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

PhD Student: Marco De Luca

Cycle: XXXVII - 37°

Training and Research Activities Report

Year: First

Tutor: prof. Anna Rita Fasolino _tutor signature_

Co-Tutor: Pasquale Cimmino (Micron Semiconductor Italia)

Date: October 31, 2022

Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII   **Author: Marco De Luca**
_____

1. Information:

- **PhD student: Marco De Luca**
- **DR number: DR995860**
- **Date of birth: 15/12/1994**
- **Master Science degree: Computer Engineering   University: Università degli Studi di Napoli Federico II**
- **Doctoral Cycle: XXXVII**
- **Scholarship type:** *Funded by Micron Semiconductor Italia S.R.L.*
- **Tutor: Prof. Anna Rita Fasolino**
- **Co-tutor: Pasquale Cimmino (***Micron Semiconductor Italia S.R.L.***)**

2. Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Software Testing	**Courses**	**48**	**6**	**27/02/2021**	**Prof. Porfirio Tramontana**	**Y**
Possible Quantum Machine Learning Approaches in HEP	**Seminar**	**2**	**0.4**	**12/11/2021**	**Dr. Michele Grossi**	**Y**
Connecting the dots: Investigating an APT campaign using Splunk	**Seminar**	**2**	**0.4**	**26/11/2021**	**Dr. Antonio Forzieri**	**Y**
Single cell omics leverage Machine Learning to dissect tumor microenvironment and cancer immuno editing	**Seminar**	**2**	**0.4**	**02/12/2021**	**Dr. Raoul J. P. Bonnal**	**Y**
Threat Hunting Use-Cases	**Seminar**	**2**	**0.4**	**13/12/2021**	**Vladimir Kurdin**	**Y**
GDPR basics for computer scientists	**Seminar**	**2**	**0.4**	**14/12/2021**	**Dr. Rigo Wenning**	**Y**
All roads lead to WebRTC: an introduction to Janus,	**Seminar**	**2**	**0.4**	**16/12/2021**	**Dr. Lorenzo Miniero**	**Y**
Designing Quantum Algorithms	**Seminar**	**2**	**0.4**	**16/12/2021**	**Prof. Michele Amoretti**	**Y**
Computational analysis of cancer genomes	**Seminar**	**1**	**0.2**	**16/02/2022**	**Nùria Lòpez-Bigas**	**Y**

Project Vāc: Can a Text-to-Speech Engine Generate Human Sentiments?	**Seminar**	**1**	**0.2**	**28/02/2022**	**Prof. V.K. Gubani**	
Explainable Natural Language Inference	**Seminar**	**1.5**	**0.3**	**13/04/2022**	**Dr. Marco Valentino**	**Y**
An Introduction to Deep Learning for Natural Language Processing	**Seminar**	**1**	**0.2**	**13/04/2022**	**Dr. Marco Valentino**	**Y**
Probing and infusing biomedical knowledge for pre-trained language models	**Seminar**	**1.5**	**0.3**	**07/06/2022**	**Dr. Zaiqiao Meng**	**Y**
Machine Learning e Big Data per la Salute	**Course**	**72**	**9**	**08/03/2021 - 11/06/2021**	**Prof. Vincenzo Moscato**	**Y**
Big Data Architecture and Analytics	**Course**	**16**	**5**	**06-08-22-27-29/04 - 06-11/05/2022**	**Prof. Giancarlo Sperli**	**Y**
Imprenditorialità Accademica,	**Course**	**14**	**4**	**26/05 - 14/06 2022**	**Prof. Pierluigi Rippa**	**Y**
Natural Language Processing,	**Course**	**48**	**6**	**08/03/2021 - 11/06/2021**	**Prof. Francesco Cutugno**	**Y**
Teaching activities regarding practical lectures/seminars during the courses of "Ingegneria del Software"	**Tutorship**	**40**	**0.48**	**12-26/04/2022 04-13/05/2022 01/06/2022**	**Prof. Anna Rita Fasolino**	**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	6	2.8	5	-	13.8
Bimonth 2	-	0.4	6	-	6.4
Bimonth 3	-	0.5	6	0.24	6.74
Bimonth 4	14	0.3	6	0.24	20.54
Bimonth 5	11	-	5	-	16
Bimonth 6	-	-	6	-	6
Total	31	4	34	0.48	69.48
Expected	30 - 70	10 - 30	80 - 140	0 – 4.8	

3. Research activity:

During my first year of PhD I focused my study on the research field of the *Software Engineering*. I carried out three main research, namely *Software Repository Mining based on Network Representation Learning*, *Software Testing* and *Safety Critical System in the automotive domain*. A new field of research emerging in the literature is the application of artificial intelligence (AI) techniques in the field of software engineering. The AI techniques can support the entire software development process: formation of the team, requirements specification phase, testing phase and so on. *Software Repository Mining* (SRM) aims to extract knowledge from software repositories, such as version control system, to support and improve software development. SRM techniques can be used to perform various types of tasks, such as community detection or expert finding. Being able to identify communities of developers within a software repository, or identifying developers with well-defined characteristics (expert finding), can help in the project team formation phase.

Continuing in the field of software engineering, my studies have also focused in the area of software testing. I expanded my knowledge in this area by attending the Software Testing course and with practical research. Research activities focused on *model based testing techniques* for augmented reality application and fragility of test cases in E2E.

The enhancement of my knowledge in the latter two fields of research, are then applied in software development processes in the industrial environment for safety critical systems. My third research field, focuses on the challenges and issues that need to be addressed in the software development and documentation process in the industrial environment to achieve compliance with safety standards.

- **Software Repository Mining:** My first research activity is about a proposal of a novel *heterogeneous graph-based model* to capture and handle, all the useful, complex, and strongly-correlated information found inside a *Developers Social Network* (DSN). Many systems of the real world can be modeled in terms of entity and their interaction and represented as a network: i) world-wide web, ii) the social media iii) the citation system iv) images and text. Understand the user relationship and interaction inside a DSN can be useful for a lot of analytic task: i) node classification, ii) link prediction iii) community detection. My work focused on a network that can be derived from *software repository mining system* like GitHub. To overcome huge graph-size issue, I studied and applied different graph embedding techniques, like R-GCN, ComplEx and HIN2Vec. To validate the proposed approach I carried out experiment in order to study the

effectiveness and the performance of the proposed community detection techniques by considering different instance of DSN.

- **Software Testing**: The first research activity focused on *model based testing techniques* for augmented reality application. The popularity of Augmented Reality (AR) applications has strongly been increased with the worldwide success of the Pokémon Go videogame released by Niantic in 2016. However, AR offers benefits in many further areas beyond entertainment, such as advertisement, education, navigation, maintenance, health, and so on. With the growing spread and success of AR applications in these fields, there has also been a necessity for approaches and technologies for assuring the quality of these applications, such as testing. A few technologies and frameworks have been recently proposed supporting the implementation and execution of test scripts that can be used to exercise the applications, but there still is a lack of effective techniques and tools for the automatic generation of executable test cases. In my research activities I investigate the possibility of using Model Based Testing techniques to generate executable test scripts from Finite State Machines modeling the behavior of the GUI of AR applications. I have applied several model coverage criteria to design test suites and I have shown the feasibility of this approach by testing two small example applications involving Unity3D and Vuforia technologies.

  My second ongoing research concerns the *fragility of test cases* in E2E scenarios while exercising the application through its GUI. Capture&Replay techniques are a well-known solution for E2E testing of Web applications, allowing test scripts to be automatically generated from real sequences of user interactions with the application under test. These techniques suffer from the problem of fragility of the produced test cases, which can break even if minor changes are made to the GUI, without changing the functionality of the application. To overcome this problem, several approaches have been proposed for generating robust test cases or automatically repairing failed test cases. In my ongoing research, I am investigating an alternative solution for generating robust test cases, for web applications developed with model-based technologies. The solution is based on automatically refactoring the template source code to inject identifiers (called "hooks"). These injected hooks will be used to define locations that allow for unambiguous retrieval of the corresponding Web page elements, improving the robustness of E2E test cases and, thus, the testability of the Web application.

- **Safety Critical System in the automotive domain:** Complexity of automotive system has increased in recent years. Nowadays cars are complex system composed by several sensor, actuators, electrical and electronic component, computer resources and so on. The ISO 26262 is a standard that deals with the functional safety of the E/E (Electric and Electronic) component of road vehicles. The standard defines a functional safety development V-process model that automotive manufacturing must follow and document to achieve compliance with the standard, otherwise the produced device want be suitable to run in commercial vehicles. My research activities conducted in collaboration with Micron focused on Chapter 6 of the ISO 26262 "*Product Development at the software level*". My work is investigating new ways and/or improvement of the documentation process conducted at the "product development at software level" for the standard compliance.

Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII   **Author: Marco De Luca**
_____

4. Research products:

- P. Tramontana, **M. De Luca**, A.R. Fasolino; *"An Approach for Model Based Testing of Augmented Reality Applications", QUAMES@RCIS, CEUR Workshop Proceedings, published, 2022*
- **M. De Luca**, *A.R. Fasolino, A. Ferraro, V. Moscato, G. Sperlì, P. Tramontana; "A community detection approach based on Network Representation Learning for repository mining", Expert Systems with Applications, submitted, 2022*
- D. Amalfitano, **M. De Luca**, *A.R. Fasolino; "Documenting Software Architecture Design in Compliance with the ISO 26262: an Industrial Case Study in the Automotive Domain", International Conference On Software Architecture (ICSA), in progress, 2023*

5. Conferences and seminars attended

16th International Conference on Research Challenges in Information Science held in Barcelona, during May 17-20, 2022. I attended this conference as presenting author for the paper "An Approach for Model Based Testing of Augmented Reality Applications".

6. Activity abroad:

-

7. Tutorship

12 hours of teaching activities regarding practical lectures/seminars during the course "Ingegneria del Software", Bachelor Degree in Computer Engineering