

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

July 19, 2017

Contents

1	Introduction	1
1.1	Introduction to Cryptography	2
1.1.1	Secret-Key Cryptography	2
1.1.2	Public-Key Cryptography	2
1.2	Secure Hardware and Embedded Cryptography	2
1.2.1	The Example of the Smart Card	2
1.2.2	Certification of a Secure Hardware	2
1.2.3	Embedded Cryptography Vulnerabilities	2
1.3	Introduction to Side-Channel Attacks	2
1.3.1	Historical Overview	2
1.3.2	Terminology and Generalities	2
	Target and Leakage Model	2
	Points of Interest	2
	Simple vs Advanced SCAs	2
	Vertical vs Horizontal SCAs	2
	Profiled vs Non-Profiled SCAs	2
	Distinguishers	2
	SCA Metrics	2
1.3.3	Side-Channel Attacks vs Machine Learning	2
	Distinguishers vs Classifiers	2
1.4	Main Side-Channel Countermeasures	2
1.4.1	Masking	2
1.4.2	Shuffling	2
1.4.3	Blinding	2
1.4.4	Random Delays and Jitter	2
2	Points of Interest and Dimensionality Reduction	3
2.1	Motivations	3
2.1.1	The Curse of Dimensionality	3
2.2	Selection on Points of Interest: Classical Statistics	3
2.3	Related Issues: Leakage Detection and Leakage Assessment	3
2.4	Dimensionality Reduction Approach	3
2.4.1	Feature Selection as a Machine Learning Task	3
3	Linear Dimensionality Reduction	5
3.1	Introduction	5
3.1.1	Principal Component Analysis	5
3.1.2	Linear Discriminant Analysis	5
3.1.3	Projection Pursuits	5
3.2	Principal Component Analysis	5
3.2.1	Statistical Point of View	5
3.2.2	Geometrical Point of View	5
3.3	Application of PCA in SCAs	5

3.3.1	Original vs Class-Oriented PCA	5
3.3.2	The Choice of the Principal Components	5
3.4	Linear Discriminant Analysis	5
3.4.1	Statistical Point of View	5
3.4.2	Geometrical Point of View	5
3.5	Application of LDA in SCAs	5
3.5.1	The Small Sample Size problem	5
4	Kernel Dimensionality Reduction	7
4.1	Motivation	7
4.1.1	Higher-Order Attacks	7
	Higher-Order Version of Projection Pursuits	7
4.2	Kernel Function and Kernel Trick	7
4.2.1	Local Kernel Functions as Similarity Metrics	7
4.3	Kernel Discriminant Analysis	7
4.4	Experiments over Atmega328P	7
4.4.1	The Regularization Problem	7
4.4.2	The Multi-Class Trade-Off	7
4.4.3	Multi-Class vs 2-class Approach	7
4.4.4	Asymmetric Preprocessing/Attack Approach	7
	Comparison with Projection Pursuits	7
5	Machine Learning Approach	9
5.1	Motivation	9
5.2	Introduction to Machine Learning	9
5.2.1	The Task, the Experience and the Performance	9
5.2.2	Supervised, Semi-Supervised, Unsupervised Learning	9
5.2.3	Training, Validation and Test Sets	9
5.2.4	Underfitting, Overfitting and Regularization	9
5.2.5	Data Augmentation	9
5.2.6	No Free Lunch Theorem	9
5.3	Machine Learning Applications in Side-Channel Context	9
5.3.1	Profiled Attack as a Classification Problem	9
	Support Vector Machine	9
	Random Forest	9
5.4	Artificial Neural Networks	9
5.4.1	Motivations Leading from Kernel Machines to Deep Learning	9
5.4.2	The Multi-Layer Perceptron	9
5.5	Simulated Experiment for Profiled HO-Attacks	9
5.5.1	The Simulations	9
5.5.2	Comparison between KDA and MLP	9
5.6	Real-Case Experiments over ARM Cortex-M4	9
6	Convolutional Neural Networks against Jitter-Based Countermeasures	11
6.1	Misalignment of Side-Channel Traces	11
6.1.1	The Necessity and the Risks of Applying Realignment Techniques	11
6.1.2	Analogy with Image Recognition Issues	11
6.2	Convolutional Layers to Impose Shift-Invariance	11
6.3	Data Augmentation for Misaligned Side-Channel Traces	11
6.4	Experiments against Software Countermeasures	11

6.5	Experiments against Artificial Hardware Countermeasures	11
6.6	Experiments against Real-Case Hardware Countermeasures	11
7	Neural Networks: Back to Dimensionality Reduction	13
7.1	Motivation	13
7.2	Stacked Auto-Encoders	13
7.2.1	The Same Issues of Classic PCA	13
7.3	Siamese Neural Networks	13
7.3.1	Distances and Loss Functions	13
7.3.2	Relation with Kernel Machines	13
7.4	A Experimental comparison between KDA and Siamese NNs	13
7.5	Collision Attacks with Siamese NNs	13
7.5.1	Experimental Results	13

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 Embedded Cryptography Vulnerabilities

1.3 Introduction to Side-Channel Attacks

1.3.1 Historical Overview

1.3.2 Terminology and Generalities

Target and Leakage Model

Points of Interest

Simple vs Advanced SCAs

Vertical vs Horizontal SCAs

Profiled vs Non-Profiled SCAs

Distinguishers

SCA Metrics

1.3.3 Side-Channel Attacks vs Machine Learning

Distinguishers vs Classifiers

1.4 Main Side-Channel Countermeasures

1.4.1 Masking

1.4.2 Shuffling

1.4.3 Blinding

1.4.4 Random Delays and Jitter

Chapter 2

Points of Interest and Dimensionality Reduction

2.1 Motivations

2.1.1 The Curse of Dimensionality

2.2 Selection on Points of Interest: Classical Statistics

2.3 Related Issues: Leakage Detection and Leakage Assessment

2.4 Dimensionality Reduction Approach

2.4.1 Feature Selection as a Machine Learning Task

Chapter 3

Linear Dimensionality Reduction

3.1 Introduction

3.1.1 Principal Component Analysis

3.1.2 Linear Discriminant Analysis

3.1.3 Projection Pursuits

3.2 Principal Component Analysis

3.2.1 Statistical Point of View

3.2.2 Geometrical Point of View

3.3 Application of PCA in SCAs

3.3.1 Original vs Class-Oriented PCA

3.3.2 The Choice of the Principal Components

3.4 Linear Discriminant Analysis

3.4.1 Statistical Point of View

3.4.2 Geometrical Point of View

3.5 Application of LDA in SCAs

3.5.1 The Small Sample Size problem

Chapter 4

Kernel Dimensionality Reduction

4.1 Motivation

4.1.1 Higher-Order Attacks

Higher-Order Version of Projection Pursuits

4.2 Kernel Function and Kernel Trick

4.2.1 Local Kernel Functions as Similarity Metrics

4.3 Kernel Discriminant Analysis

4.4 Experiments over Atmega328P

4.4.1 The Regularization Problem

4.4.2 The Multi-Class Trade-Off

4.4.3 Multi-Class vs 2-class Approach

4.4.4 Asymmetric Preprocessing/Attack Approach

Comparison with Projection Pursuits

Chapter 5

Machine Learning Approach

5.1 Motivation

5.2 Introduction to Machine Learning

5.2.1 The Task, the Experience and the Performance

5.2.2 Supervised, Semi-Supervised, Unsupervised Learning

5.2.3 Training, Validation and Test Sets

5.2.4 Underfitting, Overfitting and Regularization

5.2.5 Data Augmentation

5.2.6 No Free Lunch Theorem

5.3 Machine Learning Applications in Side-Channel Context

5.3.1 Profiled Attack as a Classification Problem

Support Vector Machine

Random Forest

5.4 Artificial Neural Networks

5.4.1 Motivations Leading from Kernel Machines to Deep Learning

5.4.2 The Multi-Layer Perceptron

5.5 Simulated Experiment for Profiled HO-Attacks

5.5.1 The Simulations

5.5.2 Comparison between KDA and MLP

5.6 Real-Case Experiments over ARM Cortex-M4

Chapter 6

Convolutional Neural Networks against Jitter-Based Countermeasures

6.1 Misalignment of Side-Channel Traces

6.1.1 The Necessity and the Risks of Applying Realignment Techniques

6.1.2 Analogy with Image Recognition Issues

6.2 Convolutional Layers to Impose Shift-Invariance

6.3 Data Augmentation for Misaligned Side-Channel Traces

6.4 Experiments against Software Countermeasures

6.5 Experiments against Artificial Hardware Countermeasures

6.6 Experiments against Real-Case Hardware Countermeasures

Chapter 7

Neural Networks: Back to Dimensionality Reduction

7.1 Motivation

7.2 Stacked Auto-Encoders

7.2.1 The Same Issues of Classic PCA

7.3 Siamese Neural Networks

7.3.1 Distances and Loss Functions

7.3.2 Relation with Kernel Machines

7.4 A Experimental comparison between KDA and Siamese NNs

7.5 Collision Attacks with Siamese NNs

7.5.1 Experimental Results