PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

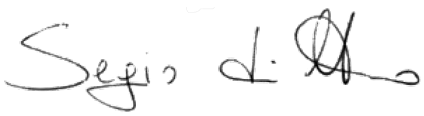
PhD Student: Luigi Libero Lucio Starace

Cycle: XXXV

Training and Research Activities Report

Year: First

Tutor: Prof. Sergio Di Martino



Co-Tutor: Prof. Adriano Peron

Date: October 21, 2020

Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXV **Author: Luigi Libero Lucio Starace**

1. Information:

- **PhD student:** Luigi Libero Lucio Starace
- **DR number:** DR993893
- **Date of birth:** 25/02/1991
- **Master Science degree:** Computer Science **University:** UNINA
- **Doctoral Cycle:** XXXV
- **Scholarship type:** *Funded by NetCom Group S.p.A.*
- **Tutor:** Prof. Sergio Di Martino
- **Co-tutor:** Prof. Adriano Peron

2. Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
A dynamic and probabilistic orienteering problem	Seminar	1	0.2	08/11/19	Prof. Claudio Sterle	Y
Flexible two-echelon location-routing for supply networks	Seminar	1	0.2	08/11/19	Prof. Claudio Sterle	Y
Lo spazio cibernetico come dominio bellico	Seminar	2	0.4	15/11/19	Prof. Gugielmo Tamburrini	Y
Accelerated computing with CUDA C/C++	Course	2	0.4	25/11/19	DIETI	Y
Marked Point Processes For Object Detection And Tracking In High Resolution Images: Application To Remote Sensing Data	Seminar	1	0.2	02/12/19	Prof. Giuseppe Scarpa	Y
Intelligenza artificiale ed etica: la ricerca in IA alla prova delle sfide etiche	Course	8	1.6	06/12/19	DIETI	Y
Safety Critical Systems for Railway Traffic Management	Course	18	3.3	10/01/20 to 27/01/20	DIETI	Y
Cybersecurity and Fuzzing for Robots, Blockchain, and more	Seminar	1	0.2	13/01/20	Dr. Roberto Natella	Y
Matlab Fundamentals	Course	20	2	20/02/20 to 27/03/20	DIETI / SPSB	Y

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Computational Biology: Large scale data analysis to understand the molecular bases of human diseases	Seminar	1	0.2	09/04/20	DIETI	Y
Elettromagnetismo e salute	Seminar	1	0.2	09/04/20	DIETI	N
How to get published with IEEE	Seminar	2	0.4	20/04/20	Dr. Alessandra Scippa	Y
Virtualization technologies and their applications	Course	20	4	06/04/20 to 15/05/20	Prof. D. Cotroneo, DIETI	Y
Innovation management, entrepreneurship, and intellectual property	Course	18	5	05/05/20 to 05/06/20	Prof. P. Rippa - StartCup Campania 2020	Y
Design and Implementation of Augmented Reality Software Systems	Course	20	4	03/06/20 to 23/06/20	ITEE - ICTH	Y
Metodi Formali	M.Sc. Course	24	3	12/03/20 to 11/06/20	Prof. Valeria Vittorini	Y
Design e Nuove tecnologie. Possibili scenari per fronteggiare l'emergenza	Seminar	1	0.2	11/05/20	Innovation Village 2020	Y
La programmazione europea e la ricerca. Nuovi scenari della programmazione europea dopo il 2020 - La gestione di un progetto di ricerca	Seminar	2	0.4	13/05/20	Innovation Village 2020	N
SAS Analytics	Seminar	2	0.4	14/05/20	SAS Academic Program Manager	N

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Planning 5G under EMF constraints: challenges and opportunities	Seminar	2	0.4	18/05/20	Prof. L. Chiaraviglio, Univ. of Rome Tor Vergata - Dr.ssa A. Cacciapuoti, Dr. M. Caleffi - DIETI	N
Joint Design of Optics and Post-Processing Algorithms Based on Deep Learning for Generating Advanced Imaging Features	Seminar	2	0.4	19/05/20	IEEE Computational Imaging Technical Committee	N
Virtual Seminars on 'Sensing'	Seminar	4	0.8	20/05/20	Plasmonica, Prof. Carlo Forestiere, DIETI	Y
Bias from the wild	Seminar	2	0.4	26/05/20	CVPL - Associazione Italiana per la ricerca in Computer Vision, Pattern recognition e machine Learning	N
Software Safety for Aerospace Applications	Seminar	1.5	0.3	27/05/20	On-line course offered on the IEEE Xplore platform	Y
Amazon EC2 and S3 Hands-on	Seminar	1.5	0.3	02/06/20	On-line course offered on the IEEE Xplore platform	Y
Large Scale Training of Deep Neural Networks	Seminar	2	0.4	06/05/20	DIETI	N

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Noninvasive Mapping of Electrical Properties using MRI	Seminar	1.5	0.3	11/06/20	Prof. R. Massa Dip. Fisica UNINA, Prof. G. Ruello Dieti UNINA	N
Exploring Autonomy in Robotic Flexible Endoscopy	Seminar	2	0.4	12/06/20	Prof. Fanny Ficuciello	Y
"Linear regression in PyTorch" and "Convolutional Neural Networks"	Seminar	2	0.4	29/06/20	CINI AIIS Labs.	N
"Efficient Data Loading with DALI" and "Mixed Precision Training using Apex"	Seminar	1	0.2	01/07/20	CINI AIIS Labs.	N
Machine Learning	Course	20	4	06/07/20 to 17/07/20	ITEE - ICTH	Y
Wearable Brain-Computer Interface for Augmented Reality-based Inspection in Industry 4.0	Seminar	1	0.2	29/07/20	Prof. P. Arpaia, DIETI	Y
Strategic Orientation for STEM Research & Writing	Course	18	3.6	16/07/20 to 17/09/20	ITEE - ICTH	Y
Algorithmic Accountability – Affidabilità e responsabilità degli algoritmi	Seminar	2	0.4	24/09/20	Fondazione Ugo Bordoni	N

1) Courses, Seminar, Doctoral School, Research, Tutorship
2) Choose: Y or N

2.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	1.8	1	5	0	7.8
Bimonth 2	3.3	0.2	6	0	9.5
Bimonth 3	2	0.8	7	0	9.8
Bimonth 4	16	5.1	5	1.6	27.7
Bimonth 5	4	0.4	4	0	8.4
Bimonth 6	3.6	0.4	5	0	9
Total	30.7	7.9	32	1.6	72.2
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

3. Research activity:

The main research topic I focused on during the first year is software testing. In particular, following the business interests of *NetCom Group S.p.A*, which funded my Ph.D. fellowship, I studied techniques and methodologies for web application testing. After an extensive preliminary study of the state-of-the-art, I started working on web application model inference, i.e., the task of automatically abstracting a model for a given web application in a black-box fashion. Currently, I am working on a crawler-based model inference approach which also uses tree kernel functions for detecting near-duplicate web pages. The intuition behind this approach is to firstly generate a web-page reachability graph for the web application using a crawler, and then to cluster similar (near-duplicate) web pages representing the same functionality into a single macro-state. I plan to use *tree kernels*, a class of kernel functions designed to evaluate similarity between tree objects, to assess similarity between the DOM of web pages during the clustering phase. As for the evaluation of the proposed approach, I plan to compare it against other state-of-the-art near-duplicate detection algorithms such as the ones recently evaluated by Yandrapally et Al. in [1]. After dealing with model inference, further research could focus on test case or test artifacts (e.g.: page objects) generation using the inferred model.

Within the main topic of software testing, I also worked on regression test prioritization and GUI testing of Android applications. For what concerns regression test prioritization, I worked on defining prioritization strategies leveraging code churn information, i.e., information on the changes in source code between two subsequent versions of a software. This work resulted in the conference paper "I*nspecting code churns to prioritize test cases*", accepted at the 32nd IFIP International Conference on Testing Software and Systems (ICTSS 2020). Moreover, I am collaborating with my research group on designing and evaluating more advanced churn-based regression test prioritization strategies that not only consider changes in source code between subsequent versions, but also take into account the nature of said changes. Indeed, not all code changes are characterized by the same likelyhood of introducing regression faults. For example, it is reasonable to assume that the simple renaming of a local variable in a method is less likely to introduce errors than a change in the condition of an iteration construct. The intuition behind the approach we are currently exploring is to prioritize test cases covering parts of code affected by more fault-prone changes.

As for GUI testing of Android applications, I worked on an empirical study comparing state-of-the-art fully-automated Android GUI testing solutions with a Capture and Replay tool used by twenty novice practitioners. In particular, I worked on running the fully automated tools and on analysing the coverage results. This study highlighted that novice human practitioners outperformed the automatic tools, and also

provided useful insight on the limitations of fully-automated GUI testing approaches. This work resulted in the journal paper "*Comparing the effectiveness of capture and replay against automatic input generation for Android graphical user interface testing*", published on Software Testing Verification and Reliability (STVR). Currently, I am further analysing the results of this empirical study to gain additional insight on the influence of the number of human testers on the achieved coverage, with the goal of finding the best trade-off between the costs of human testers and the achieved coverage.

I also participated in a study on the optimal (i.e., cheapest) allocation of software components on cloud and local resources for industrial IoT systems, taking into account not only computation-related constraints but also information security ones. In this study, I worked on formalizing the allocation problem, and on devising heuristic algorithms to approximate a solution. In particular, I implemented a genetic algorithm, a linear programming solution and a greedy algorithm to solve the problem, generated a large set of realistic problem instances, and used them to benchmark the proposed solutions. This work resulted in the journal paper "*Security-aware deployment optimization of cloud-edge systems in industrial IoT*", published in IEEE Internet of Things Journal.

Moreover, I partook on a study on a novel randomized routing algorithm to increase spatio-temporal road-network coverage in vehicle crowd-sensing systems. In this study, the spatio-temporal coverage achieved by the proposed algorithm is compared against that achieved by a standard implementation of the A* algorithm, using real-world taxi trajectories from the cities of San Francisco (USA) and Porto (PT). In particular, I worked on dataset preparation and on the analysis of the spatio-temporal coverage results. This work resulted in the journal paper "*Vehicular crowd-sensing: a parametric routing algorithm to increase spatio-temporal road network coverage*", which is currently under the second round of review for acceptance in the International Journal of Geographical Information Science (IJGIS).

Lastly, I am also continuing the study started with my M.Sc. Thesis work on model checking for hierarchical systems. In detail, I'm working with Prof. Adriano Peron, Prof. Massimo Benerecetti, and Prof. Fabio Mogavero on proving the correctness of the model checking algorithm we sketched during my thesis work. With the formal proofs in place, we plan to prepare a journal paper and submit it to the "Logical Methods in Computer Science" (LMCS) journal.

3.1. References

[1] Yandrapally, Rahulkrishna, Andrea Stocco, and Ali Mesbah. "Near-duplicate detection in web app model inference." Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering. 2020.

4. Research products:

4.1. Journal papers

Casola, V., De Benedictis, A., Di Martino, S., Mazzocca, N., & **Starace, L. L. L.** (2020). Security-aware deployment optimization of cloud-edge systems in industrial IOT. *IEEE Internet of Things Journal*, 1–1. DOI: https://doi.org/10.1109/JIOT.2020.3004732. Status: published.

Di Martino, S., Fasolino, A. R., **Starace, L. L. L.**, & Tramontana, P. (2020). Comparing the effectiveness of capture and replay against automatic input generation for Android graphical user interface testing. *Software Testing Verification and Reliability* (STVR). DOI: https://doi.org/10.1002/stvr.1754. Status: published.

Asprone, D., Di Martino, S., Festa, P., **Starace, L. L. L.** (2020). Vehicular crowd-sensing: a parametric routing algorithm to increase spatio-temporal road network coverage. *International Journal of Geographical Information Science* (IJGIS). Status: under 2nd round of review.

4.2. Conference papers

Altiero, F., Corazza, A., Di Martino, S., Peron, A., **Starace, L. L. L.** (2020). Inspecting code churns to prioritize test cases. 32nd IFIP International Conference on Testing Software and Systems (ICTSS 2020). Status: accepted.

5. Conferences and seminars attended

Attended virtual sessions for the *Empirical Software Engineering and Measurement* (ESEM 2020) conference, from 05/10/2020 to 07/10/2020.

6. Activity abroad:

None.

7. Tutorship

Four two-hours lectures within the "*Ingegneria del Software II*" M.Sc. course held by Prof. Sergio Di Martino. The lectures had the following topics: "*Unit Testing with Junit 5, Hamcrest and Mockito*", "*Test-driven development*", "*Practical session on TDD*", "*Formal Methods for Software Engineering*".