PhD in Information Technology and Electrical Engineering

Università degli Studi di Napoli Federico II

PhD Student: Francesco Caturano

XXXIV Cycle

Training and Research Activities Report – First Year

Tutor: Simon Pietro Romano

Information

My name is Francesco Caturano. I graduated in Computer Science Engineering in 2018 at University Federico II of Napoli, with a Thesis entitled “Automated discovery of CoAP-enabled IoT devices”. The Thesis activity was carried out at NomadicLab, Ericsson Finland.
I am a first year ITEE PhD student at University Federico II of Napoli, belonging to the XXXIV cycle.
I conduct my research activity using external fundings coming from a GARR (Gruppo Armonizzazione Reti della Ricerca) scholarship, as my PhD fellowship type does not provide any departmental funds.
My tutor is Professor Simon Pietro Romano.

Study and Training activities

In the first year of PhD I attended several courses and seminars which cover the three areas of interest that are shaping my research activity: Cyber-Security, Computer Networks and Artificial Intelligence.
The very first module I attended, called “cyber-conflicts”, tackled various well-known issues about modern attacks, both from a technical and a jurisdictional point of view.
The second module was called “Data Science and Optimization”. It concerned the mathematical fundaments upon which modern Machine Learning and Deep Learning techniques are based, showing how they can be considered as powerful tools to solve classical Optimization problems.
Then, I attended a short course which presented state of the art technologies in the field of Big Data, which become handy when it comes to manage huge amounts of data both in the Computer Networks and Artificial Intelligence areas.
A very interesting course which turned out to be fundamental for my research activity was the one called “Advanced techniques for software robustness and security testing”. It addressed several best-practices to test robustness of complex software systems, including fuzzing, model checking and symbolic execution.
Then I attended “Programming II”, a Master Thesis course which I hadn’t had the opportunity to include in my preceding course of studies. In order to acquire the credits, I did one of the lectures, talking about REST, the architectural design pattern of the Web, which was also one of the main topics discussed in my master Thesis.
In the same period I became familiar with censorship on the Internet, thanks to a PhD ad-hoc module. The interesting fact was getting to know that the techniques used to avoid censorship are the same used by cyber-attackers to ensure anonymity.
In the meantime I attended seminars explaining fundaments of Neural Networks and how to apply Machine Learning inside a computer network.

Next, I enrolled the “Lipari School on Network and Computer Sciences”, a PhD school which focused on Machine Learning in the field of Computer Networking. There I became familiar with the state of the art of the modern Artificial Intelligence technologies and their application to several areas which are covered by my research activity, such as monitoring networks and securing communications. On that occasion I also became familiar with python-based frameworks that help solving aforementioned issues.

In this direction, the course “Deep Learning for Image Processing” improved my knowledge about frameworks to solve dynamic programming problems and also introduced me to a cutting-edge technology in the field of Generative Models, called Generative Adversarial Networks, which I have continued to study from several other sources (Books: “Unsupervised Learning with Python”, “Generative Adversarial Networks with Python”).

In the last two months I attended another PhD School, named "Machine Learning and Security", mainly focused on techniques to exploit Machine Learning algorithms.

I also attended a seminar named "On Reinforcement Learning for computing channel capacity with feedback", as I'm diving into Reinforcement Learning techniques for testing security vulnerabilities in Web Applications, even though the application discussed in the seminar was not necessarily related to my working area.

Research activity

My research activity revolves around two main topics:

- Security Automation and Virtualization
- Artificial Intelligence for Security testing

The first area focuses on the study of cutting-edge virtualization technologies that allow for the design of emulated attack scenarios. They help building complex virtualized environments, commonly known as "cyber ranges", used for training purposes of security professionals.
In fact, my GARR fellowship is based upon the development of the Docker Security Playground, a framework that allows for the implementation of attack scenarios on virtualized network infrastructures. It is organized as a set of publicly available interactive laboratories, each one focused on a specific security issue to tackle. The users can both use the application for creating their own laboratories, as well as play those which are already designed by others and available online.
My contribute to the project goes in a minor part to the development of the architecture and in a much major part to the design of laboratories which are used as support training material for the students of the course of Network Security. They use the application in combination with my laboratories in order to tackle, from a more practical point of view, each topic discussed in the course. The ability to design ad-hoc laboratories comes handy also for research purposes, whenever it is necessary to develop distributed testbeds to be used for specific experimentations.
As a result of this activity, the paper "The role of microservices in security playgrounds", is ready for submission. We show how to enable integration of different virtualization techniques using microservices enabling technologies such as Docker, when developing cyber ranges.

The second area of research focuses on the development of innovative methodologies to test software robustness and security, with a black-box approach. The target, in the beginning of the activity, are Web Application vulnerabilities. The idea is to provide new vulnerability models, based on the behaviour that the attacker assumes while trying to detect them. So far, modelling Web Application vulnerabilities has always dealt with tools like static or dynamic analysis of the code, which is an assumption that is almost never satisfied: the code is not necessarily available. Another approach tries to structure the best practices in order to prevent/mitigate the attacks. This is not enough to provide security for web application, but we need new and easy-to-use tools that help developers designing robust web application, without any deep knowledge of the domain.
For an attacker who does not have insights of the architecture of the target system, the outcome of the decisions made is only partly under his own control. He can only rely on inference depending on the results of the inputs that he supplies. Such "decision making" issues are well described by mathematical frameworks known as "Markov Decision Processes". They become useful when it comes to study optimization problems solved with dynamic programming or reinforcement learning.
The final output of this research activity must be an intelligent agent, which performs actions and learns, based on the outcome of these actions (the "reward"), how to move forward towards the detection of the vulnerability, in an environment that provides feedback for its actions ("observations"). This is a well-known Reinforcement Learning problem, which perfectly matches the "Trial&Error" behaviour carried out by an attacker.

Products

The research activity conducted so far did not produced accepted publications yet, because it's towards the end of the first year that my work started acquiring a specific shape.

There are several papers in preparation though:

- "The role of microservices in security playgrounds", concerning the ways to aggregate different virtualization technologies using docker containers as a glue;
- "Collaborative platform for penetration testing", about a platform that leverages artificial intelligence techniques to automate security testing on Web Applications
- "Modeling XSS vulnerabilities through Reinforcement Learning", about Reinforcement Learning techniques to help producing unit tests for Web Applications vulnerabilities such as Cross Site Scripting.

One accepted conference publication:

- "Automated Discovery of CoAP-enabled IoT devices", presented at the 11th International Conference on Ubiquitous and Future Networks (ICUFN 2019). Evolution of my Master Thesis work. Not strictly related to security, but natural bridge to the world of virtualization and Web Application.

Conferences and Seminars

- Conference name: IEEE 11th International Conference on Ubiquitous and Future Networks (ICUFN 2019)
- Place: Zagreb
- Dates: 2 – 5 July 2019
- Accepted Paper: ""Automated Discovery of CoAP-enabled IoT devices"
- Presented on the 3rd day of conference

- Conference name: Net Makers (Workshop GARR)
- Place: Rome
- Dates: 08-10 October 2019
- Invited speaker
- "Docker Security Playground" presented on the 2nd day

Student: Francesco Caturano **Tutor: Simon Pietro Romano** **Cycle XXXIV**

francesco.caturano@unina.it spromano@unina.it

	Credits year 1								Credits year 2								Credits year 3								
	Estimated	bimonth 1	bimonth 2	bimonth 3	bimonth 4	bimonth 5	bimonth 6	Summary	Estimated	bimonth 1	bimonth 2	bimonth 3	bimonth 4	bimonth 5	bimonth 6	Summary	Estimated	bimonth 1	bimonth 2	bimonth 3	bimonth 4	bimonth 5	bimonth 6	Summary	Total
Modules	18	0,8	1,2	6	8	3	5	**24**	9							**0**								**0**	**24**
Seminars	13	0	0	0	0,8	6	0,2	**7**	6							**0**								**0**	**7**
Research	34	9,2	8,8	4	1,2	1	4,8	**29**	42							**0**								**0**	**29**
	65	10	10	10	10	10	10	**60**	57	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	**0**	**60**

Check	Year	Lecture/Activity	Type	Credits	Certification	Notes
30-70	1	Cyber Conflicts	Ad hoc module	0,8	x	
10-30	1	Data Science and Optimization	Ad hoc module	1,2	x	
80-140	1	Big Data	Ad hoc module	3	x	
180	1	Advanced techniques for software robustness and security testing	Ad hoc module	3	x	
	1	Programmazione II	Master Course	6	x	
	1	Internet Censorship	Ad hoc module	2	x	
	1	Deep Learning for Image Processing	Ad hoc module	3	x	
	1	Machine Learning and Security	Doctoral School	5	x	
	1	Lipari School on Network and Computer Sciences	Doctoral School	6	x	
	1	Neural Networks	Seminar	0,4	x	
	1	In-Network Machine Learning for Networks	Seminar	0,4	x	
	1	On Reinforcement Learning for computing channel capacity with feedback	Seminar	0,2	x	