UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

DOTTORATO DI RICERCA / PhD PROGRAM IN
INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Activities and Publications Report

PhD Student: Francesco Vitale

Student DR number: DR995865

PhD Cycle: XXXVII
PhD Cycle Chairman: Prof. Stefano Russo

PhD program student's start date: 01/11/2021
PhD program student's end date: 31/10/2024

Supervisor: Prof. Nicola Mazzocca

e-mail: nicola.mazzocca@unina.it

PhD scholarship funding entity: Hitachi Rail STS

General information

Francesco Vitale received in year **2021** the Master Science degree in **Computer Engineering** from the **University of Napoli Federico II**. **He** attended a curriculum in **digital technology for industry 4.0, embedded systems, and federated learning for railway applications** within the PhD program in Information Technology and Electrical Engineering. **He** received a grant from **Hitachi Rail STS**.

Study activities

Attended Courses

Year	Course Title	Type	Credits	Lecturer	Organization
1st	Modeling Complex Systems	Ad hoc course	6	Prof. M. Di Bernardo	Scuola Superiore Meridionale
1st	Real-Time Industrial Systems	MSc course	6	Prof. M. Cinque	University of Naples Federico II
1st	Virtualization technologies and their applications	Ad hoc course	5	Prof. L. De Simone	ITEE
1st	Statistical Data Analysis for Science and Engineering	Ad hoc course	4	Prof. R. Pietrantuono	ITEE
1st	Big Data Analytics and Architectures	Ad hoc course	5	Prof. G. Sperlì	ITEE
2nd	IoT Data Analysis	Ad hoc course	4	Prof. R. Della Corte	ITEE
2nd	Using Deep Learning properly	Ad hoc course	4	Dr. A. Apicella	ITEE
3rd	Strategic Orientation for STEM Research & Writing	Ad hoc course	5	Chie Shin Fraser, B.Sc.	ITEE

Attended PhD Schools

Year	School title	Location	Credits	Dates	Organization

Attended Seminars

Year	Seminar Title	Credits	Lecturer	Lecturer affiliation	Organization
1st	Hitachi Rail Workshop	1.6	R. Nappi, M.Sc. and P. Sannino, M.Sc.	Hitachi Rail STS	Hitachi Rail STS
1st	Complexity And The City: Transitioning Towards The Smart Cities Of The Future	0.2	Prof. L. Bettencourt	University of Chicago	Scuola Superiore Meridionale
1st	Introduction to Intellectual	0.4	A. Marroni		ITEE

	Property Management				
1st	Privacy-Preserving Machine Learning	0.4	V. Prodomo, M.Sc.	University Carlos III of Madrid	ITEE
2nd	Privacy and Data Protection	0.4	S. Mele		ITEE
2nd	Cybercrime and Information Warfare: National and International Actors	0.4	P. Paganini		ITEE
2nd	Conformance Checking & Incremental Process Discovery Using Trace Fragments	0.4	D. Schuster, M.Sc.	RWTH Aachen University	RWTH Aachen University
2nd	Insights by Comparison	0.4	T. Brockhoff, M.Sc.	RWTH Aachen University	RWTH Aachen University
2nd	Process discovery supporting a desirable event log while avoiding an undesirable event log	0.4	A. Norouzifar, M.Sc.	RWTH Aachen University	RWTH Aachen University
2nd	POWL: Partially Ordered Workflow Language	0.4	H. Kourani, M.Sc.	RWTH Aachen University	RWTH Aachen University
2nd	Translucent Process Mining	0.4	H. Beyel, M.Sc.	RWTH Aachen University	RWTH Aachen University
2nd	Object-Centric Simulation	0.4	B. Knopp, M.Sc.	RWTH Aachen University	RWTH Aachen University
2nd	Process Mining and Large Language Model: An Outlook	0.4	A. Berti, M.Sc.	RWTH Aachen University	RWTH Aachen University
3rd	Deep learning for Railway Safety and Maintenance: Methodologies and Applications	0.3	L. De Donato, M.Sc.	University of Naples Federico II	ITEE
3rd	Energy-Efficient Data Science	0.2	Prof. C. Ordonez	University of Houston	ITEE
3rd	The Generative Power of Deep Learning Variational Auto-Encoders and Generative Adversarial Networks for scenario generation	0.5	Prof. V. Messina	University of Milano Bicocca	University of Campania Luigi Vanvitelli
3rd	Edoardo Giusto research past, present and future	0.2	E. Giusto, Ph.D.	University of Naples Federico II	ITEE
3rd	Analytic center selection of optimization-based controllers for robot ecology	0.2	Prof. G. Notomista	University of Waterloo	ITEE

3rd	Exploring the frontiers of modern cryptography	0.3	Prof. A. Michalas	Tampere University	ITEE
3rd	Real-time Resource Management for Adaptive Embedded Systems and Applications	0.2	Prof. Dipl.-Ing. Dr. G. Fohler	Technical University of Kaiserslautern-Landau	ITEE
3rd	Using support vector machines for feature selection and outlier detection	0.2	Prof. M. Baldomero-Naranjo	University of Cadice	ITEE

Research activities

Francesco Vitale carried out research on anomaly detection with process mining and machine learning for industrial cyber-physical systems, dealing with the shortcomings of applying process mining for anomaly detection.

In the first year, Francesco Vitale mainly dealt with pre-processing low-level sensor data to convert them into event data manageable by process mining. He applied his pre-processing solutions to a railways case study from Hitachi, his funding company, and a scale replica of a manufacturing plant in collaboration with University of Messina.

In the second year, under the supervision of Prof. van der Aalst at RWTH Aachen University during his research period abroad, Francesco Vitale studied the opportunity to integrate machine learning with process mining to improve anomaly detection effectiveness. He applied his solution to railways and healthcare case studies.

In the third year, Francesco Vitale developed several research directions stemming from other shortcomings pertaining the application of process mining for behavior characterization of industrial cyber-physical systems. He conducted this research in collaboration with University Campus Bio-Medico of Rome and University of Verona.

Tutoring and supplementary teaching activities

Seminars:
- Title: "Virtualizing CPUs: Intel x86-64 architectures". Course/organization: "Computer Systems Design" (held by Prof. Nicola Mazzocca).
- Title: "Process Mining use cases for Cyber-Physical Systems". Course/organization: University Campus Bio-Medico of Rome.
- Title: "Virtualization for edge computing: A top-down view". Course/organization: "Architettura e progetto dei calcolator" (held by Prof. M. Barbareschi).

Lectures:
- Title: "Progettazione di macchine complesse: Prodotto scalare". Course/organization: "Architetture dei Sistemi Digitali" (held by Proff. N. Mazzocca and A. De Benedictis).

M.Sc. thesis tutorship:

- "Detection of performance anomalies through Process Mining." M.Sc. student: Carmine Marra. Supervisors: Proff. N. Mazzocca and F. Flammini.
- "Definition and development of a hypervisor based microkernel architecture for avionic applications." M.Sc. student: Asia Di Vincenzo. Supervisor: Prof. N. Mazzocca.
- "Anomaly Detection in Industrial Systems: an Approach based on Process Mining and Deep Learning." M.Sc. student: Fabio Martone. Supervisor: Prof. N. Mazzocca.
- "Anomaly Detection in Hypervisor-based Architectures: an approach based on Convolutional Neural Networks." M.Sc. student: Nunzio Tarallo. Supervisor: Prof. N. Mazzocca.
- "AI-based Anomaly Detection in Railway Systems." M.Sc. student: P. Catapano. Supervisor: Prof. N. Mazzocca.

Credits summary

PhD Year	Courses	Seminars	Research	Tutoring / Supplementary Teaching
1st	30.0	2.6	51.0	0.4
2nd	14.0	3.6	44.0	0.4
3rd	5.0	2.1	36.0	0.4
1st+2nd+3rd	49.0	8.3	131.0	1.6

Research periods in institutions abroad and/or in companies

PhD Year	Institution / Company	Hosting tutor	Period	Activities
1st	RWTH Aachen University	Prof. W. van der Aalst	01/02/23 - 31/07/23	Research on anomaly detection by process mining

PhD Thesis

Title: Anomaly Detection by Process Mining and Machine Learning for Industrial Cyber-Physical Systems

Abstract (304 words): The ever-increasing spread and complexity of Industrial Cyber-Physical Systems (ICPSs) allow the seamless connection of the physical and cyber worlds through advanced physical plants, sensor networks and software, improving the quality of service of several industrial applications. Unfortunately, such complexity makes ICPSs vulnerable to faults and troubles the development of accurate models capturing their real behavior. Owing to the dependability issues of ICPSs, the scientific literature proposed many data-driven solutions based on Machine Learning (ML) for characterizing ICPSs and detecting disruptive anomalies. However, ML algorithms are typically process-agnostic and may be affected by explainability issues,

especially those employing deep learning. Besides, the specific ML algorithm should be tailored to the data at hand, the types of anomalies to discover, and the learning paradigm to use.

To provide process-based characterizations of ICPSs while maintaining a high level of explainability, this dissertation investigates PM for anomaly detection due to its ability to discover reference process models from ICPS data and check for deviations from such models. To develop PM-based anomaly detection in ICPSs while accounting for the shortcomings of existing approaches, this thesis puts forward a methodology that guides ICPS development toward implementing PM-based anomaly detection. The methodological steps provide insight into the ICPS application type, requirements, and data. These insights are used to drive the development of a specific PM-based anomaly detection technique through a flexible and explainable framework that integrates ML to enhance data pre-processing and improve the classification accuracy of anomalous behaviors.

During the evaluation, several framework techniques were developed and applied to different industrial case studies, including railways, manufacturing, and healthcare. The results demonstrate that these techniques achieve high levels of explainability and detection effectiveness. These findings underscore the potential for integrating PM techniques in modern industries to enable interpretable, data-driven modeling of ICPS behavior and effectively detect deviations from nominal behavior.

Research products

Research results appear in *6* papers published in international journals. There are other papers either under review or in progress.

List of scientific publications

International journal papers

F. Vitale, S. Guarino, F. Flammini, L. Faramondi, N. Mazzocca, R. Setola, "Process Mining for Digital Twin Development of Industrial Cyber-Physical Systems," IEEE Transactions on Industrial Informatics, 2024, DOI: 10.1109/TII.2024.3465600

S. Guarino, F. Vitale, F. Flammini, L. Faramondi, N. Mazzocca, R. Setola, "A Two-Level Fusion Framework for Cyber-Physical Anomaly Detection," IEEE Transactions on Industrial Cyber-Physical Systems, 2023, DOI: 10.1109/TICPS.2023.3336608

F. Vitale, F. De Vita, N. Mazzocca, and D. Bruneo, "A Process Mining-based unsupervised Anomaly Detection technique for the Industrial Internet of Things," Internet of Things, 2023, DOI: 10.1016/j.iot.2023.100993

M. Cinque, L. De Simone, N. Mazzocca, D. Ottaviano, F. Vitale, "Evaluating virtualization for fog monitoring of real-time applications in mixed-criticality systems," Real-Time Systems, 2023, DOI: 10.1007/s11241-023-09410-4

F. Vitale, M. Pegoraro, W. M. P. van der Aalst, N. Mazzocca, "Control-flow anomaly detection by process mining-based feature extraction and dimensionality reduction," submitted to Knowledge-Based Systems (under review)

F. Vitale, F. Flammini, M. Caporuscio, N. Mazzocca, "Combining Process Mining and Unsupervised Machine Learning for Monitoring Resilient Computer Systems," submitted to IEEE Transactions on Dependable and Secure Computing (awaiting first decision)

S. Guarino, F. Vitale, E. Del Prete, L. Faramondi, N. Mazzocca, R. Setola, "d-TV-DBN: A Hierarchical and Distributed Architecture for Scalable Cyber-Physical Anomaly Detection with Bayesian Networks," submitted to IEEE Transactions on Emerging Topics in Computing (awaiting first decision)

F. Vitale, N. Dall'Ora, S. Gaiardelli, E. Fraccaroli, N. Mazzocca, F. Fummi, "Process Mining-Driven Fault Diagnosis and Simulation for Cyber-Physical Systems" (in progress)
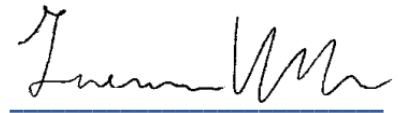
Patents and/or spin offs

Neither patents nor spin-offs were developed.

Awards and Prizes

No awards were achieved.

Date 15/10/24

PhD student signature _____

Supervisor signature _____