1. Record Nr. UNINA9910886990203321

Autore	Garcia-Alfaro Joaquin
Titolo	Computer Security – ESORICS 2024 : 29th European Symposium on Research in Computer Security, Bydgoszcz, Poland, September 16–20, 2024, Proceedings, Part I / / edited by Joaquin Garcia-Alfaro, Rafa Kozik, Micha Chora, Sokratis Katsikas
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2024
ISBN	3-031-70879-2
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (411 pages)
Collana	Lecture Notes in Computer Science, , 1611-3349 ; ; 14982
Altri autori (Persone)	KozikRafa ChorasMicha KatsikasSokratis
Disciplina	005.8
Soggetti	Data protection Cryptography Data encryption (Computer science) Computer networks - Security measures Computer networks Computer systems Data and Information Security Cryptology Security Services Mobile and Network Security Computer Communication Networks Computer System Implementation
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Organization -- Contents - Part I -- Security and Machine Learning -- Attesting Distributional Properties of Training Data for Machine Learning -- 1 Introduction -- 2 Background -- 3 Problem Statement -- 4 Distributional Property Attestation Mechanisms -- 5 Experimental Setup -- 6 Experimental Evaluation -- 6.1 Inference-Based Attestation -- 6.2 Cryptographic Attestation -- 6.3 Hybrid Attestation -- 7 Related Work -- 8 Discussions -- A Details for

Cryptographic Attestation -- References -- Towards Detection-Recovery Strategy for Robust Decentralized Matrix Factorization -- 1 Introduction -- 2 Background and Related Work -- 2.1 Decentralized Matrix Factorization -- 2.2 Threats and Remedies in Distributed Learning -- 3 The Vulnerability of DMF -- 3.1 Threat Model -- 3.2 The Tampering Attack on DMF -- 4 Our Approach -- 4.1 The Decentralized Detection -- 4.2 The Recovery Strategy -- 4.3 Comprehensive Framework -- 5 Experiment -- 5.1 Experimental Setup -- 5.2 The Threat of the Tampering Attack -- 5.3 Effective Defense with the Detection-Recovery Strategy -- 5.4 Adaptive Attack -- 5.5 More Results -- 6 Conclusion and Discussion -- A Technical Proofs -- References -- Bayesian Learned Models Can Detect Adversarial Malware for Free -- 1 Introduction -- 2 Background and Related Work -- 3 Problem Definition -- 3.1 Threat Model -- 3.2 Adversarial Malware Attacks -- 4 Measuring Uncertainty -- 4.1 Bayesian Machine Learning for Malware Detection -- 4.2 Uncertainty Measures -- 5 Experiments and Results -- 5.1 Experimental Setup -- 5.2 Clean Performance (No Attacks) in Android Domain -- 5.3 Robustness Against Problem-Space Adversarial Android Malware -- 5.4 Robustness Against Feature-Space Adversarial Android Malware -- 5.5 Generalization to PDF Malware -- 5.6 Generalization to Windows PE Files -- 6 Identifying Concept Drift. 7 Model Parameter Diversity Measures -- 8 Threat to Validity -- 9 Conclusion -- References -- Resilience of Voice Assistants to Synthetic Speech -- 1 Introduction -- 2 Voice Assistants -- 3 Related Work -- 3.1 Deepfake Speech Synthesis -- 3.2 Spofing Attacks on Biometrics Systems -- 3.3 Spoofing Voice Assistants -- 4 Experiments -- 4.1 Used Speech Synthesisers -- 4.2 Environment Description -- 4.3 Details of the Setup -- 5 Experimental Evaluation -- 6 Threat Analysis -- 7 Discussion -- 7.1 Observations -- 7.2 Mitigation Methods -- 8 Conclusions -- References -- Have You Poisoned My Data? Defending Neural Networks Against Data Poisoning -- 1 Introduction -- 2 Background -- 2.1 Feature Collision -- 2.2 Convex Polytope and Bullseye Polytope -- 2.3 Gradient Matching -- 3 System and Threat Models -- 3.1 System Model -- 3.2 Threat Model -- 4 Our Approach -- 4.1 Formal Description of the Approach -- 5 Experimental Setup -- 5.1 Dataset -- 5.2 Poison Generation Algorithms and Defenses -- 6 Evaluation -- 6.1 Poisons vs Clean Samples: A Characteristic Vector Perspective -- 6.2 Poison Detection -- 7 Related Works -- 8 Conclusions and Future Work -- A Implementation Details -- B Additional Experimental Results -- References -- Jatmo: Prompt Injection Defense by Task-Specific Finetuning -- 1 Introduction -- 2 Background -- 2.1 LLM-Integrated Applications -- 2.2 Prompt Injections -- 2.3 Examples -- 3 Related Works -- 3.1 Types of Attacks -- 3.2 Pitfalls of Traditional Defenses -- 4 Jatmo -- 4.1 Synthetic Input Generation -- 5 Results -- 5.1 Experimental Methodology -- 5.2 Main Results -- 5.3 Training with Less Data -- 5.4 Synthetic Dataset Generation -- 6 Discussion -- 7 Summary -- A Appendix -- A.1 Detailed Task Parameters -- References -- PointAPA: Towards Availability Poisoning Attacks in 3D Point Clouds -- 1 Introduction -- 2 Related Work.
2.1 Adversarial Attacks of 3D Point Clouds -- 2.2 Backdoor Attacks of 3D Point Clouds -- 2.3 Availability Poisoning Attacks in 2D Images -- 3 Methodology -- 3.1 Threat Model -- 3.2 Motivation and Challenges -- 3.3 Inspiration and Exploration -- 3.4 PointAPA: Point Cloud Availability Poisoning Attack -- 3.5 Why Does PointAPA Work? -- 4 Experiments -- 4.1 Experimental Settings -- 4.2 Evaluation on PointAPA -- 4.3 Evaluation Under Overlapped Rotation Angles -- 4.4 Robustness to Defense Schemes -- 4.5 Hyper-parameter Analysis -- 5 Conclusion --

A Appendix -- References -- ECLIPSE: Expunging Clean-Label Indiscriminate Poisons via Sparse Diffusion Purification -- 1 Introduction -- 2 Related Work -- 2.1 Clean-Label Indiscriminate Poisoning Attacks -- 2.2 Defenses Against Poisoning Attacks -- 3 Methodology -- 3.1 Threat Model -- 3.2 Motivation for Studying Defenses Against CLBPAs -- 3.3 Key Intuition and Theoretical Insight -- 3.4 Challenges and Approaches -- 3.5 Our Design for ECLIPSE -- 4 Experiments -- 4.1 Experimental Settings -- 4.2 Evaluation of ECLIPSE -- 4.3 Purification Visual Effect -- 4.4 Resistance to Potential Adaptive Attacks -- 4.5 Hyper-Parameter Analysis -- 4.6 Ablation Study -- 4.7 Analysis of ECLIPSE -- 5 Conclusion and Limitation -- A Appendix -- References -- MAG-JAM: Jamming Detection via Magnetic Emissions -- 1 Introduction -- 2 MAG-JAM Overview, Scenario and Adversary Model -- 2.1 MAG-JAM Overview -- 2.2 Scenario and Adversary Model -- 3 Jamming Detection Using Magnetic Sensor -- 3.1 DRV425 Magnetic Sensor Setup -- 3.2 Magnetic Sensor Results -- 3.3 Early Jamming Detection -- 4 MAG-JAM Evaluation -- 4.1 Experimental Setup - Magnetic Probe -- 4.2 Magnetic Emissions Collection Using the Magnetic Probe -- 4.3 Dataset Description -- 4.4 Features Extraction -- 4.5 Jamming Detection Using Autoencoder -- 5 Discussion -- 6 Related Work.
7 Conclusion -- References -- Fake or Compromised? Making Sense of Malicious Clients in Federated Learning -- 1 Introduction -- 2 Types of Byzantine-Robust Aggregation Rules -- 3 Distinguishing Fake And Compromised Adversary Models -- 3.1 Adversary with Fake Clients -- 3.2 Adversary with Compromised Clients -- 4 Our Proposed Hybrid Adversary Model -- 4.1 Comparing the Costs of Different Adversaries -- 5 Experimental Setup -- 5.1 Datasets and Hyperparameters -- 5.2 Evaluation Metric -- 5.3 Generating Synthetic Data Using DDPM -- 6 Experiments -- 6.1 Attacking Agnostic Robust AGRs -- 6.2 Attacking Adaptive Robust AGRs -- 7 Conclusions -- A Auxiliary Results of Model Poisoning Attacks Against Aware AGRs -- References -- Beyond Words: Stylometric Analysis for Detecting AI Manipulation on Social Media -- 1 Introduction -- 2 Related Work -- 2.1 Pervasiveness and Influence of Social Bots -- 2.2 Evaluation and Detection of Social Bots and AI-Text -- 3 Study Design -- 3.1 Data Generation and Preparation -- 3.2 Stylometric Analysis -- 3.3 Analysis Methods -- 4 Results -- 5 Threats to Validity -- 6 Conclusions -- References -- FSSiBNN: FSS-Based Secure Binarized Neural Network Inference with Free Bitwidth Conversion -- 1 Introduction -- 1.1 Related Work -- 1.2 Our Contributions -- 2 Preliminaries -- 2.1 Binarized Neural Networks -- 2.2 Additive Secret Sharing -- 2.3 Function Secret Sharing -- 3 Secure BNN Inference Framework -- 3.1 The FSSiBNN Overview -- 3.2 Bitwidth-Reduced Parameter Encoding Scheme with Free Bitwidth Conversion -- 3.3 Online-Efficient Secure Non-linear BNN Layers via FSS -- 4 Secure BNN Inference Protocol -- 4.1 Secure Fully Connected and Convolutional Layers -- 4.2 Secure Batch Normalization and Binary Activation Layers -- 4.3 Secure Max Pooling Layers -- 5 Theoretical Analysis and Experiment -- 5.1 Theoretical Analysis.
5.2 Experimental Results and Analysis -- 6 Conclusion -- A Proof of Sign Function Gate in Sect.4.2 -- B Analysis of Computation Complexity -- C Evaluation and Analysis of Inference Accuracy -- References -- Optimal Machine-Learning Attacks on Hybrid PUFs -- 1 Introduction -- 1.1 Problem Statement and Related Work -- 1.2 Contributions -- 1.3 Paper Organisation -- 2 Mathematical Representations of Hybrid PUFs -- 2.1 XOR Arbiter PUF -- 2.2 OR-AND-XOR-PUF -- 2.3 Homogeneous and Heterogeneous Feed-Forward XOR Arbiter PUF -- 2.4 Other Hybrid PUFs -- 2.5 State-of-Art

Modelling Structures -- 3 Methodology -- 3.1 Local Minima Problem -- 3.2 Modelling PUFs Using Miture-of-Experts -- 3.3 Routine Algorithm -- 3.4 Proposed Transition Theorem -- 4 Experiments and Evaluation -- 4.1 Modelling Hybrid PUFs Using the Generic Model -- 4.2 Modelling Hybrid PUFs Using the Proposed Transition Theorem -- 5 Conclusion -- A Transition Theorem and Proofs -- A.1 OAX-PUF -- B Feed-Forward PUF -- References -- Outside the Comfort Zone: Analysing LLM Capabilities in Software Vulnerability Detection -- 1 Introduction -- 2 Related Work -- 2.1 SAST-Based Vulnerability Detection -- 2.2 Task-Specific DL Models for Vulnerability Detection -- 2.3 LLM-Based Vulnerability Detection -- 3 Methodology -- 4 Experiments -- 4.1 Prompt Engineering and Hardware Setup -- 4.2 Datasets -- 5 Results and Discussion -- 6 Conclusions -- References -- ZeroLeak: Automated Side-Channel Patching in Source Code Using LLMs -- 1 Introduction -- 2 Background -- 3 Related Work -- 4 Threat Model and Scope -- 5 Methodology -- 5.1 Ensuring Constant-Time Execution -- 5.2 Mitigating Spectre-v1 -- 6 Evaluation -- 6.1 Patching Spectre-v1 Gadgets -- 6.2 Patching a Real World Spectre-v1 Gadget -- 6.3 Patching Real-World Javascript Libraries for Constant-Timeness -- 6.4 Comparison of LLMs -- 7 Discussion and Limitations.
8 Conclusion.

Sommario/riassunto

This four-volume set LNCS 14982-14985 constitutes the refereed proceedings of the 29th European Symposium on Research in Computer Security, ESORICS 2024, held in Bydgoszcz, Poland, during September 16–20, 2024. The 86 full papers presented in these proceedings were carefully reviewed and selected from 535 submissions. They were organized in topical sections as follows: Part I: Security and Machine Learning. Part II: Network, Web, Hardware and Cloud; Privacy and Personal Datat Protection. Part III: Software and Systems Security; Applied Cryptopgraphy. Part IV: Attacks and Defenses; Miscellaneous.