

UNIVERSITY NAME

DOCTORAL THESIS

Thesis Title

Author:

John SMITH

Supervisor:

Dr. James SMITH

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

Research Group Name
Department or School Name

August 24, 2017

Contents

1	Introduction	1
1.1	Introduction to Cryptography	1
1.1.1	Secret-Key Cryptography	1
1.1.2	Public-Key Cryptography	1
1.2	Secure Hardware and Embedded Cryptography	1
1.2.1	The Example of the Smart Card	1
1.2.2	Certification of a Secure Hardware	1
1.2.3	Modern More Complex Devices to Certify	1
1.2.4	Embedded Cryptography Vulnerabilities	1
2	Introduction to Side-Channel Attacks	3
2.1	Introduction to Side-Channel Attacks	4
2.1.1	Historical Overview	4
2.1.2	Terminology and Generalities	4
	Target and Leakage Model	4
	Points of Interest	4
	Simple vs Advanced SCAs	4
	Vertical vs Horizontal SCAs	4
	Profiled vs Non-Profiled SCAs	4
	Side-Channel Algebraic Attacks	4
	Distinguishers	4
	SCA Metrics	4
2.1.3	Side-Channel Attacks vs Machine Learning	4
	Distinguishers vs Classifiers	4
2.2	Main Side-Channel Countermeasures	4
2.2.1	Masking	4
2.2.2	Shuffling	4
2.2.3	Random Delays and Jitter	4
2.3	Higher-Order Attacks	4
2.3.1	Higher-Order Moments Analysis and Combining Functions	4
2.3.2	Profiling Higher-Order Attacks	4
	Profiling with Masks Knowledge	4
	Profiling without Masks Knowledge	4
2.4	Thesis Contribution and Organization	4
2.4.1	Foreword of this Thesis: Research of Points of Interest	4
2.4.2	Dimensionality Reduction Approach	4
	Linear Methods for First-Order Attacks	4
	Kernel Methods for Higher-Order Attacks	4
2.4.3	Neural Network Approach	4
	Toward Getting Rid of Information-Losing Preprocessing	4

3	Points of Interest and Dimensionality Reduction	5
3.1	Motivations	5
3.1.1	The Curse of Dimensionality	5
3.2	Selection on Points of Interest: Classical Statistics	5
3.3	Related Issues: Leakage Detection and Leakage Assessment	5
3.4	Observations Leading to Take a Dimensionality Reduction Approach	5
4	Linear Dimensionality Reduction	7
4.1	Introduction	7
4.1.1	Principal Component Analysis	7
4.1.2	Linear Discriminant Analysis	7
4.1.3	Projection Pursuits	7
4.2	Principal Component Analysis	7
4.2.1	Statistical Point of View	7
4.2.2	Geometrical Point of View	7
4.3	Application of PCA in SCAs	7
4.3.1	Original vs Class-Oriented PCA	7
4.3.2	The Choice of the Principal Components	7
4.4	Linear Discriminant Analysis	7
4.4.1	Statistical Point of View	7
4.4.2	Geometrical Point of View	7
4.5	Application of LDA in SCAs	7
4.5.1	The Small Sample Size problem	7
5	Kernel Dimensionality Reduction	9
5.1	Motivation	9
5.1.1	Higher-Order Attacks	9
	Higher-Order Version of Projection Pursuits	9
5.2	Kernel Function and Kernel Trick	9
5.2.1	Local Kernel Functions as Similarity Metrics	9
5.3	Kernel Discriminant Analysis	9
5.4	Experiments over Atmega328P	9
5.4.1	The Regularization Problem	9
5.4.2	The Multi-Class Trade-Off	9
5.4.3	Multi-Class vs 2-class Approach	9
5.4.4	Asymmetric Preprocessing/Attack Approach	9
	Comparison with Projection Pursuits	9
6	Machine Learning Approach	11
6.1	Motivation	11
6.2	Introduction to Machine Learning	11
6.2.1	The Task, the Experience and the Performance	11
6.2.2	Supervised, Semi-Supervised, Unsupervised Learning	11
6.2.3	Training, Validation and Test Sets	11
6.2.4	Underfitting, Overfitting and Regularization	11
6.2.5	Data Augmentation	11
6.2.6	No Free Lunch Theorem	11
6.3	Machine Learning Applications in Side-Channel Context	11
6.3.1	Profiled Attack as a Classification Problem	11
	Support Vector Machine	11
	Random Forest	11

6.4	Artificial Neural Networks	11
6.4.1	Motivations Leading from Kernel Machines to Deep Learning	11
6.4.2	The Multi-Layer Perceptron	11
6.5	Simulated Experiment for Profiled HO-Attacks	11
6.5.1	The Simulations	11
6.5.2	Comparison between KDA and MLP	11
6.6	Real-Case Experiments over ARM Cortex-M4	11
7	Convolutional Neural Networks against Jitter-Based Countermeasures	13
7.1	Misalignment of Side-Channel Traces	13
7.1.1	The Necessity and the Risks of Applying Realignment Techniques	13
7.1.2	Analogy with Image Recognition Issues	13
7.2	Convolutional Layers to Impose Shift-Invariance	13
7.3	Data Augmentation for Misaligned Side-Channel Traces	13
7.4	Experiments against Software Countermeasures	13
7.5	Experiments against Artificial Hardware Countermeasures	13
7.6	Experiments against Real-Case Hardware Countermeasures	13
8	Neural Networks: Back to Dimensionality Reduction	15
8.1	Motivation	15
8.2	Stacked Auto-Encoders	15
8.2.1	The Same Issues of Classic PCA	15
8.3	Siamese Neural Networks	15
8.3.1	Distances and Loss Functions	15
8.3.2	Relation with Kernel Machines	15
8.4	A Experimental comparison between KDA and Siamese NNs	15
8.5	Collision Attacks with Siamese NNs	15
8.5.1	Experimental Results	15
9	Conclusions and Perspectives	17
9.1	Summary	17
9.2	Strengthen Embedded Security: the Main Challenge for Machine Learning Applications	17

List of Figures

List of Tables

List of Abbreviations

SCA Side Channel Attack

List of Symbols

Chapter 1

Introduction

1.1 Introduction to Cryptography

1.1.1 Secret-Key Cryptography

1.1.2 Public-Key Cryptography

1.2 Secure Hardware and Embedded Cryptography

1.2.1 The Example of the Smart Card

1.2.2 Certification of a Secure Hardware

1.2.3 Modern More Complex Devices to Certify

1.2.4 Embedded Cryptography Vulnerabilities

Chapter 2

Introduction to Side-Channel Attacks

2.1 Introduction to Side-Channel Attacks

2.1.1 Historical Overview

2.1.2 Terminology and Generalities

Target and Leakage Model

Points of Interest

Simple vs Advanced SCAs

Vertical vs Horizontal SCAs

Profiled vs Non-Profiled SCAs

Side-Channel Algebraic Attacks

Distinguishers

SCA Metrics

2.1.3 Side-Channel Attacks vs Machine Learning

Distinguishers vs Classifiers

2.2 Main Side-Channel Countermeasures

2.2.1 Masking

2.2.2 Shuffling

2.2.3 Random Delays and Jitter

2.3 Higher-Order Attacks

2.3.1 Higher-Order Moments Analysis and Combining Functions

2.3.2 Profiling Higher-Order Attacks

Profiling with Masks Knowledge

Profiling without Masks Knowledge

2.4 Thesis Contribution and Organization

2.4.1 Foreword of this Thesis: Research of Points of Interest

2.4.2 Dimensionality Reduction Approach

Linear Methods for First-Order Attacks

Chapter 3

Points of Interest and Dimensionality Reduction

3.1 Motivations

3.1.1 The Curse of Dimensionality

3.2 Selection on Points of Interest: Classical Statistics

3.3 Related Issues: Leakage Detection and Leakage Assessment

3.4 Observations Leading to Take a Dimensionality Reduction Approach

Chapter 4

Linear Dimensionality Reduction

4.1 Introduction

4.1.1 Principal Component Analysis

4.1.2 Linear Discriminant Analysis

4.1.3 Projection Pursuits

4.2 Principal Component Analysis

4.2.1 Statistical Point of View

4.2.2 Geometrical Point of View

4.3 Application of PCA in SCAs

4.3.1 Original vs Class-Oriented PCA

4.3.2 The Choice of the Principal Components

4.4 Linear Discriminant Analysis

4.4.1 Statistical Point of View

4.4.2 Geometrical Point of View

4.5 Application of LDA in SCAs

4.5.1 The Small Sample Size problem

Chapter 5

Kernel Dimensionality Reduction

5.1 Motivation

5.1.1 Higher-Order Attacks

Higher-Order Version of Projection Pursuits

5.2 Kernel Function and Kernel Trick

5.2.1 Local Kernel Functions as Similarity Metrics

5.3 Kernel Discriminant Analysis

5.4 Experiments over Atmega328P

5.4.1 The Regularization Problem

5.4.2 The Multi-Class Trade-Off

5.4.3 Multi-Class vs 2-class Approach

5.4.4 Asymmetric Preprocessing/Attack Approach

Comparison with Projection Pursuits

Chapter 6

Machine Learning Approach

6.1 Motivation

6.2 Introduction to Machine Learning

6.2.1 The Task, the Experience and the Performance

6.2.2 Supervised, Semi-Supervised, Unsupervised Learning

6.2.3 Training, Validation and Test Sets

6.2.4 Underfitting, Overfitting and Regularization

6.2.5 Data Augmentation

6.2.6 No Free Lunch Theorem

6.3 Machine Learning Applications in Side-Channel Context

6.3.1 Profiled Attack as a Classification Problem

Support Vector Machine

Random Forest

6.4 Artificial Neural Networks

6.4.1 Motivations Leading from Kernel Machines to Deep Learning

6.4.2 The Multi-Layer Perceptron

6.5 Simulated Experiment for Profiled HO-Attacks

6.5.1 The Simulations

6.5.2 Comparison between KDA and MLP

6.6 Real-Case Experiments over ARM Cortex-M4

Chapter 7

Convolutional Neural Networks against Jitter-Based Countermeasures

7.1 Misalignment of Side-Channel Traces

7.1.1 The Necessity and the Risks of Applying Realignment Techniques

7.1.2 Analogy with Image Recognition Issues

7.2 Convolutional Layers to Impose Shift-Invariance

7.3 Data Augmentation for Misaligned Side-Channel Traces

7.4 Experiments against Software Countermeasures

7.5 Experiments against Artificial Hardware Countermeasures

7.6 Experiments against Real-Case Hardware Countermeasures

Chapter 8

Neural Networks: Back to Dimensionality Reduction

8.1 Motivation

8.2 Stacked Auto-Encoders

8.2.1 The Same Issues of Classic PCA

8.3 Siamese Neural Networks

8.3.1 Distances and Loss Functions

8.3.2 Relation with Kernel Machines

8.4 A Experimental comparison between KDA and Siamese NNs

8.5 Collision Attacks with Siamese NNs

8.5.1 Experimental Results

Chapter 9

Conclusions and Perspectives

9.1 Summary

9.2 Strengthen Embedded Security: the Main Challenge for Machine Learning Applications