

Data-Driven Modeling of Cyber-Physical Systems using Side-Channel Analysis

Sujit Rokka Chhetri
Mohammad Abdullah Al Faruque

Data-Driven Modeling of Cyber-Physical Systems using Side-Channel Analysis

Sujit Rokka Chhetri
Department of Electrical Engineering
and Computer Science
University of California
Irvine, CA, USA

Mohammad Abdullah Al Faruque
Department of Electrical Engineering
and Computer Science (EECS)
University of California
Irvine, CA, USA

ISBN 978-3-030-37961-2 ISBN 978-3-030-37962-9 (eBook)
<https://doi.org/10.1007/978-3-030-37962-9>

© Springer Nature Switzerland AG 2020

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

We dedicate this book to our family.

Preface

Cyber-physical systems (CPS) consist of a unique integration of discrete cyber-domain processes and continuous physical domain components. Current modeling approaches use extensive first-principle approaches to derive the various components of CPS. However, it is difficult to model some of the stochastic phenomenon (such as environmental variation, physical process variation, etc.) of CPS using the first-principle approach. Hence, in this book we have explored a data-driven modeling approach for CPS and present various methodologies for modeling security and creating virtual replica or digital twin of the physical system. Furthermore, we will also present new algorithms to handle complex non-Euclidean data for modeling the CPS. More specifically, the book will present and exploration of the unintended emissions from the physical domain of the CPS to infer various cyber-domain states.

This book provides a new perspective on modeling the cyber-physical system using a data-driven approach. It covers the use of state-of-the-art machine learning and artificial intelligence algorithms for modeling various aspects of the CPS. It provides insight on how a data-driven modeling approach can be utilized to take advantage of the relation between the cyber and the physical domain of the CPS for aiding the first-principle approach in capturing the stochastic phenomenon affecting the CPS.

The book provides a practical use case of the data-driven modeling approach for securing the CPS, presenting novel attack models, building and maintaining the digital twin of the physical system. Furthermore, it also provides novel data-driven algorithms to handle non-Euclidean data. In summary, this book presents a novel perspective for modeling the CPS.

The first-principle based approach for modeling the CPS is complex and time-consuming. Below we present three major reasons for proposing a new book in this area:

- Due to the advancement of machine learning and artificial intelligence algorithms, there has been a huge leap in performing data-driven modeling. However, to the best of our knowledge, there are no books covering the data-driven modeling of the CPS to aid in capturing the stochastic phenomenon affecting CPS.

- The book presents some practical application for securing the CPS as well as building the digital twin of the physical twin of CPS. The digital twin is expected to be one of the pillars for next generation of CPS. Hence, this book provides timely coverage of building and maintaining the digital twins of CPS.
- The book also provides novel algorithms for handling not just Euclidean data but also non-Euclidean data. These algorithms will thus demonstrate how the next generation of digital twins may be made more cognitive by allowing it to process and extract information from complex and higher dimensional data.

Some of the unique features of the book can be listed as follows:

- Only book covering the data-driven modeling of the CPS utilizing the unique relation between the cyber and the physical domain.
- Coverage of machine learning and artificial intelligence algorithms for data-driven modeling of the CPS.
- Practical use case of the data-driven modeling approach for security and building digital twin of the CPS.
- Well-structured and comprehensive book chapters covering the breadth and depth in data-driven modeling of CPS.

Irvine, CA, USA
Irvine, CA, USA

Sujit Rokka Chhetri
Mohammad Abdullah Al Faruque

Acknowledgments

We would like to thank all the current and past members of Advanced Integrated Cyber-Physical Systems (AICPS) lab at the University of California of Irvine for contributing to the research and content of this book. We are really grateful for your help and support.

Contents

1	Introduction	1
1.1	Cyber-Physical System	1
1.2	Data-Driven Modeling	2
1.3	Side-Channel Analysis	3
1.4	Book Sections	3
1.4.1	Part I: Data-Driven Attack Modeling	3
1.4.2	Part II: Data-Driven Defense of Cyber-Physical Systems	4
1.4.3	Part III: Data-Driven Digital Twin Modeling	4
1.4.4	Part IV: Non-Euclidean Data-Driven Modeling of Cyber-Physical Systems.....	5
1.5	Summary.....	6
	References.....	6
 Part I Data-Driven Attack Modeling		
2	Data-Driven Attack Modeling Using Acoustic Side-Channel.....	11
2.1	Introduction.....	11
2.1.1	Research Challenges and Contributions	12
2.2	Background and Related Work	12
2.3	Sources of Acoustic Emission	14
2.3.1	System Description	14
2.3.2	Equation of Motion	14
2.3.3	Natural Rotor Oscillation Frequency	15
2.3.4	Stator Natural Frequency	16
2.3.5	Source of Vibration	16
2.4	Acoustic Leakage Analysis	18
2.4.1	Side-Channel Leakage Model.....	18
2.4.2	Leakage Quantification	20
2.4.3	Leakage Exploitation	20

2.5	Attack Model Description	21
2.5.1	Attack Model	21
2.5.2	Components of the Attack Model.....	22
2.5.3	Attack Model Training and Evaluation	31
2.6	Results for Test Objects	34
2.6.1	Speed of Printing	34
2.6.2	The Dimension of the Object	35
2.6.3	The Complexity of the Object	35
2.6.4	Reconstruction of a Square.....	36
2.6.5	Reconstruction of a Triangle	37
2.6.6	Case Study: Outline of a Key	37
2.7	Discussion	39
2.7.1	Technology Variation.....	39
2.7.2	Sensor Position	39
2.7.3	Sensor Number	39
2.7.4	Dynamic Window	40
2.7.5	Feature Separation during Multiple Axis Movement and Noise	40
2.7.6	Target Machine Degradation	40
2.8	Summary.....	40
	References.....	41
3	Aiding Data-Driven Attack Model with a Compiler Modification	43
3.1	Introduction.....	43
3.2	Attack Model Description	44
3.3	Compiler Attack	47
3.3.1	Profiling Phase	48
3.3.2	Attack Phase	49
3.3.3	Compiler Modification	50
3.3.4	Transformations for Leakage Maximization	51
3.4	Experimental Results.....	53
3.4.1	Accuracy Metric	55
3.4.2	Mutual Information.....	56
3.4.3	Partial Success Rate	58
3.4.4	Total Success Rate	60
3.5	Discussion	61
3.5.1	Countermeasures	62
3.6	Summary.....	63
	References.....	63

Part II Data-Driven Defense of Cyber-Physical Systems

4	Data-Driven Defense Through Leakage Minimization	67
4.1	Introduction.....	67
4.1.1	Motivation for Leakage-Aware Security Tool	67
4.1.2	Problem and Challenges.....	68
4.1.3	Contributions	69

4.2	System Modeling	69
4.2.1	Data-driven Leakage Modeling and Quantification	70
4.2.2	Attack Model	71
4.2.3	Formulation of Data-Driven Leakage-Aware Optimization Problem	73
4.2.4	Success Rate of the Adversary	76
4.3	Experimental Results.....	78
4.3.1	Mutual Information.....	79
4.3.2	Test with Benchmark 3D Models	81
4.4	Case Study with an Attack Model	85
4.4.1	Success Rate Calculation	85
4.4.2	Test Case with Reconstruction	86
4.5	Discussion	87
4.6	Summary.....	89
	References	89
5	Data-Driven Kinetic Cyber-Attack Detection	91
5.1	Introduction.....	91
5.1.1	Motivation	92
5.1.2	Problem and Challenges.....	92
5.1.3	Contributions	93
5.2	Kinetic Cyber-Attack Adversary Model	93
5.3	KCAD Method	94
5.3.1	Mutual Information.....	95
5.3.2	KCAD Architecture	96
5.3.3	Acoustic Analog Emissions	99
5.3.4	Performance Metrics	101
5.4	Experimental Results	101
5.4.1	Experimental Setup.....	101
5.4.2	Mutual Information Calculation	102
5.4.3	Model Function Estimation	102
5.4.4	Results for Detection of Kinetic Attack	104
5.4.5	Test Case: Base Plate of a Quad Copter	106
5.5	Discussion	107
5.6	Summary.....	108
	References	108
6	Data-Driven Security Analysis Using Generative Adversarial Networks.....	111
6.1	Introduction.....	111
6.1.1	Research Challenges.....	112
6.1.2	Preliminaries	113
6.1.3	Novel Contributions	114
6.2	CGAN-Based CPPS Security Model	114
6.3	CGAN Model Generation.....	116

6.4	Case Study and Analysis	118
6.4.1	G_{CPPS} Generation	119
6.4.2	Experimental Data Collection.....	121
6.4.3	CGAN Modeling	121
6.4.4	Security Analysis Results	122
6.5	Summary.....	125
	References.....	125

Part III Data-Driven Digital Twin Modeling

7	Dynamic Data-Driven Digital Twin Modeling	129
7.1	Introduction.....	129
7.1.1	Research Challenges.....	130
7.1.2	Contributions	130
7.1.3	Digital Twin Model.....	131
7.2	Digital Twin of Cyber-Physical Additive Manufacturing System	131
7.2.1	Key Performance Indicators (KPIs)	132
7.3	Keeping Digital Twin Updated	135
7.4	Building Digital Twin	136
7.4.1	Sensor/Emission Modality Selection	136
7.4.2	Feature Engineering	136
7.4.3	Sensor Positioning	137
7.4.4	Data-Driven Models	138
7.5	Experimental Setup	138
7.5.1	The Test-Bed.....	139
7.5.2	Test 3D Objects	142
7.5.3	Data Collection	142
7.5.4	Data Segmentation.....	144
7.6	Simulation and Results for Digital Twin Models	146
7.6.1	Digital Twin Models.....	146
7.6.2	Aliveness	146
7.7	Summary.....	151
	References.....	152
8	IoT-Enabled Living Digital Twin Modeling.....	155
8.1	Introduction.....	155
8.1.1	Research Challenges.....	156
8.1.2	Contribution.....	156
8.1.3	Motivational Case Study for Multi-Sensor Data Analysis	157
8.1.4	Related Work	158
8.2	Background.....	159
8.2.1	Concept Definition	159
8.2.2	IoT Sensor Data as Side-Channels.....	160
8.2.3	Metric for Quality Measurement.....	161

8.3	Building the Digital Twin	161
8.3.1	$DT_{product}$ Parsing	163
8.3.2	Feature Extraction	163
8.3.3	Synchronize and Segment	163
8.3.4	Clustering Algorithm	164
8.3.5	Anomaly Localization Algorithm	165
8.3.6	Digital Twin Update Algorithm	165
8.3.7	Quality Inference Model	166
8.4	Experimental Setup	167
8.4.1	IoT Sensors	167
8.4.2	Digital Twin Parameters	169
8.4.3	Sensor Position Analysis	169
8.4.4	Performance of Clustering Algorithms	171
8.4.5	Anomaly Localization Accuracy	171
8.4.6	System Degradation Prediction Analysis	174
8.4.7	Quality Inference	176
8.4.8	Comparative Analysis	177
8.5	Discussion	179
8.6	Summary	180
	References	181

Part IV Non-Euclidean Data-Driven Modeling of Cyber-Physical Systems

9	Non-euclidean Data-Driven Modeling Using Graph Convolutional Neural Networks	185
9.1	Introduction	185
9.2	Related Work	186
9.3	Graph Learning Using Convolutional Neural Network	187
9.3.1	Knowledge Graph Extraction	188
9.3.2	Attribute Embedding	189
9.3.3	Neighbor Nodes Aggregation	189
9.3.4	Structural Graph Convolutional Neural Network Layers	192
9.3.5	Classification for Engineering Design Abstraction	197
9.3.6	Graph Learning Algorithm Hyper-Parameters	197
9.4	GrabCAD Dataset	198
9.5	Results	199
9.5.1	Activation Functions	201
9.5.2	Kernel Size	201
9.5.3	Dropout	202
9.5.4	Layers	203
9.6	Discussion	204
9.7	Summary	205
	References	205

10	Dynamic Graph Embedding	209
10.1	Introduction	209
10.1.1	Research Challenges	210
10.1.2	Contribution	210
10.2	Related Work	210
10.2.1	Static Graph Embedding	211
10.2.2	Dynamic Graph Embedding	211
10.2.3	Dynamic Link Prediction	212
10.3	Motivating Example	212
10.4	Methodology	215
10.4.1	Problem Statement	215
10.4.2	dyngraph2vec Algorithm	215
10.4.3	Optimization	218
10.5	Experiments	218
10.5.1	Datasets	219
10.5.2	Baselines	220
10.5.3	Evaluation Metrics	220
10.6	Results and Analysis	220
10.6.1	SBM Dataset	221
10.6.2	Hep-th Dataset	222
10.6.3	AS Dataset	222
10.6.4	MAP Exploration	223
10.6.5	Hyper-Parameter Sensitivity: Lookback	224
10.6.6	Length of Training Sequence Versus MAP Value	225
10.7	Discussion	226
10.8	Summary	227
	References	227
	Index	231