PhD Program in Security, Risk and Vulnerability

Course Offerings

Core- Curriculum Courses

Cybersecurity and Reliable AI

Academic Year 2023-24

Università
di Genova

Title: Theory and Practice of Learning from Data

Teachers: Luca Oneto

Email: luca.oneto@unige.it

Duration: 20 hours

Credits: 5 CFU

When: TBD

Where: TBD

CV: Cybersecurity & Reliable AI

Link: www.lucaoneto.it/teaching/tpld

Exam: Small presentation (max 30 min) on how the concepts presented in the course can be used/extended by the student.

Abstract: This course aims at providing an introductory and unifying view of information extraction and model building from data, as addressed by many research fields like Data Mining, Statistics, Computational Intelligence, Machine Learning, and Pattern Recognition. The course will present an overview of the theoretical background of learning from data, including the most used algorithms in the field, as well as practical applications.

Program:

- Inference: induction, deduction, and abduction
- Statistical inference
- Machine Learning
- Deep Learning
- Model selection and error estimation
- Implementation and Applications

References:

- C. C. Aggarwal "Data Mining - The textbook" 2015
- T. Hastie, R. Tibshirani, J. Friedman "The Elements of Statistical Learning: Data Mining, Inference, and Prediction" 2009.
- S. Shalev-Shwartz, S. Ben-David "Understanding machine learning: From theory to algorithms" 2014
- I. Goodfellow, Y. Bengio, A. Courville "Deep learning" 2016
- L. Oneto "Model Selection and Error Estimation in a Nutshell" 2020

Title: Information Hiding

Teacher: Luca Caviglione

Email: luca.caviglione@cnr.it

Duration: 20 hours

Credits: 5 CFU

When: Late Spring/Summer 2024 (i.e., June to July), 5 half-days from 9:00 to 13:00 CEST

Where: preferred venue is via Skype or Microsoft Teams to reach a wide audience. Otherwise, University of Genova – DIBRIS, Via Dodecaneso, Genova.

CV: Cybersecurity & Reliable AI

Link: https://www.linkedin.com/in/lucacaviglione/

Exam: A small presentation (about 15-30 minutes) on a real threat utilizing steganographic techniques or covert channels.

Abstract: Information hiding techniques are increasingly used in investigative journalism to protect the identity of sources or by malware to hide its existence and communication attempts. Therefore, understanding how information hiding can be used to empower privacy of users or endow malicious software with the ability of staying "under the radar" are essential to fully assess the modern cybersecurity panorama. In this perspective, the course introduces the use of information hiding in modern threats and privacy-enhancing architectures with emphasis on two different research areas, specifically: i) techniques for creating network covert channels for communicating with a remote command & control facility, exfiltrate sensitive information and or enforce privacy ii) how to create and detect a covert channel implementing an abusive local path between two colluding applications to bypass the security framework of mobile devices.

To give a comprehensive overview on information hiding and steganography, the course will also cover the use of information hiding and steganographic techniques for watermarking purposes. For instance, it will showcase the main mechanisms for watermarking images, sounds and network flows for management, retrieval, metadating, authentication and copyright

enforcement. The course will also discuss possible countermeasures or mitigation methodologies for facing the risks of the increasing amount of steganographic threats observed in the wild.

Program:

- Course introduction and a general view on information hiding.
- Information hiding as a cybersecurity threat: malware and colluding applications.
- Network covert channels (including air-gapped covert channels).
- Information hiding for watermarking, privacy enhancing, and metadating.
- Countermeasures (e.g., detecting obfuscated malware or removing ambiguities in protocols).

References:

- W. Mazurczyk, L. Caviglione, “Steganography in Modern Smartphones and Mitigation Techniques”, IEEE Communications Surveys & Tutorials, IEEE, Vol. 17, No.1, First Quarter 2015, pp. 334 - 357.
- L. Caviglione, W. Mazurczyk, “Never Mind the Malware, Here’s the Stegomalware”, IEEE Security & Privacy, Vol. 20, No. 5, pp. 101-106, Sept.-Oct. 2022.
- W. Mazurczyk, L. Caviglione, Information Hiding as a Challenge for Malware Detection, IEEE Security & Privacy, Vol. 13, No. 2, pp. 89-93, Mar.-Apr. 2015.
- L. Caviglione, M. Podolski, W. Mazurczyk, M. Ianigro, “Covert Channels in Personal Cloud Storage Services: the case of Dropbox”, IEEE Transactions on Industrial Informatics, IEEE, Vol. 13, No. 4, pp. 1921 - 1931, August 2017.
- L. Caviglione, M. Gaggero, J.-F. Lalande, W. Mazurczyk, M. Urbanski, “Seeing the Unseen: Revealing Mobile Malware Hidden Communications via Energy Consumption and Artificial Intelligence”, IEEE Transactions on Information Forensics & Security, IEEE, Vol. 11, No. 4, pp. 799 – 810, April 2016.
- W. Mazurczyk, L. Caviglione, “Cyber Reconnaissance Techniques”, Communications of the ACM, Vol. 64, No. 3, pp. 86-95, March 2021.
- Steg-in-the-wild (https://github.com/lucacav/steg-in-the-wild): a curated list of attacks observed in the wild taking advantage of steganographic or information-hiding-capable techniques.
- Steg-tools (https://github.com/lucacav/steg-tools): a list of software tools and resources for learning and experimenting with steganography and information hiding.

Title: Network monitoring and inspection

Teachers: Matteo Repetto

Email: matteo.repetto@ge.imati.cnr.it

Duration: 20 hours

Credits: 5 CFU

When: Winter (January – March 2024)

Where: TBD

CV: Cybersecurity & Reliable AI

Link:

Exam: Use of the tools introduced in the course to find cyber-attacks in network traffic traces.

Abstract: The Internet is the main carrier for cyber-attacks, so it is not surprising that most detection techniques build on network flow monitoring and packet inspection. There is a huge amount of information that potentially can be gathered from the network, but deep packet inspection at line rate is extremely challenging even in hardware, especially in case of high-speed links (1 Gbps and upward).

This course will give a basic understanding of common tools for flow monitoring and packet inspection, with specific emphasis on how to extract custom information that is ever more needed to detect modern attacks. Besides, the eBPF framework provided by the Linux kernel will be introduced as a power mechanism to build efficient, custom and portable inspection and enforcement processes.

Program:

- Introduction to network monitoring
- Overview of common network cyber-attacks
- Network protocols for flow monitoring

- Interactive inspection tools: wireshark, tshark, tcpdump for wired/wireless network monitoring
- Passive network monitoring and deep packet inspection: nProbe, PacketBeat, Zeek
- Programmable monitoring: Zeek scripting and eBPF

References:

- L. Caviglione, W. Mazurczyk, M. Repetto, A. Schaffhauser, M. Zuppelli. Kernel-level tracing for detecting stegomalware and covert channels in Linux environments. Computer Networks, Volume 191, May 2021. DOI: 10.1016/j.comnet.2021.108010
- M. Repetto, A. Carrega, R. Rapuzzi. An architecture to manage security operations for digital service chains. Future Generation Computer Systems. Volume 115, February 2021, Pages 251-266. DOI: 10.1016/j.future.2020.08.044
- M. Repetto, G. Bruno, J. Yusupov, G. Lamanna, B. Ertl, and A. Carrega. Automating Mitigation of Amplification Attacks in NFV Services. IEEE Transactions on Network and Service Management. *Early access.* DOI: 10.1109/TNSM.2022.3172880
- M. Zuppelli, M. Repetto, A. Schaffhauser, W. Mazurczyk, L. Caviglione. Code Layering for the Detection of Network Covert Channels in Agentless Systems. IEEE Transactions on Network and Service Management. *Early access.* DOI: 10.1109/TNSM.2022.3176752

Title: Verification of Neural Networks

Teachers: Stefano Demarchi – Armando Tacchella

Email: stefano.demarchi@edu.unige.it – armando.tacchella@unige.it

Duration: 20 hours

Credits: 5 CFU

When: From Monday June 10th to Friday June 14th 2024, from 14.00 to 18.00

Where: TBD

CV: Cybersecurity and Reliable AI

Link: https://dibris.unige.it/armando.tacchella@unige.it

Exam: Small project using the tools and techniques presented during the course

Abstract: In this course we present the tool NeVer2, an integrated environment for designing, learning and verifying (deep) neural networks. NeVer2 borrows its design philosophy from NeVer, the first package proposed in 2010 that integrated learning, automated verification and repair of (shallow) neural networks in a single tool. The goal of NeVer2 is to provide a similar integration for deep networks by leveraging a selection
of state-of-the-art learning frameworks and integrating them with verification algorithms to ease the scalability challenge and make repair of faulty networks possible.

Program:

- Verification of Neural Networks: problem definition, state of the art and current challenges
- Computing output images of ReLU-based neural networks with star-sets to obtain sound and complete verification algorithms
- Abstraction supported by star-sets to obtain sound verification algorithms
- Refinement and counter-example finding techniques
- An introduction to NeVer2 by example

References:

1. Luca Pulina, Armando Tacchella: An Abstraction-Refinement Approach to Verification of Artificial Neural Networks. CAV 2010: 243-257
2. Luca Pulina, Armando Tacchella: Challenging SMT solvers to verify neural networks. AI Commun. 25(2): 117-135 (2012)
3. Francesco Leofante, Nina Narodytska, Luca Pulina, Armando Tacchella: Automated Verification of Neural Networks: Advances, Challenges and Perspectives. CoRR abs/1805.09938 (2018)
4. Dario Guidotti, Francesco Leofante, Luca Pulina, Armando Tacchella: Verification of Neural Networks: Enhancing Scalability Through Pruning. ECAI 2020: 2505-2512
5. Dario Guidotti, Luca Pulina, Armando Tacchella : pyNeVer a Framework for Learning and Verification of Neural Networks. ATVA 2021.
6. Stefano Demarchi, Dario Guidotti, Andrea Pitto, Armando Tacchella: Formal Verification Of Neural Networks: A Case Study About Adaptive Cruise Control. ECMS 2022: 310-316

Title: Blockchain technology and DeFi security Teachers: Meriem Guerar

Email: meriem.guerar@unige.it

Duration: 20 hours (5 half-days)

Credits: 5 CFU

When: May 2024

Where: TBD (online, via MS Teams and/or in presence at DIBRIS, Via Dodecaneso 35)

CV: Cybersecurity & Reliable AI

Link: -

Exam:
Conduct a security audit of a provided Solidity smart contract. Students must identify potential vulnerabilities covered during the course within the contract, write a report summarizing their findings, and present them.

Abstract:
The growth of the decentralized finance (DeFi) ecosystem, based on blockchain technology and smart contracts, has led to an increased demand for secure and reliable smart contract development. To date, attacks against smart contracts and DeFi protocols have resulted in billions of dollars in financial losses, posing a significant threat to the security of the entire DeFi ecosystem. The course offers a detailed exploration of the knowledge necessary to understand smart contracts and DeFi security. This involves in-depth discussions on the fundamental aspects of blockchain technology, with a particular emphasis on Ethereum. It also delves into the mechanics of smart contract development, covering the fundamentals of the Solidity programming language. Furthermore, the course provides foundational insights into the world of DeFi, shedding light on its protocols and the principles underlying this decentralized financial ecosystem. By the conclusion of this course, students will have acquired a profound understanding of the security principles that distinguish blockchain technology from conventional systems, as well as insights into its applications in real-world scenarios, such as food supply chain tracking and Self-Sovereign Identity (SSI). Additionally, students will gain expertise in developing and deploying smart contracts, with hands-on experience creating non-fungible tokens (NFTs). Moreover, students will attain fundamental knowledge of DeFi and the ability to conduct audits of Solidity-based smart contracts.

Program:

- Introduction to blockchain technology: origin and purpose, cryptography, consensus mechanisms, blockchain types, real-world use cases.

- Ethereum and decentralized applications: Ethereum Virtual Machine, Solidity Basics, Fungible and Non-fungible tokens, compiling and deploying smart contracts.
- Smart contract security: Reentrancy Attack, arithmetic overflow and underflow, force-feeding, accessing private data, unsafe DelegateCall, Denial of Service (DoS), phishing with TX.origin, hiding malicious code, etc.
- Introduction to DeFi (Decentralized Finance) and its applications: Decentralized Exchanges (DEX), Stable Coins, lending and borrowing, Automated Market Makers (AMMs), etc.
- DeFi security: Flash Loan Attacks, front-running, sandwich attacks, oracle manipulation, network and consensus attacks, etc.

References:
Course material (i.e. slides)

Title: Virtualization and Containers: An Introduction

Teacher: Enrico Russo

Email: enrico.russo@unige.it

Duration: 20 hours (5 half-days)

Credits: 5 CFU

When: Summer 2024 (ask the teacher)

Where: TBD

Curriculum: Cybersecurity and Reliable AI

Exam: TBD

Abstract: The course offers a fundamental knowledge of virtualization and containerization technologies, covering both theoretical concepts and practical applications. Specifically, it focuses on the comprehension of various virtualization types and their underlying mechanisms, emphasizing the network and storage levels. An exploration of the infrastructure as a code paradigm and orchestration will demonstrate the design of reproducible environments, along with managing high availability and fault tolerance in deployed applications. At the end of the course, students will be able to take advantage of such technologies by simplifying the creation of testbeds, simulating infrastructures and software ecosystems, and sharing environments for research experiments.

Program:

- Introduction to Virtualization
- (Virtual) Networking
- Automation and Infrastructure as Code (IaC)
- Storage
- Containers
- Containers Orchestration and Clustering

References:

- E. Russo, G. Costa, G. Longo, A. Armando and A. Merlo, "LiDiTE: a Full-Fledged and Featherweight Digital Twin Framework," in IEEE Transactions on Dependable and Secure Computing, doi: 10.1109/TDSC.2023.3236798.
- Longo, G., Orlich, A., Musante, S., Merlo, A., and Russo, E. (2023). MaCySTe: A virtual testbed for maritime cybersecurity. *SoftwareX*, *23*, 101426.
- Docker Docs - https://docs.docker.com/
- Podman Documentation - https://docs.podman.io/en/latest/
- Vagrant Documentation - https://developer.hashicorp.com/vagrant/docs
- Ansible Documentation - https://docs.ansible.com/

Title: Mobile Security

Teacher(s): Alessio Merlo, Luca Verderame

Email: alessio.merlo@ssuos.difesa.it, luca.verderame@unige.it

Duration: 20 hours (5 half-days)

Credits: 5 CFU

When: Summer 2024

Where: ONLINE

Curriculum: Cybersecurity and Reliable AI

Exam: each student will prepare a short report detailing a security analysis of a mobile app made using the concepts and tools presented in the course. At the completion of the project, students are required to present the main results included in the report (max 5 min).
Abstract: The course provides an overview of the main topics related to the security of mobile devices and applications. The course offers an insight into the leading mobile operating systems (i.e., Android and iOS) and their security issues. Moreover, the course provides a discussion on emerging mobile security technologies (e.g., Host-Based Card Emulation, and Trusted Execution Environment), security threats, and possible countermeasures. The second part of the course will cover the security of Mobile Applications, with a particular focus of state-of-the-art frameworks and methodologies for the vulnerability assessment of Android applications. Finally, the course will provide specific hands-on sessions with tools and techniques for the vulnerability assessment of Android applications.

Program:

- Introduction to mobile devices: history, features, and evolution;
- Architecture and security features of the principal mobile OSes (Android, iOS);
- Emerging mobile security technologies, security threats, and countermeasures;
- OWASP for Mobile Application Security Analysis;
- Reverse Engineering of Android Apps and Reversing Countermeasures;
- Static and Dynamic analysis of Android Apps.

During the hands-on sessions, students will get acquainted with a number of static and dynamic analysis tools, including

- ObfuscAPK - https://github.com/ClaudiuGeorgiu/Obfuscapk
- Apktool - https://ibotpeaches.github.io/Apktool/
- Jadx - https://github.com/skylot/jadx
- Inspeckage - https://github.com/ac-pm/Inspeckage
- Approver - https://approver.talos-sec.com
- Charles Proxy - https://www.charlesproxy.com

References:

- Romdhana, A., Merlo, A., Ceccato, M., & Tonella, P. (2022). Deep reinforcement learning for black-box testing of android apps. ACM Transactions on Software Engineering and Methodology.
- Merlo, A., Ruggia, A., Sciolla, L., & Verderame, L. (2021). ARMAND: Anti-repackaging through multi-pattern anti-tampering based on native detection. Pervasive and Mobile Computing, 76, 101443.
- Mayrhofer, R., Stoep, J. V., Brubaker, C., & Kralevich, N. (2021). The android platform security model. ACM Transactions on Privacy and Security (TOPS), 24(3), 1-35.
- OWASP Mobile Security Project. Mobile Security Testing Guide (MSTG) and Mobile Application Security Verification Standard (MASV) https://owasp.org/www-project-mobile-app-security/

Title: Adversarial Machine Learning

Teachers: Luca Demetrio

Email: luca.demetrio@unige.it

Duration: 12 hours (3 half-days)

Credits: 3 CFU

When: 3 – 4 – 5 July 2024

Where: online on MS Teams

CV: CSRAI

Link: -

Exam: 2 written assessments (one with multiple choice questions, one hands-on assessment using the SecML software library, https://secml.readthedocs.io/en/v0.15)

Abstract: Today machine-learning algorithms are used for many real-world applications, including image recognition, spam filtering, malware detection, biometric recognition. In these applications, the learning algorithm can have to face intelligent and adaptive attackers who can carefully manipulate data to purposely subvert the learning process. As machine learning algorithms have not been originally designed under such premises, they have been shown to be vulnerable to well-crafted attacks, including test-time evasion and training-time poisoning attacks (also known as adversarial examples). In particular, the security of cloud-based machine-learning services has been questioned through the careful construction of adversarial queries that can reveal confidential information on the machine-learning service and its users. This course aims to introduce the fundamentals of the security of machine learning, the related field of adversarial machine learning, and some techniques to assess the vulnerability of machine-learning algorithms and to protect them from adversarial attacks. We report application examples including object recognition in images, biometric identity recognition, spam and malware detection, with hands-on on attacks against machine learning and defences of machine-learning algorithms using the SecML software library, https://secml.readthedocs.io/en/v0.15/.

Program:

1. Introduction to adversarial machine learning: introduction by practical examples from computer vision, biometrics, spam, malware detection.

2. Design of learning-based pattern classifiers in adversarial environments. Modelling adversarial tasks. The two-player model (the attacker and the classifier). Levels of reciprocal knowledge of the two players (perfect knowledge, limited knowledge, knowledge by queries and feedback). The concepts of security by design and security by obscurity
3. System design: vulnerability assessment and defense strategies. Attack models against machine learning. Vulnerability assessment by performance evaluation. Taxonomy of possible defense strategies.
4. Hands-on classes on attacks and defences of machine-learning algorithms using the SecML open-source Python library for the security evaluation of machine learning algorithms (https://secml.readthedocs.io/en/v0.15/).
5. Summary and outlook. Current state of this research field and future perspectives

References:

- B., Battista, F. Roli. "Wild patterns: Ten years after the rise of adversarial machine learning." Pattern Recognition 84 (2018): 317-331.
- B. Biggio, F.Roli, Wild Patterns, Half-day Tutorial on Adversarial Machine Learning: https://www.pluribus-one.it/research/sec-ml/wild-patterns
- Biggio, B., Corona, I., Maiorca, D., Nelson, B., Srndic, N., Laskov, P., Giacinto, G., Roli, F. Evasion attacks against machine learning at test time. ECML-PKDD, 2013.
- Biggio, B., Fumera, G., Roli, F. Security evaluation of pattern classifiers under attack. IEEE Trans. Knowl. Data Eng., 26 (4):984–996, 2014.

Title: Machine Learning Crash Course (MLCC) 2024

Teachers: Lorenzo Rosasco, DIBRIS (lorenzo.rosasco@unige.it), Silvia Villa, DIMA (silvia.villa@unige.it), Giovanni Alberti, DIMA (giovanni.alberti@unige.it), Simone Di Marino, DIMA (simone.dimarino@unige.it), Matteo Santacesaria, DIMA (matteo.santacesaria@unige.it)

Duration: 20 hours

Credits: 6 CFU

When: from Tuesday the 25th to Friday the 28th of June 2024

Where: DIBRIS, Via Dodecaneso 35

CV: CSRAI

Link: TBA

Exam:

The test will consist in completing remotely the notebooks that the class will work on during the labs, and a writing a report commenting on the numerical results obtained. The school will take place exclusively in person, it will not be streamed online. Active attendance will be part of the evaluation.

Abstract:
Machine Learning is key to develop intelligent systems and analyze data in science and engineering. Machine Learning engines enable intelligent technologies such as Siri, Kinect or Google self-driving car, to name a few. At the same time, Machine Learning methods help deciphering the information in our DNA and make sense of the flood of information gathered on the web, forming the basis of a new “Science of Data”. This course introduces the fundamental methods at the core of modern Machine Learning. It covers theoretical foundations as well as essential algorithms. Classes on theoretical and algorithmic aspects are complemented by practical lab sessions.

Program:
Tue - 9.30-11.00 - Class 1: - Introduction to Statistical Machine Learning
Tue - 11.30-13.00 - Class 2: - Local Methods and Model Selection

Tue - 14.30-16.30 - Lab 1: - Local Methods for Classification
Wed - 9.30-11.00 - Class 3: - Empirical Risk Minimization with Linear Models
Wed - 11.30-13.00 - Class 4: - Optimization and SGD
Wed - 14.30-16.30 - Lab 2: - ERM with Linear Models
Thu - 9.30-11.00 - Class 5: - Kernel Methods
Thu - 11.30-13.00 - Class 6: - Neural Networks
Thu - 14.30-16.30 - Lab 3: - Kernel Methods and Neural Networks
Fri - 9.30-11.00 - Class 7: Sparsity and variable selection
Fri - 11.30-13.00 - Class 8: Dimensionality Reduction and PCA
Fri - 14.30-16.30 - Lab 4: - Sparsity and PCA

References:
TBA