CW1173) Two-Part Version  CW305 Artix FPGA Target Board | $250.00  $325.00  $800.00 | http://store.newae.com/chipwhisperer-lite-cw1173-basic-board/  http://store.newae.com/chipwhisperer-lite-cw1173-two-part-version/  http://store.newae.com/cw305-artix-fpga-target-board/ |
| **SAKURA**  SAKURA-G FPGA | $1437.10 | http://www.troche.com/sakura/order.html |
| **SASEBO**  SASEBO-G-Ⅱ | Not Available | In order to purchase this board, Due to governmental restrictions regarding the export of cryptographic hardware, we must submit an Export Compliance Certificate before they can process our order. |

Table. 1: Cost Comparison of Existing Boards

As the first phase of this implementation, an algorithm will be used to convert analog data into digital format that able to take the power consumption reading to a computer. The device is based on the board Arduino and board FPGA, and by connecting these parts, an AtMega8-16PU chip, AVR programmer, and couple of 22pF 680uF capacitors and 100 ohm resistors which will be able get the power consumption readings from the target chip. After that implementing algorithm on Arduino environment [13], we can get the readings to a computer in a digital form. The final output of this component is to get a visual representation of the power consumption which indicate the power spike patterns of that particular cryptographic algorithm.

* 1. **Data classification, analysis and module training using machine learning**

Implementing a training module using machine learning techniques [15] is also plays a major part in this research. The objective of this component is to come up with a system which is fine-tuned to identify potential vulnerabilities in a cryptographic algorithm and suggests appropriate countermeasures which adds randomness to the identified patterns by analyzing the meter readings of the side channel information.

Extracted power consumption measurements are first classified by selecting the subset of all available data in a linear manner procedure. All the data is preprocessed by formatting, fixing wrong data and sampling methods through many iterations. Data preparation takes more significant role in this sector, because to have a successful responses, meaningful and well organized dataset must be inputted. Preprocessed data is transformed to appropriate machine learning styles using techniques such as scaling, attribute decomposition and attribute aggregation. [16] Using supervised and unsupervised learning styles, system is trained fluently to obtain the most accurate and reliable solution for the identified vulnerability. Supervised learning enables to get solutions from past learn data which means by deriving them from previously analyzed power consumption data in cryptographic algorithms. [17] After using essential algorithms to train the model, necessary steps should be taken to utilize the performance of the process in order to retrieve meaningful information in a short time period. The final output is to improve the precision of the countermeasures for the identified vulnerabilities.

* 1. **Automate Code embedding mechanism using machine learning**

The objective of this phase is to develop an automated mechanism to add suitable modifications to the analyzed algorithm according to the suggested countermeasures from the trained data analysis module. This is more important as all the modifications should be appended to the algorithm in code level [18] without creating and impact of the output of the algorithm and its security features.

For this task we use machine learning techniques with the help of free and open source machine learning libraries [19]. Both supervised and unsupervised techniques will be used, supervised will enables us to get solutions from past learn data which means by deriving them from analyzed power consumption data in the past, while unsupervised will derive inferences from datasets and also to apply those suggested modification to the algorithm automatically. For the automation part, Artificial intelligence techniques [20] are used to generate the required code to append to the algorithm. The expected outcome of the component is to give out an algorithm which will be able to resist side-channel-attacks while providing the same output (encrypted data) from the algorithm same as before.

According to the finding we have done so far, we identified that present implemented solutions don’t have a device that only focus on retrieving power based side channel details. Using this suggested solution, an end user can modify his algorithm to resist power based side channel attacks. Considering to the hardware devices in present stage, suggested hardware device is less complex and cost effective along with the ability to take reliable countermeasures for side channel attacks. As this product is free and open source, at the initial stage we are going to promote this through social media and then pricing will be added to the upcoming services afterwards.

1. **DESCRIPTION OF PERSONAL AND FACILITIES**

|  |  |  |
| --- | --- | --- |
| Member | Component | Task |
| Pathirana K.P.A.P | Device implementation for extracting power consumption information | * Designing the plan to implement the board. * Exploring for required parts for building device. * Implementing the device according to the design. * Implementing algorithm to accurately gather power consumption data. |
| Hangawaththa N.H.A.D.A | Data classification, analysis and module training using machine learning | * Classifying and analyzing power consumption measurements. * Pre-processing gathered information to come up with a relevant data set. * Choosing appropriate machine learning algorithms to train the module * Transforming gathered dataset in to the chose algorithms. * Enhancing utilization and performance of the ML algorithms |
| Lankarathne L.R.M.O | Automate Code embedding mechanism using machine learning and Artificial intelligence | * Identifying necessary facilities to come up with machine learning techniques. * Learning necessary techniques of automation techniques. * Choosing appropriate machine learning algorithms to train the module. * Implementing automate code generation. * Enhancing utilization and performance of the ML algorithms |

Table. 2: Description of Personal and facilities

1. **Budget and Budget Justification**

Estimated cost for Board implementation Rs.10, 000 /=

**REFERENCES**

[1] Hagai Bar-Ei, “Introduction to Side Channel Attacks” [Online], Available: <http://gauss.ececs.uc.edu/Courses/c653/lectures/SideC/intro.pdf> [Accessed: March 12, 2017]

**APPENDICES**

***[Appendix – 1: List of Figures]***

**List of figures**

Fig. 1: SASEBO-GIII 7

Fig. 2: Chipwishperer 8

Fig. 3: SAKURA-G 9

Fig. 4: High-Level Diagram 13

***[Appendix – 2: List of Tables]***

**List of Tables**

Table. 1: Cost Comparison of Existing Boards 14

Table. 2: Description of Personal and facilities 17

***[Appendix –3: List of Abbreviations]***

**List of abbreviations**

Abbreviation Description

SCA Side Channel Attack

SPA Simple Power Analysis

DPA Differential Power Analysis

CRT Chinese Remainder Theorem

EM Electro Magnetic

IC Integrated Circuits

ADC Analog to Digital Converter

DAC Digital to Analog Converter

PLL Phase Lock Look

LNA Low Noise Amplifier

FIA Fault Injection Attacks

PUF Physical Unclonable Function

FPGA Field Programmable Gate Array

ML Machine Learning

AI Artificial Intelligence

LSI Large Scale Integration