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 – Second 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 second year ITEE PhD student at University Federico II of Napoli, belonging to the XXXIV cycle.

My PhD fellowship type does not provide departmental funds. I conducted the research activity using external fundings coming from a GARR (Gruppo Armonizzazione Reti della Ricerca) scholarship, for the whole first year, as well as for the major part of the second year. Currently, I am holder of a DIETI fellowship, which focuses on the topic “Automating penetration testing procedures through Reinforcement Learning”.

My tutor is Professor Simon Pietro Romano.

Study and Training activities

To complete the ad-hoc module credit requirements, the second year I attended two courses:

- Machine Learning for Health;
- Virtualization Technologies and their applications.

They both fit well with my research activity. In fact, I am dealing with the application of Reinforcement Learning techniques to automate the procedures of Web Application Penetration Testing, as well as using lightweight virtualization technologies to emulate real network attack scenarios. More details in the next section.

I did not attend seminars during the second year. I plan to acquire the remaining seminar credits during the third year.

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.

During the second year of my GARR fellowship, I continued 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.

As a result of this activity, the paper “Capturing flags in a dynamically deployed microservices-based heterogeneus environment” has been published and presented at the conference named IPTcomm 2020 (Principles, Systems and Applications of IP Telecommunications). In the paper, we tackle, from a practical point of view, some of the limitations that arise from the employment of OS virtualization as well as propose a solution that enables integration of different virtualization technologies in order to solve some of the discussed issues.

The second area of research focuses on the development of intelligent methodologies to test Web Applications security, using a black-box approach. In fact, the majority of the second year was spent developing a model that frames the behavior of a hacker as a sequential decision making problem, under uncertainty. The problem is seen as a Markov Decision Process and solved using Reinforcement Learning The developed approach has been tested on a particular vulnerability exploit model, known as Cross-Site Scripting.

To the purpose, a novel Multiobjective Reinforcement Learning environment was developed. The term “Multiobjective” signifies that there are several optimization targets for the reinforcement learning agent to pursue.

In order to provide users with a practical tool, useful during actual penetration testing campaign, the first result has been a Suggester, a semi-automated platform that recommends the best actions to a human in the loop. Such system has shown very good results in terms of accuracy, compared to state of the art XSS detection tools. The work is discussed in the paper “Discovering Cross-Site Scripting vulnerabilities using a Multiobjective Reinforcement Learning environment”, currently under review by Computers & Security.

Recent developments, produced a fully automated platform, providing a completely autonomous agent that can discover Cross-Site Scripting flaws in a web application.

Also, the scalability of the proposed approach is currently under evaluation for other Web Application Injection vulnerabilities, such as SQL Injection.

At the same time, the platform known as WAVSEP (Web Application Vulnerability Scanner Evaluation Project), which has ben used to train the Reinforcement Learning agent, is being modified adding new test cases to support the constantly changing security scenario.

Web Application Penetration Testing is studied not only from an exploit model point of view, but also as a complex activity, that encompasses several stages. The application of intelligent models is already under consideration, using the framework of Knowledge Graphs, that allows to create knowledge bases for decision support systems/expert systems. The design of this approach is described in the position paper "Hacking Goals: a goal-centric attack classification framework", soon to be presented at ICTSS conference.

Moreover, a completely automated platform that allows integration of external tools to perform the separate phases of a penetration test has been designed and implemented. The result of this work is described in a paper currently under preparation, entitled "An automated approach to offensive security".

Products

Accepted Conference publications:

- "Capturing flags in a dynamically deployed microservices-based heterogeneous enivironment", presented at IPTcomm 2020 (Principles, Systems and Applications of IP Telecommunications).
- "Hacking Goals: a goal-centric attack classification framework" (short paper), to be presented (December 10th 2020) at ICTSS 32ND IFIP International conference on testing software and systems.

Currently under review:

- "Discovering Cross-Site Scripting vulnerabilities using a Multiobjective Reinforcement Learning environment", Computers & Security (Elsevier).
- "A Distributed Security Tomography Framework to assess the exposure of ICT Infrastructures to Network Threats", Journal of Information Security and Applications (Elsevier).

Currently under preparation:

- "An automated approach to offensive security", which presents a completely automated platform that allows integration of external tools to perform the separate phases of a penetration test.

Conferences and Seminars

- Conference name: IPTcomm 2020 (Principles, Systems and Applications of IP Telecommunications);
- Place: Virtual (originally Chicago)

- Dates: 13 – 15 October 2020
- Accepted paper: “Capturing flags in a dynamically deployed microservices-based heterogeneous enivironment”

Student: Francesco Caturano francesco.caturano@unina.it

Tutor: Simon Pietro Romano spromano@unina.it

Cycle XXXIV

	Credits year 1								Credits year 2								Credits year 3								
		1	2	3	4	5	6			1	2	3	4	5	6			1	2	3	4	5	6		
	Estimated	bimonth	bimonth	bimonth	bimonth	bimonth	bimonth	Summary	Estimated	bimonth	bimonth	bimonth	bimonth	bimonth	bimonth	Summary	Estimated	bimonth	bimonth	bimonth	bimonth	bimonth	bimonth	Summary	Total
Modules	18	0,8	1,2	6	8	3	5	**24**	15	0	0	0	4	4	0	**8**	21							**0**	**32**
Seminars	13	0	0	0	0,8	6	0,2	**7**	6	0	0	0	0	0	0	**0**	12							**0**	**7**
Research	34	9,2	8,8	4	1,2	1	4,8	**29**	39	10	10	10	6	6	10	**52**	30							**0**	**81**
	65	10	10	10	10	10	10	**60**	60	10	10	10	10	10	10	**60**	63	0	0	0	0	0	0	**0**	**120**

Check
30-70
10-30
80-140
180

Year	Lecture/Activity	Type	Credits	Certification	Notes
2	Virtualization Technologies and their applications	Ad-hoc Module	4	x	
2	Machine Learning for Health	Ad-hoc Module	4	x	