

Elisa Bertino · Sonam Bhardwaj ·
Fabrizio Cicala · Sishuai Gong ·
Imtiaz Karim · Charalampos Katsis ·
Hyunwoo Lee · Adrian Shuai Li ·
Ashraf Y. Mahgoub

Machine Learning Techniques for Cybersecurity

Synthesis Lectures on Information Security, Privacy, and Trust

Series Editors


Elisa Bertino , Purdue University, West Lafayette, IN, USA

Elena Ferrari, University of Insubria, Como, Italy

The series publishes short books on topics pertaining to all aspects of the theory and practice of information security, privacy, and trust. In addition to the research topics, the series also solicits lectures on legal, policy, social, business, and economic issues addressed to a technical audience of scientists and engineers. Lectures on significant industry developments by leading practitioners are also solicited.

Elisa Bertino · Sonam Bhardwaj ·
Fabrizio Cicala · Sishuai Gong · Imtiaz Karim ·
Charalampos Katsis · Hyunwoo Lee ·
Adrian Shuai Li · Ashraf Y. Mahgoub

Machine Learning Techniques for Cybersecurity

Elisa Bertino 
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Sonam Bhardwaj
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Fabrizio Cicala
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Sishuai Gong
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Imtiaz Karim
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Charalampos Katsis
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Hyunwoo Lee
Department of Energy Engineering
(Energy AI Track)
Korea Institute of Energy Technology
(KENTECH)
Naju, Korea (Republic of)

Adrian Shuai Li
Department of Computer Science
Purdue University
West Lafayette, IN, USA

Ashraf Y. Mahgoub
Department of Computer Science
Purdue University
West Lafayette, IN, USA

ISSN 1945-9742 ISSN 1945-9750 (electronic)
Synthesis Lectures on Information Security, Privacy, and Trust
ISBN 978-3-031-28258-4 ISBN 978-3-031-28259-1 (eBook)
<https://doi.org/10.1007/978-3-031-28259-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

This work is subject to copyright. All rights are solely and exclusively licensed 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

Preface

The protection of information and information infrastructures from unauthorized access, use, disclosure, disruption, modification, or destruction is today more critical than ever as they represent attractive targets for a diversity of malicious actors. Those actors have different motivations, such as financial gains, sabotage, espionage, intellectual property theft, and data theft. They may be supported by enemy nations and can leverage an arsenal of attack toolkits, zero-day vulnerabilities, and compromised passwords available on the dark web.

Research and design of defense techniques have, however, greatly progressed over the past 30 years, and there is an increased general awareness of threats and attacks in cyberspace and the need for better defenses by public and private organizations and governmental agencies as well as by the general public.

In order to devise more effective defenses, recent security solutions leverage machine learning techniques, which are today quite effective because of the huge and diversified technical advances in the area of machine learning combined with big data collection and analysis capabilities. We observed that in the last 10 years, the use of machine learning techniques for security tasks has been steadily increasing in research and also in practice. Many recent papers have proposed approaches for specific tasks, such as software security analysis and anomaly detection. However, these approaches differ in many aspects, for example, with respect to the types of features used in machine learning models and the datasets used for training the models. Also, the use of machine learning for security tasks is not trivial. For example, suppose one would like to use machine learning techniques for network intrusion detection. In that case, one has to understand the features to extract from network flows for proper training and use machine learning models able to classify the flows as benign or malicious. To date, however, there is no book or survey article that systematically covers the entire area of machine learning techniques for cybersecurity. This monograph aims to address such a gap.

A comprehensive discussion and analysis of the various machine learning techniques require, however, a proper taxonomy of these techniques. We decided to organize the discussion around the following main cybersecurity functions: security policy learning, software security analysis, hardware security analysis, detection, and attack management.

For some of those functions, many approaches have been proposed, such as the ones for intrusion detection. For others, approaches are still very limited—for example, for attack management. For topics on which many approaches have been proposed, we selected the approaches that we considered most interesting for the discussion. The monograph also covers challenges in using machine learning for cybersecurity—many of which are common to other domains; however, we try, whenever possible, to discuss these challenges from a cybersecurity perspective. Throughout the discussion, we also point out research directions based on our analysis of existing approaches, techniques, and tools. The book also includes a chapter that can be interesting from an educational point of view. This chapter covers three case studies—each related to a well-known cyber attack; for each attack, we discuss which machine learning technique(s) (if any) would have prevented/mitigated which steps of the attack. The case studies are interesting as they show that attacks are typically multi-steps, so one must deploy many different defense techniques to enhance security.

Writing this monograph has been an exciting journey for us as we had several interesting discussions and also identified new ideas for future research. We hope you will enjoy learning about machine learning for cybersecurity as much as we have!!

West Lafayette, IN, USA
West Lafayette, IN, USA
West Lafayette, IN, USA
West Lafayette, IN, USA
West Lafayette, IN, USA
West Lafayette, IN, USA
Naju, Korea (Republic of)
West Lafayette, IN, USA
West Lafayette, IN, USA
December 2022

Elisa Bertino
Sonam Bhardwaj
Fabrizio Cicala
Sishuai Gong
Imtiaz Karim
Charalampos Katsis
Hyunwoo Lee
Adrian Shuai Li
Ashraf Y. Mahgoub

Acknowledgements This work has been partially funded by NSF under Grants DGE-2114680 and CNS-2112471.

Contents

1	Introduction	1
1.1	Artificial Intelligence, Machine Learning, and Deep Learning	2
1.2	Security Functions	3
1.2.1	Security Policy Learning	4
1.2.2	Software Security Analysis	4
1.2.3	Hardware Security Analysis	4
1.2.4	Detection	5
1.2.5	Attack Management	5
1.3	Security Life Cycle	5
1.4	Organization of This Monograph	7
2	Background on Machine Learning Techniques	9
2.1	Preliminary Notions	9
2.2	Neural Networks	10
2.3	Autoencoders	12
2.3.1	Denoising Autoencoders	12
2.3.2	Variational Autoencoders	13
2.4	Recurrent Networks and Long Short-Term Memory	14
2.5	Attention Mechanism	15
2.6	Reinforcement Learning	17
2.7	Transfer Learning	18
2.7.1	Notations and Definitions	18
2.7.2	Fine-Tuning	19
2.7.3	Domain Adaptation	19
2.8	Embedding Techniques	21
3	Security Policy Learning	23
3.1	Access Control Policies	24
3.1.1	Learning Access Control Policies	25
3.1.2	Policy Transfers Across Domains	30

3.1.3	DL Models for Access Control Decisions	31
3.1.4	Model-Independent Policy Mining	32
3.2	Network Security Policies	34
3.2.1	Firewall Rule Miners	35
3.2.2	ML-Based Firewall Systems	37
3.2.3	Network Security Policies for Traditional Networks	38
3.2.4	Network Security Policies for IoT	39
3.3	Privacy Policy Contradiction Identification	40
3.4	Adaptive Security Policy Learning Systems	41
3.5	Research Directions	44
4	Software Security Analysis	47
4.1	Static Analysis	48
4.1.1	A Survey on Machine Learning Techniques for Source Code Analysis	48
4.1.2	Recent Approaches	50
4.2	Fuzzing Techniques	52
4.2.1	Fuzzing Steps	53
4.2.2	ML-Based Fuzzing	54
4.3	NLP-Based Techniques for Specification Analysis	59
4.3.1	Finite State Machine Extraction	59
4.3.2	Zero-Shot Protocol Information Extraction	60
4.3.3	4G LTE Testcase Generation	62
4.3.4	Semantic Information Analysis of Developer's Guide	63
4.3.5	Security-Specific Change Request Detection	65
4.3.6	Capturing Privacy-Related Settings in Android	66
4.4	Supporting Techniques	66
4.4.1	Neural Network-Based Function and Type Identification	66
4.4.2	Reverse Engineering	67
4.5	Research Directions	69
5	Hardware Security Analysis	71
5.1	ML-Based Hardware Test Input Generation	72
5.2	ML-Based Detection of Hardware Trojans	75
5.3	Research Directions	77
6	Detection	79
6.1	Types of Malware	80
6.2	ML-Based Anomaly Detection	81
6.2.1	Networks	82
6.2.2	IoT Systems	84
6.2.3	Cyber-Physical Systems	86
6.2.4	Ransomware	89

6.3	Malware Detection and Classification	91
6.3.1	Portable Executable File Format	92
6.3.2	Analysis and Detection Techniques	94
6.3.3	Data Preparation and Labeling for ML-Based Malware Analysis	95
6.3.4	Malware Detection and Analysis: Features for Specific Platforms	98
6.3.5	Malware Representation	101
6.4	Research Directions	104
7	Attack Management	105
7.1	Attack Mitigation	106
7.2	Defense Enhancement	107
7.3	Digital Forensics	109
7.3.1	NLP-Based Attack Analysis	110
7.3.2	Transformer-Based Contextual Analysis of Security Events	110
7.3.3	GNN-Based Memory Forensic Analysis	111
7.3.4	An Explanation Method for GNNs Models	112
7.4	Research Directions	113
8	Case Studies	115
8.1	The Target Data Breach	115
8.2	The SolarWinds Attack	121
8.3	The WannaCry Ransomware	124
9	Challenges in the Use of ML for Security	131
9.1	Data Availability and Quality	131
9.2	Selection of Models, Hyperparameters, and Configurations	133
9.2.1	Selecting the Right Model	133
9.2.2	Hyperparameter and Configuration Tuning	133
9.3	Ethics	134
9.3.1	Explainability	136
9.3.2	Fairness	137
9.3.3	Robustness	139
9.3.4	Transparency	140
9.3.5	Privacy	141
9.4	Security of ML	141
9.5	Research Directions	143
10	Concluding Remarks	145
	Appendix: Publicly Available Datasets	147
	References	151

Acronyms

ABAC	Attribute-based Access Control
AI	Artificial Intelligence
ANN	Artificial Neural Network
ARM	Advanced RISC Machine
APT	Advanced Persistent Threats
ARuM	Association Rule Mining
AST	Abstract Syntax Tree
BO	Bayesian Optimization
CVE	Common Vulnerabilities and Exposures
CNN	Convolutional Neural Network
CFG	Control Flow Graph
DA	Domain Adaptation
DDG	Data Dependency Graphs
DGCNN	Deep Graph Convolutional Neural Network
DL	Deep Learning
DM	Data Mining
DNN	Deep Neural Network
DoS	Denial of Service
eBPF	Extended Berkeley Packet Filter
ELF	Executable and Linkable File
FSM	Finite State Machine
GAN	Generative Adversarial Networks
GNN	Graph Neural Network
HI	Hazard Indicator
IoT	Internet of Things
IDS	Intrusion Detection System
IVFG	Inter-procedural Value Flow Graph
LSTM	Long Short-Term Memory
LTL	Linear Temporal Logic
ML	Machine Learning

MM	Memory Management
MOS	Memory Operation Synopsis
MAB	Multi-Armed Bandit
MLP	Multi-Layer Perceptron
NLP	Natural Language Processing
NN	Neural Networks
NSF	Network Security Function
PCA	Principle Component Analysis
PE	Portable Executable
RBAC	Role-based Access Control
RF	Random Forest
RL	Reinforcement Learning
RNN	Recurrent Neural Network
SDN	Software-Defined Networks
SFS	Sequential Forward Search
SGD	Stochastic Gradient Descent
SR	Security Requirement
SR-CR	Security Related Change Request
SVM	Support Vector Machine
SVR	Support Vector Regression
TF-IDF	Term Frequency-Inverse Document Frequency
TL	Transfer Learning
t-SNE	t-Distributed Stochastic Neighbor Embedding
VAE	Variational Auto Encoder
VGG	Visual Geometric Group



Attacks to computer, information, and communication systems, collectively referred to as *cyberspace*, are on a dramatic increase. Attacks have a variety of goals, such as data ransoms, denial of service, critical infrastructure sabotage, data theft, and information tampering, and are carried out by many different actors with motivations that include financial gains, cyberwar, misinformation, and disinformation. Security of the cyberspace, referred to as *cybersecurity* (security, for short), is more critical than ever for our society that increasingly relies on cyberspace for all services, functions, and processes we may think of.

It however is well known that there is no system that can be 100% secure from all adversaries. Critical systems, protocols, and software considered secure are constantly analyzed by intelligent adversaries with sufficient resources, leading to the identification of vulnerabilities allowing these adversaries to craft exploits for breaking into computer and network systems. Vulnerabilities, unknown to the creators or users of a system, are called zero-day vulnerabilities, and the exploits that take advantage of them are called zero-day exploits [213]. Recent attacks are increasingly more sophisticated in the vulnerabilities they exploit, and are supported by the availability on the dark web of attack toolkits, detailed zero-day vulnerability information, and compromised security credentials.

The huge expansion of cyberspace due to Internet of Things (IoT) devices and systems, robots, autonomous vehicles, and new wireless and cellular technologies, each with different security postures, has also substantially increased the attack surface. Consequently, the number of adversaries attempting to find new ways of breaking into these systems has skyrocketed. It is clear that protecting the cyberspace requires an array of advanced technical defenses as well as their systematic deployment based on a security life cycle.

In order to devise more effective defenses, recent security solutions leverage machine learning (ML) techniques, which are today widely applied because of their technical significant advances combined with big data collection and analysis capabilities. However, a major problem is that the application of ML techniques to cybersecurity is not trivial. For

example, if one would like to use ML techniques to classify malware, one has to understand the features to extract from malware for properly training and using ML classification models.

1.1 Artificial Intelligence, Machine Learning, and Deep Learning

Today, terms such as artificial intelligence (AI), machine learning (ML), and deep learning (DL) are widely used not only in the technical literature but also in the media, popular culture, advertising, and more [41]. These terms are often used interchangeably. There are however differences among the respective areas that we outline in what follows.

AI is a field that started in 1956. The goal then, as now, was to use computer systems to perform tasks requiring human intelligence [124]. The initial focus was on tasks like playing checkers and solving logic problems. AI then specialized based on specific application areas, such as robotics, natural language processing, and computer vision [13]. Early AI approaches were mainly based on declarative knowledge provided by humans, for example, in terms of logical rules and ontologies. Such knowledge would then be used as input by inference mechanisms, often based on some formal logic. Today, AI encompasses a broad set of technology solutions that can learn on their own.

A major problem of early AI approaches was the lack of scalability because of their reliance on human inputs. ML techniques, which started to be widely used in the 1980s, address this problem by relying on data, instead of explicit human input. They apply statistical methodologies to identify patterns occurring in data. They improve their prediction tasks every time they acquire new data. A special category of ML techniques is represented by data mining (DM), which basically addressed the problem of identifying patterns on very large datasets. Most research in DM was initiated by the database community, which for example introduced the pioneering concept of association rule mining [10] and designed efficient algorithms to minimize scans on data stored in secondary storage. However, even though ML techniques can improve their prediction accuracy, “they only explore data based on programmed data feature extraction; that is, they only look at data in the way we program them to do so. They do not adapt on their own to look at data in a different way” [41].

DL techniques represent an important category of ML techniques that address the shortcoming of early ML techniques. DL essentially refers to algorithms that adapt, when exposed to different situations or data patterns. Vaguely inspired by biological neural networks, DL algorithms try to learn various characteristics from data and use them for decision-making/prediction on similar unseen data. DL techniques have gained interest because of the increased amounts of data available and their various algorithmic innovations as well as significant improvements in computing capabilities enabled by GPUs, which have made fast training and deployment of DL models possible [169]. DL has been tremendously successful at tasks such as image classification, object detection, and text and voice recognition.

1.2 Security Functions

A detailed and comprehensive discussion of ML-based techniques for cybersecurity is best based on a taxonomy of ML-based *security functions*, that is, security techniques and processes for which ML approaches have been proposed. The taxonomy we refer to is shown in Fig. 1.1. As we can see from the taxonomy, the top five categories correspond to major ML-based security functions that we briefly discuss in what follows.

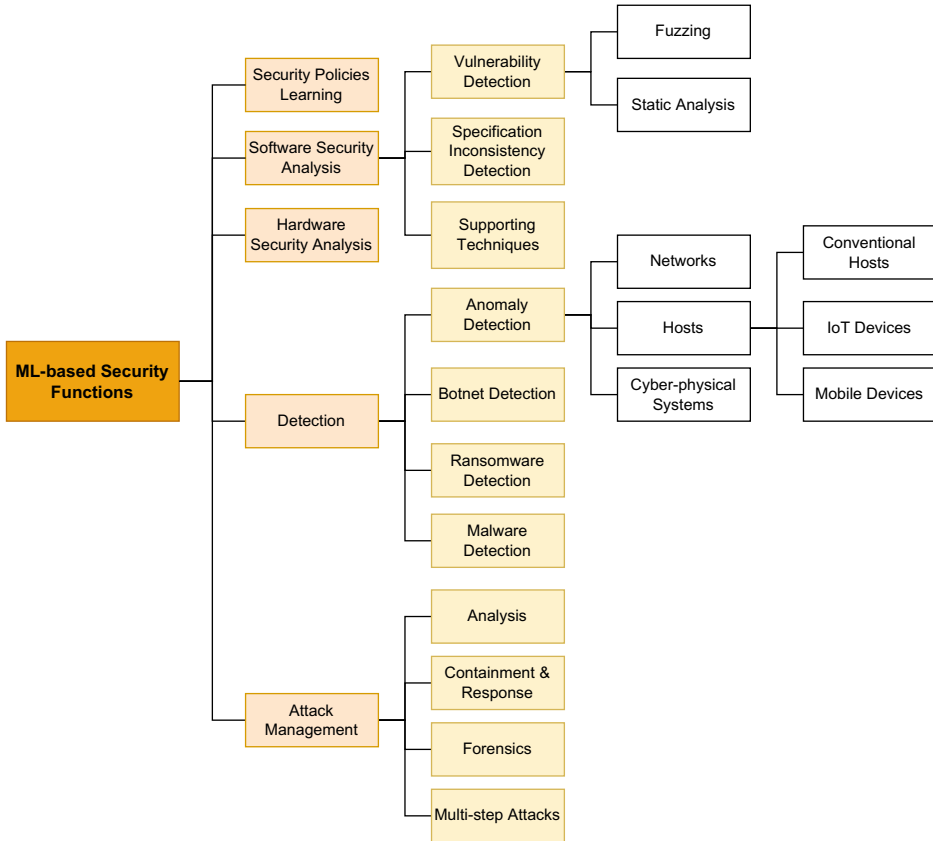


Fig. 1.1 Taxonomy of ML-based security functions

1.2.1 Security Policy Learning

Security policies are critical for configuring security tools and appliances, including access control systems, authentication systems, and network firewalls. As a manual specification of policies is time-consuming and not scalable, it has been one of the first areas to which ML techniques have been applied. The relevance of security policy learning will be increasing given the recent zero-trust architectures [192] and frameworks [133], which will require the specification, deployment, and testing of very large number of policies.

1.2.2 Software Security Analysis

Software systems are key components of all infrastructure and application domains we may think of. However, software systems are still insecure, despite the fact that the “problem of software security” had been known to the industry and research communities for decades [31]. Therefore, it is not surprising that software security analysis has recently become one relevant application area for ML techniques. ML-based approaches range from enhancing fuzzing to ensure better coverage [187] to predicting the effects of different combinations of control parameter values for drones [95] and making static analysis scalable for large code bases [119]. Such initial approaches show that ML techniques can make software security analysis more effective. We can expect that this area will see many novel ML-based approaches to be developed, given the pressing problem of software security.

1.2.3 Hardware Security Analysis

Hardware is commonly assumed to be the root-of-trust for computer systems, in that trust is established by committing functionality to silicon, which represents a stronger security foundation compared to the flexible but more vulnerable software [223]. However, hardware can be attacked, via for example side channels [188], and can even include malicious components (e.g., hardware Trojans). The major use of ML has been for the security evaluation of cipher implementations against side-channel attacks [100] and the construction of attack models against physical unclonable functions [101]. More recent applications of ML include the characterization of faults that can be exploited by attackers [206] and security-aware design flow for chip design [118, 188]. However, ML techniques will undoubtedly enable the design of novel approaches expanding the faults and vulnerabilities that can be detected.