

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 25, 2017

Contents

I	Context and State of the Art	1
1	Introduction	3
1.1	Introduction to Cryptography	3
1.1.1	Secret-Key Cryptography	3
1.1.2	Public-Key Cryptography	3
1.2	Secure Hardware and Embedded Cryptography	3
1.2.1	The Example of the Smart Card	3
1.2.2	Certification of a Secure Hardware	3
1.2.3	Modern More Complex Devices to Certify	3
1.2.4	Embedded Cryptography Vulnerabilities	3
2	Introduction to Side-Channel Attacks	5
2.1	Introduction to Side-Channel Attacks	6
2.1.1	Historical Overview	6
2.1.2	Terminology and Generalities	6
	Target and Leakage Model	6
	Points of Interest	6
	Simple vs Advanced SCAs	6
	Vertical vs Horizontal SCAs	6
	Profiled vs Non-Profiled SCAs	6
	Side-Channel Algebraic Attacks	6
	Distinguishers	6
	SCA Metrics	6
2.2	Main Side-Channel Countermeasures	6
2.2.1	Random Delays and Jitter	6
2.2.2	Shuffling	6
2.2.3	Masking	6
2.3	Higher-Order Attacks	6
2.3.1	Higher-Order Moments Analysis and Combining Functions	6
2.3.2	Profiling Higher-Order Attacks	6
	Profiling with Masks Knowledge	6
	Profiling without Masks Knowledge	6
2.4	Thesis Contribution and Organization	6
2.4.1	Foreword of this Thesis: Research of Points of Interest	6
2.4.2	Dimensionality Reduction Approach	6
	Linear Methods for First-Order Attacks	6
	Kernel Methods for Higher-Order Attacks	6
2.4.3	Neural Network Approach	6
	Toward Getting Rid of Information-Losing Preprocessing	6

3	Introduction to Machine Learning	7
3.1	Basic Concepts of Machine Learning	7
3.1.1	The Task, the Experience and the Performance	7
3.1.2	Supervised, Semi-Supervised, Unsupervised Learning	7
3.1.3	Training, Validation and Test Sets	7
3.1.4	Underfitting, Overfitting and Regularization	7
3.1.5	Data Augmentation	7
3.1.6	No Free Lunch Theorem	7
3.2	Machine Learning Applications in Side-Channel Context	7
3.2.1	Profiled Attack as a Classification Problem	7
	Support Vector Machine	7
	Random Forest	7
	Neural Networks	7
II	Contributions	9
4	Points of Interest	11
4.1	Motivations	11
4.1.1	The Curse of Dimensionality	11
4.2	Selection on Points of Interest: Classical Statistics	11
4.3	Related Issues: Leakage Detection and Leakage Assessment	11
4.4	Generalized SNR for Multi-Variate Attacks	11
4.5	Observations Leading to Take a Dimensionality Reduction Approach	11
5	Linear Dimensionality Reduction	13
5.1	Introduction	13
5.1.1	Principal Component Analysis	13
5.1.2	Linear Discriminant Analysis	13
5.1.3	Projection Pursuits	13
5.2	Principal Component Analysis	13
5.2.1	Statistical Point of View	13
5.2.2	Geometrical Point of View	13
5.3	Application of PCA in SCAs	13
5.3.1	Original vs Class-Oriented PCA	13
5.3.2	The Choice of the Principal Components	13
5.4	Linear Discriminant Analysis	13
5.4.1	Statistical Point of View	13
5.4.2	Geometrical Point of View	13
5.5	Application of LDA in SCAs	13
5.5.1	The Small Sample Size problem	13
6	Kernel Dimensionality Reduction	15
6.1	Motivation	15
6.1.1	Higher-Order Attacks	15
	Higher-Order Version of Projection Pursuits	15
6.2	Kernel Function and Kernel Trick	15
6.2.1	Local Kernel Functions as Similarity Metrics	15
6.3	Kernel Discriminant Analysis	15
6.4	Experiments over Atmega328P	15
6.4.1	The Regularization Problem	15

6.4.2	The Multi-Class Trade-Off	15
6.4.3	Multi-Class vs 2-class Approach	15
6.4.4	Asymmetric Preprocessing/Attack Approach	15
	Comparison with Projection Pursuits	15
6.5	Drawbacks of Kernel Methods	15
	Misalignment Effects	15
	Memory Complexity and Actual Number of Parameters	15
	Two-Phases Approach: Preprocessing-Templates	15
7	Convolutional Neural Networks against Jitter-Based Countermeasures	17
7.1	Moving from Kernel Machines to Neural Networks	17
7.2	Misalignment of Side-Channel Traces	17
7.2.1	The Necessity and the Risks of Applying Realignment Techniques	17
7.2.2	Analogy with Image Recognition Issues	17
7.3	Convolutional Layers to Impose Shift-Invariance	17
7.4	Data Augmentation for Misaligned Side-Channel Traces	17
7.5	Experiments against Software Countermeasures	17
7.6	Experiments against Artificial Hardware Countermeasures	17
7.7	Experiments against Real-Case Hardware Countermeasures	17
8	KDA vs Neural Networks Approach for HO-Attacks	19
8.1	Simulated Experiment for Profiled HO-Attacks	19
8.1.1	The Simulations	19
8.1.2	Comparison between KDA and MLP	19
8.2	Real-Case Experiments over ARM Cortex-M4	19
9	Siamese Neural Networks for Collision Attacks	21
9.1	Introduction	21
9.2	Siamese Neural Networks	21
9.2.1	Distances and Loss Functions	21
9.2.2	Relation with Kernel Machines	21
9.3	Collision Attacks with Siamese NNs	21
9.3.1	Experimental Results	21
10	Conclusions and Perspectives	23
10.1	Summary	23
10.2	Strengthen Embedded Security: the Main Challenge for Machine Learning Applications	23

List of Figures

List of Tables

List of Abbreviations

SCA Side Channel Attack

List of Symbols

Part I

Context and State of the Art

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

2.2 Main Side-Channel Countermeasures

2.2.1 Random Delays and Jitter

2.2.2 Shuffling

2.2.3 Masking

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

Kernel Methods for Higher-Order Attacks

2.4.3 Neural Network Approach

Chapter 3

Introduction to Machine Learning

3.1 Basic Concepts of Machine Learning

3.1.1 The Task, the Experience and the Performance

3.1.2 Supervised, Semi-Supervised, Unsupervised Learning

3.1.3 Training, Validation and Test Sets

3.1.4 Underfitting, Overfitting and Regularization

3.1.5 Data Augmentation

3.1.6 No Free Lunch Theorem

3.2 Machine Learning Applications in Side-Channel Context

3.2.1 Profiled Attack as a Classification Problem

Support Vector Machine

Random Forest

Neural Networks

Part II

Contributions

Chapter 4

Points of Interest

4.1 Motivations

4.1.1 The Curse of Dimensionality

4.2 Selection on Points of Interest: Classical Statistics

4.3 Related Issues: Leakage Detection and Leakage Assessment

4.4 Generalized SNR for Multi-Variate Attacks

4.5 Observations Leading to Take a Dimensionality Reduction Approach

Chapter 5

Linear Dimensionality Reduction

5.1 Introduction

5.1.1 Principal Component Analysis

5.1.2 Linear Discriminant Analysis

5.1.3 Projection Pursuits

5.2 Principal Component Analysis

5.2.1 Statistical Point of View

5.2.2 Geometrical Point of View

5.3 Application of PCA in SCAs

5.3.1 Original vs Class-Oriented PCA

Remark. Stacked Auto-Encoders...

5.3.2 The Choice of the Principal Components

5.4 Linear Discriminant Analysis

5.4.1 Statistical Point of View

5.4.2 Geometrical Point of View

5.5 Application of LDA in SCAs

5.5.1 The Small Sample Size problem

Chapter 6

Kernel Dimensionality Reduction

6.1 Motivation

6.1.1 Higher-Order Attacks

Higher-Order Version of Projection Pursuits

6.2 Kernel Function and Kernel Trick

6.2.1 Local Kernel Functions as Similarity Metrics

6.3 Kernel Discriminant Analysis

6.4 Experiments over Atmega328P

6.4.1 The Regularization Problem

6.4.2 The Multi-Class Trade-Off

6.4.3 Multi-Class vs 2-class Approach

6.4.4 Asymmetric Preprocessing/Attack Approach

Comparison with Projection Pursuits

6.5 Drawbacks of Kernel Methods

Misalignment Effects

Memory Complexity and Actual Number of Parameters

Two-Phases Approach: Preprocessing-Templates

Chapter 7

Convolutional Neural Networks against Jitter-Based Countermeasures

7.1 Moving from Kernel Machines to Neural Networks

7.2 Misalignment of Side-Channel Traces

7.2.1 The Necessity and the Risks of Applying Realignment Techniques

7.2.2 Analogy with Image Recognition Issues

7.3 Convolutional Layers to Impose Shift-Invariance

7.4 Data Augmentation for Misaligned Side-Channel Traces

7.5 Experiments against Software Countermeasures

7.6 Experiments against Artificial Hardware Countermeasures

7.7 Experiments against Real-Case Hardware Countermeasures

Chapter 8

KDA vs Neural Networks Approach for HO-Attacks

8.1 Simulated Experiment for Profiled HO-Attacks

8.1.1 The Simulations

8.1.2 Comparison between KDA and MLP

8.2 Real-Case Experiments over ARM Cortex-M4

Chapter 9

Siamese Neural Networks for Collision Attacks

9.1 Introduction

9.2 Siamese Neural Networks

9.2.1 Distances and Loss Functions

9.2.2 Relation with Kernel Machines

9.3 Collision Attacks with Siamese NNs

9.3.1 Experimental Results

Chapter 10

Conclusions and Perspectives

10.1 Summary

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