MECHANICAL ENGINEERING | PHYSICS | PRESERVATION OF THE ARCHITECTURAL HERITAGE | SPATIAL PLANNING AND URBAN DEVELOPMENT | STRUCTURAL, SEISMIC AND GEOTECHNICAL ENGINEERING | URBAN PLANNING, DESIGN AND POLICY | AEROSPACE ENGINEERING | ARCHITECTURAL COMPOSITION | ARCHITECTURE, BUILT ENVIRONMENT AND CONSTRUCTION ENGINEERING | ARCHITECTURE, URBAN DESIGN, CONSERVATION OF HOUSING AND LANDSCAPE | ARCHITECTURAL, URBAN AND INTERIOR DESIGN | BIOENGINEERING | DESIGN | ELECTRICAL ENGINEERING | ENERGY AND NUCLEAR SCIENCE AND TECHNOLOGY | ENVIRONMENTAL AND INFRASTRUCTURE ENGINEERING | INDUSTRIAL CHEMISTRY AND CHEMICAL ENGINEERING | INFORMATION TECHNOLOGY | INTERIOR ARCHITECTURE AND EXHIBITION DESIGN | MANAGEMENT ENGINEERING | MATERIALS ENGINEERING | MATHEMATICAL MODELS AND METHODS IN ENGINEERING

PhD Yearbook | 2018

DOCTORAL PROGRAM IN INFORMATION TECHNOLOGY

Chair:
Prof. Andrea Bonarini

Introduction
The PhD program in Information Technology (IT) goes back to the year 2001, when the two traditional programs in Automation- Computer Engineering and Electronics-Telecommunications were merged. As such, the unified course covers the research interests in four scientific areas, namely: Computer Science and Engineering, Electronics, Systems and Control, and Telecommunications. This broad variety of research activities is focused in the ICT area, and perfectly corresponds to the core mission of the Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB). However, following the historical development of the Department, and the new trends of the modern society, some cross-related research fields are also developed, such as robotics, machine learning, big data, intelligent data analysys, Industry 4.0, Internet of things, ecology, environmental modelling, operations research, and transportation systems. The PhD program in IT is the largest at the Politecnico in terms of number of students. There are about than 50 first year students and about 190 in total. The students are subject to an examination every year to evaluate the progress achieved in their research and course work.

Topics
The research carried out in the Department in the field of Information Technology is supported by 35 laboratories, and is organized in 4 main areas, listed here below. *Computer Science and Engineering* (Vice-Coordinator: Prof. Cristiana Bolchini): Information systems, Database management, Information design for the web, Methods and applications for interactive multimedia, Embedded systems design and design methodologies, Dependable systems: performance, security and reliability, Autonomous robotics, Artificial intelligence, Computer vision and image analysis, Machine learning, Dependable Evolvable Pervasive Software Engineering, Compiler Technology, Natural Language Processing and Accessibility. *Electronics* (Vice-Coordinator: Prof. Angelo Geraci): Circuits and systems: theory and applications, Single-photon detectors and applications, Radiation detectors and low noise electronics, Electronic circuit design, Electron devices. *Systems and Control* (Vice-Coordinator: Prof. Paolo Bolzern): Control systems, Robotics and industrial automation, Optical measurements and laser instrumentation, Dynamics of complex system, Planning and management of environmental systems, Operations research and discrete optimization. *Telecommunications* (Vice-Coordinator: Matteo Cesana): Networking, Applied electromagnetics, Information transmission and radio communications, Optical communications, Wireless and space communications, Remote sensing, Signal processing for multimedia and telecommunications.

Industrial collaborations
Due to its intrinsic technological nature, the PhD curriculum is corroborated by many industrial collaborations. About 50% of the total number of scholarships are funded by companies or by international research projects involving industrial partners. In the school vision, the collaboration between university and companies is ideally based on the challenge of turning invention and scientific research into technological innovation. This shapes new technology frontiers and builds a fertile atmosphere for a mixture of world-class research at universities and in private and public companies. This also contributes to create a common terrain of friendly culture, to size the risk, and to believe in strong basic research. The external referee board is composed by members of public and private companies, working in management, production, and applied research. The board is in charge of monitoring the activities of the PhD program and giving suggestions for its development. The board meets once a year to point out the new emerging research areas worth to be investigated and to monitor the visibility of the course in the industrial and service world.

Educational aspects
The teaching organization and subject of the courses reflect the scientific interests of DEIB faculties. The curricula include a wide choice of courses (about 30 per year) of different nature offered by this PhD Program, and more than 30 courses for basic skill formation offered by the Polimi PhD School. Access to external courses and summer schools is also encouraged. The challenge is to promote interdisciplinary research while offering advanced courses that spur innovative and cutting edge research. Therefore, particular attention is devoted to help students to make their best choice according to an internal regulation scheme.

Internationalization
Every year, several courses are delivered by foreign professors. Moreover, the PhD program encourages joint curricula through agreements with foreign institutions. We have several Double Degree and Joint Degree agreements with countries in all continents. Every year we receive more than 150 applications from foreign countries and about 20% of our PhD candidates come from there.

Conclusions
The core mission of this PhD Program is to offer an excellent PhD curriculum, through high quality courses, a truly interdisciplinary advanced education, cutting-edge research, international collaborations, and collaborations with companies.

PHD BOARD OF PROFESSORS		
ANDREA BONARINI	MATTEO CESANA	RAFFAELA MIRANDOLA
CESARE ALIPPI	FABIO DERCOLE	ANDREA MONTI GUARNIERI
FRANCESCO AMIGONI	LORENZO FAGIANO	BARBARA PERNICI
LUCA BASCETTA	GIANCARLO FERRIGNO	MATTEO PRADELLA
GIUSEPPE BERTUCCIO	CARLO FIORINI	IVAN RECH
CRISTIANA BOLCHINI	SIMONE GARATTI	ALESSANDRO SOTTOCORNOLA SPINELLI
PAOLO BOLZERN	NICOLA GATTI	MASSIMO TORNATORE
ANDREA CASTELLETTI	ANGELO GERACI	
STEFANO CERI	MAURIZIO MAGARINI	

PHD ADVISORY BOARD	
Paolo Amato	Micron
Claudio Bartolini	Cloud4Wi
Mario Caironi	IIT
Cristina Cremonesi	The European Ambrosetti
Riccardo De Gaudenzi	European Space Agency
Giuseppe Fogliazza	MCE Srl
Massimo Leoni	IBM
Renato Lombardi	Huawei
Renato Marchi	Gruppo PAM
Massimo Valla	TIM
Luisa Venturini	Vodafone
Stefano Verzura	Huawei
Roberto Villa	IBM

Prizes and awards
In 2017 the following awards have been obtained by PhD students:

Chorafas Foundation Award: **Abdulrahman Kaitoua, Rafael Jan Pablo Schmitt**
Young Researcher Award of the International Hydropower Association (IHA): **Rafael Schmitt** former PhD student
IEEE Computer Society Italy Section Chapter 2016 PhD Thesis Award: **Amir H. Ashouri**
17th International Conference on Web Engineering BEST STUDENT PAPER award: **Carlo Bernaschina**
Maxeler Open Dataflow Design Competition – 2nd place: **Emanuele Del Sozzo, Marcello Pogliani**
CAiSE Best Paper Award: **Marouan Mizmizi**
Premio Francesco Carassa 2017: **Alberto Bernardini**
XILINX Open Hardware 2017 PhD Award**: Emanuele Del Sozzo**, **Marco Rabozzi**, **Marco Nanni**
"Switch2Product Innovation Challenge" – 1st place: **Rolando Brondolin**
Innovation Design Contest 2017 – 3rd place **Andrea Ragni**
Microsoft AI for Earth Azure Research award: **Darian Frajberg**
HIPEAC Tech Transfer Award per il progetto HyPPO: **Rolando Brondolin, Marco Arnaboldi**
IEEE Computer Society Lance Stafford Larson Writing Award: **Davide Quarta**
IEEE MEMS 2018 Outstanding Student Paper Award: **Cristiano Marra**

SAFER-HRC: A METHODOLOGY FOR SAFETY ASSESSMENT THROUGH FORMAL VERIFICATION IN HUMAN-ROBOT COLLABORATION

Mehrnoosh Askarpour - Supervisor: Dino Mandrioli

Robots usually function in workspaces divided (e.g., by fences) from those of human operators. However, novel robotic and cyber-physical systems have evolved in size and functionality to include the collaboration with human operators within common workspaces. This new application field is often referred to as Human-Robot Collaboration (HRC) and is increasingly prominent in people's lives and in industrial domains, for example in manufacturing applications. However, HRC raises new challenges to guarantee system safety, due to the presence of operators. Close proximity and frequent physical contacts between operators and robots, and intrinsic non-determinism in operators' behavior make it difficult for safety assessors and analyzers to cope with the dynamism of collaborative applications. Yet, formal verification techniques can help in this regard through the exhaustive state-space exploration of system models, which can identify unwanted situations early in the development process. In fact, this thesis proposes to use formal verification techniques for analyzing risks in HRC, through a methodology, which is compatible with well-known standards in the area of collaborative robotics. In particular, the methodology relies on temporal logic based models to describe the different possible ways in which tasks can be carried out, and on fully automated formal verification techniques to explore the corresponding state-space to detect and modify the hazardous situations at early stages of system design. This innovative methodology, called SAFER-HRC, is centered on TRIO temporal logic language and Zot bounded satisfiability checker, to assess the safety risks in an HRC application. This work focuses on mechanical hazards which unlike other types (e.g., thermal, electrical, etc.), are results of human-robot interaction. The methodology starts from a generic modular formal model and customizes it for the target system; It then iteratively checks the model against certain safety properties, to study the safety of the collaborative environment for human operators. The generic model contains a nondeterministic formal model of operator behavior, which captures the hazardous situations resulting from human errors. This method has an iterative and incremental nature, and allows safety engineers to refine their designs until all plausible erroneous behaviors are considered and mitigated. Finally, this thesis introduces a toolsupported approach for the automated generation of formal models from a semiformal language, in particular UML. The tool prototype is based on Papyrus UML modeler that provides mechanisms to allow safety engineers, who are typically not experts in formal methods, to automatically generate and formally verify logical models from an UML-based notation, which is more attuned with their background.

MULTIROBOT EXPLORATION OF COMMUNICATION-RESTRICTED ENVIRONMENTS

Jacopo Banfi - Supervisor: Prof. Francesco Amigoni

In the last two decades, research in mobile robotics has shown that exploiting a team of cooperative robots can be a valid alternative to the employment human operators in carrying out repetitive, difficult, or hazardous tasks. Applications like warehouse management, information gathering, search and rescue, and patrolling are some representative examples. Such teams of robots operating in the same environment to achieve common goals are customarily referred to as *multirobot systems*.
From a general perspective, different challenges are involved in the deployment of a multirobot system for a particular application: these range from the choice (or development) of appropriate robotic platforms to deal with low-level issues (sensing and actuation) to the development of *decision-theoretic planning tools* aimed at enabling the robots to accomplish the assigned tasks with the required autonomy. In the context of mobile robotics, decisions happening at the highest levels of abstraction imply the definition of a *navigation strategy*, which specifies high-level directives for driving the robots in an environment to achieve some goals, like visiting the most "interesting" locations.

This thesis is concerned with the development of novel navigation strategies for teams of autonomous mobile robots in the context of a particular kind of information gathering task, namely, the *exploration of unknown environments*. In a nutshell, the problem that the robots have to face in this context is the following. Imagine some robots placed at the entrance of an unknown environment, as shown in Figure 1. They are equipped with sensors able to perceive their surroundings (e.g., laser range scanners and/or cameras) and can communicate with their teammates and, possibly, with a supervising control station either by means of an *ad hoc* network, or through a pre-deployed communication infrastructure. In the context of "plain" map building, the objective of the exploring robots is simple: construct a representation (map) of the unknown environment in an efficient way. However, if we think of search and rescue applications, map building is only functional to enabling the robots to find as many victims as possible in a short period of time.
Throughout the years, several efforts have been devoted to study coordinated multirobot exploration strategies assuming that communication is possible between any two locations. However, this is a reasonable assumption only in a very limited number of real settings. In fact, real operational conditions may require to deploy a team of robots only equipped with local limited-range communication modules: for instance, think of a search and rescue mission in a collapsed building, where the pre-existing WiFi infrastructure has been destroyed. Methodologies and techniques to devise multirobot systems for effective exploration in presence of communication constraints, despite their remarkable practical relevance, have not yet reached a level of maturity comparable to that of their "ideal" counterparts. Therefore, our work aims at improving the current state-of-the-art in multirobot exploration of communication-restricted environments by addressing some of its key challenges that, to the best of our knowledge, are still unsolved. The main contributions of the thesis can be summarized as follows.

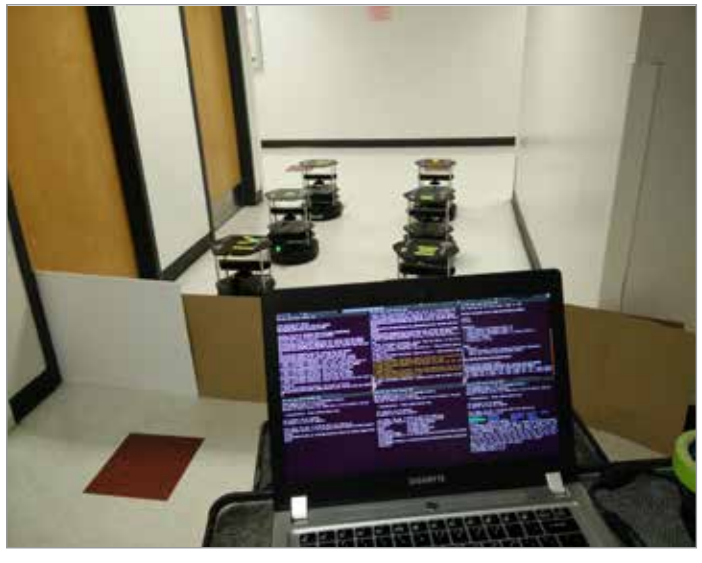

Fig. 1 - Six robots ready to explore an unknown environment.

First, we provide a general framework for modeling multirobot exploration in presence of communication constraints, where we theoretically justify some modeling choices regarding the construction of robots' paths by means of an original theorem concerning the intractability of time-optimal multirobot path planning even on 2D grid graphs with holes (a typical environment discretization used in mobile robotics).
Second, we present a novel exploration strategy operating under what we call recurrent *connectivity constraints* to a fixed base station, which are often imposed in search and rescue settings. In short, our strategy enforces the robots to be connected with the base station each time they gather new information from the exploration frontiers. By allowing to submit new plans to arbitrary groups of robots, provided that they are "ready" (informally, they have sent all the required data to the base station), we are able to obtain a good trade-off between rapidity of the mapping process and situational awareness at the base station. Plans prescribe the robots to reach locations from which new information on the environment can be acquired and locations that form a relay chain with the base station, according to the predicted presence of communication links.
Third, we study a possible approach to overcome the widespread adoption in the literature of overly conservative or unrealistic assumptions on the robots' communication capabilities by proposing coordinated navigation strategies aimed at mapping the unknown communication features of a (fully or partially) known environment. These strategies work by collecting samples for building a Gaussian Process (GP) representing a map of the signal strength between any two locations of the environment. Such maps can then be used to infer the existence of communication links in a non-conservative, yet reliable, way. Being specifically designed for an online setting, our strategies for building communication maps could be employed along with several multirobot exploration strategies presented in the literature by simply pairing a first team of exploring robots with an additional dedicated team. However, nothing prevents their use also for building communication maps for applications when the environment is fully known in advance, such as surveillance.
Finally, we propose an algorithmic framework for computing and handling the execution of backup plans aimed at dealing with the absence of predicted communication links in generic multirobot information gathering missions (hence including exploration as a particular case). In particular, we formalize what we call the *multirobot reconnection problem*, study its computational complexity, and provide optimal and heuristic algorithms for practical settings. These algorithms can be used to improve the robustness of all the information gathering strategies which exploit predictions of the availability of a communication link between two locations (including ours) which, in general, could be affected by false-positives.

EFFICIENT ENERGY-AWARE MODELS FOR CLOUD COMPUTING SYSTEMS AND NETWORKS

Amine Barkat - Supervisor: Prof. Antonio Capone

Developments in communication networks gave the birth to cloud computing which is revolutionizing the use of IT services in companies as electricity did in its time. Once, companies had to produce their own energy to operate before the arrival of electricity networks. Today, companies must manage their IT assets with the constraints related to their consumption and maintenance. With cloud computing, companies can use on demand IT services without worrying about the management, security and maintenance of the infrastructure.
However, the energy consumption of cloud system is not negligible. Indeed, currently the energy consumed by ICT is estimated to be more than 4% of the worldwide consumption and it is expected to double in the next few years. This consumption has effects not only on economies of governments and companies, but also on planet environment through its carbon footprint, which makes energy efficiency of cloud systems one of today's major challenges.
In this thesis we address cloud energy issues by developing optimization models based on operational research techniques by taking into account the economic perspective of the cloud providers and users. Unlike most of existing methods that handle with different related problems separately, in our works we have tried to enlarge the vision, and jointly consider different problems in single frameworks. Two types of Cloud systems were considered, and for each type joint optimization has proved to be more effective and more energy efficient with respect to the nature of considered problems.
More specifically, the first type considered, is a Cloud system providing Platform as a Service (PaaS) to its users, where load balancing between different servers can be used to save energy, however, it may be very expensive in terms of network traffic, as it may lead to congestions and excess of energy consumption in routers if too much workload is exchanged between servers. Therefore, jointly managing servers and their interconnecting network was a targeted task.
We have proposed a holistic framework for geo-distributed data centers and their interconnection network.
This solution is based on the feature offered by virtualization technology to migrate virtual machines between different physical machines for load balancing. We could define a multi objective model, that we can separate its sub-problems to three distinct problems: Data centers energy management, Network energy management, and Green energy management.
The defined model has allowed to study different phenomena related to the problem, as well as the interactions between different parameters. This solution is formulated using mixed integer linear programming based on a set of linear constraints. The goal of the model is to optimize the three objectives concurrently with the aim of minimizing the total energy consumption and expenses of the Cloud system and its network.
The proposed model takes advantage of the differences between energy prices of various geographical locations of Data centers, as well as the availability of green energy resources.
Basically, different cloud services are migrated between data centers through virtual machine migration technology. VM migration is performed when the price of energy is lower in the destination DCs or to explore more green energies. While migrating virtual machines, we take into consideration network traffic constraints, and also the routers consumption. Alongside, the model tends to balance cloud energy consumption between green and brown energies using storage technology. Green power can be stored in low traffic time periods for a later use when the price of electricity is so high.
The amount of stored power can be used to address the requests received by the same data center, or received from other data centers from which services are migrated.
The obtained results by running the model under different scenarios show the importance of the joint optimization comparing to the separate optimization. Gains in term of energy expenses can be up to 34% comparing to the same proposed solution but considering a separate optimization, while energy savings can be up to 70% comparing to solutions in which network is not optimized.
Regarding energy consumption, savings can be really significant up to 43% beside the better exploitation of the environment friendly power.

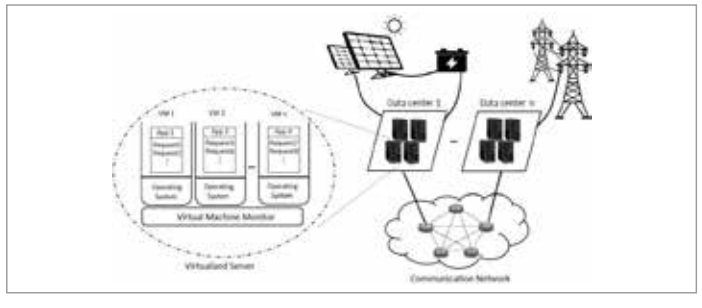

Fig. 1

In the second part of this thesis, we have considered a new emerging type of Cloud computing which is Cloud Radio Access Network or C-RAN. Our goal is to fill the gap and to develop efficient optimization framework for Cloud Radio Access Networks.
Following the same principle of wider vision is more optimal, in this type of systems we have tried to minimize energy consumption at a very early stage which is network planning stage. We have introduced a mixed integer linear programming model in which we jointly optimize the planning and energy management simultaneously, with the aim of minimizing both installation and operational costs (CAPEX and OPEX). In this multi objective model, we have tried to target three problems: equipment installation, energy management of RRHs and BBUs, and quality of service.
Basically, among a set of candidate sites for RRHs and BBUs pools, the objective of the model is to find the set of equipment to install in order to guarantee a maximum coverage of the targeted area, as well as to reduce the OPEX during the management phase. Power saving is achieved by turning on and off RRHs depending on the traffic demands.
Inside the BBU pool, the number of running BBUs is the minimum that ensures the optimal serving for the connected RRHs.
Regarding the Quality of Service, the model tries to connect users to the closest RRH for a better connection.
Weight parameters are used in the objective function for each sub-problem. We have tested the model using realistic data, and with variation of weight parameters. The obtained results show that significant energy savings can be achieved using joint optimization of planning and energy management of C-RAN.
Energy savings with an additional CAPEX can reach 74% based on the tested scenarios. Also we observed that small cells are more energy efficient and easy to manage comparing to the big cells.

Fig. 2

SILICON DRIFT DETECTORS AND READOUT ELECTRONICS FOR HIGH THROUGHPUT SPECTROSCOPY APPLICATIONS

Giovanni Bellotti - Supervisor: Prof. Carlo Ettore Fiorini

X-ray Absorption Fine Structure (XAFS) spectroscopy studies the absorption coefficient of materials in the region close to the absorption edges. After-edge oscillations of the absorption coefficient contain information about the chemical and physical structure that surrounds the excited atom. Fluorescence mode XAFS spectroscopy makes it possible to obtain the XAFS for thick or very diluted samples measuring their emission spectrum that, in the dilute limit, is proportional to the absorption. In this technique, the sample is excited by a beam of photons, whose energy can be finely tuned using a monochromator. The energy of the beam is changed at small steps starting in the region immediately before the absorption edge and ending 1 keV after it.
The fluorescence produced by the sample is acquired by an X-ray radiation detector. Solid State Detectors (SSDs) are generally used for XAFS as they guarantee high energy discrimination.
Among SSDs, the best performances in terms of count rate and energy resolution are provided by Silicon Drift Detectors (SDDs). SDDs, that have been introduced by E. Gatti and P. Rehak in 1983, are characterized by a signal-charge collecting electrode with a very low capacitance, independent from the device active area. These detectors have the high energy resolution required by the applications (in a 55-Fe spectrum, the Ka line should have a FWHM < 200 eV) to distinguish among the different spectral lines of interest. As in SSDs each photon is revealed independently and its processing takes a finite amount of time, there is a finite dead time after each pulse in which the detector is not able to acquire another pulse. This sets a limit on the maximum count rate the detector can achieve. Synchrotron beamlines are experiencing in these years an increment in brilliance (i.e. in fluorescence flux) that asks for an improvement of the existing detectors in terms of count rate.
My doctoral activity took place in the framework of the ARDESIA experiment. ARDESIA (ARray of DEtectors for Synchrotron radIation Applications) is an experimental project funded by the INFN National Scientific Committee V, whose goal is the realization of an X-ray spectrometer for synchrotron experiments based on arrays of SDDs. The spectrometer has been designed and manufactured in all its parts, from the detection module to the mechanical structure and the readout electronics.
ARDESIA detection unit is the main building block of a flexible and adaptable platform for high-count rate high-energy resolution soft X-rays spectroscopy. It is composed by a monolithic array of 4 SDDs, with an area of 5×5 mm^2 each, read by a 4-channel CUBE preamplifier. The SDD arrays are manufactured by Fondazione Bruno Kessler (FBK, Trento, Italy). The window of the SDDs is characterized by a thin dielectric layer and a very shallow junction that optimize the collection efficiency of carriers generated close to the surface. The SDDs are produced with a low-leakage process, that keeps the leakage current at room temperature lower than 200 pA/cm^2, making it possible to obtain good energy resolution even at relatively high temperatures. CUBE CMOS charge preamplifier is connected to the SDD's anode through chip-to-chip wire bonding. The high transconductance of the input MOSFET makes this technology suitable for high-count rate applications. The detection unit has an area of 16×16 mm^2 and a fully modular structure. Multiple units can be arranged together without dead space among them. The detection unit has an energy resolution higher than 126 eV at moderate count rate (some kcps per channel) and it can sustain input count rates higher than 1Mcps per channel.
DaFne-Light DXR1 is a is a soft X-ray beam-line source installed at DaFne synchrotron at the National Laboratories of Frascati (LNF), near Rome. The beamline is characterized by the production of a high flux X-ray beam in the 1-3 keV energy range. A single detection module has been used to build a complete spectrometer for DaFne-Light DXR1 beamline. Up to now X-ray absorption fine structure spectroscopy (XAFS) measurements at DaFne-Light DXR1 were performed in transmission mode. Using ARDESIA detector it is possible to obtain the XAFS in fluorescence geometry, which is preferable when dealing with diluted or thick samples. The spectrometer interfaces the detection unit with the scattering chamber, assuring that it does not interfere with the vacuum of the chamber and it is not interfered by the environment of the chamber. The detection unit is kept in vacuum at low temperature during operations. All the bias voltages for the unit are generated inside the instrument. An o-ring connection between the flange and the detector structure allows the latter to move inside the scattering chamber while keeping vacuum condition.
TERA (Throughput Enhanced Readout ASIC) is an embedded shaping amplifier for ARDESIA detector. It is a 4-channel analog pulse processor for SDDs equipped with CUBE preamplifier with fully digital output and high-count rate capability. Each channel implements a 7-th order time continuous semigaussian filter with high output dynamic range, AC-coupled with CUBE preamplifier. There are 4 different programmable pulse-widths, from 200 ns (very high-count rate) to 2 ms (high energy resolution), so it is suitable for both high count rate and high energy resolution, and 4 programmable energy ranges, from 5 keV to 40 keV. A derivative-based events recognition circuit optimizes the throughput achievable with a 3-phase peak stretcher circuit while rejecting piled-up events. The channels are multiplexed in couples, and each channel is equipped with a 4-elements wide analog memory bridge that matches the random events distribution with the fixed acquisition rate. The output of the multiplexer is sampled by a 12-bits 4 Msps on-chip ADC with serial output interface.

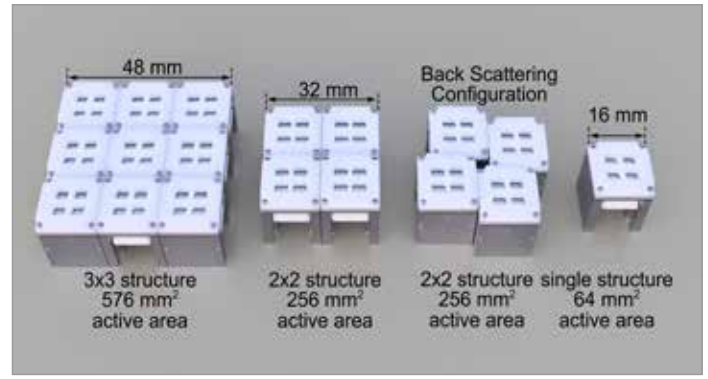

Fig. 1 - ARDESIA detection unit in single element and multi element configurations

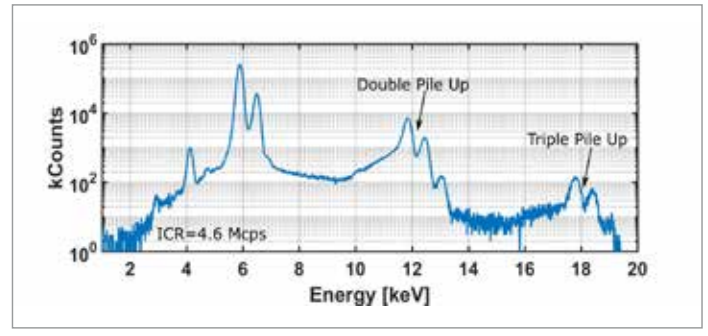

Fig. 2 - 55-Fe spectrum obtained with ARDESIA detection unit (sum of the 4 channels) with 4.6 Mcps input count rate.

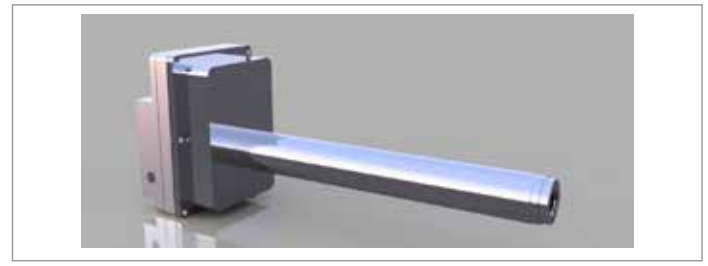

Fig. 3 - ARDESIA spectrometer mechanical structure.

SYSTEM SUPPORT FOR TRANSIENTLY-POWERED EMBEDDED SENSING SYSTEMS

Naveed Anwar Bhatti - Supervisor: Prof. Luca Mottola

Transiently powered embedded systems are becoming popular because of their self-sustainable, no maintenance and easily deployable nature. However, there is an intrinsic challenge with these systems: they can be unpredictably interrupted, as energy harvesting by no means can ensure a predictable supply of energy. Reboots will frequently happen, which translates into a waste of resources, including energy, as applications need to re-initialize and re-acquire the state. As a result, the overall performance inevitably suffers. To allow an application to cross the boundaries of periods of energy unavailability, prior solutions, either save only a portion of program memory (avoiding the heap) limiting developers to employ sophisticated programming techniques, or resort to hardware modifications by replacing SRAM with FRAM, that may not only impact cost but also processing speed.

This thesis aims to design software techniques for transiently-powered embedded devices, allowing an application to make progress, with a minimum possible energy spent on saving the system state and without resorting to hardware modifications. In the first part of the thesis, we present the detailed analysis of the existing energy harvesting and wireless energy transfer solutions for wireless sensor networks (WSNs). We define desirable properties, classify existing solutions, and argue about their applicability in different deployment environments. Later, we conduct a comprehensive survey of the state of the art for transiently-powered embedded systems. We discuss challenges, define goals and classify transiently-powered embedded system solutions into different categories based on the techniques they use to ensure forward progress of the application.

In the second part of the thesis, we develop three different techniques for saving system state quickly and in an energy-efficient manner, exploiting different properties of non-volatile memory. Key to their efficiency is the way the state information is organized on non-volatile memory. Our results, through extensive evaluation, crucially indicate that there is no “one-size-fits-all” solution. It is the application’s memory characteristics that will make one technique preferable over another. These evaluation results also lead us to design an additional technique, DICE, in which, instead of reading non-volatile memory to compute changes in the main memory (which is energy-hungry operation and used in previous techniques), we track changes in the main memory through just code instrumentation. This makes DICE not only further reduce the amount of data to write onto non-volatile memory to ensure forward progress of the application, but also helps existing system support complete a given workload with better energy efficiency and reduced execution latency.

Finally, we present HarvOS that decides when to save the system state by looking at the worst-case energy cost required to reach the next opportunity to save system state, depending on the program structure as represented in the control-flow graph. HarvOS allows the system to make an informed decision, at every opportunity to save system state, on whether to continue with the normal execution or save the system state.

Our evaluation indicates that HarvOS allows transiently-powered embedded systems to complete a given workload with 68% fewer restarts, compared to existing literature.

OPTIMIZATION MODEL FOR RESOURCE ALLOCATION OF MULTIPLE APPLICATIONS IN VIRTUAL SENSOR NETWORKS

Sonda Bousnina - Supervisor: Prof. Matteo Cesana

The past few years have witnessed an enormous progress in the field of Wireless Sensor Networks (WSNs) which attracted growing attention from both industrial and research sides. These networks are currently used in the scientific, medical, commercial, and military domains covering a wide range of applications like home automation, industrial control, security, healthcare, etc.
WSNs have become one of the fundamental components of empowering the vision of the Internet of Things (IoT), where massive numbers of smart objects interacting with the surrounding environment have existence on the Internet. Actually, the IoT paradigm considers that real world objects can be equipped with sensing capabilities to gather information on their environment, processing/storage capabilities to locally filter and store data, and wired or wireless connections to deliver the collected/processed data to the Internet. This data transfer can be completed either directly, or by wireless multi-hop paths leveraging the cooperation of other smart objects for traffic relaying and creating WSNs.
The classical design approach of WSNs is highly vertical with network deployments tightly customized on the needs of one or very few specific applications (one application; one network design approach), in which the hardware and network resources are customized to the specific application requirements.
On one hand, such design paradigm permits to have optimal performance on the specific application, nevertheless, on the other hand, it prevents resources (hardware and software) reuse when other applications and services must be contemplated. However, there is an increasing need for developing general purpose WSNs that can easily be configured to support multiple applications with heterogeneous requirements. At this point, novel approaches are recently being investigated targeting this kind of networks which often go under the names of Virtual Sensor Networks (VSN) or Software Defined Sensor Networks (SDSN).
The vision of VSNs calls, on one side, for software platforms to abstract away the complexity and diversity of the available physical resources, and, on the other side, it requires effective algorithms to dynamically allocate such physical resources to multiple incoming applications requesting for service. Consequently, VSNs feature to a more efficient resource utilization, lower cost and increased flexibility and manageability in WSN deployments.
While network virtualization is already a reality in many communication networks, research on sensor network virtualization is still in its early stages and comprehensive solutions still need to be found to cope with the specific characteristics of WSNs in terms of limited node capabilities and communication bandwidth.
In this context, we propose the design of a novel virtualization engine for WSNs to effectively optimize, adapt and orchestrate the physical resources of the underlying WSNs across multiple applications. Given the background, this is the first approach to model, analyze and allocate the physical resources of the general purpose WSN to multiple concurrent applications while accounting for the limitations in the physical resources (processing, storage, available bandwidth, limited communication range) and the specific application requirements for different applications in virtual sensor networks from an optimization point of view.
This thesis begins by focusing first on the optimization problem of coordinating the shared resources in VSNs while taking into account the specific application requirements and the constraints posed by the physical wireless network.
We start by giving an analysis of a virtual sensor network where different kinds of applications and sensor nodes coexist and cooperate. After that, we formulate mathematically the optimization problem of maximizing the overall revenue out of the application deployment process while minimizing the cost related to activating sensor nodes. Constraints regarding sensor nodes capabilities (memory, computation, energy) and network limitations (topology, shared bandwidth) are included and the complexity analysis of the problem to optimally allocate physical network resources to applications in VSNs is provided.
Later on, the proposed model is evaluated and validated by simulation based on realistic parameters for the set of applications and sensor nodes. We mainly focus on showing the benefits of virtualization and we analyze thoroughly the impact of varying the main model parameters (number of scalar and multimedia nodes; number of sinks; lifetime; type of routing) in the system performance.
Henceforth, due to the high computational complexity of the resulting optimization model, we introduce two heuristics to reduce the processing time while keeping the optimality gap low: one based on the linear relaxation of the problem formulation and another one leveraging a greedy approach. Finally, experimental results for both heuristic have been validated via numerical analysis. In this part of work, we start with a scenario completely static. In other words, we assume in this case that the whole set of applications was known in advance and constant during the analysis time period. Since the reference problem is proved to be NP-Complete and as a consequence the time required to solve the problem increases very quickly, we propose two heuristic algorithms in reduced computation time: The first solution approach exploits the Linear Relaxation of the problem formulation, whereas the second one is formulation independent and it follows a Greedy strategy.
To encompass more realistic scenarios, we extend our model to consider the dynamic management of applications deployment. Thus addressing the case where the set of applications to be deployed is not known from the beginning but rather applications “appear” in the system according to specific arrival processes.
Correspondingly to the static case, heuristic algorithms are introduced to obtain nearly optimal solutions. Moreover, to reflect the inherent heterogeneity of sensor nodes hardware platforms in common Internet of Things applications, the network topologies used for performance evaluation include sensor node hardware with different characteristics in terms of cost and available resources.
All the way through this thesis, we continuously evaluate and validate our models considering realistic parameters from actual sensor nodes and deployed applications to provide a detailed performance evaluation and to assess the gain involved in letting multiple applications share a common physical network with respect to one-application, one-network vertical design approaches.
In a nutshell, the thesis sheds light on the various ways of developing the needed intelligence for resource sharing in VSN. Consequently, this intelligent framework can be applied to many heterogeneous smart environments to decrease the deployment costs, reduce the energy consumption and increase the usability and efficiency of the network.

DISTRIBUTED RANDOMIZED MODEL SELECTION FOR NONLINEAR IDENTIFICATION AND SUPERVISED MACHINE LEARNING

Aida Branković - Supervisor: Prof. Luigi Piroddi

Today thousands of variables or features are often used in classification problems. It is therefore crucial to select the most relevant ones in order to obtain robust, reliable, and easily interpretable models, not to mention storage space and classification time issues. Feature Selection (FS) aims precisely at selecting features that allow a good discrimination among samples of different classes. Suitable criteria are required to remove irrelevant and redundant features. Similar issues are encountered in nonlinear identification. For example when identifying polynomial NAR[MA]X models from data one is faced with the task of selecting the most appropriate model structure to represent the underlying system. This task, denoted Model Structure Selection (MSS), is akin to FS.

Both mentioned tasks configure combinatorial optimization problems aiming at selecting the combination of features or model terms that result in the most accurate classifier or model. The objective of this thesis is to investigate the possibility of employing some recent randomized techniques, originally developed in the nonlinear identification area, to FS problems, and to extend those techniques in both contexts, in order to deal with large-size problems. Indeed, the difficulty of these subset selection techniques increases rapidly with the size, given the exponential complexity of the underlying combinatorial problems.

The first outcome of the research is a novel classification approach (denoted RFSC, for Randomized Feature Selection and Classification), adapted from the nonlinear model identification framework, which jointly addresses the feature selection and classifier design tasks. The classifier is constructed as a polynomial expansion of the original features and a selection process is applied to find the relevant model terms. The selection method progressively refines a probability distribution defined on the model structure space, by extracting sample models from the current distribution and using the aggregate information obtained from the evaluation of the population of models to reinforce the probability of extracting the most important terms. The performance of the RFSC was found to be quite satisfactory on small/medium size problems.

To address large size problems, a distributed scheme is here proposed, which employs a vertical partitioning on the features and operates the selection in parallel on different feature subsets. The method alternates the parallel selection phase with a partial information exchange among the different processors, which reinforces the probability that promising terms be selected. The proposed scheme is applicable to both nonlinear identification and FS problems and in both frameworks it resulted in significant improvements in performance and efficiency. Moreover, the method has a tendency to produce small models, easily amenable to interpretation. While capable of addressing much larger problems than the non-distributed approach developed previously, the distributed scheme was found to be ineffective when dealing with extra-large search spaces (as are encountered, e.g., with micro-arrays), due to computational issues associated with parameter estimation and classifier design. An alternative version of the distributed scheme was then developed to target micro-arrays in particular, which employs a non-parametric multivariate filter algorithm and population extraction using the distance correlation index (dCor) as a criterion.

Finally, while analyzing the behavior of the RFSC, it was noticed that structurally different classifiers may result in equivalent performance due to the discrete nature of the 0 - 1 loss function in classification problems. The randomized characteristic of the RFSC was then exploited to generate ensembles of classifiers. In most cases the results demonstrate an improved accuracy when ensembles of classifiers are employed with respect to the 'single classifier' case.

All proposed methods have been evaluated and compared to other well-known FS and MSS methods on standard benchmarks for classification/nonlinear identification problems. The results show the effectiveness of the proposed methods with respect to competitor methods both in terms of prediction accuracy and model complexity.

STABILITY CONTROL AND ANALYSIS OF TWO-WHEELED VEHICLES OUT OF PLANE DYNAMICS

Fabio Busnelli - Supervisor: Prof. Sergio M. Savaresi

This thesis deals with the study and the design of a semi-active stability control of out-of-plane dynamics for two-wheeled vehicles. The thesis is divided in three parts: the first part is devoted to introduce two-wheeled vehicle dynamic, together with the actuators and sensors commonly installed on production motorcycle; the second part is focused on the implementation of a multi-body model, the design of semi-active stability control algorithm and the design of an optimal steering damping for handling purpose; the third part deals with the estimation of the side-slip angle for two-wheeled vehicles.

The control design problem considered in the second part correspond to two different problems:

1. stability control system based on semi-active steering damper, in particular based on an electro-hydraulic steering damper;
2. optimal steering damping design to enhance handling of the motorbike when the vehicle is not affected by stability issues.

In the third part the side-slip estimation problem is carried out and two approaches are considered: a kinematic based estimator and a black-box approach, based on an artificial Neural Network, are analyzed and tested. Both the problems, the control design algorithm and the side-slip estimation, are firstly studied in a simulation environment and then their efficacy is tested with an equipped motorbike.

At first a dynamic multi-body model has been developed. The model has been used to make sensitivity respect to longitudinal speed, rolling angle, longitudinal acceleration and steering damping coefficient. Special focus of sensitivity is given to the out-of-plane modes of weave and wobble. Validity of the model is confirmed by the fact the similar conclusions are given by multi-body models presented in the literature. Furthermore, the model parameters have been identified by means of an optimization procedure based on experimental data, where modes of interest have been excited. The most interesting results from the control view point is given by the sensitivity to the steering damper coefficient. It turned out that when the steering damper coefficient increases weave mode is worse damped while wobble mode is better damped. The control algorithm developed in this work faced and solved this trade off with the algorithm presented in the next chapter.

Results given by the simulation are exploited to develop the control algorithm implemented and tested to improve stability and handling. The key feature of the algorithm is given by the Mode Detection algorithm that processes the steering damping signal to determine whether a mode is exited or not. When a mode is exited the Mode Detection algorithm defines which mode is excited and optimally damp in a closed loop fashion the oscillations. In the case no oscillations are detected an open loop control, or adaptive control, is engaged to allow the driver the best handling of the motorbike. Both the stability control algorithm and the adaptive control strategy have been studied and analyzed in simulation to find the optimal strategy respect to the given objective. In the first case the objective is to optimally damp the oscillations to avoid instability issues, while handling objective is to require to the driver the minimum force to accomplish a curve. To prove the results experimental tests have been carried out. The effectiveness of the stability control is presented through ad-hoc experiments that excite weave and wobble modes. At last, adaptive control experiments are carried out and drivers' feedback confirm the efficacy of the strategy.

Stability of motorbikes are strictly related with the side slip angle, but its measure it is critical. This work presents a black-box method to estimate the sideslip angle of a powered two wheeled vehicle. In particular the approach exploits the neural network framework to identify a non linear static relation between the available signal and the sideslip angle. The training data are achieved exploiting the most common on-board signals. Moreover, it is shown which are the most significant signals for the sideslip estimation. It has presented how the network has been designed and proves its effectiveness both in simulation and through real experimental data. Simulation is carried out with the multi-body simulator BikeSim. Several kind of manoeuvres have been simulated as well as several test manoeuvres have been used to verify the network effectiveness. This is has been achieved by defining different handling conditions changing both the speed profiles and the road conditions (bank percentage of the curves). Results shown that the RMS estimation error of the side slip angle never exceed 0.63 degrees. This results has been improved when the roll and steering angle are included to train the network: in these cases the RMS estimation error drops respectively at 0.12 degrees and 0.07 degrees. Finally, a sensitivity analysis respect to the signals of a 6-axis IMU shows that the most significant axis is the roll rate. Experimental data have been used to prove how this method performs with real and various maneuvers. To this purpose, a rich dataset has been registered and several maneuvers have been considered e.g. handling, slides, slaloms, U-turns, WOT. 6-axis inertial signals, rear wheel speed and vehicle speed are used to train the Neural Network. Sideslip angle is given by an optical sensor. All this maneuvers have been used to train a Neural Network and similar maneuvers have been used to test the designed network. Results are compared to an available commercial solution. The proposed approach showed its effectiveness and it outperforms the commercial solution with a mean improvement of about 13% and a RMSE of 0.7 degrees. This result improves even more when roll angle is included in the regressor vector. In this case, the improvement respect to the commercial solution is about 38% meaning 0.51 degrees RMS estimation error. At last, the sensitivity analysis respect to the signals of a 6-axis IMU confirms that the most significant axis for the sideslip estimation is the roll rate. In the literature, sideslip estimation requires a model, therefore the proposed approach overcomes the following main issues: define a proper model that can be used to design the estimator and identify the parameters of the model. This is not straightforward because parameters that are usually needed i.e. cornering stiffness are not trivial to identify and may vary a lot depending on the road and tire conditions.

At last, a kinematic approach for sideslip estimation is presented. The method was developed for four-wheeled vehicles, where the sideslip is referred to CoG of the vehicle and the kinematic model is given by one point moving in a 2D space. In this work, kinematic relations have used to be able to write the 3D kinematic of the motorbike in a 2D space. This is possible because measures of acceleration of the motorbike, angular velocities and attitude angles are usually available on modern motorbikes equipped with sensors like inertial platforms. The validity of this approach is shown in simulation and the sideslip has been estimated with an RMS error of about 0.1 degrees. Moreover, the designed observed confirm its efficacy with a completely different circuit, and sensitivity analysis has been carried respect to the sensor positioning and road banking. The last sensitivity showed that the estimation error rapidly increases when road banking is unknown, leading to an absolute error of 90 % with 30 % of road banking. At last, a stability analysis has been done through the Lyapunov theory because the observer depends on different parameters. The results of the stability analysis showed that stability is guaranteed when these parameters are positive and bounded.

STATEFUL DATA PLANE ABSTRACTIONS FOR SOFTWARE-DEFINED NETWORKS AND THEIR APPLICATIONS

Carmelo Cascone - Supervisor: Prof. Antonio Capone

Software-Defined Networking (SDN) enables programmability in the network. Unfortunately, current SDN limits programmability only to the control plane. Operators cannot program data plane algorithms such as load balancing, congestion control, failure detection, etc. These capabilities are usually baked in the switch via dedicated hardware, as they need to run at line rate, i.e. 10-100 Gbit/s on 10-100 ports.

In this work, we present two data plane abstractions for stateful packet processing, namely OpenState and OPP. These abstractions allow operators to program data plane tasks that involve stateful processing. OpenState is a simple but powerful extension to OpenFlow that permits the definition of forwarding rules as finite state machines. OPP is a more flexible abstraction that generalizes OpenState by adding computational capabilities, opening for the programming of more advanced data plane algorithms. Both OpenState and OPP are amenable for high-performance hardware implementations by using commodity hardware switch components. However, they are both based on a problematic design choice: to use a feedback-loop in the processing pipeline. This loop, if not adequately controlled, can represent a harm for the consistency of the state operations. Memory locking approaches can be used to prevent inconsistencies, at the expense of throughput. We present simulation results on real traffic traces showing that feedback-loops of several clock cycles can be supported with little or no performance degradation, even with near-worst case traffic workloads.

To further prove the benefits of a stateful programmable data plane, we introduce two novel applications: Spider and FDPA. Spider permits to detect and react to network failures at data plane timescales, i.e. micro/nanoseconds, also in the case of distant failures. By using OpenState, Spider provides functionalities equivalent to legacy control plane protocols such as BFD and MPLS Fast Reroute, but without the need of a control plane. That is, detection and rerouting happen entirely in the data plane.

FDPA allows a switch to enforce approximate fair bandwidth sharing among many TCP-like senders. Most of the mechanisms to solve this problem are based on complex scheduling algorithms, whose feasibility becomes very expensive with today's line rate requirements. FDPA, which is based on OPP, trades scheduling complexity with per-user state. FDPA works by dynamically assigning users to few (3-4) priority queues, where the priority is chosen based on the sending rate history of a user. Experimental results on a 10 Gbit/s testbed show that FDPA is able to provide fairness and throughput comparable to scheduling-based approaches.

DEVELOPMENT OF CUSTOM-TECHNOLOGY SINGLE-PHOTON AVALANCHE DIODE ARRAYS FOR HIGH-PERFORMANCE APPLICATIONS

Francesco Ceccarelli - Supervisor: Prof. Angelo Gulinatti

Recent years have seen the rise of Single-Photon Avalanche Diodes (SPADs) as the solid-state alternative to Photo Multiplier Tubes (PMTs) in many Single-Photon Counting (SPC) and Time-Correlated Single-Photon Counting (TCSPC) applications: high Photon Detection Efficiency (PDE), compactness, high reliability and the compatibility with the fabrication of detector arrays are just few of the main advantages of SPADs. Furthermore, silicon custom technologies have widely contributed to obtain large-area detectors, with excellent performance in terms of Dark Count Rate (DCR), afterpulsing probability, time jitter and detection efficiency over the visible range and up to 1-mm wavelength. In particular, a significant breakthrough was the introduction of the Red-Enhanced SPAD (RE-SPAD), the first SPAD able to attain a remarkable PDE in the near infrared range (e.g. 40% at a wavelength of 800 nm), while maintaining a good time jitter of less than 100 ps FWHM. These features have been already proven to be crucial in single-photon applications like single-molecule Forster Resonant Energy Transfer (smFRET), a powerful tool used to study the conformation of diverse biomolecules, like proteins, nucleic acids, etc. Nevertheless, this application suffers from long measurement times, strongly limiting the time scale on which the biological phenomena can be observed. Multi-spot excitation/detection schemes are a promising way to increase the throughput of single-molecule analysis, but this requires the use of a complex setup in which a suitable SPAD array has to be employed.

The main focus of this doctoral work was the development of high-performance RE-SPAD arrays, using silicon custom technologies. The first prototypes of RE-SPAD arrays (Fig. 1) have been fabricated at the Cornell Nanoscale Science and Technology Facility (CNF) and part of my work was focused on the experimental characterization of these devices, aimed both at quantifying their performance and at understanding their behaviour. This led to some significant results. In particular, with the new experimental set-ups (both wafer-level and chip-level) that I designed I was able to demonstrate that the electrical isolation has been fully recovered by means of deep trenches, with no detrimental effect on the detector DCR. Furthermore, I showed also that additional n+ structures implemented in the deep trenches are sufficient to inhibit the direct optical crosstalk between pixels. Therefore, deep Phosphorus diffusion are no longer needed, with obvious advantages in terms of compactness and thermal budget. Overall, these results demonstrate that RE-SPAD arrays can be employed in photon counting applications with the same performance of single-pixel devices. A 32x1 RE-SPAD array has been exploited in the development of the first complete RE photon detection module, able to attain a PDE as high as 70% at 650 nm; this module is currently employed in a multi-spot smFRET set-up, in the context of a fruitful collaboration with the University of California, Los Angeles (UCLA).

Even though the attained results are remarkable, the current design suffers some important limitations. Indeed, the onset of the edge breakdown strongly limits the applicable excess bias and, in order to overcome this, a guard ring structure is present in each pixel. Unfortunately, the possibility to increase the operating overvoltage comes at the expenses of an increased complexity in the biasing of the device and in a reduction of the attainable fill factor. Both these limitations prevent the fabrication of dense arrays and completely preclude the scaling toward arrays with thousands of pixels. Therefore, with my PhD work I propose a new design, that resorts to a high-energy Boron implantation to obtain at the same time a fully optimized electric field profile and a higher edge-breakdown voltage. Furthermore, I also show how both standard and RE detectors can greatly benefit from the use of a standard Phosphorus implantation instead of a predeposition, both in terms of PDE and temporal response. Detectors with the new design will be fabricated and experimentally investigated in the next months, in order to verify the predicted results and to exclude any detrimental effect on the device noise of both the new technological approaches.

A second important contribution that I gave was on the experimental characterization of standard SPADs belonging to the previous generation and fabricated by the National Research Council of Italy - Institute for Microelectronics and Microsystems (IMM-CNR sez. Bologna). These devices allowed me to participate in the development of a complete detection module based on a 8x8 standard SPAD array, getting acquainted also with the issues that are typical of the system integration. In particular, the module has been envisioned also for use in an alternative operating mode, that combines the output coming from the 64 pixels in order to enhance the maximum count rate to more than 2 Gcps, to provide photon-number resolving capabilities and to extend the dynamic range to 141 dB, a value higher than reported so far in literature. In particular, this value is currently limited by the speed at which the Active Quenching Circuits (AQCs) can be operated and by the presence of some high-DCR SPADs in the array. To the aim of investigating the future perspective of the 8x8 array, I also developed a detection head based on a low-noise, single-pixel SPAD, operated with a new AQC able to attain a dead time as low as 8.3 ns. The result was the best dynamic range ever reported for a SPAD: 152 dB, that can become even higher if this configuration is employed in a parallel module like the aforementioned one. Furthermore, I studied also the effect of the introduction of a n- epitaxial layer on timing performance: thanks to the reduction of the parasitic elements, the jitter shows a strong reduction of the dependence on the detection threshold. This opens new prospects in the design of high-threshold detection systems with a high number of pixels and reduced electrical crosstalk among them

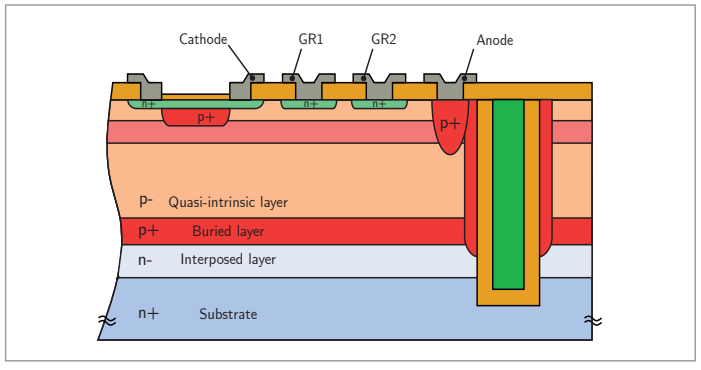

Fig. 1 - Cross section of the RE-SPAD pixel used to demonstrate the capabilities of RE-SPAD arrays

FULLY PRINTED ORGANIC IMAGERS ON FLEXIBLE SUBSTRATES FOR LARGE AREA APPLICATIONS AND NOVEL RADIATION DETECTORS

Matteo Cesarini - Supervisor: Prof. Dario Andrea Nicola Natali

In the last decades organic semiconductors have gained considerable interest in electronics because of their mechanical compliance and flexibility, the possibility of being processed and deposited via solution in ambient conditions, and to be chemically engineered in their optical properties. The prospect of depositing these materials adopting additive techniques derived from the graphical arts represents a fundamental asset for these emerging electronic technologies. As an example, inkjet printing, screen printing, roll-to-roll coating and other methods show compatibility with a wide selection of substrates, are upward scalable and thus capable of addressing large area deposition, and are compatible with already assessed industrial processes. Interestingly, these deposition techniques can now be used to deposit an increasing amount of different inorganic materials, which have complementary properties and enlarge the possibilities for printed devices. A novel concept of cheap, entirely printed "intelligent" system could become a game changer in numerous applications, paving the way to cost-effective solutions that were not even thinkable before. Examples can be envisioned in smart packaging, electronic skin, large-area scanning for border and customs security, diagnostic imaging, smart packaging, point-of-care disposable systems and wearable technology. In the field of light-sensing or light emitting applications, organic semiconductors are particularly promising for their outstanding absorption coefficients and luminescence in the visible range. Organic LED-based screens and lighting systems are now integrated with standard technology, gaining outstanding market share. Moreover, significant improvements were recently demonstrated in the field of organic, solution processed solar cells. Solution processed inorganics, in turn, become decisive when detecting higher energy photons, showing remarkable performances.
The focus of this work is the development of entirely printed opto-electronic systems and devices for large area imaging and radiation detection, using fully scalable, low material waste techniques that retain all the aforementioned (actual and potential) advantages.
The main adopted techniques are inkjet printing and screen printing. After an introductive chapter, where a summary of the fabrication techniques and an overview of the state of art of fully-printed photodetectors are presented, this document is organized in two parts. The first part focuses on fully-inkjet printed organic devices developed on flexible substrates. Chapter 2 investigates process optimization for performance reproducibility and yield in inkjet-printed vertical photodiodes. In these devices, photon to charge transduction is exerted by stacked organic photodetectors with PEDOT:PSS semitransparent contacts and P3HT:PCBM active layer. The bottom contact is functionalized by printing amine rich polymers to adjust the electrode workfunction. In this study, it is found that the most crucial process step to obtain reproducible, well performing devices is the printing of the nanometer thin interlayer used as a surface modifier for the bottom electrode. Tuning solution composition, deposition parameters, and ambient relative humidity for the deposition of this layer, an enhanced process control, yield and performance is achieved for devices, paving the way for the results described in the next chapter. A new technique to improve top contact coverage on the underlying hydrophobic active layer is also discussed. In chapter 3 the fabrication and functionality of a fully inkjet-printed organic imager on flexible polyethylene naphthalate is demonstrated. Large area image sensors are much attractive for indirect X-ray imaging, but a fully-printed prototype has not been demonstrated yet. The most common choice for the addressing element of a passive pixel, viz. a transistor, requires the integration of the transistor process and of the photodetector process, which turns out to be non-trivial in the fully printed approach. To overcome this issue, an alternative scheme is adopted, with a diode as an addressing element. The developed scheme and its critical parameters are studied in order to optimize performance. The performances of the pixel are compared to the requirements for indirect X-Ray imaging, showing noteworthy results. Printed imager prototypes on flexible substrates, reaching 8 x 8 (64 pixels) dimension, only limited by the capability of the developed custom readout system, are proven capable of reconstructing letter patterns. These results are a strong suggestion of the possibility to apply this technique to develop true low cost, large-area scanners.
The second part of the thesis concerns printed inorganic based devices for ultraviolet and gamma detection. In chapter 4 a novel kind of simple, low-voltage, printed ultraviolet detector is presented. In contrast to photodetectors based on simple charge pair photogeneration, here information is stored in a property of the material, exploiting a property of titanium dioxide, namely UV light-induced conductance switching. This effect grants devices a number of peculiarities. The first presented version of the detector has a vertical structure developed by sandwiching the screen printed active material between a top metal contact and a bottom semi-transparent contact on glass. For this detector, a process tunable memory effect is demonstrated, which could enable its use in a smart monitoring tag for light sensitive goods. The work was brought further by demonstrating imagers as simple crossbars with no need for addressing elements, in prototypical 5 x 5 arrays on glass, potentially useful for forensic reflective imaging. UV detectors were also implemented in a fully-inkjet printed, lateral configuration on flexible PEN substrates. This second implementation is a suitable candidate for whole-day sunlight exposure monitoring for skin protection. A peculiar effect is also discovered in the lateral implementation, when the active material is mixed with PEDOT:PSS. This last version of the devices shows a reverse current response to UV light, and provides a markedly faster time response and a high current switching against small amounts of dose. Chapter 5 presents preliminary tests performed on printed vertical titanium dioxide sensors, demonstrating their capability to directly sense gamma radiation. Furthermore, a strong, induced sensibilization to the visible range of light, absent in non irradiated devices, is observed after exposure. The presented results will be object of further studies. Possible strategies towards further optimization of devices for their use in gamma dosimetry are discussed.

ANALYSIS AND DESIGN OF SUSPENSION CONTROL SYSTEMS FOR OFF-HIGHWAY VEHICLES

Tommaso Colombo - Supervisor: Prof. Sergio M. Savaresi

The study, design and control of vehicle vertical dynamics is one of the most important themes in vehicle design as it radically influences the driving experience. Commonly two objectives are considered for suspensions design: the driver's comfort and the tires road-holding. The first one is the suspension capability of filtering the road asperities, reducing the accelerations the driver perceives. The second one is suspension ability to maximize the vertical contact forces between the vehicle's tires and the ground, guaranteeing the best grip. The vehicle's suspension's parameters are designed to match these two objectives. Anyway, it is well known that passive suspensions have intrinsic limits that, even considering one goal per time, do not allow to optimize the performances in all conditions. For example, considering the comfort objective, it is known that in common road vehicles a large damping guarantees better attenuation of the body resonance, while a low damping provides a better high frequency filtering. To overcome the limits of passive suspensions, in the last decades, controllable electronic suspension systems have been widely studied and used in the automotive and motorcycles field.

According to literature electronically controlled suspension systems be ideally classified depending on the control variable, the control bandwidth and the power request. The most commonly used electronic suspension classes are:

- *Load leveling suspensions:* this kind of suspensions has the capability of generating a desired quasi-static vertical force in order to balance the load forces acting on the suspension. It can be considered as an active suspension with very slow dynamics (<0.05 Hz) and its main purpose is to control the suspension equilibrium point at a desired position, rejecting the disturbances caused by load variations. The power request for a light road vehicle usually is some hundreds of watts. This kind of technology is usually found on passengers transportation vehicles (e.g. busses) which present large load variations, and luxury cars, where changing the height can be used to decrease the air drag consumption while driving, and to help the passengers exit from the car when it is parked. Not much literature is present on load leveling control.
- *Active damping suspensions:* this class features a controllable shock absorber whose damping can be varied run-time. Depending on their bandwidth they are called *adaptive* suspensions (0.5-5 Hz) or *semi-active* suspensions (10-50 Hz). As the control variable is the suspension damping, no energy can be introduced in the system. Semi-active suspensions technology is without doubt the most used and studied class of electronically controlled suspensions as it is considered the best compromise between energy consumption and control performances. Wide literature is present about semi-active suspension control systems, for both comfort and road-holding control. They are widely used in luxury cars and motorcycles to guarantee the best possible comfort and driving feeling, controlling the vertical dynamics along its complete bandwidth.
- *Active suspensions:* this kind of suspension presents a large bandwidth linear actuator which can fastly generate a vertical force along the suspension in order to control the all the vehicle vertical dynamics. Depending on the bandwidth can be internally classified as *slow active* (1-5 Hz) or *fully active* (20-30Hz). The actuator can be in parallel of two passive spring and shock absorbers or can totally substitute the passive hardware. This class of suspension is the most performing in terms of vertical dynamics controllability, and therefore can bring mayor benefits for both comfort and road holding. Anyway, due to its large power demand (1-10 kW), their usage on commercial vehicles has been limited. They have been used on F1 cars for some years, until they were banned in the early 1990s.

Most of literature and development about these electronic suspensions systems has been driven by the great "research engine" of the automotive field, which led and economically supported the research, innovation and application of electronic systems in general. For this reason, the main focus of the researches in electronic suspension systems was the vertical dynamics of road vehicles, and problems related to them.

Anyway, the evolution of electronic control system has recently started to affect also off-highway vehicles, in particular agricultural machines. In this context, the interest in vehicle vertical dynamics control is continuously growing due to its mayor potential impact on driving experience. Furthermore off-highway vehicles present different critical issues w.r.t to classical road vehicles, which makes this theme even more interesting. In particular:

- *Driver's long term health*: off-highway vehicles are driven for long periods of time (even entire days) on very rough where the driver is subjected to large accelerations and suspension oscillations. In these conditions reducing the vibration's perceived by the driver is not only related to comfort, but also to driver's long term health, and therefore it becomes a mayor suspension objective.
- *End-stroke impact*: vehicle's suspensions presents very large oscillations due to high level of roughness of off-roads ground profiles. This radically increases the risk of impacting the suspension end-stop which has two main drawbacks: first it would dramatically deteriorate the driver's comfort, second it may damage the suspension hardware. For this reason, controlling the suspension such as the end-stroke risk is minimized becomes fundamental.
- *Steep slopes and load variations*: in off-road driving, very steep slopes can be encountered. In this conditions the suspension is subjected to large slow varying load transfers which can get the suspension equilibrium point in a very asymmetrical position w.r.t suspension limits, increasing the possibility of impacting the end-stop. Furthermore, working vehicles such as agricultural machines, are often loaded of ballasts and large weights which have the same effect on the suspension. It is therefore clear that a load-leveling suspension systems becomes fundamental in this field, and a proper control system design is necessary.

Between the many off-highway vehicles, agricultural machines are of mayor interest because: they present complex non-standard vertical dynamics w.r.t to other off-road vehicles; they are used in such a wide range of road conditions and environments that they present all the previously cited critical issues; finally, as the introduction of electronic suspension systems in the agricultural field is relatively recent, not much literature is present. In fact, some research can be found about the analysis and control of agricultural vehicles vertical dynamics, but there is still much space for investigation and innovation. In particular, problems related to load leveling control design and end-stroke avoidance control, at the bests of author's knowledge has not been properly covered and present many open issues.

This thesis work is focused on the analysis, modelling and control of agricultural vehicles' vertical dynamics and the application of *load-leveling* and *semi-active* electronic suspensions technologies to them. This is done through the analysis of case studies related to the control of the cabin and front axle suspension system, which are the most classical suspension configurations that are found in tractors. All the analysis and proposed control systems focus on the driver's comfort, which as previously described is of mayor interest in this field.

DEFENDING FROM FINANCIALLY-MOTIVATED SOFTWARE ABUSES

Andrea Continella - Supervisor: Prof. Stefano Zanero

Software is involved in every aspect of our world, from our homes to large enterprises, and, in particular, it manages our data. As a consequence, software abuses can drastically impact our lives, for instance causing substantial financial losses or affecting people's privacy. This raised the attention of cybercriminals, who found in this scenario a lucrative business. In fact, in the past twenty years the motivation behind the cybercriminals' modus operandi has changed. No longer searching only for notoriety and fame, they have turned their attention to financial gain. Indeed malicious software, "malware," is one of the most dangerous Internet threat nowadays. In this dissertation, we detail our research on the analysis and detection of the current software abuses, with the aim of defending users from such threats. In particular, we concentrate on three main threats, which caused billion dollars losses in the past years. First, we focus on information-stealing malware, also known as "banking Trojans," a class of malware that steals victims' private information (e.g., banking credentials) by taking control of the victims' browser — *Man in the Browser* attacks— in order to perform financial frauds. Second, we focus on ransomware, another class of malware that encrypts victims' files, preventing legitimate access until a ransom is paid. Third, we focus on mobile privacy leaks. Mobile apps collect a wealth of users' private information, which is particularly attractive. In fact, cybercriminals are known to sell users' private information on the underground markets, and advertisement libraries massively gather such data to illicitly increase their profits.

Todays' Security Threats

Every year, new threats are discovered and, while attackers take advantage of them until effective countermeasures are found, researchers and security experts continuously implement new defense mechanism to protect users. Symantec detected 357 millions of new malware variants in 2016. The number of financial Trojan detections decreased by 36 percent in 2016 (73 percent in 2015) and mobile banking malware targeted more than 170 apps for credential stealing. Extortion-based schemes turned out to be particularly effective for cybercriminals in the last years. Ransomware, malware that encrypts users' files and asks for a ransom to release the decryption key(s), has been the most prevalent class of malware in the last two years. From 2015 to 2016, the number of ransomware families increased from 30 to 100, and the average ransom amount raised from 294 USD to 1,077 USD. Such a great diffusion made experts define 2016 as the "year of extortion." Furthermore, the last years have been also dominated by high-profile data breaches. In the last 8 years more than 7.1 billion identities have been exposed in data breaches. This proves how cybercriminals are interested in users' private data. This is due to the fact that people's information are particularly profitable on the underground markets, which today run a very proficient business: everyone can buy credit card information, full identities, or rent a scam hosting solution.

Original Contributions

In the aforementioned threat landscape, our main research area focuses on financially-motivated software abuses. In particular, we focus on generic approaches to detect these malicious activities and protect users from such threats. Our contributions focus on the mitigation of three main threats that have been widely spread and caused billion dollars losses: banking Trojans, ransomware, mobile privacy leaks.

Banking Trojans Analysis and Detection

Banking Trojans can be detected by static signatures that precisely identify malicious binaries. However, this approach is not generic and strongly depends on the implementation details of the malware sample. In addition, new families and new versions of such Trojans are constantly released. For these reasons, we propose a novel, generic, and effective approach to analyze and detect the common behavior of this malware. Modern Trojans are in fact equipped with a common functionality, called *WebInject* , used by cybercriminals to silently modify web pages on the infected hosts. In summary:

- We proposed a tool, *Prometheus*, that, based on web-page differential analysis, characterizes WebInject mechanisms in an implementation-independent fashion, without needing a-priori knowledge about the API hooking method used by the malware, and generates robust, behavioral signatures of the WebInject behavior.
- We combined the web page differential analysis with a memory forensics inspection technique to validate the generated signatures.
- We performed experiments on a dataset of real, active Trojans, and provided insights from a data analysis point of view (i.e., classification of the URLs where injections occur typically) that is used for validating our approach.
- We developed a prototype tool, *Iris* , that leverages the signatures produced by *Prometheus* to check, on the client side, whether a web page is rendered on an infected machine.

Protection from Ransomware Attacks

Preventive and reactive security measures can only partially mitigate the damage caused by modern ransomware attacks. In fact, pure-detection approaches (e.g., sandboxes or pipelines) are not sufficient, because, when luck allows a sample to be isolated and analyzed, it is already too late for several users. We believe that a forward-looking solution is to equip modern operating systems with generic, practical self-healing capabilities against this serious threat. In summary:

- We performed the first, large-scale data collection of I/O request packets generated by benign applications in real-world conditions. Our dataset contains about 1.7 billion IRPs produced by 2,245 different applications.
- We proposed a ransomware-detection approach that enables a modern operating system to recognize the typical signs of ransomware behaviors.
- We proposed an approach that makes a modern filesystem resilient to malicious encryption, by dynamically reverting the effects of ransomware attacks.
- We implemented these approaches in *ShieldFS* as a drop-in, Windows kernel module that we showed capable of successfully protecting from current ransomware attacks.

Mobile Privacy Leaks Detection

Despite significant effort from the research community in developing privacy leak detection tools, it is still unclear whether apps and ad libraries can hide the fact that they are leaking private information. In fact, all existing analysis tools have limitations: data flow tracking suffers from imprecisions; on the other hand, network traffic analysis cannot handle encryption or custom encoding. We propose a new approach that is not affected by such limitations, and it is also resilient to obfuscation techniques, such as encoding, encryption, or any other kind of transformation performed on private information before it is leaked. In summary:

- We developed a tool, *Agrigento* , that performs root cause analysis of nondeterminism in the network behavior of Android apps.
- We showed that, in most cases, non-determinism in network behavior can be explained and eliminated. This key insight makes privacy leak detection through differential black-box analysis practical.
- The results of our empirical study provide new insights into how modern apps use custom encoding and obfuscation techniques to stealthily leak private information and to evade existing approaches.

DEVELOPMENT OF SCINTILLATION DETECTORS BASED ON SILICON PHOTOMULTIPLIERS FOR HIGH-ENERGY GAMMA-RAY APPLICATIONS

Giulia Cozzi - Supervisor: Prof. Carlo Fiorini

The Dissertation represents a new challenge of Silicon PhotoMultiplier (SiPM) application to high-energy gamma-ray detection with scintillators in nuclear physics experiments and Prompt Gamma Imaging (PGI) in proton therapy. Although SiPMs are considered a promising alternative to the PhotoMultiplier Tubes (PMTs) in many medical imaging applications, more efforts are mandatory to introduce SiPM technology also into the field of high-energy gamma-ray detection. The main challenges are linked to the high energy range, within saturation effects can be present, and the need of several units to cover the large area of detection required by the applications. Therefore, an optimization of the detector and the front-end electronics is mandatory to face the high energy range and to get high spectroscopic performances and imaging capability comparable to that measured with PMTs. In this context, aim of the Dissertation is to develop an innovative multichannel spectrometer and imaging module based on scintillators and SiPMs which can be innovative for the two high-energy applications. The SiPM-based detection module has been designed with a modular structure to be compatible, except for the scintillator and the photodetector array, for both nuclear physics and PGI applications. The photodetector chosen for nuclear physics experiments, where monolithic and large (>1") $LaBr_3$:Ce scintillators are used, is composed by NUV-HD SiPM (FBK, Trento, Italy). This SiPM shows high PhotoDetection Efficiency (PDE > 45% at 380 nm), low Dark Count Rate (DCR < 100 kcps/mm^2) and small microcell size (30 x 30 um^2), allowing to improve the energy resolution and the dynamic range. The SiPM module consists of an array of 5 by 6 SiPMs, each one having an active area of 6 x 6 mm^2. The single array is used for the 1" x 1" crystal readout and it is assembled in a 2 x 2 format to read the 2"x 2" scintillator. On the other hand, the photodetector for PGI applications has a pixelated geometry, which is a possible solution to decrease the count rate/pixel requirements thanks to the detector segmentation. The photodetector for PGI applications is produced in RGB-HD SiPM technology (FBK, Trento, Italy) and designed to improve the linearity at high energies thanks to a smaller microcell size. The photodetector is composed by an array of 8 x 8 SiPMs: a single SiPM has an active area of 4 x 4 mm^2, 15 x 15 um^2 microcells, for a total of 70000 microcells/SiPM. Each SiPM is coupled 1:1 to a 4 x 4 x 22 mm^3 LYSO crystal.
In both photodetector geometries, signals are processed by two ASICs, improving the compactness of the system, and acquired by an external Data AcQuisition system (DAQ). Moreover, the module also implements a real-time stabilization of the SiPM gain to compensate for the gain drift with temperature. The ASIC chip used in this prototype is ANGUS, a 36-channel front-end ASIC in standard CMOS 0.35 um technology. Although it cannot provide a suitable dynamic range for a nuclear physics experiments up to tens of MeV because it was developed for low-energy applications, it has allowed to test the nuclear physics detection module up to 1.3 MeV. This has been enough to fulfil the primary objective of the work, i.e. the comparison between SiPM and PMT performances for $LaBr_3$:Ce readout. A current filter has been designed on the PGI detection module, which has allowed to decrease the input current of the chip to acquire up to 13 MeV of gamma-ray energy, with the drawback of the introduction of more variability on the readout channels. This modification can't be done for $LaBr_3$:Ce readout because of the high energy resolution target.

An experimental characterization of the two photodetector modules has been then performed in terms of spectroscopic performances and position sensitivity. The gamma-ray response linearity and energy resolution of the1" and 2" $LaBr_3$:Ce crystals readout in an energy range of 122 keV to 1.33 MeV is presented. The energy resolution achieved with the NUV-HD SiPMs shows a significant improvement over earlier works and nearly matches the resolution obtained using a super bialkali PMT (Hamamatsu R6233-100).
In fact, a 3.19 ± 0.01% FWHM at 662 keV was obtained at a room temperature of 30°C with SiPM readout, similar to the measured 3.07 ± 0.03% FWHM at 662keV achieved coupling the same crystal to the Hamamatsu R6233-100 PMT. For the 1" crystal an energy resolution of 3.70 ± 0.01% FWHM was measured with SiPMs at 662 keV, as slightly worse performance with respect to the 2" detector due to an inferior crystal. This assumption was confirmed by the 3.57 ± 0.04% FWHM energy resolution measured at 662 keV with the same scintillator coupled to the same Super Bialkali PMT. An analysis of the energy resolution terms based on the measurements performed communicates an important achievement in development of SiPMs for scintillator readout, demonstrating, for the first time in literature, the equivalence between SiPM and PMT for high-resolution spectroscopy. The experimental characterization of the 2" $LaBr_3$:Ce detection module in terms of position sensitivity is also performed, because it is an another fundamental requirements to correct the Doppler broadening in nuclear physics experiments. A spatial coherence has been successfully demonstrated and a preliminary 2 cm spatial resolution has been estimated with the centroid method for image reconstruction. Although at the time of writing it was not possible to perform beam measurements in a proton therapy facility, the reported experimental results demonstrate that the designed photodetection module is suitable also for future prompt gammas detection. Spectroscopic measurements between up to 1.3 MeV show a satisfactory energy resolution of 13%, 9% and 8% FWHM at 662 keV, 1173 keV and 1333 keV, respectively. Results are the same of the state-of-the-art PGI gamma camera, and it is correct to assume that they would allow an accurate energy calibration of the system, and, consequently, a correct setting of the two energy thresholds within reconstruct the prompt gammas profile. Moreover, an adequate position sensitivity necessary for profile reconstruction, coherent with the pixel dimension of 4 mm, has been proved. Finally, thanks to the use of a LED set-up which simulates the absorption of gamma-rays from 800 keV to 13 MeV, the module shows a proper response in the energy range of the application.
Future activities will include the test of the SiPM-based photodetection module in the energy range of nuclear physics experiments to make a compete comparison in terms of achievable performances and linearity with PMT-based readout. Moreover, a characterization in a proton therapy facility is necessary. In order to make successful these measurements, there are two aspects that need to be considered for future improvements of the prototype: the design of an ASIC for high dynamic range for nuclear physics experiments and the development of a fast electronics to face the high count rate required by the next-generation accelerators which are spreading in proton therapy facilities.

CONFLICT MITIGATION IN WATER RESOURCES SYSTEMS BY INFORMED DATA-DRIVEN CONTROL AND FINANCIAL HEDGING TOOLS

Simona Denaro - Supervisor: Prof. Andrea Castelletti

Conflicts over shared water resources at the national, international or sectoral level are common to many large river basins worldwide. Conflicts may arise over water quality requirements, water allocation in space and time, or conflicting operating management policies. Aggravating this context, accelerated population rise along with improved living standards have led to increased water demand, while urbanization and concentration of people exacerbated weather disaster vulnerability. Climate instability and uncertainty preclude reliable forecasts of extreme weather events, which leads to more conservative and polarized management practices. As a result, global freshwater quality and availability, as well as the space for negotiation, have been dramatically reduced.In today complex and fast changing world, last century traditional resource management strategies, characterized by "command-and-control" policies and structural large scale actions, prove inadequate to cope with uncertain socio-economic and climatic conditions. Uncertainty also complicates the evaluation process in planning new infrastructures, as past conditions no longer reflect future ones. A transition towards more flexible and integrated approaches has started to emerge in recent years. The focus is on increasing the overall productivity of water by working toward more efficient management policies, rather than seeking new sources of supply. Distributed and coordinated management, better informed and participated decision making and smart economics are all strategies that fall within this path. Following these premises, the main goal of this thesis was to contribute to novel cooperative management strategies able to mitigate conflict and improve system-wide efficiency and equity. Specifically, we focused on two areas of improvement: **better informed control** strategies and novel **financial hedging tools**.One important aspect we considered was power distribution, because it usually plays a significant role in shaping shared management policies in water systems with multiple water users. When the agents do not share equal power positions, a system is affected by power asymmetry. In this thesis we focused on geographic power asymmetry and its role in defining upstream/downstream water disputes. In these systems, conflicts resolution can be even more challenging, because the status quo is beneficial to the upstream part who may be reluctant to cooperate. In this context, we developed our work around three main **research questions**: (1) how relevant is the use of exogenous information at the system-wide scale in increasing the overall water system efficiency? (2) Can financial risk hedging strategies, particularly index-based insurance contracts, act as a tool to foster cooperation and facilitate negotiation among conflicting agents in a power asymmetrical context? (3) In case of extreme weather events and high uncertainty, how can insurance solutions still be viable and affordable to agents being affected by different risks? To address these questions, we explored novel approaches to foster a cooperative attitude in shared river basins, on two representative case studies in the Italian Lake District characterized by conflict in power asymmetric conditions, which is exacerbated under extreme weather conditions. The first strategy we evaluated, namely informed partially data-driven control (green area in Figure 1), addressed the disproportion between the amount of data at the catchment scale that is made available to decision makers by abundant different sources, and the limited set of information that water managers still rely on to drive their operating rules. In this phase, our research goal was to assess the actual improvement generated by expanding the information system the operation is conditioned upon to include additional hydro-meteorological variables. Specifically, we used raw observational data collected at the system-wide scale, to describe the enlarged state of the controlled system. This study also evaluates the operational value of coordination as it assumes information sharing between the upstream and the downstream agents. In this study, we adopted a recently developed methodology called Information Selection Assessment (ISA) which addresses information processing and assimilation into the system management. Results suggest a combination of seasonal snow data and information sharing can lead the overall system performance to a 20% improvement. The findings were useful to highlight the importance of establishing cooperative actions, in this case information sharing, among the different users in the catchment. Another relevant conclusion that we could draw was that control policies can benefit from data on snow dynamics which carries seasonal information and useful insights on future water availability in the catchment.Beside informed control, we investigated the potential for innovative economic and financial tools to foster cooperation in conflicting, power asymmetric context (blue area in Figure 1). We focused on financial instruments and specifically on index-based insurance contracts. This particular type of insurance bases indemnities on a weather variable, thus translating hydrological uncertainty into a financial one which can be hedged by purchasing the desired level of risk coverage. In the first phase, a state of the art review of financial tools applications to water management was performed. We addressed research question (2) by designing an index-based portfolio as a tool in the hand of a third institutional party wanting to redefine water rights in a power asymmetric water system. Findings indicate index-based insurance contracts can offer cheaper solutions for redefining water rights and consistently shifting unfair trade-offs in power asymmetric water systems. We also explored the issue of insurability of extreme weather events like catastrophic floods and droughts (research question (3)). This type of events characterized by low probability/high risk can become very difficult to insure in the absence of subsidies, reinsurance or a significant portfolio diversification. This latter approach is adopted in this thesis by designing unprecedented risk pooling strategies combining risks from different, competing users at the basin scale. The proposed approach proved beneficial to both promoting insurance premium affordability, by reducing premiums up to 7%, and facilitating collaboration schemes at the catchment scale. Overall, we think that the methods proposed and developed in this research, as well as its findings, provide important and novel insights for an effective and efficient mitigation of conflict in shared river basins. As a general take-home conclusion we can assert that fostering cooperation in conflicting multi-agents water systems is a good strategy to improve the overall system efficiency and equity. In this context, financial tools can be adopted in a variety of ways to provide bargaining solutions.

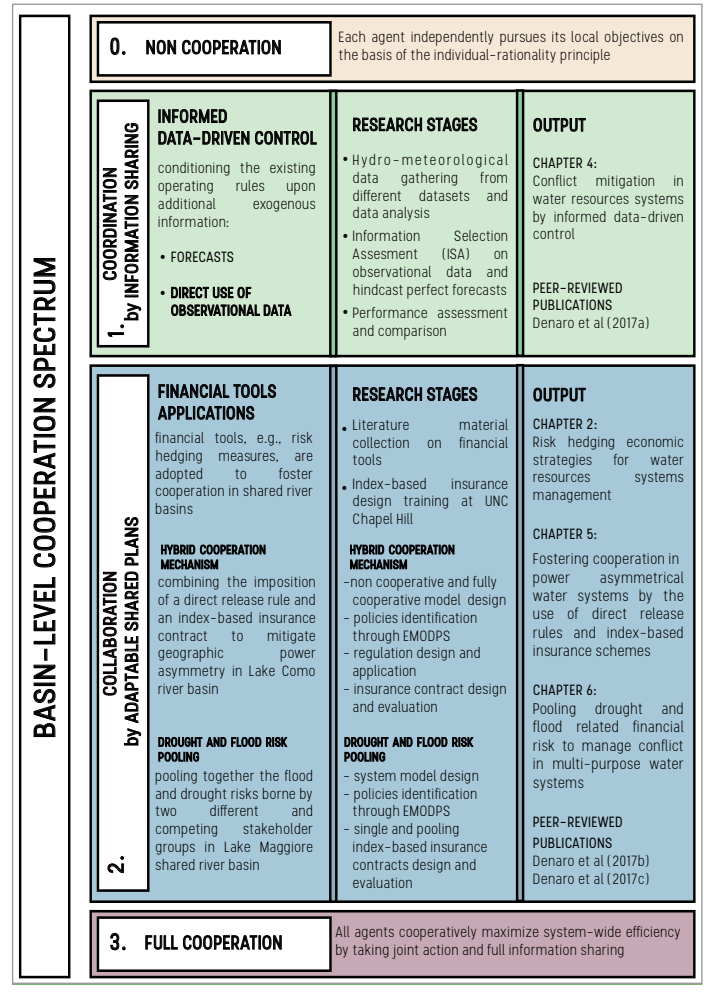

Fig. 1 - Thesis research framework.

PATROLLING ADVERSARIAL ENVIRONMENTS EXPLOITING AN ALARM SYSTEM

Giuseppe De Nittis - Supervisor: Prof. Nicola Gatti

Physical security is one of the most important challenges of our times. Due to the terrible events happened in the last decades all around the world, especially nowadays in Europe, novel techniques and methods are being developed to face new threats and dangers. But security means also helping people and saving lives, e.g., detecting and rescuing desperate migrants trying to cross the Mediterranean Sea.
Algorithmic Game Theory allows us to scientifically investigate these phenomena, modeling such interactions as mathematical problems and designing suitable algorithms to deal with these threats.
When patrolling large environments or infrastructures, a crucial issue is to guarantee some level of protection to each area without being able to constantly surveil them. A common countermeasure is the usage of cheap but wide-ranged sensors, able to detect malicious events that may occur.
In this thesis, we propose the first Security Game model with the presence of an alarm system able to trigger alarm signals, which carry the information about targets that can be under attack. Specifically, we focus on the exploitation of such information to improve the effectiveness of patrolling strategies.

The dissertation is structured in three parts, according to the research lines along which the contributions are developed.

Uncertainties of the Alarm System
First, we study the uncertainties that may affect the alarm system. We start considering the scenario in which the Defender can control a single patroller and the alarm system is affected by spatial uncertainty, i.e., the signal sent to the Defender communicates that something suspicious is happening in an area, without specifying the exact location. We divide the problem into two phases: signal response and patrolling. For the signal response phase, we provide a complexity analysis and design two exact algorithms and two approximation algorithms. We show that, without false positives and missed detections, the best patrolling strategy reduces to stay in a place, wait for a signal, and respond to it at best.
Then, we introduce a significant positive missed detection rate, i.e., no alarm signal is generated even though an attack is occurring. This scenario is in favor of the Attacker, who can exploit such flaw in the alarm system. This is why such new element puts the problem in a different perspective and requires a new approach to be solved. We deeply analyze security games in which the alarm system is both characterized by detection uncertainty and spatial imperfection, tackling the challenge of designing tractable algorithms for real-life scenarios. In particular, we show that standing still and waiting for a signal is no more the best response, and provide the Defender with the best patrolling strategy to move her resource. We prove that Markovian strategies are arbitrarily worse than optimal non-Markovian ones, and thus we resort to a deterministic approach.

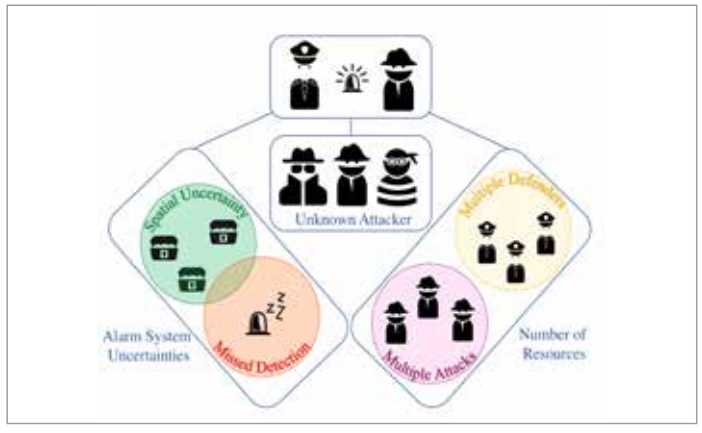

Fig. 1 - Research lines explored in the thesis.

A Coordinated Defense and Multiple Attacks
The second direction we investigate is on the dimension of the problem, namely, the number of resources both the Defender and the Attacker can control.
We study how the Defender should behave if she can control multiple resources. First, we tackle the problem of finding the minimum number of defending resources assuring non-null protection to every target, and then we study how the Defender should move them. The challenge is designing algorithms able to scale up with the number of resources. We prove that the problem of finding the best Defender's strategy, when the number of defensive resources is given, is hard. We also show that the problem of finding the minimum number of resources assuring non-null protection to every target cannot be approximated in polynomial time within a constant factor on arbitrary graphs. We design an algorithm to find the best strategy to respond to any alarm signal once an allocation of resources in the environment is given, according to different degrees of coordination among the resources, each described by an adversarial team game with different forms of strategies.
Then, we investigate the opportunities the Attacker can take when she can perform multiple attacks, simultaneously or sequentially. The challenge is due to the high interaction level among the players, e.g., the Attacker can use resources to make the patroller to move away from some valuable targets and, subsequently, she can attack those targets. When the number of resources is a fixed parameter, the problem admits an algorithm capable of finding the strategies on the equilibrium path, requiring polynomial time in the size of the graph. Conversely, we show that there is no algorithm requiring polynomial time in the number of Attacker's resources, even in the simplified case in which the Attacker uses all her resources simultaneously. Unfortunately, computing the equilibrium strategies requires the knowledge on the number of Attacker's resources. Since it is unlikely to have this information, we study the robustness of a Defender's strategy when the guess such number, evaluating the worst-case inefficiency of this strategy, showing that can be arbitrary even when the guess is a wrong estimate—both over and under—for just a single resource. Furthermore, we investigate the use of online algorithms to adopt when no information is available to the Defender. We provide a tight upper bound over the competitive factor when non-stochastic online algorithms are used, and we show that the factor can be improved by resorting to randomization.

Facing the Unknown
Finally, we introduce the notion of uncertainty in the type of the Attacker. We tackle the problem of facing an unknown adversary, whose profile is just known to be in a list of possible profiles. Here, a different approach is required: we learn the profile of the Attacker and exploit such information to prevent possible future attacks. We show that state-of-the-art bandit and expert algorithms suffer from a linear and logarithmic regret, respectively, in the length of the time horizon. Thus, we present two novel approaches, bridging together game-theoretical techniques and online learning tools. In the first approach, the Defender has a belief about the follower and updates it during the game, and we provide a finite-time analysis showing that the regret of the algorithm is constant in the length of the time horizon. In the second approach, the learning policy is driven by the estimated expected regret and is based on a backward induction procedure. We provide a thorough experimental evaluation in concrete security settings, comparing our algorithms with the main algorithms available in the state of the art of the online learning field and showing that our approaches provide a remarkable improvement in terms of expected pseudo-regret minimization.

COMPILER TECHNIQUES FOR BINARY ANALYSIS AND HARDENING

Alessandro Di Federico - Supervisor: Prof. Giovanni Agosta

Despite the growing popularity of interpreted or byte-compiled languages, C/C++ and other languages targeting native code are still dominantly used for system programming. Programs compiled to native code present a set of challenges compared to alternatives. In particular, in this work we focus on how they can be efficiently analyzed, how existing security measures (known as binary hardening techniques) perform, and how new ones can be introduced to secure features that have received little attention. We propose rev.ng a binary analysis framework based on QEMU, a popular dynamic binary translator and emulator, and LLVM, a mature and flexible compiler framework. rev.ng can easily handle a large number of architectures and features a set of analyses to recover basic blocks locations, function boundaries and prototypes in an architecture- and ABI-independent way. rev.ng can be used for instrumentation, debugging, decompilation, retrofitting of security features and many more purposes. Our prototype encompasses about 17 kSLOC of C++ code and has been publicly released under a Free Software license. The core component of rev.ng is revamb: a static binary translator which can accurately identify all the basic blocks, and, in particular, the targets of indirect jumps for switch statements. Along this work, we will make heavy use of analysis techniques popular in the compiler literature, such as Monotone Frameworks, to recover an accurate control-flow graph, identify function boundaries and the number and location of function arguments and return values. We will also discuss how rev.ng can handle native dynamic libraries, how it can be easily employed for instrumentation purposes, how it can be extended to handle even more architectures and how its performance compares to tools with analogous purposes such as QEMU, Valgrind, Pin and angr. We also study two often overlooked features of C/C++ programs: variadic functions and the RELRO link-time protection mechanism. We propose HexVASAN, a sanitizer for variadic functions to ensure that the number and type of arguments used by the variadic function match those passed by the caller, and leakless, an exploitation technique to bypass the RELRO protection in its several forms.

MODELLING HARMONIC AND RHYTHMIC COMPLEXITY FOR APPLICATIONS IN MUSIC INFORMATION RETRIEVAL

Bruno Di Giorgi - Supervisor: Prof. Augusto Sarti
Co-Supervisor: Dr. Massimiliano Zanoni

What exactly defines "complexity"? It can be argued that obeying or violating the common patterns and expectations will affect the complexity of a work of art. Music is no exception, with its many languages that unfold through time, such as harmony, rhythm, melody, orchestration, timbre.

In this work we focus specifically on harmony and rhythm and analyze some of the characteristics that influence complexity.
In doing so, we first study the relevant music descriptors, such as chords, keys and beats, proposing new models to automatically extract these properties from the audio signal.
Successively, we propose data-driven and model-based methods for estimating complexity from symbolic representations of harmony and rhythm.
Signal processing, machine learning techniques and musical theory are used throughout this work to achieve such goals.

The main contributions of the thesis are subdivided into two parts: the first addresses harmony and contains our works on chord and key extraction as well as the estimation of harmonic complexity; the second part is devoted to rhythm analysis and includes our works on the estimation of beat instants and rhythmic complexity.

As far as harmony is concerned, we begin by describing our chord and key recognition system. We focus on one particular aspect of Western pop and rock music that is arguably overlooked in the related literature: the diatonic modes. Two modes in particular, called Dorian and Mixolydian, complement and connect the two well-known opposite poles, the major (Ionian) and minor (Aeolian) modes. We incorporate these modes and provide a novel, musically meaningful parameterization of a known dynamic Bayesian Network approach. These variations increase the accuracy of the system, as shown by the results.

We then move forward and analyze how the expectations formed when listening to tonal chord sequences can influence perceived harmonic complexity. We model expectations by training three different language models, i.e. prediction by partial matching, hidden Markov model and recurrent neural networks, on a novel large dataset containing half a million annotated chord sequences. We then train a compound model and use it to generate a set of chord sequences that we included a listening test. Results show a strong relation between negative log probability of the chord sequences, given by our language model, and the subjects' complexity ratings.

As far as the rhythmic information is concerned, we focus on beats, which are often localized given an onset detection function extracted from audio signal, as well as the tempo path that is estimated from it. The tracking strategy required for estimating the sequence of beat instants from such descriptors is usually based on dynamic programming algorithms. We propose a novel strategy based on an efficient generation and joint steering of multiple simple trackers. Although the method performs an heuristic search as opposed to the full search of the dynamic programming approach, this solution is shown to lead to improved computational efficiency. The method is also compared with a broader set of state-of-the-art solutions, in order to offer a more general analysis.

Finally, we review some of the models for estimating rhythmic complexity from symbolic representations. We focus on the class of rhythms with unusual time signature, which are common in some Western genres and non-Western musical cultures. We propose a novel model, generalizing concepts as beat induction, almost maximal evenness and weighted metrical hierarchy to such class of rhythms. In order to validate our model, we conducted a test where subjects were asked to tap along with the rhythms, while their performance were recorded and measured. Early results from the performance test show that our model estimates rhythmic complexity more accurately than the other models found in the literature.

The overall contribution of the thesis is composed of models for the estimation of musically relevant information, such as chord, key and beat, from the audio signal; as well as models for the estimation of perceived harmonic and rhythmic complexity, given the symbolic representation of related musical elements, such as chord and onset sequences.

Accessory contributions include the release of the annotations of chords, keys and beats created for our experiments. The annotations have received support from other researchers and have been used for international contests such as the Music Information Retrieval Evaluation eXchange (MIREX) contest. Also, the source code of some of our models have been shared with related publications or through public repositories.

DISTRIBUTED DECISION MAKING WITH APPLICATION TO ENERGY SYSTEMS

Alessandro Falsone - Supervisor: Prof. Maria Prandini

This thesis addresses decision making problems in systems composed of multiple agents with computation capabilities, which can communicate with their neighbors and aim at optimizing performance, subject to some coupling in their decisions. Our focus is on problems that can be formulated as a mathematical program involving decision variables, objective functions, and constraints of the agents. Centralized solutions of the global optimization program pose several difficulties, ranging from being computationally intensive to forcing agents to disclose their local information to some central unit. This calls for distributed resolution strategies exploiting the agents computation and communication capabilities. Throughout the thesis, we focus on different global optimization problems, which are all characterized by a decomposable structure, but present a feature that couples the agents decisions. More precisely, we consider structures characterized by an objective function that is the sum of the local objective functions, with the agents decisions possibly subject to local constraints, and a coupling element given by either a common decision variable or some network-wide constraint. For each global problem, a distributed algorithm is introduced, which consists of an iterative procedure where agents solve a local optimization problem, share their decisions with their neighbors, and repeat the process based on the received information. The algorithms are introduced first for the deterministic case, and then extended to account for the presence of uncertainty affecting local objective functions and/or constraints, as it is typically the case when agents have only partial knowledge on the environment. We also consider the case of discrete decision variables and address the issue of the exponential increase of the problem complexity via its decomposition in sub-problems with a smaller number of decision variables.

We provide a unifying framework to cope with the main complexity features of cooperative multi-agent decision making problems: heterogeneity, as we allow the agents to have different performance criteria and constraints; privacy, as we do not require agents to disclose their local information; uncertainty, as we take into account uncertainty affecting the agents locally and/or globally; and combinatorial complexity, as we address the case of discrete decision variables.

Chapter 1

We introduce the problem of decision making in multi-agent systems motivating the interest in this research area through some relevant applications in the energy system domain. We then describe the thesis structure and synthesize its contribution.

Chapter 2

We propose a new distributed algorithm for solving convex global optimization programs over time-varying networks, where the agents have to agree on a common decision vector. In contrast with other approaches in the literature, we are able to handle the case when agents have different local constraint sets and we relax some regularity assumptions regarding their local cost functions. The proposed procedure is tested on the power allocation problem in cellular networks, where the agents are the base stations serving mobile users within a cell, the common decision vector is the power to be assigned to the mobile users, and the base stations aim at maximizing the throughput of the entire network while accounting for local power limitations and minimizing the interferences among users sharing the same transmission channel. Simulations show that the proposed approach avoids oscillatory behaviors of the tentative solution across iterations, in contrast with an alternative approach in the literature.

Chapter 3

We exploit dual decomposition together with the algorithm developed in Chapter 2 for the solution of convex programs, where the agents local decisions are coupled via a network-wide constraint. In contrast with other approaches in the literature, we achieve optimality of both the dual and the original decision variables. Our approach requires a lower amount of information to be exchanged and no prior coordination among the agents regarding the computation of a specific strictly feasible candidate solution. The proposed algorithm is tested on the plug-in electric vehicles charging problem, where the vehicles are the agents, the coupling constraint is given by the network power limits, and the objective is to set the overnight charging schedule of each vehicle so as to ensure a desired battery state of charge for the morning after while minimizing the cost of the electricity. An interesting feature of the approach is that feasibility of the solution is preserved when some permanent failure of the communication network occurs, causing the communication graph not to be connected anymore.

Chapter 4

We consider the linear case of the set-up of Chapter 3, i.e., separable objective function with local decision variables and a coupling constraint, but allow some decision variables to take integer values. This results in a (non-convex) mixed integer linear program which involves the whole network. Inspired by a recent work in the literature, we propose a decentralized algorithm that is able to compute, in a finite number of iterations, a solution that is feasible for the global problem, with precise guarantees regarding its suboptimality level. Decentralized schemes require the presence of a central unit, which however is typically required to perform some simple computation and coordination tasks only. In our set-up the central unit is in charge of gathering some information about the agents contribution to the coupling constraint and updating the associated dual variable. The proposed approach is tested on an integer version of the plug-in electric vehicles charging problem, where the vehicles can charge and discharge their battery, but they cannot set the charging rate, which is fixed. At the end of the chapter we suggest a possible way to derive a distributed version of the algorithm.

Chapter 5

We revisit the problems introduced in Chapters 2 and 3 treating the case in which the agents are affected by uncertainty. We develop a sample-based approach to decision making that guarantees convergence to a solution that is feasible for almost all uncertainty instances except for a predefined fraction. The key novel feature of our approach is that agents use a finite number of uncertainty realizations, which can be available locally and do not need to be exchanged with the other agents, thus preserving the private nature of the agents local dataset and avoiding to overload the communication links. Compared to other approaches in the literature, we are able to provide guarantees for a finite number of samples, as opposed to asymptotic guarantees in the number of samples; we do not need any description of the uncertainty, as we rely only on samples; and we are able to take into account correlation of the uncertainty affecting the overall multi-agent system using only the agents local information. The developed techniques are tested on the energy management problem of a building cooling district. All buildings in the district are connected to a common chilled water circuit through which they can exchange cooling energy. Each building has to guarantee a certain thermal comfort for its occupants and, to this end, it can decide to produce cooling energy with its own chiller plant or to draw energy from the cooling network. The goal consists in minimizing the overall electrical energy cost of the district while accounting for the uncertainty in each building cooling energy demand, which is characterized through a (private) dataset of past cooling energy demand profiles per building.

Chapter 6

We draw some conclusion, illustrate future research directions, and briefly describe other interesting results that we developed and are related in terms of methodology or application to the thesis, but do not really fit its main topic.

EXPLOITING PUBLIC WEB CONTENT TO ENHANCE ENVIRONMENTAL MONITORING

Roman Fedorov - Supervisor: Prof. Piero Fraternali

The quantity of public content, generated by users or by sensors, available on the web nowadays is reaching unprecedented volumes. This massive collection of data contains an enormous amount of latent knowledge, which can be used for a variety of purposes, such as event detection and predictive modeling.
The goal of this thesis is to explore the methods for extracting such knowledge and building useful applications using low-cost, publicly available, multimedia web content, with reference to the field of environmental monitoring, which often suffers from the lack of significant and exhaustive input data. The thesis aims to answer this question by illustrating a use case in which the problem has been successfully addressed. This includes the complete path from retrieving and analyzing the data up to using it for environmental purposes and proving its usefulness.
Given a massive amount of unstructured and non-authoritative visual content publicly available on the social web, we devise techniques for the automatic analysis of such content, so as to extract environmentally relevant spatio-temporal data and objectively prove the utility of such data.
We specifically focus on monitoring snow cover in mountainous regions, that is, the spatial extent of terrain surface covered by snow. Snow processes are traditionally observed by means of ground measurements stations, which can either be manned or fully automated. In both cases, measurements are accurate and capture different aspects, including the snow depth and density (possibly at different altitudes). However, the number of ground measurement stations is limited, thus enabling only a sparse sampling of the snow cover over large areas. Moreover, the high variability of snow processes, which depend on temperature, elevation, exposure, slope, winds, etc., is such that it is difficult to extrapolate snow depth and density at different locations.
We chose public photographs taken in mountain regions as the use case of visual content. A large fraction of them contains the skyline defined by mountain peaks, slopes, ridges, crests, etc., both as main subject and as background. Such photographs implicitly contain precious information related to snow cover phenomena, which can complement the traditional measurements and has not been fully exploited so far. We argue that social web content might represent an additional source that can complement and enrich the traditional ones, due to its unique characteristics in terms of spatio-temporal coverage resolution and cost.
We describe an acquisition pipeline that continuously retrieves new images containing mountain slopes from photo-sharing platforms and public webcams; a set of algorithms for mountain image geo-registration (Figure 1) that, given a geolocated photograph as input, infer photograph landscape skyline (Figure 2) and its geographical properties; a novel approach for snow/non-snow pixel level classification and propose several virtual snow cover measures.

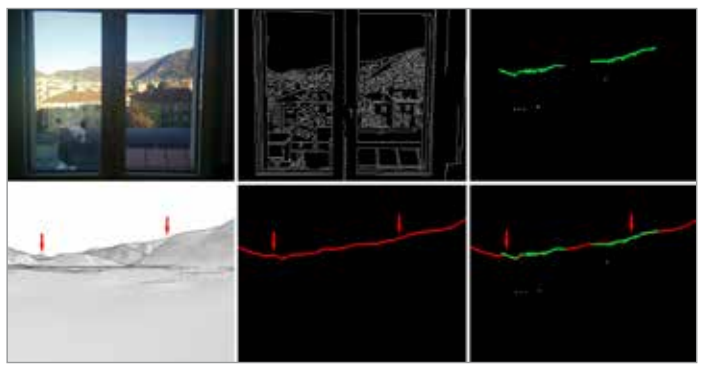

Fig. 1 - Example of the peak identification in presence of many noise edges

Furthermore, we present a supervised learning data-driven water management model that, among other inputs, relies on the authoritative snow measurements. We define the performance metric of the model and test how the performance varies if we complement or replace the authoritative snow measurements with the virtual snow indexes computed by the proposed system. We argue that the data acquired from one single touristic webcam is able to replace the authoritative snow measurements without a performance drop and to complement the authoritative snow measurements, improving the performance.
Finally, we discuss the potential of the crowdsourcing in systems that exploit unstructured content for environmental monitoring with two use cases: a web portal that allows users to explore the acquired photographs, contribute their own content and help the geo-registration process by correcting the errors made by automated tools, and a real-time augmented reality mobile application that identifies mountain peaks, engaging the users to contribute with their photographs (Figure 3).

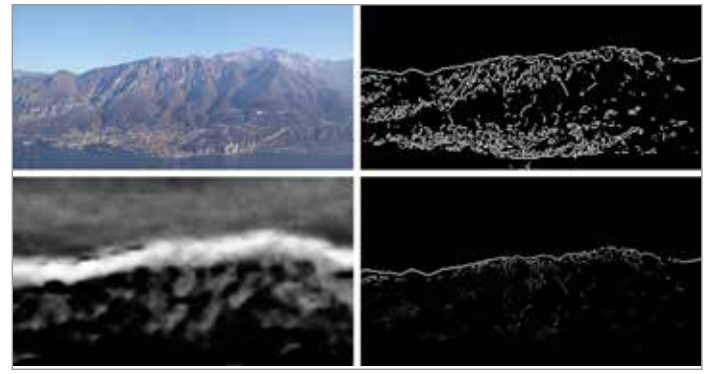

Fig. 2 - Example of a landscape skyline identification.

Fig. 3 - A screenshot of the mountain identification mobile application.

DISCREPANCY ANALYSIS: A METHODOLOGY FOR AUTOMATED BUG DETECTION IN HARDWARE DESIGNS GENERATED WITH HIGH-LEVEL SYNTHESIS

Pietro Fezzardi - Supervisor: Prof. Fabrizio Ferrandi

Field Programmable Gate Arrays (FPGA) are steadily becoming more appealing in computing. They provide reconfigurability and flexibility like software, while guaranteeing low power consumption and massive parallelism close to what is possible with Application Specific Integrated Circuits. This is very promising in the current general struggle to find new paradigms that can cope with the end of Moore's Law. For these reasons, FPGAs are not under investigation just for prototyping, but they are increasingly used in datacenters, High-Performance Computing and irregular parallel applications. One of the main obstacles to the adoption of such devices is that the skillset necessary to program them effectively is very broad. Digital circuits designers are harder to find than software engineers, and the development of a dedicated digital circuit to execute a given task usually requires considerably more time than a software implementation of the same functionality. FPGAs are traditionally programmed using so called Hardware Description Languages (HDL), that allow designers to describe every component of the digital electronics that compose the system. For their nature, HDLs are very tightly related to the underlying devices and electronics, forcing designers to focus at the same time on the low-level details of the electronics and on the high-level algorithmic level of their implementations. This close relationship with the underlying hardware also means that the same HDL design cannot be easily ported from an FPGA device to another, without significant modifications. These three components – scarceness of skilled engineers, languages that place multiple heavy responsibilities on the shoulders of designers, and lack of portability of hardware designs – clearly represent a scalability problem that has to be overcome to enable FPGA computing to really go mainstream.

In recent years, a promising approach in this field that has received much attention is High-Level Synthesis (HLS). The main insight behind High-Level Synthesis is that the key to removing all the obstacles to FPGA design is to use a high-level software programming language, instead of HDL, as a starting point for hardware synthesis. With this approach software engineers could start programming FPGAs, ignoring the details of the underlying electronic designs to focus on algorithms, while at the same time creating designs that are easily portable on new devices. Today, many different academic and commercial tools are able to generate HDL design starting from a variety of programming languages: C, C++, Java, Python, Haskell, Erlang, and many others. However, the most common and well supported by FPGA vendors are C and C++. These languages are at an higher level compared to HDL and allow designers to write portable code, while still giving programmers the capability to control low-level details that might be relevant in hardware. Moreover, C and C++ are the basis of a large number of standard libraries and programming language extension for HPC and multithreading, like POSIX threads (pthreads), OpenMP, OpenCL, and CUDA. Commercial HLS tools and academic projects often provide support for High-Level Synthesis starting from these multi-threaded language extensions, providing programmers with familiar tools to easily exploit the massive physical parallelism available on FPGAs. These solutions are also seen as a promising trend to provide well-known programming paradigms for FPGAs in the cloud.

However, for a programming paradigm to succeed, there is more to be taken care of besides the programming language. One of the fundamental aspects that will determine the success or the failure of High-Level Synthesis in the long term is the ecosystem of development and verification tools that will be built around it. Indeed, in hardware design, testing and verification typically constitute a significant portion of the whole effort for a project. Another important factor is the support for the integration of multiple components in System-on-Chip (SoC) design. These components can be either generated with HLS, hand-written by designers, or provided as Intellectual Property (IP) blocks by third-parties. According to ITRS prediction future SoC architectures will be characterized by heavy reuse (more than 90% by 2020) of IP blocks for reducing design cost and time-to-market. To increase productivity and tackle design complexity, designers will need to raise the abstraction level and use Electronic System Level (ESL) methodologies based on High-Level Synthesis to automatically generate and integrate the IP descriptions in a suitable HDL design. This will result in a proliferation of IP vendors specialized in the optimization of specific functionalities, while system designers will focus on the integration of the different components, posing new threats for the design and verification of complex architectures. At the same time HLS compilers are growing in complexity, adding more optimizations passes to generate more efficient accelerators in term of frequency, area, or power consumption. This complexity is hidden to users and managed by the tools, but it can become a real burden during testing, debugging and verification. Given that in SoC design up to more than 50% of the overall time can be spent on verification, the risk is that the speedup HLS gives to development could be negatively compensated by a slowdown in testing and debugging. HDL generated by HLS is not intended to be human-friendly, because that is not the purpose of HLS. This may become a problem if the HLS users are software engineers with little previous exposure to HDL and to hardware design. In order to avoid this risk, or even to improve the testing and debugging experience as much as the development phase has been improved, it is critical for High-Level Synthesis tools to integrate techniques and workflows to also manage verification. In particular, the tools need to keep track of the additional complexity introduced and managed by HLS during the design stage, in order to be able to reason about it later during verification, helping users to unravel the details of what went wrong in case of bugs. In this way, it is possible to extend the support that these tools give to the designers beyond implementation phase, up to the testing, debugging and verification steps.

This thesis describes the definition, implementation, and evaluation of a methodology for automated bug detection, called Discrepancy Analysis, targeted at hardware designs generated with High-Level Synthesis. Discrepancy Analysis is based on a notion of equivalence between the execution of the hardware generated with High-Level Synthesis and the execution of the software obtained from the original high-level source code used to generate that hardware. Using this notion of equivalence, the thesis describes how to compare automatically the two executions, and how to detect and isolate the first mismatch if present. All these operations are executed without human interaction, relieving users from the time-consuming and error-prone tasks to select the necessary signals for debugging, analyzing the signal traces to identify the malfunction, and backtracking it to the original high-level source code. The methodology is tightly integrated with the High-Level Synthesis process. As a consequence, it supports all compiler optimizations available during High-Level Synthesis. This coupling with the High-Level Synthesis tool also allows to automatically select in the generated designs the signals necessary for automated bug detection. Despite the tight coupling with the High-Level Synthesis tool, the discussion is kept as general as possible and only relies on common features that are present in all the known commercial and academic tools. The thesis also describes two extensions of Discrepancy Analysis: one to support automated bug detection in hardware generated with High-Level Synthesis of multithreaded code; one to support automated bug detection on pointers and memory accesses.Two bug detection flows based on Discrepancy Analysis are presented. The first is based on simulation of the hardware at the Register Transfer Level and performs the automated bug detection process offline after execution. The second flow is for on-chip bug detection. The generated hardware is instrumented with dedicated checker components, that analyze the execution on the fly, halting the circuit if a mismatch occurs and notifying it to users. Both the debug flows have been implemented and tested with BAMBU , an open source research framework for High-Level Synthesis developed at Politecnico di Milano.The results have been evaluated in terms of performance, coverage, and other advantages brought to the overall debugging experience, like the considerable reduction of the size of the waveforms files that can be achieved with a heuristic for automated signal selection. This evaluation showed Discrepancy Analysis to be fast, accurate, and effective in identifying several different classes of bugs, coming from the original high-level code, from external libraries of components, and even subtle bugs injected by the High-Level Synthesis tool itself. A thorough and extensive analysis of these classes of bugs has been carried on, both on the baseline version and on the presented extensions for multithreaded code, for pointers, and for on-chip debugging. The technique used to compress the execution traces for On-Chip Discrepancy Analysis, based on Efficient Path Profiling, also showed reductions of the memory consumption necessary for on-chip debugging up to 95% compared to previous state-of-the-art.

EMPOWERING INTERACTIVE TECHNOLOGIES FOR CHILDREN WITH NEURO-DEVELOPMENTAL DISORDER AND THEIR CAREGIVERS

Mirko Gelsomini - Supervisor: Prof. Franca Garzotto

Neuro-Developmental Disorder (NDD) is an umbrella term for a group of disorders arising during the developmental period that is characterized by severe and often co-occurring deficits in the cognitive, social, communicative, motor, behavioral and emotional spheres.
In this regard, the use of digital interactive technology is considered as a promising approach that does not replace current therapies but can be incorporated into them and support care-givers in their daily routine.
Disability nowadays no longer means a condition, an incapacity, or lack that belongs to a body, but rather a product of the interactions between self, society, body, and the variety of interactions from political economies to personal and socio-economical commitments that they engender. From a socio-economical point of view, the rapid growth of these technological resources requires a carefully planned response by health institutions and their administrators. In addition to relevant staff development and training, this response urges to address the cost of technological resources applicable; the individuals' unique needs with which those resources will be most useful; the benefits and risks of using them; the methods for preparing care-givers (clinicians and families) to accept, adopt and appropriate these resources; and an organizational commitment to empirically evaluating their effectiveness and utility.
This work addresses the previous issues by developing the "Empowerment Framework" in order to offer individuals in the care-giving sector the technological tools, services, abilities, and opportunities to feel empowered in using innovative interactive technologies, adapting them to meet their needs in a process called "appropriation", and share their acquired knowledge, original technological approaches, and evidence-based results through a knowledge-based system.
My contributions lie within the "Empowerment Framework", a framework that offers a space so that health institutions, administrators, clinicians, educators, parents and eventually external stakeholders can contribute in it. Using the framework, they will get access to:
a broad set of evidence-based design guidelines, carefully organized by interaction paradigm and category, informed by the literature and by the experience gained through the years;
1) a portable, accessible, modular and scalable software platform that will enable them to collaborate with partners in designing, developing, configuring, evaluating and sharing the activities through the web, building a new model of social crowd-sourcing;
2) a psycho-pedagogical model of intervention for teachers, parents and novice care-givers; co-developed with specialists and informed by macro-psychological theories and many tech-based studies;
3) adoption and sustainability guidelines to provide administrators and care-givers with the information they need to integrate and expand the use of technologies in their contexts.

Over the last decades, through many policies and efforts, public awareness of NDD has become widespread, especially for children in the Autism Spectrum Disorder. NDD includes a wide umbrella of disabilities and unique abilities. This highly prevalent condition presents itself in a wide variety of ways, engendering interest from a vast array of professional care-givers, researchers, and other stakeholders. Thus, people interested in NDD and those with specific backgrounds in medicine, biology, genetics, psychology, education, physiology, speech, language, and technologies alike have become interested in the use of technologies to support these individuals.
The media attention to children with NDD, their teachers and families has fueled what some may consider to be a tech bubble in the NDD space-myriad applications and devices, many of which have not been rigorously tested. Therefore, more and more researchers are working to find evidence of the efficacy of these technologies though, in an environment in which families and educators may be desperate to try new therapies, before their early intervention support runs out, waiting for research evidence may not be desirable. Thus, there is currently a great opportunity to build on early research and commercial successes, contribute to evidence-based practices, and develop holistic empirically grounded interventions.
This thesis is my echo to this opportunity, an attempt to build a framework to sustain the entire community to do their best, eventually translating our science into real-world applications: involving both tech design, development and transfer of tools, interventions and policies that apply what is known in the scientific space to what is needed in the clinical, educational and familiar space. The Empowerment Framework is thus my firm response to the many research questions that raised during my PhD Career.
Concluding, this work clarifies that technology has a significant role to play in supporting and empowering care-givers and children with neuro-developmental disorder.
Technology-based solutions have the potential to lighten the burden that falls on care-givers by helping them to coordinate and apply the demanding tasks and the complex networks of relationships involved with caring for and educating others, collecting empirical data.
But technology-based products are not likely, by themselves, to be a complete solution. Other social support systems are needed to ensure adoption, appropriation, and diffusion.
To catalyze greater innovation adoption, a number of things need to happen which are included in my agenda: to generate better information about the realities of care-giving; to develop business cases that identify what kinds of returns on investment new technology-based applications can provide; to evaluate and gain empirical effectiveness of these tools for children and users in general, respecting ethical and legal requirements.
Researchers can help deepen the understanding of care-givers' and children's needs, while foundations, employers and public agencies can provide the support needed to carry out this re-search.
Once this knowledge is available, I am confident that entrepreneurs and application developers will step forward to devise new ways of meeting these needs.
With this work I began to lay the groundwork for a technological solution informed by the initial studies and state of art. My aim is that this step will result in care-givers and children benefiting from the full potential of technologies to ease their learning effort, improve their health and empower their lives.

Fig. 1 - Child playing with Teo, the robot

Fig. 2 - Children playing in the Magic Room

Fig. 3 - Child playing with Polipo, a smart toy

A METHODOLOGICAL APPROACH TO DATA DESIGN, ANALYSIS AND EXPLOITATION

Angela Geronazzo - Supervisor: Prof. Cristiana Bolchini

During the last decades, improving energy efficiency in buildings has risen as an important societal issue and research area, motivated by a pressing quest to design, develop and implement effective and affordable energy demand reduction strategies. The ultimate goal is to optimize the trade-off between energy consumption and the occupants' comfort, aiming at reducing the high-energy demand and carbon footprint without compromising the occupants' quality of life. In this perspective, current research trends focus on Information and Communication Technologies (ICT) to enhance energy efficiency in buildings and communities.
Buildings, especially existing, non-residential ones, and occupants are the two key elements of the stated problem, to be opportunely monitored, dynamically controlled and made aware in order to pursue the desired optimization goal. In this context, ICT supports the following challenges: it empowers the adoption of Building Management Systems (BMSs) and Energy Management Systems (EMSs), equipped with sensors and actuators, with the mission to optimize the available resources and operational goals thus improving occupants comfort while saving energy; it enables the integration between building systems, e.g., lighting, ventilation, Heating Ventilation And Cooling (HVAC) systems; it supports the collection and storage of real-time information, e.g., indoor conditions, energy consumption, as well as the modeling of information about the building and its systems, e.g., constructive features, sensor positions, system technical details; and it contributes to the definition of an integrated infrastructure useful to extract from the available heterogeneous, multi-facet data the information needed to provide an overarching picture of the building ecosystem thus empowering dynamic operation strategies. Nonetheless, it introduces the need to master the heterogeneous, multi-facet data collected from the building ecosystem and to effectively exploit it. This challenge requires to address data quality issues, by designing ad-hoc processing approaches and to develop effective analysis methods able to support the discovery of useful knowledge and to enhance results interpretation. Nonetheless, data harvested from the building ecosystem represents a unique opportunity to develop new methods to address data quality issues, such as timeliness and consistency, and to design new analysis approaches, oriented to solve challenges concerning energy efficiency, occupants' comfort assessment and improvement, building control and fault detection.
In this context, there is an evident quest to define a methodological framework that identifies the main choices that are available in designing a monitoring and exploitation process, possible alternatives for each identified choice and their expected effects on the collected data. These choices need to be evaluated with respect to the goals of the monitoring activity. However, excepts for the broad guidelines provided by ASHRAE, there is no recommendations to support the design of monitoring campaigns and the process is typically planned and deployed on the basis of experience of domain experts and/or off-the-shelf kits, thus making each building a stand-alone, unique case. Hence, case studies in literature adopt customized solutions and the description usually focus on innovative aspects rather than discussing the monitoring solution, which, however, has an impact on the data exploration process and its results.
The main contribution of this thesis is the introduction of the Building Monitoring and Exploitation (BuildME) framework for the design of monitoring and exploitation processes in building science, comprising a wide set of applications ranging from energy audit, to comfort analysis and models calibration. The BuildME framework aims at leading the expert through these steps by distilling the relevant aspects that needs to be tackled and by providing a method to compare different case studies possibly pursuing a similar goal. The methodological framework identifies a set of so-called dimensions, e.g., goal, sampling strategy, aggregation policy, missing values and outliers handling, useful to define a monitoring solution with respect to a specific application scenario. For each of these dimensions, various alternatives are outlined. The framework was designed in a flexible and extensible way allowing one to introduce or remove both new dimensions and/or additional values. It is worth mentioning that dimensions are not orthogonal and influence each other, and an accentuate analysis needs to be carried out in order to tune each dimension. The framework systematically considers the relevant aspects to be analyzed in order to design an effective and efficient monitoring and exploitation process. Moreover, it supports the comparison among different existing solutions reported in literature, in order to determine what elements actually affect the final results and what solutions could be reasonably applied to other case studies pursuing the same goals.
The BuildME framework was defined in the context of a field study, that provides an experimental setup, a set of requirements for the monitoring campaign and multiple exploitation goals, with reference to the mission of the EMS and occupants thermal comfort assessment. The case study supports the definition of the methodology and offers the opportunity to define and investigate dimensions and their values by providing an experimental setup and a set of goals to pursue. Case study B and C are discussed to investigate the framework applicability to complex case studies and to assess its extensibility with reference to a wider range of goals and requirements. The study merges two different research projects, namely Smart Campus as Urban Open Labs (SCUOLA) and Zero Energy Buildings in Smart Urban Districts (EEB). The first focuses on the efficient management of a university campus in a Smart Grid scenario, the second oriented to occupants' thermal comfort and behavior with reference to building performance and refurbishment approaches. The monitoring campaign planned by domain experts emphasizes the lack of a methodological approach to monitor and exploit building data, thus leading to the design of the presented BuildME framework. The case study focuses on occupants' thermal comfort analysis and supports the comparison of dimension values and approaches in real-time and non real-time scenarios.
A further contribution to the definition process spawned from field studies reported in literature that were analyzed to compare monitoring strategies and analysis approaches tailored on different goals including energy audit, models calibration, data visualization and building dynamic control.
Furthermore, the framework were applied to two case studies, focused respectively on passive buildings performance and heat loss analysis, to discuss its flexibility in dealing with different goals and complex real scenarios.
Future work comprises the application of the framework to new case studies, to include a wider set of characteristics related to the scenario, such as intended usages, building features, and pursuing a broader group of goals, such as occupants' behavior and model calibration. Moreover, it would be beneficial a deeper comparison of the BuildME framework with international standards, e.g., ASHRAE guidelines and procedures, and certifications, e.g., LEED v4.
Nonetheless, the introduction of more rigorous monitoring strategies would enhance the quality of the harvested data thus enabling a more effective data exploitation towards the improvement of building operational strategies and the optimization of resource usage.

DESIGN OF ULTRA-FAST FRONT-END ELECTRONICS FOR NEW PIXEL DETEC-TORS FOR THE EUROPEAN XFEL

Andrea Grande – Supervisor Prof. Carlo Fiorini

The recent development of Free Electron Lasers based on the Self-Amplified Spontaneous Emission technique has allowed a considerable reduction of the emitted wavelength. SASE FELs sources can provide coherent light pulses in the X and γ-regime with remarkable brilliance and short pulse duration. These short wavelengths are useful in a wide range of applications, from physical studies to medical diagnostic to biology.
The European X-ray Free Electron Laser (European XFEL) facility is the latest FEL based on the SASE process and is currently under construction in the Hamburg area, Germany. The laser is going to be ready for the first lasing operation at the end of the 2017. It will cover an energy range from 206.6 eV keV, relating to a wavelength span from 0.1 nm to 6 nm. Thanks to its super-conducting linear accelerator, European XFEL will be able to provide the highest pulses rate ever obtained, with X-ray pulses only 220 ns apart and grouped in bunches of 2700 flashes each. The Full Width at Half Maximum (FWHM) of the pulses will be less than 100 fs, hence providing an extremely high peak brilliance.
Due to the high energy range and the wide range of application, several imagers are under development for the X-ray photon detections, including single-point, 1-D and 2-D detectors. Among the 2-D imagers, the DEPFET Sensor with Signal Compression (DSSC) project, thanks to its innovative single pixel structure and its low noise architecture, aims to provide wide energy coverage together with the single photon detection capability. To achieve this result, a novel DEPFET structure is under development by the DSSC consortium. The detector has a special property, which is the signal compression. In the DEPFET Sensor with Signal Compression, the internal gate is extended beyond the channel into the source region. In this way, in the case of small signals, the charge released in the substrate is stored exactly under the channel, while for large signals, the charge is pushed to the region under the source. Therefore, the sensor has a non-linear characteristic capable to guarantee both the single photon resolution and a large dynamic range at the same time.
The target of the project is the design of a 1Mpixels camera (1024×1024) composed by 256 ASICs of 64x64 pixels. Each pixel is bump bonded to a dedicated read out channel. Each channel will contain a fast, low noise analog front-end, an 8-bit analog-to-digital converter and a static RAM to store the information. Data will be sent out from the focal plane during the long gaps between subsequent macro bunches.
In spite of its capability in high dynamic range and low noise performance, the manufacturing of the DEPFET sensor needs a sophisticated processing technology which requires relatively long manufacturing time. As a backup solution, the Mini-SDD approach was proposed. Although, the Mini-SDD does not exhibit the same performance of DEPFET, it has the advantage of being available in a shorter time, therefore allowing first beam tests in the early stage of the experiments. The alternative sensor is made of mini Silicon Drift Detector (mini-SDD) and the compression behavior is obtained at front-end level on the readout ASIC and not by the silicon sensor, as in the DEPFET.
In the first version of Mini-SDD solution, the DEPFET has been removed and replaced still by a PMOSFET transistor, integrated in the ASIC chip and realized with the same IBM technology of the following stages of the ASIC. The PMOSFET is designed in a way to provide low noise and a compression characteristic close to the one of the DEPFET-based solution. This goal has been achieved by forcing the input PMOSFET, which is not equipped by an intrinsic non-linear gain like the DEPFET, to operate in a non-linear mode with a signal compression induced by the signal itself. A simple way to achieve this mechanism is to put a resistor in series with the PMOSFET and to connect the other side of the resistor to the virtual ground of the filter. Based on the same idea, a second prototype has been designed to improve both the compression behavior and noise performance of the Front-End (FE). In particular, thanks to the introduction of an NMOSFET as non-linear resistor and the inclusion of three input branches, the dynamic range has been increased from 500 photons of the previous prototype up to 2800 photons at 1 keV.
The use of the FE in matrix application, like the 64x64 channels of the DSSC project, urged the study on the yield and robustness of the circuit from power supply fluctuations and crosstalk between pixels. For these scope, in my research, the design, realization and testing of new Mini-SDD front-ends has been discussed. Both open loop and closed loop solutions have been demonstrated through both analytical evaluation and experimental characterization. All the solutions have been designed to keep the same filter, ADC and memory as in the DEPEFET sensor scheme.
A bulk control circuit to correct the chip to chip process variations of the open-loop non-linear FE has been presented. The circuit is capable to stabilize the FE response in presence of threshold voltage variations, acting on the bulk voltages of the Front-End transistors and exploiting the body effect. The control circuit does not affect the noise performances of the Front-End since it is a low frequency loop. It acts only on the transistors biasing point.
Moreover, a novel circuit aimed to increase the robustness of the open-loop FE has been presented. The aim of the circuit is to allow the use of the open-loop FE in matrix applications with many pixels. Thanks to the introduction of cascode transistors, it is possible to decouple the FE from the power line and virtual ground of the filter. The proposed FE is a good com-promise between noise performance (worse of only 10-15 %) and robustness.
Beside these open-loop FE solutions, a closed loop solution has been also investigated, due to its well know properties of robustness and stability. A classical architecture, the one of the Charge Sensitive Amplifier (CSA), have been implemented, in a baseline configuration with linear gain (i.e. without gain compression as in the open-loop FE). Furthermore, to couple the CSA with following FCF filter a voltage to current converter has been introduced. During the programming phase a negative feedback has been designed to fix the CSA output at the same voltage as the filter virtual ground before the burst phase of the European XFEL. The mechanism exploits the output stage of the CSA. The FE has been integrated in the 64x64 (F2) channels ASIC.
In linear mode, the CSA cannot achieve the single photon resolution and a large dynamic range at the same time. To reach this target a CSA with gain compression mechanism has been also studied. The circuit exploits in the feedback network a non-linear PMOS. Furthermore, to better control the threshold voltage variations of the non-linear capacitor for a more robust transition between high and low gain, a bulk control loop has been used. With the proposed circuit it is possible to achieve a dynamic range up to 3000 photons at 1 keV, still keeping the same noise performance as the linear version of the CSA.
In conclusion, the use of the Mini-SDD solution as an alternative approach to DEPFET sensor in the DSSC project has been successfully demonstrated through both analytical evaluation and experimental characterization. Although the DEPFET approach has the best performance, the improved version of the Mini-SDD FEs have good noise performance, robustness and dynamic range. F2 is going to be integrated in the Mega Pixel X-ray camera of the DSSC collaboration for different scientific experiments such as: analysis of microscopic structures, study of very fast processes, cluster experiments, generation and study of extreme states of the matter etc.

RELIABILITY STRATEGIES FOR NEXT GENERATION CLOUD NETWORKS

Ali Hmaity - Supervisor: Prof. Massimo Tornatore

This work addresses the network reliability challenges of cloud-based services with special focus on two specific use-cases. The former considers the resiliency of Cloud Service Providers (CSPs), which exploit cloud networks to provide ubiquitous access to contents, while the latter concerns Infrastructure Providers (InPs) and targets the case where Network Functions (NFs) are implemented as software modules and run in virtualized environment, spread across the network.

This dissertation is motivated by the growing need of resiliency in cloud-based networks. In particular, as the Internet is shifting from being an end-to-end communication model, to an end-to-content communication model, CSPs are investigating novel resiliency mechanisms to maintain the service continuity, especially in case of disaster-based failures. In case of such wide-spread failures, CSPs might not be able to guarantee the **Network Connectivity** (i.e., the reachability of all nodes from any node in the network). Hence, researchers proposed the concept of **Content Connectivity** (i.e., the reachability of the content/service from any node in the network) that can be satisfied even if the CSP's virtual network is disconnected.

On the other hand, the increase of cloud-based service adoption urges InPs to investigate new strategies to contrast the competition of CSPs and keep stable profit margins.

In this context, Network Function Virtualization (NFV) was proposed as a new architectural paradigm to improve the flexibility of network service provisioning, reduce Capital Expenditure (CapEx) and Operational Expenditure (OpEx) and decrease the time to market.

Figure. 1 shows a detailed overview of this set-up running on top of a cloud-based infrastructure. The figure illustrates also the clear separation between CSP and InP domains. In particular, the CSP handles the virtualization of applications and management of hardware clusters that can be whether centralized or distributed across the edge of network (i.e., edge computing).

NFV can revolutionize how InPs design their infrastructure by leveraging virtualization to separate software instances from hardware appliances, and decoupling functionalities from locations for faster service provisioning. NFV supports the instantiation of Virtual Network Functions (VNFs) through software virtualization techniques and runs them on Commercial-Off-The-Shelf (COTS) hardware. Hence, the virtualization of network functions opens the way to the provisioning of new services without the installation of new equipment.

NFV simplifies service deployment by exploiting the concept of service chaining: a Service Chain (SC) is a sequential concatenation of VNFs and/or hardware appliances to provide a specific Internet service (e.g., VoIP, Web Service, etc.) to the users. However, deploying NFV solutions in operational networks require solving multiple issues related to performance, availability, security and survivability. One important key design in an NFV framework is the ability of the NFV-Management and Orchestration (NFV-MANO) component to ensure service continuity. Such objective translates into many requirements that the Network Function Virtualization Infrastructure (NFVI) must satisfy, among which resiliency and geo-redundancy requirements, which are the focus of this work.

In light of the above, this dissertation is motivated by the growing need of resiliency in cloud-based networks. In particular, as the Internet is shifting from being an end-to-end communication model, to an end-to-content communication model, CSPs are investigating novel resiliency mechanisms to maintain the service continuity, especially in case of disaster-based failures. In case of such widespread failures, CSPs might not be able to guarantee the **Network Connectivity** (i.e., the reachability of all nodes from any node in the network). Hence, researchers proposed the concept of **Content Connectivity** (i.e., the reachability of the content/service from any node in the network) that can be satisfied even if the CSP's virtual network is disconnected.

On the other hand, the increase of cloud-based service adoption urges InPs to investigate new strategies to contrast the competition of CSPs and keep stable profit margins.

In this context, Network Function Virtualization (NFV) was proposed as a new architectural paradigm to improve the flexibility of network service provisioning, reduce Capital Expenditure (CapEx) and Operational Expenditure (OpEx) and decrease the time to market.

Thesis contribution

In this dissertation, we address some open challenges in the resiliency of next-generation cloud-based networks. The main contributions of this thesis work are as follows:

- New optimization model to solve the problem of providing CSP's virtual networks with Network Connectivity (NC) in case of single-link failures and maintaining Content Connectivity (CC) in case of double-link failures, considered as disaster zones.
- New optimization models to design novel protection strategies for VNFs placement and SCs provisioning with the objective to provide resiliency against single-node and/or single-link failures and maximize consolidation of VNFs. New heuristic approach to solve the problem of placement and routing of VNFs, when latency and/or capacity are taken into consideration, with the objective to improve consolidation, increase acceptance ratio and capacity utilization. Further analysis is carried out to assess the benefit of the convergence of the fixed and mobile architectures. New optimization models and heuristic algorithms to evaluate the potential power savings introduced by the combination of multiple power reduction strategies in protected IP-over-WDM networks, with focus on dynamic traffic conditions.

Figure. 2 summarizes the overall contributions and shows their respective area of application in a cloud network. Three out of the four proposed contributions are applied in core networks, while the fourth is applied in metro networks. The contributions in Chapters 3, 4 and 5 benefit to the InP, while our contribution in Chapter 2 assumes that the CSP and InP constitute a single provider, which corresponds to the case where an InP has its proper Cloud infrastructure and hence can play the role of a CSP as well (e.g., Vodafone)

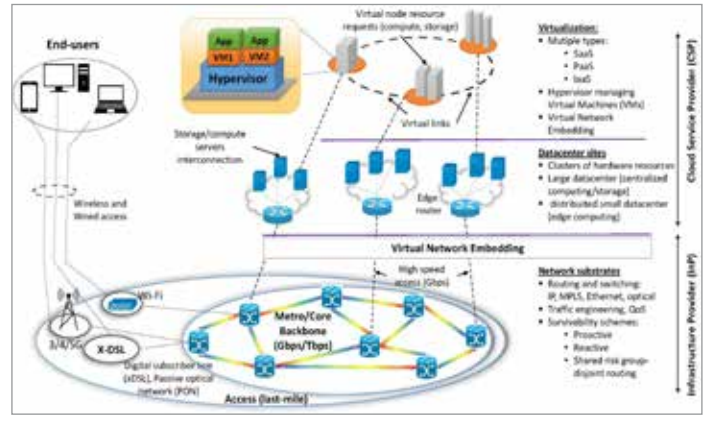

Fig. 1

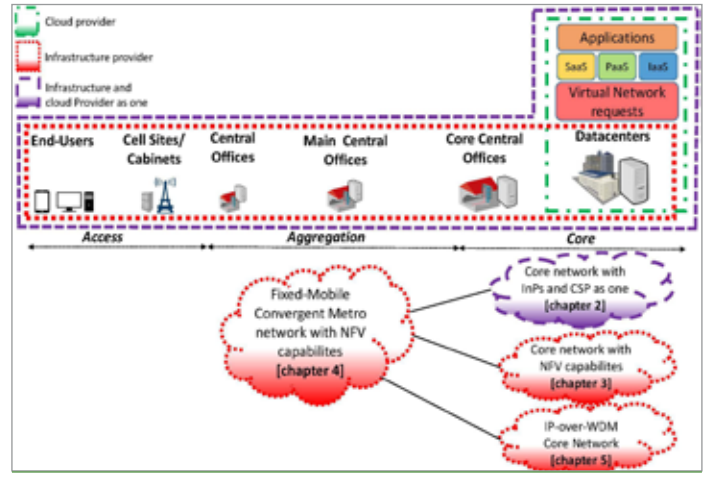

Fig. 2

STUDY OF LEVEL MEASUREMENT PULSE RADAR SYSTEMS

Misagh Khosronejad - Supervisor: Prof. Giuseppe Macchiarella

The technical benefits of radar as a level measurement technique are clear. Radar provides a non-contact sensor that is virtually unaffected by changes in process temperature, pressure or the gas and vapor composition within vessel. In addition, the measurement accuracy is unaffected by changes in density, conductivity and dielectric constant of the product being measured or by air movement above the product. These benefits have become more significant to the process industry since the advent of low costs, high performance, two wire loop powered radar level transmitters. This breakthrough, in the summer of 1997, produced an unprecedented boom in the use of non-contact microwave radar transmitters for liquids and solids process level application. Currently, there are two FM-CW and Pulse radar techniques in common use for level measurement applications. Design theory for both types of radars are well known, while the developing process can be a challenge and usually requires original techniques to experimentally verify the operational performance of the radars. In radar technology, to make a contribution in developing process of front-end RF circuits, it is required to have competitive system with the performance and cost of radar systems in use today. The major objective of the dissertation was to design and fabricate a front-end RF system for level measurement pulse radars in order to push the radar capability a step further by higher frequencies (K-band). Since radar design techniques are well-known theoretically, challenges of the work were under experimental phase. The aim was achieved while there were limits regarding budget, available components, laboratory instruments and etc. The front-end microwave module took advantage of smaller size, low cost and ease of fabrication in comparison with many other level measurement radars. Each block of the system (oscillator, filter, mixer, directional coupler, antenna) was individually designed and realized and the simulation and measurement results were reported. All blocks came together and form the front-end microwave module circuit.

The microwave module circuit as a section of level measurement pulse radar device went under a real case scenario test to measure distance and results were presented.

As long as compactness of electronic devices is requested, the antenna designer is pushed to reduce the antenna size as well. Planar antennas are widely used in variety of applications due to their compact shape, light weight, less complexity, low cost and ease of implementation and integration with microwave integrated circuits. Despite these significant advantages, planar antennas suffer low impedance bandwidth, low gain and poor radiation efficiency. However, we have designed and realized two types of novel planar array antennas with high directivity and gain. These two antennas could be used for level measurement process due to their high radiation efficiency in comparison with classic horn antennas which are most common types of antennas used in level measurement applications. Several other research activities that were carried out during the PhD cycle, are listed following.

A design method to realize multilayer polarization converter devices in transmission: The method was validated by realization of two multilayer full cross polarization converter devices. These devices are capable to rotate the polarization 90-degree independently from the direction of the input linear polarization.

Design of two perfect linear polarization converter devices in reflection: These devices are capable to rotate the polarization 90-degree for normal incident waves.

Implemented an efficient numerical method in order to use direct mode-matching technique to analyze step-discontinuity in elliptical waveguide: This method allowed us to improve analysis performance of elliptical-shaped aperture in horn antennas.

Proposing a numerical analysis that simply allow us to study the effects of hemispherical and elliptical dielectric lenses on the directivity of horn antennas: This analysis illustrated a reliable enhancement in radiation performance of horn-radome system.

Design of a novel dual-band partial reflective surface (PRS) in order to enhance the directivity of a dual-band planar dipole antenna: This periodic structure was effectively enhancing the directivity of the basis antenna, and it took advantages of simplicity, low cost, and ease of fabrication.

Design and implementation of a planar Archimedean spiral antenna suitable for UHF remote sensing applications: This antenna had application particularly in detecting pulse discharges in power transformers. Antenna took advantages of compactness, low cost and ease of fabrication in comparison with other similar antennas in market.

Extraction a novel formula to calculate the effective dielectric constant of a structure consisting of antipodal planar dipole element: This formula was used to study the radiation performance of planar log-periodic dipole array (PLPDA) antennas. Introducing the proposed formula into classic design procedure of log-periodic dipole array antenna allows us to calculate the design parameters of a PLPDA antenna accurately.

Design and implementation of a novel multilayered dual-polarized planar antipodal dipole antenna: This antenna allows us to effectively apply polarization diversity technique in WLAN applications. Results illustrated -20dB isolation between ports for 25% bandwidth at 5.5GHz.

STUDY OF NOVEL DEVICES FOR CROSSPOINT MEMORY AND NEUROMORPHIC APPLICATIONS

Mario Laudato - Supervisor: Prof. Daniele Ielmini

Nowadays, the amount of digital information is increasing exponentially due to the large diffusion of consumer electronics. All over the world an ever-growing number of people has the possibility to purchase smartphones and tablets able to save a large amount of photo, video, music and applications in a single portable device. This social phenomenon has caused an enormous request of non-volatile memories capable to store a high quantity of data in a very small volume, maintaining the stored information even without voltage supply. In the last ten years the Flash technology has been the leader in the non-volatile memory market, but has now reached the scaling limits and the most important semiconductor companies have to face crucial issues related to atomistic doping and charge trapping. Moreover, the Von Neumann bottleneck leads to require a memory able to reduce latency time during CPU operations, combining the high-performance of SRAM and DRAM and high density of Flash Technology and HDD. To this aim, in the last years both industrial and academic research groups have put their great effort on the design of a revolutionary non-volatile memory architecture with novel materials, increased scaling capability and high-performance in order to replace the Flash memory technology. In July 2015 Intel and Micron has reached this purpose announcing 3D Xpoint non-volatile memory technology, which relies his working principle on a crossbar array architecture, where each memory element has a selector element in series located at the intersection between a row and a column electrode wire. This architecture enables multiple stacking and is able to achieve the minimum size of the device of $4F^2$, obtaining scalable bit cost and high density comparable with Flash memory. In this technology the selector element plays a fundamental role with his high non-linear characteristics and high on-off current ratio, allowing to limit sneak path currents during programming operations and maximizing the available array size thanks to a significant reduction of power consumption within the memory architecture. At the same time, alternative in-memory computing approaches are becoming increasingly attractive to develop novel logic and neuromorphic computation in order to overcome Von Neumann bottleneck issues. Indeed typical operations like image learning, pattern recognition and decision exhibit high computational cost for boolean CMOS processors, while, for human brain, they represent elementary processes. In this scenario, the development of new devices designed specifically for neuromorphic computing could enable high density and low power networks in order to properly operate learning and recognition tasks. For this reason, different industrial and academic institutions are looking for promising candidate among the emerging memories that could offer the best properties to mimic synapse behavior in our brain. Nowadays the most developed technologies for both crosspoint memory application and neuromorphic computation are represented by phase change memory (PCM) and metal-oxide resistive switching memory, commonly known as RRAM. In last years these devices have reached mature technology and outstanding feature in terms of fast switching and high endurance for PCM and low power and low-cost fabrication for RRAM, paving the way for future Storage Class Memory (SCM) with huge market opportunities. In addition, intrinsic variability for RRAM and gradual set transition for PCM, which represents constraints on storage memory application, could result in key strength for neuromorphic network synapses. This doctoral dissertation will be focused on Silicon-based selector devices and emerging memories (PCM and RRAM). The work describes working principles, physical understanding and innovative applications in the field of crosspoint memory and neuromorphic computation. The electrical characterization carried out on these different devices has played a fundamental role in this aim, providing the starting point for physical understanding and modeling and paving the way for design and implementation of neuromorphic hardware for pattern learning and recognition.

DECORRELATION PHENOMENA IN A GEOSYNCHRONOUS SYNTHETIC APERTURE RADAR: THEORY, TECHNIQUES AND PERFORMANCE

Antonio Leanza - Supervisor: Prof. Andrea Monti Guarnieri

Gesosynchronous orbit (GEO) Synthetic Aperture Radar is a novel concept which would provide significant potential advantages over the present Low-Earth Orbit (LEO) systems. In particular the nearly-zero inclined GeoSAR would grant the continuous time coverage of the observed sub continental region with 12 hours revisit time, coarse resolution imaging every 20 minutes and high integration gain to compensate the significant spread loss, limiting in this way the required physical antenna size and trasmitted power. However the long integration time (minutes to hours) introduces signal decorrelation problems. In this thesis the study of these decorrelation phenomena is tackled, providing analysis, performance evaluations and when possible, methods to counteract their effects on the focused images. After a general description of the main GeoSAR characteristics and potential applications, the nature of the main decorrelating sources is described, along with their respective effects on the Impulse Response Function of the GeoSAR system. In particular two decorrelation sources have been considered, since their contribute is more relevant for the GeoSAR systems, namely decorrelating clutter and atmospheric phase noise.

Clutter decorrelation is due to those moving targets that slowly change their electromagentic response within times comparable to radar observation interval. The phenomenon has been described through different models depending mainly on target type and timescale.

Target decorrelation issue has been addressed with the two-fold purpose of analyzing the interferometric coherence for the GeoSAR candidate bands (C and X-band) and studying the moving target decorrelation on the sub-daily timescales comparable to the GeoSAR integration times. The analysis performed on Sentinel-1 (C-band) and Cosmo-Skymed (Xband)

image stacks showed that, over vegetated areas, the decorrelation time constant results in the order of several tens of days for the C-band and rapidly falls down below one day in the X-band.

The results brought to the conclusion that, to get acceptable performance in high-resolution GeoSAR interferometry over forests or rural areas, C-band is the limit frequency for the single sensor system

while the X-band would require a constellation in order to reduce furtherly the revisit. Ground-Based radar Ku-band data have been exploited to evaluate clutter decorrelation from the mid-term (minutes) to long-term (hours). The mid-term observations showed the validity of Billingsley model, even if with significant differences between the predicted and observed parameters. In particular the observed decay rate ß values resulted higher. The reason relies in the impact

of small targets like leaves or grass, which is more relevant at Ku-band w.r.t. the lower frequencies analyzed by the Billingsley experiment, and the different portion of the spectrum analyzed in the two experiments. The long-term analysis indicated that target decorrelation, at this timescales, is better represented by the exponential model, also adopted for the interferometry. Furthermore these analysis highlighted that decorrelation is not stationary with a wide range of decorrelation time constants. This makes

difficult the prediction of the performance in terms of Signal-to-Clutter Ratio which, in general, can be evaluated only in a statistical way. This, along with filling the gap between mid-term and long-term decorelation behaviours, represents a challenging and interesting issues to address in future works. Both mid-term and long-term observations evidenced a daily and a seasonal trends that indicates how the performances improves during the night and/or during the cold months.

Atmospheric phase noise (APS) is then addressed. In the considered bands, the main contribute is due to troposphere which introduces a delay in the wave travelling time and hence a multiplicative phase noise that depends on the acquisition time and targets position.

Its statistical characterization has been exploited to derive a model of the APS affected Focused Impulse Response Function (FIRF) which allows to estimate significant effects such as the peak power loss, the mainlobe widening and the clutter noise originated by the sidelobes spreading. The optimal sub-aperture is then defined which represents the aperture interval that provides the best compromise, in terms of decorrelation, between the temporal and spatial contributes. Unlike clutter decorrelation, APS decorrelation can be (partially) compensated. Hence it has been proposed a strategy to estimate it and compensate its effects in the focused image. The processing was composed by a sub-apertures interferometric-based estimation followed by a regularized MMSE focusing. The operations were repeated iteratively widening the sub-apertures. Exploiting a Sentinel-1 image and a possible GeoSAR design, a series of simulations have been performed in order to evaluate the proposed processing performance. The results showed significant improvements but limited by the sidelobes clutter, as foreseen by the theory. In fact simulations showed that the final residual decorrelation is comparable with the one observed in the optimal sub-aperture. Jointly analyzing the decorrelation phenomena, an important consideration is that decorrelating clutter noise is scene dependent whereas APS is not. In fact, decorrelating clutter nuisance is absent when urban areas (or stable targets in general) are imaged, no matter what the system frequency is.

On the contrary, APS is always present on any kind of scene, even if with variable strength, and its impact significantly depends on the system frequency. For this reason, in the GeoSAR design, the choice of band depends mostly from the APS. According to the analysis performed, C-band would be a suitable solutions for a single-sensor COMSAT-like GeoSAR system (eccentricity = 0.0008). In the X-band (or higher frequency) APS impact becomes critical. However in this case reducing the synthetic aperture time, by the exploitation of a more eccentric orbit or by a multi-sensor system (constellation), could solve the problem and provide good performances.

MULTICORE RESOURCE MANAGEMENT A HORIZONTAL PERSPECTIVE

Simone Libutti - Supervisor: Prof. William Fornaciari

The end of Dennard's scaling has been one of the most disruptive events occurring in the evolution of computing platforms. In order to cope with the subsequent increase in chips power density, hardware designers have progressively moved towards solutions that leverage the concepts of *parallelism* and *heterogeneity*. As a consequence, modern architectures feature an increasing number of *shared computational resources* that are power-hungry, can possibly be different in nature, and can be concurrently used by multiple applications.
This increasing hardware complexity has in turn affected the software stack: task scheduling and mapping have become challenging problems due to the need of maximizing the performance of applications while minimizing power, temperature, and contention on the shared resources.
This dissertation directly tackles the above mentioned problem. In particular, we address resource management from a horizontal perspective, trying to identify the challenges and solutions that pertain the increasingly blurred area between high-end embedded and *High Performance Computing* (*HPC*) systems.
At operating system level, we studied how the Linux Control Groups (*cgroups*) perform CPU time allocation in linux-based systems. In particular, we analyzed how the *cgroups cpu controller* limits bandwidth. We discovered that, in certain scenarios, the CPU bandwidth that is effectively exploited by the applications could potentially be not only lower than the expected one, but also notably higher, which may cause troubles in multi-application scenarios.
Then, we dealt with application characterization. In particular, we employed Design Space Exploration techniques to characterize resource usage and energy consumption of applications. Experimental results shown that, by exploiting the characterization information to suitably map applications, the resource manager is indeed able to optimize scheduling decisions.
We also shown that, in order to maximize the *Quality of Service* (*QoS*) that can be squeezed out from the available resources, applications need to constantly update their software parameters in order to continuously adapt them to the system status.
We then studied the benefits that come from allowing a system-wide resource manager and an application-specific auto-tuner to work in a synergistic way. The main idea behind that study is that, whereas resource managers allocate resources to applications according to precise and known system-wide optimization goals, they are often unaware of what applications really need.
We tackled this problem by moving part of the management complexity to the applications side: each application relies on an application-specific auto-tuner, i.e., a component that is specifically configured for the target application and is able to tune the application software parameters at runtime in order to make it comply with its *QoS* goal despite a runtime-variable resource availability.
In order to better exploit the capabilities of multi-core processors, we then added another degree of complexity: *heterogeneity*. We specifically addressed *big.LITTLE* architectures, i.e., processors that feature two clusters of cores: a performing and power-hungry one (*big*), and a slower and power-efficient one (*little*). The two clusters share the same Instruction Set Architecture; therefore, threads can be migrated between the two clusters during runtime. This, in turn, allows operative systems to exploit the trade-off between performance and power consumption.
We studied how to dynamically migrate threads among the *big* and *little* clusters in order to maximize the usage of the big cluster while minimizing the performance losses that are induced by resource contention. We did so by introducing the concept of *stakes function*, which represents the trade-off between *exclusive allocation* and *sharing of resources* in multi-core processors. We introduced a co-scheduling policy that exploits *stakes functions* as a metric to take co-scheduling decisions on heterogeneous processors.
Then, we used the two clusters of cores as a heterogeneous *OpenCL* device. In this context, we presented a mechanism that forces the *OpenCL* runtime to view the *big.LITTLE* processor as a custom set of heterogeneous devices instead of viewing it as a single device.
Finally, we dealt with distributed systems, mainly focusing on *High Performance Computing*. In those scenarios, the objectives of resource management are typically different from those of embedded systems. In particular, resource management techniques for *HPC* mostly focus on minimizing power consumption and thermal hot-spots, detecting and counteracting faults and aging-induced performance variability, and exploiting heterogeneous accelerators.
First of all, we performed and interesting study of how the freeze/restore-based process migration of *MPI* applications, which is usually performed at node granularity to address faults, can be made fine-grained in order to migrate only parts of the application on a different computing node. The outcome of this study was the development of the *mig* framework, an *OpenMPI* module that allows *MPI* applications—or even just a subset of their processes—to be migrated from an *HPC* node to another one without requiring developers to change their applications' code nor performing intrusive changes to the *OpenMPI* framework.
Then, we focused on the node-level optimization of distributed computation. We presented a resource management approach that exploits the trade-off between power consumption and performance when executing *HPC* applications that must comply with runtime-variable *QoS* requirements. We applied an adaptive performance-aware execution model in the context of a real scientific application domain on a multi-core *HPC* system. The approach is based on the concept of "resource minimization via late termination". That is, we minimize the amount of resources that the applications can use so that they are barely able to comply with their *QoS* requirements.
The unused resources can be therefore used to perform system-wide optimizations such as minimizing power consumption or isolating faulty parts of the system.
The final step towards a performance, quality and power aware (but yet homogeneous) *HPC* resource management consisted in designing a feedback-based and partially decentralized resource management approach that allowed applications to comply with their *QoS* goals while minimizing resource usage; minimizing the negative effects of faults-induced performance variability; and leveling the temperature throughout the available computing cores, so that hot-spots are avoided and the effects of temperature on *Mean Time To Failure* are equally balanced on the cores.
To conclude this dissertation, we presented our "work in progress" in the context of the *MANGO* European Project, which aims at performing the first steps towards an unified runtime management support for deeply heterogeneous *HPC* systems. We discussed the goals, requirements and solutions for an *HPC* software stack that targets deeply heterogeneous architectures composed of both general purpose nodes and a variety of accelerators.
In particular, we proposed to employ a combination of resource management techniques to control the allocation of computing units and memory resources to different applications under *QoS* requirements; and a low-level runtime support to provide a minimum common base among different accelerators, thus allowing functional portability of applications and an easier porting of high-level programming models on the different accelerators.
During our work, we extensively used, modified and extended the *Barbeque Run-Time Resource Manager*, which is an Open-Source manager developed at Politecnico di Milano. You can find our contributions to this framework at *https://bitbucket.org/bosp/barbeque*.

X-RAY DETECTION SYSTEMS FOR ADVANCED LIGHT SOURCES

Chang Liu – Supervisor: Prof. Andrea Castoldi

In order to spark the light of science in the dark field of science, starting from the mid-1990s, a number of third generation synchrotron facilities have been constructed worldwide, speeding electrons around at the speed of light, and emitting extremely intensive beams of X-rays. Moreover, in recent years, fourth-generation light sources based on the Free Electron Laser (FEL) concept are being developed, which are the most powerful X-ray sources ever existed on the planet. Their peculiar characteristic is having a very large impact on many areas of science such as physics, chemists, biologists, medical applications, nanotechnology, taking us further into the future. However, these unique characteristics also pose the challenge, in parallel, of a large improvement step in the required detection systems.
Among all the possible detector solutions for these advanced sources, the silicon drift detector (SDD) has become the most widely used technology for X-ray spectroscopic applications, which demand low-noise (high energy-resolution) and high counting rate over a wide energy range. In the meanwhile, several position-sensing SDD topologies have been developed, aiming at combining 2D position resolution together with energy resolution. Our research effort has been focused on the R&D of novel silicon drift detector (SDD) topologies for both X-ray imaging applications and X-ray spectroscopy applications along three lines of activity:
First, we developed and studied a peculiar drift detector architecture --- the Multi-Linear Silicon Drift Detector (MLSDD), for X-ray imaging applications at high charge levels and at fast operating speed. The MLSDD --- by combining arrays of deep p-implants and deep n-implants alternately in the bulk --- improves significantly charge handling capability and the average drift speed, and features separate channels for drifting the signal charge to a point-like collecting anode. It enables therefore X-ray imaging together with X-ray spectroscopy. Compared to other X-ray imaging detectors with compatible spatial resolutions, this topology allows the readout nodes to be on one detector side only, which implies more efficient interconnection schemes. In our study, novel MLSDD prototypes were fabricated with different layout features (e.g. p+ strip pitch size, deep n-implant size, surface potential between p+ strips) and experimentally characterized, by IR laser scanning techniques, for different positions, different charge injection levels, and different ionization profiles. The results help understand clearly the effects of the new design features, and help qualify imaging capability and optimize the detector design.

The second line of activity is the investigation of integrated charge injection devices, an ideal approach to deal with the calibration problem for large area imaging detectors requested by the advanced X-ray facilities, where a large parameter set and extreme experimental conditions must be investigated (e.g. the European X-ray FEL in Hamburg). The charge injection device can be integrated into individual pixels or channels and can inject controllable amount of charge at precise time in order to allow real-time calibration of the pixel response (i.e. amplitude and time). Based on a well-known charge injection device --- injection through a metal-oxide-semiconductor (MOS) structure --- we have implemented several injection structures and conducted a comprehensive study including characterization and modelling of such a device. The obtained modeling clarifies the physical relationships among all the relevant parameters for future detector developments.

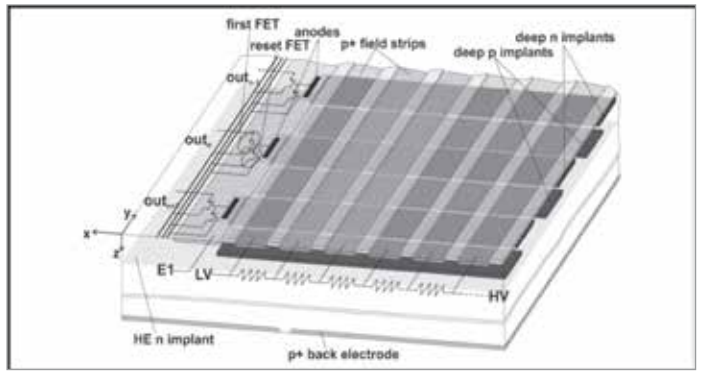

Fig. 1 - Schematic 3D view of a Multi-Linear Silicon Drift Detector

Last, a third research line, centered in advanced detection systems for X-ray spectroscopy, was carried out under the INFN project ReDSoX/ReDSoX2 and in the frame of the EUROFEL project (in collaboration with Synchrotron ELETTRA Trieste , IASF-Bologna, Universita' di Pavia, Universita' di Bologna, IAPS-Roma). The aim of the project is to develop new X-ray spectroscopy systems for the TwinMic and XAFS Beamlines at Synchrotron ELETTRA Trieste (Italy) by using monolithic arrays of SDD cells. Compared to the current system in service, the new system is required to have high energy resolution in the soft X-ray energy range (e.g. < 1keV) and a 10 times higher counting rate capability. In order to optimize the cell geometry of the first detector prototype, we have finely mapped charge transportation properties (i.e. signal amplitude, collection time and charge cloud size) in cells of different shape. Currently a second generation of SDD arrays (with optimized layout and leakage current) and new low-noise frontend ASICs have been integrated. A vacuum system with cryogenics cooling has been developed in order to study the ultimate spectroscopic performance (in terms of equivalent noise charge – ENC) at various temperatures. At room temperature we achieved 1.6 electrons r.m.s. ENC with ASIC only (i.e. zero detector capacitance).

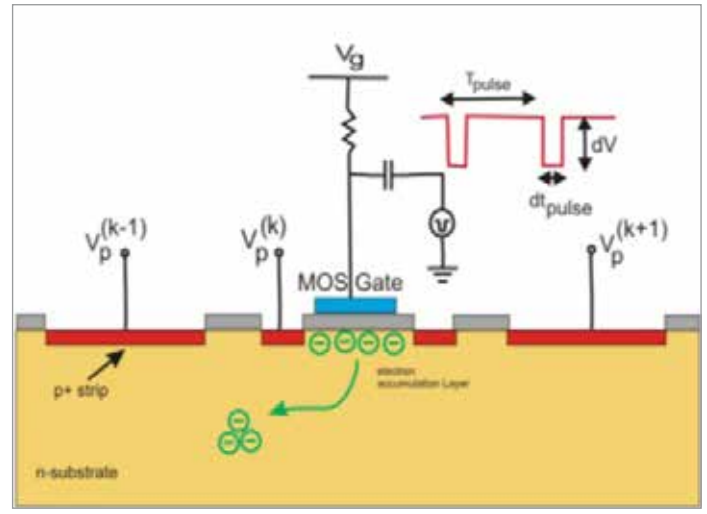

Fig. 2 - Sketch of the working principle of a MOS injector

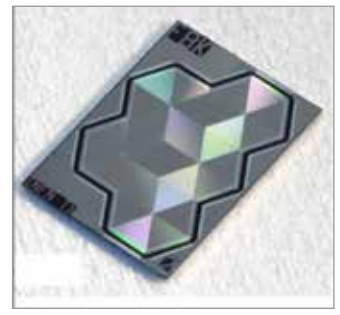

Fig. 3 - Monolithic SDD array concept: an example consists of 5 hexagonal SDD cells

WEARABLE TECHNOLOGIES FOR THE UNOBTRUSIVE INVESTIGATION OF CARDIOVASCULAR SYSTEM

Prospero Lombardi - Supervisor: Prof. Franco Zappa

So far, a better understanding of the beat-to-beat dynamics of cardiac and vascular mechanics is advisable. However, novel and unobtrusive wearable sensors and devices are needed for the data collection. In this work two monitoring systems, the MagIC-Space and the Seismote systems, were developed for the investigation of those physiological aspects in the frame of two different research projects.

MagIC-Space, the first achievement, is a textile-based system (fig.1) which allowed us to investigate for the first time aspects of sleep physiology in microgravity, including the mechanical behavior of the heart by seismocardiography (SCG), i.e. the measure of chest vibrations produced by the beating heart. The system was used with success aboard the International Space Station (ISS) for the recording of sleep data from one astronaut in seven night experiments. The development was based on the enhancement of a previously available prototype (originally measuring the electrocardiogram (ECG) and respiration) and consisted in the re-design of the system 1) to include new sensors and relevant electronics for the additional measure of SCG, body temperature and skin resistance, and 2) to make the hardware compliant with the strict NASA regulations for a safe usage aboard the ISS. As to the latter issue, specific countermeasures were adopted in the system to reduce the risk of electrical shock, thermal and chemical hazards. Laboratory tests were then carried out to verify the adherence of the system to the requirements and obtain the final NASA approval. In addition, attention was paid to provide the system with automatic recovery procedures from most likely non-fatal failures possibly occurring in the harsh ISS environment.

Seismote, the second achievement, is an architecture which enables a joint assessment of cardiac mechanics and haemodynamics by multi-point SCG and photoplethysmographic (PPG) measures. PPG gives information on the arrival time of the pressure pulse to a peripheral artery and from this signal we may derive information on vascular characteristics. The developed architecture consists of a number of miniaturized sensor nodes, placed on different body sites for multi-point measures, and a hub which collects and stores data from nodes, allows the additional measure of ECG and can be wirelessly connected to a remote device, e.g. a smartphone, for real-time data streaming. Two versions of the system were realized, one with a wired (fig.2) and one with a custom wireless link (fig.3) between nodes and hub. Notably, the wireless link allows us to connect up to 13 nodes and to collect 52 different signals sampled at 200 Hz. To realize the wireless architecture, solutions were developed to address the major and open issues of low power data transmission and time synchronization among nodes. A custom protocol was developed since none of available wireless technologies provided the required throughput with the reduced power budged of nodes. The new protocol is roughly 18% more efficient than the recently released Bluetooth Low Energy in terms of up-link data overhead. As to the synchronization, a novel procedure, based on low level timestamp broadcasting, was designed and implemented by means of the custom protocol achieving a small average error of 27.53 µs.

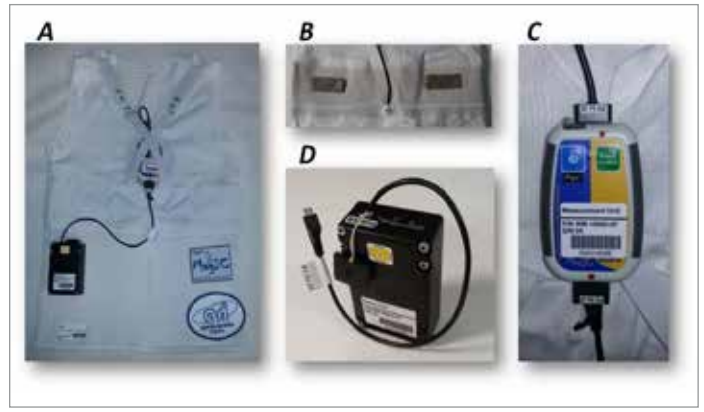

Fig. 1 - MagIC-Space: A) The system consists of a smart garment, a recording unit placed on the sternum and a battery unit. B) The garment integrates textile ECG electrodes (left and right patches), a respiration gauge and a temperature probe (small spot in the center). C) Recording unit. D) Battery unit.

Fig. 2 - Wired version of Seismote.

Fig. 3 - Wireless version of Seismote.

ADVANCED METHODS TECHNIQUES AND DIGITAL ARCHITECTURES FOR HIGH PERFORMANCE TIMING OF EVENTS

Nicola Lusardi - Supervisor: Prof. Angelo Geraci

The precise measurement of time intervals is a primary goal in a growing number of applications and the challenge to achieve increasingly higher resolutions than ever is a main topic of research. In this sense, Time-of-Flight measurements and Time-Correlated Photon Counting are two milestones. Since the intrinsic resolution of the sensors used today is in the order of tens of picoseconds, the measurement systems must guarantee performance at least of this order.
The choice of making digital a part or the totality of the measurement electronic systems exploits well-known advantages from the adaptivity, to the versatile calibration, to the easiness of implementation of powerful processing algorithms with lower power consumption, and area occupation with respect to the equivalent analog solutions.
The last generation of digital programmable devices as Field Programmable Gate Arrays (FPGAs) and System-of-Chips (SoCs) has made possible the implementation of high-accuracy Time-to-Digital Converter (TDC) architectures with performance comparable with ASIC realizations. In this way, all the well-known advantages of using programmable devices are exploited, such as totally tunable characteristics, easiness of portability, reduced time-to-market, lower migration cost from one generation of technology to another, just to name a few. This has to be done always keeping on foreground the maximum functional performance.
The dissemination of experiments in which high-performance timing measurements become increasingly important, is triggering the development of new versatile instruments, especially digital. Moreover, the concept of high-performance is not only related to resolution and precision but also indicates other features like high full-scale-range, high acquisition rate, multi-channel operation mode, real-time processing, flexibility, modularity and, last but not least, low time-to-market and implementation costs.
We presents a new fully configurable system whose structure is a hardware, firmware and software bundle for high-performance timing measures in experiments where high-performance time resolving is a primary request.
Figure 1 shows the schematic structure of the whole system. The hardware is a custom Printed Circuit Board (PCB) that is composed of an analog front-end allowing the acquisition up to 16 simultaneously time events, FPGA processing unit, various communication modules (e.g. USB 3.0 and Ethernet links), and a software that runs on a generic computer.
For each channel, the input acquisition stage is a programmable threshold comparator that converts the analog input pulses (0-3.3V) in digital signals with timing jitter less than 7 ps r.m.s. The user can set the programmable threshold with resolution of 50 μV by means of two 16 bits 8-channels I2C programmable DACs.
The FPGA module is a TE0712-02-200-2C provided by Trenz Electronic that hosts an industrial 28-nm programmable Xilinx Artix-7 XC7A200T. The core of the firmware part into the programmable logic is a multi-channel Tapped-Delay-Line Time-to-Digital Converter (TDL-TDC) that is organized as a HDL IP-Core Module (TDC IP-Core) compatible over all Xilinx Serie-7 FPGAs and SoCs. Tab.I returns a synioptic view of achievable performance.
The TDC IP-Core is characterized by an area saving structure and is composed of 16 independent channels implemented in the FPGA. Each channel has average resolution of 250fs and hardware full-scale-range of 9.45s. The precision, without front-end, is lower than 9.7ps r.m.s. per single channel that means a global precision less than 12 ps r.m.s. per single channel. The maximum event rate is 45MHz. Each channel implements an 8th-order Super Wave Union sub-interpolation over the 2222ps delay-lines that performs the time-to-digital conversion; that achieves a measurement resolution that is 8-time lower than the propagation delay between adjacent taps. In order to extend the hardware full-scale-range up to 9.45s, the Nutt-interpolation has been implemented. In this way, the measurement of the time interval between beginning and end of the interval is composed of a coarse part and a fine contribution from the TDL. We have verified that no detrimental effects of crosstalk among channels are detectable.
The system is composed by firmware-software components, whose synergy maximizes real-time performance of the instrument. The firmware is organized as a daisy chain bus with high grade of modularity, where the TDC IP-Core reads and writes data and, in parallel, other modules perform real-time data processing.
A time tagging read-out module is implemented by default that extends the hardware full-scale-range thanks to software flexibility. Moreover, the user can implement more histogram-makers in hardware and level-zero trigger modules. The hardware histogram-maker creates a 2^{15} bins real-time histogram from the time difference between two input channels in a program-defined time interval selected by the user, whereas the level-zero trigger stores the measurement that starts and stops with a programmable pattern in a defined time window within the full-scale-range set.
The instrument is not only suited for TE0712 but the PCB part can be connected to all the TE07xx Trenz Modules that host different families of FGPAs and SoCs available in the 28-nm Xilinx Serie-7 series; moreover, the hardware is completely compatible with Xilinx UltraScale FPGA and UltraScale+ SoC available in TE0841 and TE0820 modules respectively. The firmware can be easily migrated thanks to its high-modularity.
This instrument is being used in several national (2) and international (5) research activities, among which a review of some as examples is presented. Moreover, the most significant achievements accomplished during the activity are attested by 39 international publications (8 on journal), of which 13 as first author and 1 single name submission.

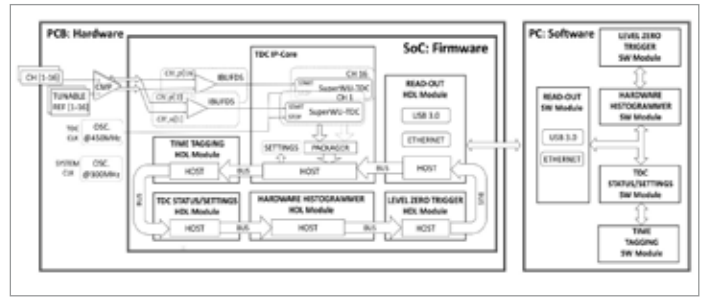

Fig.1 – Schematic description of the hardware/firmware/software bundle.

FEATURE	VALUE
Number of Channels	16
TDC clock frequency	450MHz
System frequency	100MHz
Decoding/Calibration Hardware in FPGA	1 per channel
Device Occupation per Channel	948 SLICEs, 54kbit BRAM, 1 DSP48E1s
Maximum Rate per Channel	45MHz
Processing Latency	110ns
Channel Resolution	250fs
Channel Precision	<9.7ps r.m.s.
Full-Scale-Range	9.45s
Channel Crosstalk	No spurious measure on idles and active channels
Precision Crosstalk Influence	<3ps r.m.s.on active channels
Maximum Non-Linearity Error	<4.2ps
Voltage Drift Influence	No detectable
Temperature Drift Influence	ps/°C without Peltier

DEVELOPMENT OF A FMCW SHORT RANGE GROUND BASED RADAR

Luigi Maggi - Supervisor: Prof. Andrea Monti Guarnieri

In the last decade, ground radar sensors, which are used for civilian applications, received an important technological evolution. Ground RADAR tools for monitoring and control are now an established reality in some areas such as open pit mining, and monitoring of medium-scale geological phenomena such as landslides and subsidence. However, the technology is still young and ground RADAR systems present several interesting research areas. One of the most interesting one, is certainly the analysis of advanced solutions for the antenna synthesis, in order to make imaging techniques RADAR increasingly efficient and reliable. The purpose of this activity is to define one or more solutions for the antenna synthesis of a radar system with particular attention to SAR image applications.

The an innovative FMCW RADAR in W band, more precisely at 77GHz was developed. Furthermore the Digital Beam Forming technique in order to improve the performance in terms of image ambiguity or, from other point of view, in terms of scan speed.

Hence, the work has been based on all radar imaging techniques currently implemented in ground sensors that include: 1) synthesis of antenna through the movement of the sensor along a linear slide 2) synthesis of antenna through radiating elements array 3) Azimuth Digital Beamforming. In general, the three main solutions outlined can be hybridized in order to provide trade-offs that can respond to different application requirements.

The work will provide a detailed benchmark of existing techniques highlighting the advantages and disadvantages of each application in the main bands of ground radar sensors (W), and in the particular case of SAR image, with particular attention to technological constraints and technical feasibility of the solution.
During the Ph.D. several campaigns have been carried out to evaluate the performance and reliability of the system in different scenarios and type of applications that are below reported:
SAR scan for Image Application
Ranging Application as vibration and dynamic deformation

Finally, an innovative approach has been defined and implemented in a prototype sensor RADAR currently available in ARESYS. The prototype was a based system of a product which will be lunched in 2017.

The test Campaign was done in Premazzi Cave, in Varese. The Radar was in linear actuator configuration and it was placed in the centre at a distance of about a hundred meters as in the figure below. As we can see in the figure 2 below it has been placed a corner target at a distance of 24 meters.

The acquired image at full resolution is reported in Figure.3 , pointing out the reference corner reflector and the excavator.

Fig. 1 - FMCW RADAR and scan system

Fig. 2 - Radar point of view

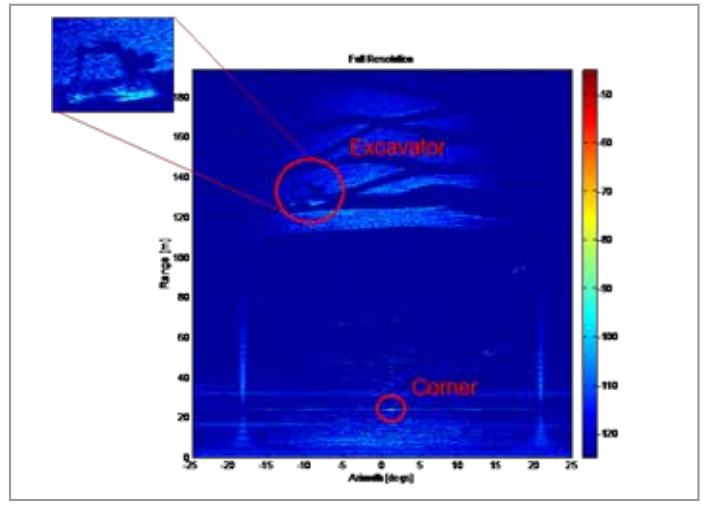

fig. 3 - image focalized

NANOFABRICATION AND CHARACTERIZATION OF HIGH-PERFORMANCE GRAPHENE FIELD-EFFECT TRANSISTORS

Aida Mansouri - Supervisor: Prof. Roman Sordan

The unique electrical, thermal, and physical properties of graphene have attracted research interest both from the electronics and materials science communities. Scaling limitations in conventional silicon technology, pushed the technology market to investigate new materials. Graphene, with exceptionally high mobility and being cheap, flexible, transparent and a true 2D material, gained much attention in the field of nanotechnology. Among industrial applications of graphene, flexible electronics is one of the most promising. In graphene, the mobility is at least two orders of magnitude higher with respect to materials used inorganic electronics. However, fabrication of an entirely flexible electronic device (including batteries and contacts) is still technologically challenging. Moreover, graphene could be an excellent material for the implementation of various sensors due to its large surface to volume ratio, excellent carrier mobility, high thermal and electrical conductivity, and unique optical properties. pH sensors, environmental contamination sensors as well as strain and pressure sensors are the most developed applications of graphene-based sensor technologies. Graphene could also be used as a material for energy generation and storage. The idea of flexible power storage and supercapacitors could be revolutionary in wearable and flexible electronics. The general applications of graphene are summarized below.

Since the discovery of graphene in 2004, graphene electronics have made significant improvements. Many devices from single transistors to more complex analog and digital circuits have been demonstrated. However, along with progress, the physical limitations of graphene have become more evident. The dominant part of this work concerns nanofabrication and characterization of graphene field effect transistors (GFETs) for high frequency applications. Using GFETs as building blocks, more complex devices were demonstrated, such as graphene inverters.

During this work the main limiting factors in RF graphene applications were diagnosed and then optimized. Some of these factors are: Contact resistance R_c, access resistance R_A, and gate resistance R_G. Transfer length method (TLM) devices were fabricated to investigate the contact resistance in graphene. Different metals such as Au, Ni and Pd have been tested to obtain metal contacts with graphene with the lowest contact resistance. Among the various metals tested, pure gold with R_c = 190 Ωμm exhibited the lowest contact resistance with graphene. Further studies on contact resistance showed that the contact geometry could affect the contact resistance. By etching holes in the graphene channel below the contacts, the edge of contact between graphene and metal contact was increased. This resulted in an outstanding decrease in device contact resistance. The lowest contact resistance achieved in TLM devices with a holey pure Au contacts was R_c = 24 Ωμm.

The access resistance was minimized by using T-gate and self-aligned layouts in which the source-gate and drain-gate distances in the GFETs were minimized These adjustments were applied to RF GFET to test and characterize their performance at high frequencies. S-parameter measurements were performed on GFETs and a small signal model was designed and optimized within a frequency range of 10 MHz to 50 GHz using advance design systems (ADS). The model shows very good agreement with S-parameter measurements. From S-parameter measurements, RF FOMs were also extracted, revealing an exceptionally high f_{max} / f_t ratio above 3. Moreover, outstanding intrinsic gain Av > 30 dB (in GFETs with channel width, W = 10 μm and gate length, L = 1 μm) and high $|S_{21}|$ > 12.5 dB (in GFETs with W = 100 μm, L = 1.1 μm) were extracted from S-parameters, which outperformed GFETs. The maximum cut-off frequency f_t = 13 GHz and maximum frequency of oscillation f_{max} = 27 GHz were measured in GFETs with holey contacts (with W = 10 μm, L = 1 μm). Thanks to ADS, the parasitic and intrinsic components of the GFET small signal model were de-embedded to characterize their influence on FOMs of GFETs. From these analyses it has been found that the gate resistance R_G is the main limiting factor suppressing f_{max} . Measured f_{max} in GFETs with R_G was 17.7 GHz and after deembedding R_G, f_{max} grew to 32.3 GHz. Further improvements were seen by passivating graphene by hexagonal Boron Nitride, the improvements in quality of the oxide-graphene interface, and by utilizing smooth and charge-trap free substrates.

Fig. 1 - Summary of graphene applications in electronics

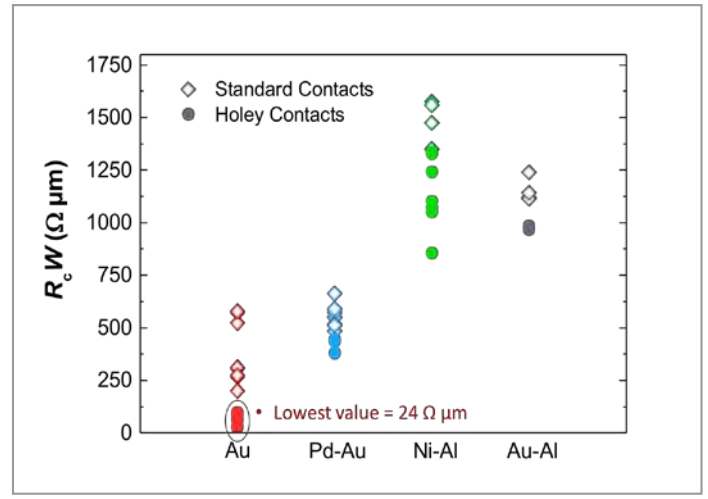

Fig. 2 - Contact resistant results with holey and standard contacts. The contact resistance was decreased by introduction of holes. Different metals were included in this investigation. The lowest contact resistance was obtained with holey pure Au contacts, Rc = 24 Ω μm

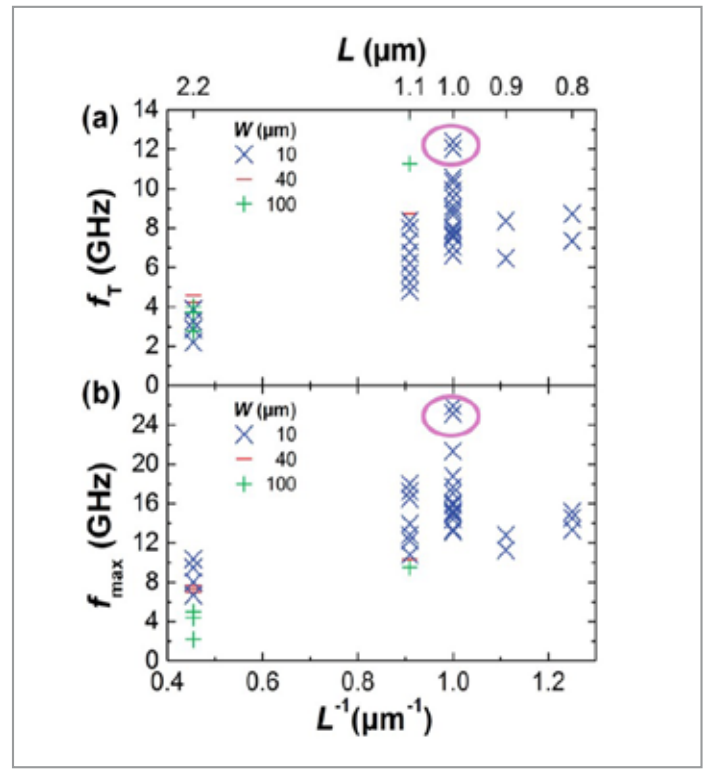

Fig. 3 - Extracted f_{max} and f_t from S-parameter measurements performed on GFETs with various *W* and *L*. The plot shows the maximum f_{max} and f_t regarding to each device. The largest f_{max} and f_t are marked by circles

BEYOND FULL RATIONALITY: MODELING TRADEOFF DYNAMICS IN MULTI-OBJECTIVE WATER MANAGEMENT

Emanuele Mason - Supervisor: Prof. Andrea Castelletti
Co-Supervisor: Dr. Matteo Giuliani

Anthropocene denotes the scale and intensity of the anthropic influence on natural processes and ecosystems. Recent trends in the scientific literature on natural resources acknowledge this issue by putting forward the concept of Coupled Human-Natural System (CHNS). It denotes systems where the human and natural components are so entangled that a correct assessment of system resilience or sustainability require a comprehensive study of both parts. Within the human component, a crucial role is played by decision making, which mediates human interactions at the various levels of system governance, ranging from institutional to operational decisions that directly impact the natural resource. Modeling of decision making has been a daunting challenge for researchers. However, they developed various approaches, among which the normative approach.

In the normative approach, it is assumed that an agent's decisions seek to rationally achieve a certain goal. The rationality hypothesis enabled a plethora of theoretical studies in various fields, e.g., economics, or optimal control theory, which supported the broad adoption of this modeling approach to provide prescriptions. However, the same hypothesis has been strongly criticized on the basis of empirical evidence of behavior deviations. In particular, full rationality have thwarted modelers' efforts to deal with systems operated for multiple objectives. In these systems, the operating policy has to balance multiple goals, by reflecting the preferences of the decision maker and/or of the stakeholders. Another issue arises from the time dynamics of the tradeoff and preferences that are unlikely to remain stationary but are instead adjusted in response to various changes. Triggers of the change may be exogenous influences that modify the conditions at the system boundary, or extreme events, such as floods or droughts, originated by the inherent variability within the system.

The objective of this thesis is to advance algorithms adopting the normative approach to develop behavioral models of system operators. The proposed algorithms are able to cope with tradeoffs among multiple objectives, and with the time evolution of preferences.

A first effort has been devoted to formalize the modeling of multiple objectives by recognizing the inherent uncertainty in their formulation. This leads us to adopt the idea of rival framings, each representing a set of objectives formulations, to rationalize the search for the candidate set of objective functions that represents the operator of the modeled system.

On this premise, we then propose two different algorithms to identify the tradeoff among the multiple objectives that best represents the historical operations in the modeled CHNS. The first algorithm adopts Inverse Reinforcement Learning to efficiently identify a set of weights that measure the preferences of the system operator as they can be presumed by a time series of system operations. This algorithm is able to achieve high quality, above 0.9 goodness-of-fit, in a synthetic application. We also applied it to a real case study, effectively improving the operator's behavioral model with respect to single-objective counterparts. Moreover, we were able to quantify the effect of an exogenous transition on the system in terms of change in the weights of operating objectives.

We also developed a second algorithm for tradeoff identification, inspired by multi-agent negotiation protocols, called Set-based Egocentric Concession protocol (SEC). Operator's behavioral models identified with this algorithm prove to be accurate, as we tested on a synthetic case study. Moreover, SEC identifies the tradeoff as a function of a set of parameters, named attitudes, that can be used to model tradeoff evolution in time. To this end, we propose an autoregressive model of attitude evolution driven by the recent system performance as they reflect the extreme variability of the system, e.g., in terms of droughts and floods. We found this model a promising start to explain the evolution of the tradeoff of a time series of decisions with dynamic preferences, developed for the synthetic case study.

More significantly, we framed the testing of the proposed model of preference evolution in a scientific approach that has significant implications for the construction of reliable projections of the future evolutions of CHNS.

ROBUST CONTROL OF AUTONOMOUS SYSTEM FOR WIND ELECTRICAL GENERATORS

Soufiane Meddouri - Supervisor: Prof. Luca Ferrarini

Facing the exhaustion of fossil energy resources and environmental problems caused by the exploitation of these resources, other alternative energy resources have been and should be developed. One alternative is to use renewable energy resources, which however show natural fluctuations of the main source. These energies include wind power which represents a huge resource with more than 54 GW of power installed across the global market in 2016, its capacity has expanded rapidly to 486.8 GW, cumulative capacity grew by 12.6%. All wind turbines installed worldwide by mid-2015 can generate 4 % of the world's electricity demand.
Faced problems and limitations on the development of wind farms

The production of electricity from wind energy is facing a significant increase, nowadays we mainly build wind farms of high power and high efficiency, since the first wind farm established in 1980. , the technology is widely developed with a legislative and tariff framework which has been set up for the development. In order to better explore the effectiveness of the wind energy production, the wind farms are usually installed in rural sites, or in off-shore where we can make the most of open spaces and windy. Nevertheless, there are few inhabitants and little electricity consumption there. So the problems of transmission of electricity over long distances become very meaningful. We need a great investment for cables, pylons, and loss during transport is relatively important, which can reach 10% . In addition, till now we do not have a good and effective storage system, especially for long term storage system. In this context, it's important to consider the possibility of development of the small production units near to the users. Moreover, the existence of big towers and wind effects caused by buildings offer favorable conditions for installing small wind turbines in cities. Indeed, they are increasingly popular for autonomous operations to ensure supply of isolated sites
The present thesis deals with the study, control and performance improvement of a chain of wind energy conversion into electrical energy in autonomous operation, including a storage system based on a flywheel.
It concerns a study, design, implementing and comparison of different control schemes with different control techniques applied to stand alone wind system for variable speed wind turbine with saturated induction generator. Based on Field Oriented Control (FOC) principle, fuzzy logic control technique (FLC), Lyapunov-Based control and Model Predictive Control (MPC) have been proposed with different strategies and situations. Namely MPC based single model, Switched MPC and Adaptive MPC strategies have been performed, the same switched principal strategy also was adopted for the Lyapunov control, where the switching strategy is based on different models according to different situations of the model in several working points. The linearized models are obtained by means of identification around equilibrium point procedure. A special variable-step smoothing technique was introduced as well which provides good performances in various situations, the effectiveness of the developed control schemes is evaluated through several comparisons with traditional control techniques in different situations and scenarios.
The above mentioned studies require the modeling of the different parts of the studied system and the determination of the different parameters needed for the control. In particular, the nonlinear model of the generator which takes the saturation effect into account of the magnetic circuit is represented by means of variable inductance. It is approximated with a polynomial function of degree 12. Using the saturated model of the induction generator allows the study of the system with good accuracy.
Due to the significant fluctuations of the wind as well as the load, which introduce an unbalance between the power production and the power needed by the load, that can cause different control problems to the system, an extension of system has been validated by introducing a storage system using flywheel, in order to improve the electric quality and contribute in the ancillary service. The flywheel is then connected by means of an induction machine which is operating as motor and generator for storing and retrieving energy respectively. The mechanical part based on a wind turbine and gearbox have been modelled as well and connected to the saturated induction generator model considered in the system, where a suitable control technique has been designed, using an optimal torque control (OTC) to track the maximum power point (MPP) of the wind turbine. An interesting association and control of the whole system (wind turbine, generator and the storage system) have been achieved with good and improved performance of the wind conversion system, in terms of DC bus voltage control, and the control of the power delivered by the wind and required by the load, regardless the different disturbances and parameters uncertainties of the system.
The results presented in this thesis are very interesting and relate to subject that continues to generate research in terms of effective control of a device with non-linear characteristics combined with uncertainties on the value of the quantities identified. The approaches adopted in the thesis are relevant in the sense that the adopted control algorithms are introduced and applied in order of increasing complexity while showing their contributions in the control, more and more effective of the DC voltage at the output of the rectifier.

TOWARDS FULLY-INTEGRATED FREQUENCY-MODULATED MEMS GYROSCOPES

Paolo Minotti - Supervisor: Prof. Giacomo Langfelder

Producing cheap and stable vibratory MEMS gyroscopes that do not require expensive and time-consuming calibration procedures has proven to be extremely challenging, due to the dependence of the scale factor (or sensitivity) of the sensor on fabrication imperfections and environmental changes (e.g., temperature variations). Conventional microgyroscopes, commonly operated as amplitude-modulated (AM) systems, are very sensitive to the accuracy of the drive-mode displacement control, to the value of the frequency mismatch between the drive and the sense modes, and to the electronic gains in the sense readout chain. The result is a significant scale factor variability over both (i) offline unavoidable process spreads, and (ii) online variations of the ambient temperature and humidity. For manufacturers, a large amount of the cost of consumer-grade MEMS gyroscopes is thus given by expensive calibration routines: (i) the scale factor must be tuned on every sample at the rate table, and (ii) its dependence across the temperature/humidity operating range must be measured on a representative number of samples, to apply digital compensation. Resulting performance limitations, e.g., the residual scale factor drifts due to imperfect calibration, generally preclude their widespread adoption in forthcoming applications, e.g., inertial grade navigation and pedestrian dead reckoning.

An alternative architecture for angular rate sensors is represented by MEMS gyroscopes where the rate information is frequency-modulated (FM) onto the resonance frequency of the micromachined structure, rather than amplitude-modulated onto the displacement along one of its mechanical modes. FM sensing benefits from a direct frequency output, the ease of interfacing with digital signal processing, and large dynamic range. In most FM implementations published in the literature, the scale factor is a dimensionless quantity that depends only on the angular gain of the structure, thus significantly lowering the scale factor sensitivity to environmental changes and process spreads.

In addition to the advantages in reducing calibration time and costs, the FM approach promises huge benefits in the aforementioned applications where stability is of great concern.

One very promising FM approach was proposed by the Berkeley Sensors and Actuators Center, in 2014. They realized a small-footprint FM gyroscope, by implementing continuous-time mode reversal, also referred to as Lissajous FM operation, whose name is inspired by the trajectory that the proof mass follows in such configuration. By letting a mechanical structure with two main orthogonal vibration modes oscillate at each mode's resonance, an angular rate oriented orthogonally to both modes is inferred from an AC resonance frequency variation measurement. By implementing continuous-time reversal, slow natural frequency drifts are not aliased as rate variations. Their published prototype was implemented with discrete electronics, and only yaw-rate sensing was demonstrated.

This thesis reports the developments towards the full integration of a tri-axial, consumer-grade, frequency-modulated MEMS gyroscope, by addressing the challenges of (i) developing a low-power integrated circuit and (ii) designing the mechanical structure of both yaw and pitch (or roll) sensing in an industrial process. Concerning point (i), in a Lissajous FM gyroscope, the required scale factor stability is obtained at the cost of a relatively-small scale factor value, lower than 1 Hz / 360 dps. Targeting, e.g., a 10 mdps/rt-Hz resolution, the required frequency noise density is lower than 30 μHz/rt-Hz, that translates in an ultra-low-phase-noise requirement on the electronic circuitry, which should be, at the same time, low-power for always-on applications, making an integrated electronic design not trivial. A custom low-power, low-phase-noise integrated circuit was then designed, optimized for frequency-modulated operation. Concerning point (ii), a challenge in the development of 3-axis FM gyroscopes is the design of in-plane sensing systems, as one of the operational modes should move out of plane: due to the planarity of the micromachining process, this breaks the symmetry of the vibration pattern, increasing the complexity of both the mechanical design and the control electronics. Within this thesis, both yaw- and pitch-rate, analog-output sensing systems based on Lissajous frequency-modulation are experimentally demonstrated for the first time, by coupling the developed integrated circuit with two novel micromachined structures fabricated with a 24-μm-thick industrial process. In operation, both gyroscopes show a repeatable and stable scale factor, with less than 0.55% of part-to-part variability, obtained without any calibration, and 35 ppm/°C of variability over a 25° to 70° C temperature range, outdoing of an order of magnitude the performance of state-of-the-art amplitude-modulated solutions. The whole system achieves a 10-mdps/√Hz resolution with a 500-μA current consumption only, providing a ±2000 dps full-scale range and a 50-Hz sensing bandwidth. The design of a frequency-digitization circuit, required to provide a digital-output sensor, is also reported, but not experimented. As a conclusive remark, a thorough system-level performance comparison between the proposed architecture and conventional amplitude-modulated MEMS gyroscopes is carried out, aimed at highlighting both the advantages and the disadvantages of such an approach.

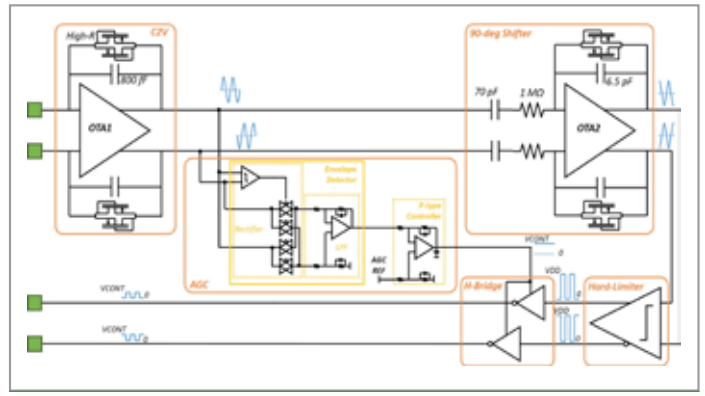

Fig. 1 - Schematic of the low-power, low-phase-noise integrated oscillator specifically designed for frequency-modulated operation of MEMS gyroscopes.

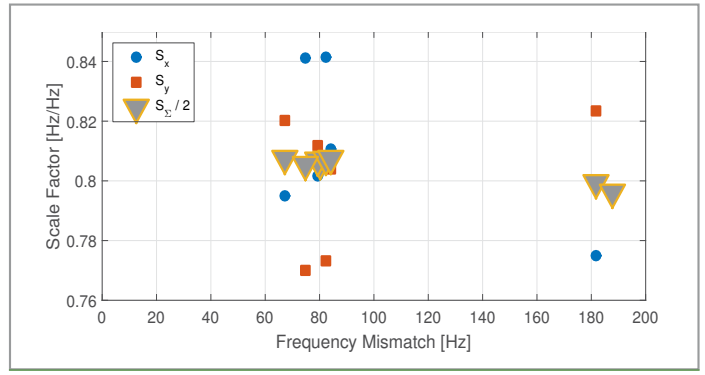

Fig.2 - Report on the scale factor measurements performed on eight different LFM yaw samples.

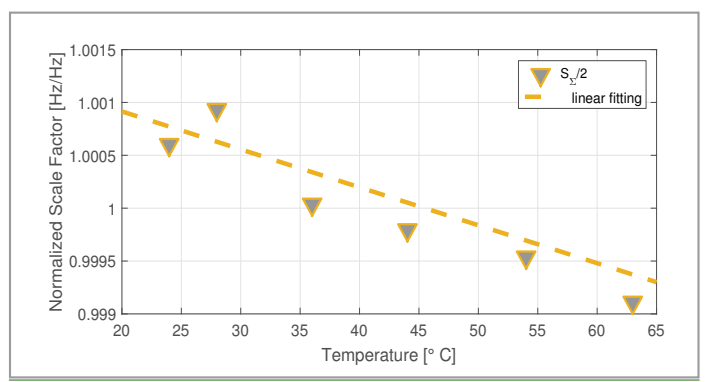

Fig. 3 - Report on the scale factor measurements on a LFM yaw sensor, performed at different operating temperatures.

PHASE-CODED BRILLOUIN OPTICAL CORRELATION DOMAIN ANALYSIS FOR FAST AND HIGHLY-RESOLVED DISTRIBUTED STRAIN AND TEMPERATURE MONITORING

Jacopo Morosi – Supervisor: Prof. Mario Martinelli

Monitoring, diagnostic and early warning are becoming key tasks for an efficient management of many industrial processes. Conventional sensing systems are based on electrical devices such as accelerometers, vibrometers, strain-gages and thermocouples, each one providing point measurements of a single specific parameter. However, optical fiber sensors can now provide an alternative sensing methodology, owing to important advantages such as resistance to harsh environments, absence of electromagnetic interference and small size, which allows them to be directly embedded in the structure to be monitored. A single optical fiber can be used both as the sensing element and the leading cable conveying information back to the receiver, which can thus be placed far away from the monitored area.
A further key feature of fiber optic sensors is their capability in providing highly-accurate multi-parameter information on the entire monitored structure in a truly distributed way, thereby replacing, with a single optical cable, a plurality of discrete sensors.
From the 80s, distributed fiber optic sensors have been the subject of extensive research. They rely on inelastic scattering phenomena, such as Raman and Brillouin and usually exploit stimulated configurations. When the stimulated backscattering configuration is used, access from both sides of the sensing fiber is required. This way, high signal to noise ratio is guaranteed, allowing for more accurate measurements. Extensive research on Brillouin-based sensing has been carried out continuously throughout the past 40 years, and many different configurations have been proposed, the vast majority of them exploiting the Stimulated Brillouin Scattering (SBS) process. The measurement principle relies on the intrinsic dependence on temperature and strain of the Brillouin Frequency Shift (BFS), that is the frequency detuning between a pump and a probe lightwave, counter-propagating inside the fiber under test (FUT), for which the Brillouin Gain Spectrum (BGS) exhibits a maximum. From the BFS, the strain/temperature distribution along the entire length of the sensing fiber can be inferred. The traditional time-domain based localization approach (OTDR), which enables to relate information given by the BFS to specific positions on the FUT, allows for spatial resolution which is typically limited to ~1 m by the acoustic-wave response time.

Yet, many predictive maintenance applications, such as thermoelectric generation, automotive and power cable manufacturing are now demanding smart diagnostic solutions able to provide a detailed information about temperature and strain distributions on machines, plants or mechanical components with a much better spatial resolution. For these applications the monitored distance might be quite limited, the main focus being more on a high number of sensing points, fast acquisition and cost-effectiveness. In this frame, phase-coded Brillouin Optical Correlation Domain Analysis (phase-BOCDA) can indeed be a potential solution fulfilling these new requirements. However, a few unsolved issues are currently limiting its actual in-field application. This research, starting form a basic phase-coded BOCDA scheme, mainly addresses these issues, proposing and integrating novel hardware and software solutions. At first, a novel two-steps estimation algorithm is applied to remove the background noise – which usually contributes for more than 90% on the received signal. This approach has proven to be extremely effective, allowing to always recover a good quality BGS from which the BFS can be precisely calculated. Experimental results have proven the ability of phase-BOCDA setup to recover the temperature and strain profiles along the fiber under test with errors always lower than 1 °C and 20 με respectively. Attention has also been given to the typical low measurement speed affecting BOCDA approach. To speed-up the measurement, single slope-assisted and double slope-assisted approaches have been applied for the very first time to a phase-BOCDA scheme. Proof of concept experiments have shown that, while retaining the 2-cm spatial resolution achieved by frequency-scanning method, the measurement time can be effectively reduced by 100 times, this value being limited by the available electronics. In particular, double slope-assisted approach proved to guarantee a higher accuracy also in presence of SBS gain fluctuations and a better sensitivity for small strain values. Finally, polarization issues affecting the efficiency of Brillouin amplification process, and thus reducing the measurement accuracy, have been addressed. Polarization scrambling has emerged as the most hardware-wise solution to avoid polarization fading issues, as it does not require consistent modifications of the BOCDA layout. However, the low scrambling speed ensured by commercial polarization scramblers based on electro-optic components can be a limiting factor when high-speed measurements are required. For this reason, we have investigated and developed a novel all-optical and completely passive device based on polarization chaos nonlinearly excited in an optical fiber through cross polarization interaction between a light wave and its intense backward replica generated by means of an amplified reflective loop. The proposed system can reach scrambling speeds as high as 610 krad/s with a Degree of Polarization (DOP) lower that 2.5%, thus clearly outperforming conventional scramblers. As the main elements involved in the scrambling setup are also required in the BOCDA scheme, a new phase-BOCDA layout has also been proposed, which directly integrates the all-optical scrambling capability.

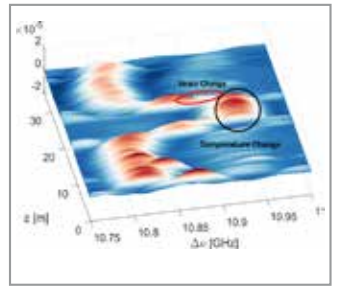

Fig. 1 - Example of a 3D BGS measured by phase-Coded BOCDA sensor. FUT sections subject to strain and temperature changes are highlited.

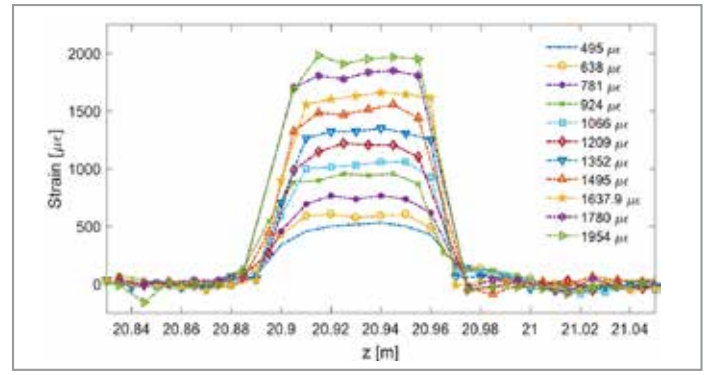

Fig.2 - Different strains measured on an 8-cm strained FUT section

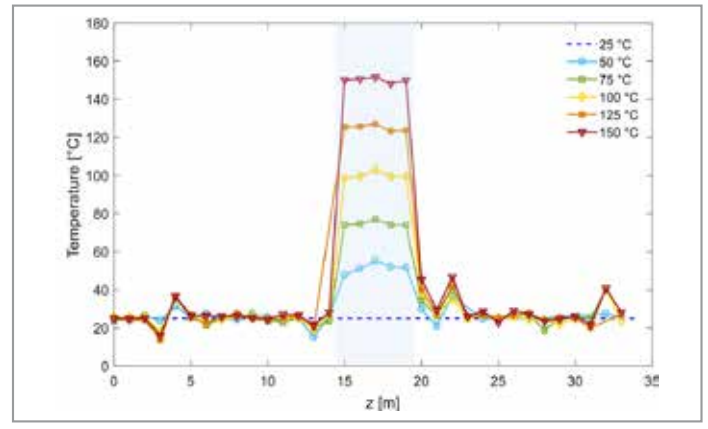

Fig. 3 - Different temperatures measured on a 5-m heated FUT section.

ALGORITHMS AND METHODS FOR THE DESIGN AND DEVELOPMENT OF INTELLIGENT, CONTEXT AWARE AND SUSTAINABLE MOBILITY SERVICES

Alessio Pagani - Supervisor: Prof. Francesco Bruschi

The quantity of data generated by public transport agencies, users and public transport oriented applications is constantly growing. This information is not fully exploited though, especially in the public transports field: just a few of the existing applications make use of this information and no service is able to offer a dynamic, real-time, individually-customized interaction. Aim of this thesis is to provide the necessary infrastructure to enable the development of innovative, context-aware services and applications in the field of the public transportation. A set of methods and tools has been designed to create the structures and the information required. These instruments have been integrated in a general framework able to reconstruct and update in real-time the public mobility state. The public mobility state can be defined as the intersection of three state dimensions: · The *public transport network state* represents the current position of each means and by their equations of motion; · The *road network state* represents the current state (i.e., travel speed, traffic flows, congestion) of the roads in a given area; · The *users state* is composed by the position of the traveling users (e.g., gathered using a smartphone application), their destination and, when on a vehicle, the matching with the means of transport they are traveling on. These three states are not directly obtainable, thus they have been reconstructed using publicly available data, such as open data, transport means timetables, actual arrival times upon each stop and users' GPS position. Each state has been elaborated in a different conceptual module and, together with the data fusion module and the API server module, they compose what in this thesis is defined as the Public Mobility Platform. The **Public Mobility Platform** manages all the information regarding the public mobility, the vehicles, the traveling users, the state of the transportation network and road network. This platform enables the design of user oriented planning and assistance services increasing the perceived quality of the public transport service and offering a better travel experience, reducing traveling time and improving perceived comfort, even in case of disruptions. Furthermore, this innovative platform can also supply public transport agencies with detailed information such as number of traveling passengers, commutation times and routes. The platform is composed of several modules, each of which deals with a conceptual task.

The **public transport network reconstruction module** is designed to reconstruct the position and the estimated motion laws of all the means of transportation in a given area, using information typically available, such as estimated time of arrival of the vehicles at the stations. This module uses a state-based Bayesian approach for the reconstruction of the transit system state from the limited publicly available knowledge. The proposed general approach has proved to be effective both in reconstructing the state of the transport network and in exploiting it to predict the position of the public transport vehicles in the near future, also in the presence of various real world kinds of noise, such as information blackouts. The **user contextualization module** integrates the location of users in the reconstructed public mobility state in order to estimate when the users are on a vehicle. Obtaining this information poses two main challenges. The first challenge concern the travel state of the users, that is when they are actually on the vehicle or not (e.g., they are just walking on the street close to a bus). The second challenge revolves around the association of these generic users to the vehicle they are riding, taking into account noise, limited precision of the sensors available (e.g., GPS, accelerometers) and possible presence of multiple means of transport in the same area. The proposed solution is based on a Particle Filter with a model that includes the probability for each user to be on a selected vehicle. The system has proved to be effective in presence of noisy data and to identify correctly when the users are actually on a vehicle or not. Thanks to underlying model, this approach is also convincing in the identification of the specific vehicle each agent is traveling on. The **road network reconstruction module** , through a knowledge discovery method that processes the real-time position of the car sharing vehicles available, estimates interesting metrics, such as travel time, congestions and vehicles flows at different times and in different days. The estimation of the road network state does not require any dedicated sensor or probe, instead it uses only information (historical and in real-time) about the car sharing vehicles movements during the day. This approach is used to reconstruct the road network in the cities where other services (with dedicated resources) are not available and can be used, as an additional source, to improve the accuracy of road network state in the cities where other services are already available. The reconstructed road network state can be provided to external applications and used to estimate the best routes for the private public transport services (e.g., car sharing). Moreover, this state is integrated with the public mobility state in order to improve the speed accuracy of the vehicles that travel on roads shared with the private transportation. A further module, the **network features classification module** , is designed to predict urban mobility scenarios and to explain them. The main topic is relative to the private public transportation (e.g., taxis and car sharing) and includes two main techniques based on explainable deep learning: the first one is focused on the design of interpretable deep neural networks, using neural sub-networks to represent engineered input features. These sub-network are used to explain how these features determine the predictions. The second technique revolves around the representation of geographical informations about travels using road graph and features selection. An *ad hoc* simulator has been developed to test the set of methods and tools and the Public Mobility Platform. The accuracy and the robustness of the approach have been demonstrated, even in case of missing data, extremely inaccurate GPS position of the users and unlikely corner cases. Afterwards, a working implementation of this platform has been developed for the city of Milan and, in collaboration with *TIM Joint Open Lab* , it has been evaluated, proving to have excellent performances even in real case scenarios. Nevertheless, the system is general an can be exported in other cities and easily adapted to the real-time data available. Summarizing, an innovative platform that enables interesting innovative public transport services is proposed. From a user point of view, the public transport applications can shift from a planning paradigm to an assistance one, where information is provided during the whole trip and not just before the beginning. These new applications can provide the maximum precision and accuracy considering the context of the users (e.g., the estimated time of arrival is constantly updated, taking into account the precise means the users are traveling on, they can play context-aware games and share news with the other passengers). From the transit agencies perspective, it is possible to evaluate metrics valuable to duly assess service quality, such as the number of connections taken, mean waiting times at connections and commutation time. If adequate relevant information is available to travel agencies in real time, the new generation of transport services could dynamically adjust service parameters, and immediately assess the effects, in order to dynamically improve service quality in closed-loop fashion.

A LEARNING APPROACH FOR PRICING IN E-COMMERCE SCENARIO

Stefano Paladino - Supervisor: Prof. Nicola Gatti

Over the past few years, there has been a significant increase in the use of e-commerce websites. Thanks to technological progress and the massive adoption of e-commerce, almost everything can be bought online, from groceries and clothing to holiday packages and cars. Market research shows that the one of e-commerce is a market with a global value of more than two trillion USD, and it will even grow in the future.

Online markets have many features that can be exploited by vendors, thus, with the advent of e-commerce, a number of new strategies became possible. For instance, prices can be easily adjusted without incurring in any cost, while, in traditional markets, price changes would often induce costs, since a new catalog had to be printed or price tags had to be replaced. Furthermore, in online markets it is possible to access historical data without substantial costs, making it easier for vendors to study customers' behavior in order to make more accurate and informed decisions.

With the spread of e-commerce, metasearch engines began to arise as well. These tools are so named as they conduct searches across multiple independent e-commerce and they aggregate the results, to allow customers to evaluate and compare the offers for a product more clearly. The scenario of online selling of travel products is a noteworthy example in which the role of metasearch engines is acquiring a great importance. In this scenario, we have Online Travel Agencies (OTAs) which provide online booking facilities for flight tickets, hotels and other travel-related services to customers. Some of the most famous OTAs are lastminute.com, Expedia and eDreams. Metasearch engines have emerged in this field as a response to the need of users to compare offerings without having to consult each OTA individually. The most famous metasearch engines are websites such as Skyscanner, Google Flights or Kayak. Their relevance has been increasing in recent years, and market analysis for the US shows that travelers, when asked about their last flight trips, were almost equally likely to have consulted metasearch engines versus OTAs websites, with roughly three-quarters of the total doing so.

Our work investigates the pricing problem in the setting of online sales of digital goods from the point of view of an e-commerce, such as an OTA, selling its products in a metasearch environment.

Metasearch engines send a lot of traffic, then resulting in sales, to OTAs' websites. From the data of one of the major European OTA, we saw that more than a half of the profits are made from the sales on metasearch engines. Thus, this scenario presents a great profitability, but also a number of characteristics which makes the problem very challenging. We have a vast catalog of items to price. We have almost no information about our customers since they do not directly use our websites but they go on metasearch engines, which actually act as middlemen: they send a lot of traffic to OTAs' websites, but they give OTAs very few information about customers' behavior and purchasing history. Furthermore, most users perform searches without the actual intent of buying but only for informational purposes, generating a huge amount of searches performed every day, but with only few of them converting into bookings. Another difficulty arises from the non-stationarity of the environment, since we have seasonal effects on the market and we have a lot of competitors which impact on the non-stationarity of the environment by changing their marketing strategies. All these characteristics make the problem of finding the optimal pricing strategy really complex and with a lot of variables to take into account. It is tough for a human operator to tackle this computational burden, considering all the facets of the problem.
In our work, we study the problem of finding the pricing strategy that maximizes the expected profit of an e-commerce. We design an automatic pricing system which uses clustering techniques to partition the catalog in contexts of items sharing similar features, and online learning techniques to learn the optimal price of each context.

First, we tackle the clustering problem by learning from historical data collected by recording the interactions with customers. We focus on the Learning from Logged Bandit Feedback (LLBF) setting. Commonly, the logs generated by the interaction between the system and a user present the structure of a sequential decision process: basing on a context, the system takes an action from a set of possible choices and, afterwards, the user provides the system with a feedback, in terms of reward. The peculiarity of this setting is that the feedback, as it happens in bandit settings, is only on the chosen action showed to the user, while no information is available about other possible actions. Some approaches had been proposed to address this problem, but they lack in some of the fundamental characteristics that make an algorithm suitable for practical purposes. Indeed, they did not provide a clear interpretability of the final model since there is no direct method to infer those features that most influence the resulting model. Moreover, in economics scenarios, the proposed algorithm should be as risk-averse as possible, but most of the theoretical guarantees available in literature are provided in terms of average value. Finally, existing approaches usually require knowledge of the behavior of the user and assume it to be stationary, which is rarely met in practice in microeconomics scenarios. We propose a novel algorithm, whose goal is to solve all the mentioned drawbacks of the literature approaches. The algorithm we propose can learn a risk-averse policy to maximize the expected profit and makes use of statistical lower confidence bounds to build a decision tree, which provides both a decisional tool over future samples and an instrument to highlight the features that influence the profit the most.

Then, we study algorithms to learn the optimal policy to follow in each context to find the price that maximizes the expected profit. We study online learning techniques, in particular the Multi-Armed Bandit (MAB) ones, which have been widely studied in literature and provided evidence to be effective also in real-world scenario. MAB problems have been tackled with two distinct approaches, the frequentist and the Bayesian ones. The goal of a frequentist algorithm is to achieve the best parameter-dependent performance, and the expected mean rewards corresponding to the arms are considered as unknown deterministic quantities. Conversely, in the Bayesian approach, each arm is characterized by a distribution corresponding to the arm parameter. Even if it is possible to use existing general-purpose algorithms to solve our problem, by exploiting the pricing structure we can improve the performance of the classical algorithms. More specifically, we exploit the monotonicity of the conversion rate in the price and the a priori information e-commerce sellers have about the customers' behavior and the maximum conversion rate. To the best of our knowledge, these two properties have never been studied before. Furthermore, we tackle both stationary and non-stationary settings, as already done in literature. Finally, we study the property of unimodality over the expected profit. We present algorithms exploiting one or more of these features at the same time, also providing theoretical guarantees for the proposed methods.

These are the techniques we used to design an automatic pricing system, deployed in collaboration with one of the major Online Travel Agencies in Europe. The two problems of clustering and online learning algorithms are interconnected and continuously communicating: the data generated from the interaction of the users with our MAB algorithms are collected and passed through our LLBF algorithm to update the contexts model and improve the performance of the system.

NOVEL CONTRIBUTIONS TO ROBOT FORCE CONTROL FOR INDUSTRIAL MANIPULATORS

Matteo Parigi Polverini - Supervisor: Prof. Paolo Rocco

Control of robot interaction with the environment, generally referred to as robot force control, is required to face the inadequacy of pure motion control for the successful execution of those robot tasks involving contact with a surface. Widely popular since the early 1980s, research on force control algorithms employing a conventional (single arm) robot has gradually lost its appeal during the last decade, despite the growing employment of robots in finishing and machining operations would strongly benefit from increased controllers' performance. At the same time, the recent diffusion of new industrial robotic platforms, like light-weight dual-arm robots, has driven research on robot force control towards the execution of complex and dexterous robotic tasks, such as bimanual automated assembly. In this regard, the research findings described in this thesis cover two main areas of robot force control: performance improvement in implicit force control for traditional industrial robots and force controlled bimanual assembly through constraint-based trajectory generation for dual-arm robots.

1-Performance in Implicit Force Control

Implicit force control is a force regulation approach developed for position controlled industrial robots, typically required in robotic polishing, deburring and machining. It is based on an external force control loop, closed around the inner robot positioning system. Improving the performance of this controller with the novel tools developed in this thesis is one of our targets. Two requirements for an improved control performance have been introduced and addressed: fast convergence speed with absence of force overshoots and avoidance of environment modeling and identification. Experimental validation has been performed on the COMAU Smart Six robot, a 6 degrees-of-freedom position controlled industrial manipulator manufactured by COMAU, see Fig. 1.

Fig. 1 - The COMAU Smart Six position controlled industrial manipulator.

In order to achieve the first performance requirement, a constrained control approach has been developed which extends the so-called invariance control by exploiting set invariance conditions for output regulation purposes. By introducing a constraint on the regulation error, it is possible to derive a control law that achieves fast converge speed and absence of output overshoots with respect to the reference, by means of set invariance and Lyapunov stability conditions. The proposed control law, formulated for single-input single-output systems and extended to multi-input multi-output systems, has been applied to the implicit force control problem and experimentally validated. Controller robustness to compliance uncertainties has been addressed first in an adaptive fashion, and later by exploiting robust set invariance for a smooth convergence to the force reference.

Fig. 2 - Snapshot from a peg-in-hole experiment.

With respect to the second performance requirement, a mixed data-driven and model-based control approach has been proposed. A model-based feedback loop (accounting for the robot configuration-dependent joint and link compliance) aims at making the robot-environment system entirely depending on the unknown environment transfer function, while a data-driven PID controller, tuned with the Virtual Reference Feedback Tuning Algorithm (VRFT), is responsible for the force regulation. A hierarchical architecture is further proposed to enhance the control performance, in terms of a high bandwidth of the closed-loop system with avoidance of force overshoots. The introduced data-driven/model-based controller, is combined with an outer Model Predictive Controller (MPC) acting as reference governor to select on-line the optimal reference to the inner closed-loop system.

2-Force Controlled Bimanual Assembly

The redundancy, dexterity and inherent compliance of bimanual light-weight robotic systems, commonly referred to as Cobots, naturally motivate their employment in autonomous assembly operations, e.g. insertion, folding, capping/tapping. However, while the lower inertia combined with the compliant structure of this type of robots provides an intrinsic degree of safety towards manipulated objects, which is beneficial in assembly operations, the consequent lower position accuracy makes the application of force control algorithms even more crucial for a successful task execution compared to traditional industrial (single-arm) robots. As an additional trade-off, the dexterity of a dual-arm system (allowing for the execution of tasks with an increased level of complexity compared to traditional positioning tasks) and its kinematic redundancy (enabling the simultaneous execution of multiple task with a given order of priority) introduce additional constraints on the planning problem of a robot trajectory. Based on these considerations, it becomes clear that, when a bimanual light-weight manipulator is employed to perform an assembly task, real-time constrained-based trajectory generation and force control are crucial elements for a successful task execution. To this end, bimanual robotic assembly is treated in this thesis as an equivalent constraint-based trajectory-generation control problem fulfilling the typical force control requirements involved in assembly operations, i.e. a complaint robot motion and a robust bounding of the interaction force. Estimation of the contact force/torque further enables sensorless execution of the assembly process. The validity of the proposed control approach is experimentally validated on a bimanual assembly use case involving both a peg-in-hole insertion and a cap rotation task, performed with the ABB YuMi dual-arm robot, see Fig. 2 and Fig. 3.

Fig.3 - Snapshots taken from an experiment of bimanual cap assembly.

TIME-CORRELATED-SINGLE-PHOTON-COUNTING SYSTEMS: CHALLENGING THE LIMITS

Pietro Peronio - Supervisor: Prof. Ivan Rech

The analysis of optical signals by means of Single-Photon Avalanche Diodes (SPADs) has undergone a huge spread in recent years, thanks to the achieved ultimate sensitivity, which raised the interest of several applications from different fields, both scientific and industrial. For instance, systems featuring single-photon detectors are effectively employed for Fluorescence Lifetime Imaging Microscopy (FLIM) and Forster Resonance Energy Transfer (FRET) in life science, for Laser Imaging Detection and Ranging (LIDAR) in remote object sensing and for Quantum Key Distribution (QKD) in quantum cryptography.
Besides providing single-photon sensitivity, SPADs can also detect the arrival time of photons with a jitter as low as few tens of picoseconds. These two features are combined by Time-Correlated Single Photon Counting (TCSPC), which is a very efficient technique for measuring weak and fast periodic signals. The remarkable timing precision that is achievable made TCSPC be acknowledged as the gold standard for FLIM and spread it in the life science field in general.
Despite the remarkable sensitivity and timing precision that TCSPC can achieve, this technique intrinsically suffers for a relative "long" measurement time, due to the fact that the measure has to be repeated many times to record enough photons and reconstruct the desired statistics. Aiming to overcome this limitation, many advanced TCSPC instruments have been developed, which feature several channels operated in parallel to reduce the measurement time. Nevertheless, if we look at state-of-the-art instruments, we clearly see that there is a strong trade-off. Indeed, on one hand, commercial systems feature few channels but with remarkable performance, especially in terms of timing precision, differential nonlinearity (DNL) and conversion frequency; whereas, on the other hand, research systems feature a significantly higher number of channels, even few tens of thousands, but with performance that is far away from the best in class.
In multichannel systems, the huge number of channels is obtained by exploiting the standard CMOS process, which allows the integration of both the sensor and processing electronics on the same die. In this way, very dense array can be developed, but the performance is limited, because the technology is not optimized for the design of SPAD detectors and the power consumption is a constraint as well.
I decided to follow an orthogonal approach to break the existing trade-off, by first defining the maximum throughput that the data link toward the PC can sustain (about 10Gbps) and then tailoring the number of acquisition chains to saturate it. Since in advanced TCSPC system the ASIC and system designs are strongly connected to get the desired performance, I contributed to the development of an innovative routing architecture and a fast TAC.
The purpose of the designed router is to connect the fired detectors to the lower number of acquisition chains available. Despite this solution's being easy at a first glance, it is anything but trivial. Indeed, the easiest way would be to divide the SPAD array into sub portions and associate each one to a single converter by exploiting static multiplexers, but it would not ensure the maximum resource utilization if the illumination were sparse. Conversely, the developed smart architecture extracts the fired detectors in a random and unbiased way, by assigning variable priorities to each pixel and selecting the highest ones. In this way, the reconstructed image is not distorted due the presence of hot pixels.
The fast TAC (F-TAC) is meant for working in conjunction with the smart router and process the received timing signals without blocking the data flow. The core of this device is a single-channel Time-to-Amplitude Converter that is parallelized 16 times to make the overall dead time negligible with respect to the laser operating frequency (typically up to 80MHz).

When operating together, the router and the F-TAC generate a data throughput of several Gbps that has to be properly managed in order not to lose information and impair the measurement speed. In particular, data should be managed both upstream, at the system level, and downstream, at the PC level, to avoid bottlenecks during the acquisition. As for the upstream management, I designed and developed hardware and firmware solutions to implement on-board histogram feature and time-tag mode. Regarding the downstream management, there are two aspects that have to be carefully considered: the physical communication with the PC and the software processing.
To this aim, I developed a USB 3.0 connection, since it is wide spread, and an Ethernet 10G link, for the most demanding applications, like High Content Screening (HCS), to deal with the data transfer and sustain the generated throughput.
As for the software processing, it is a crucial part of the system, because when running at full speed, the 10-Gbps throughput that is generated cannot be stored onto an external SSD, otherwise the storing capacity would limit the measurement duration.
Conversely, it is of utmost importance that the software receives the data and directly extracts the desired information. To this aim, a collaboration has started with another research group at Politecnico di Milano, to develop a custom software solution able to process the recorded photons without impairing the throughput.
Besides exploiting several detectors in parallel to reduce the measurement time, advanced TCSPC systems can interface with microscopes to perform a scan of the sample. Indeed, even if modern instruments can feature thousands of detectors, the number of pixels is not high enough for some FLIM applications. To this aim, I developed a general-purpose firmware interface to manage the synchronization signals and divide the acquisition into pixels and frames. The proposed solution can either receive the synchronization signals from the microscope or provide it with the properly phased signals.
I implemented some of the developed solutions on a 32-channel complete TCSPC system to test and validate them. The instrument, which features state-of-the-art performance, has been employed on a research project to distinguish various stages of aggregation of alpha synuclein (aSyn) in cells, which is a small, natively unstructured protein that can aggregate into insoluble structures that are toxic to neurons, a phenomenon closely linked to the pathology of Parkinson's disease. The system should enable the direct measurement of the efficacy of aggregation-inhibiting drugs.
The results obtained from the testing of the 32-channel system were state of the art and showed that a break of the trade-off was effectively feasible. These premises pushed me toward the design of advanced TCSPC systems that could feature a higher and higher number of channels. As a first technological step on this way, I worked on a 32x32 instrument whose main core is a SPAD array developed exploiting a custom technology. The main goal of this work was to analyze the main issues that affect the design of large TCSPC instruments in order to find possible solutions. In the presented system, the detector array is directly connected to the smart router, which selects five fired detectors at each excitation cycle and connects them to five acquisition chains containing the fast TACs. The communication with the PC is handled through the implemented USB 3.0 connection and the Ethernet 10G link, whereas the presented custom interface manages the communication with the microscope. This system is meant to fully exploit the bandwidth of the data link toward the PC and maximize the utilization of the acquisition chains. As already mentioned, the designed instrument is only a technological step toward the development of large multichannel TCSPC systems; indeed, when a faster data link toward the PC is enabled, also larger systems will be developed.

INTEGRATED OPTICAL PLATFORM FOR BIOSENSOR APPLICATIONS

Nicola Peserico - Supervisor: Prof. Andrea Melloni

Biosensors are essential tools for the daily life of everybody, used in various environments from clinical exams for patients up to control of water quality in distribution sites. There are several technologies available for implementing biosensors, but research in this field is always looking for new solutions, aiming for higher sensitivity, faster responses, wide dynamics range, solution suitable for low cost implementations, etc.
The three years of Ph.D. have been mostly dedicated to the research on integrated optics circuits for biosensing application. The field of biosensors is a wide multidisciplinary field regarding different competences to successfully arrive to a prototype. The research has mainly concentrated on the implementation of an innovative, robust, and reliable integrated optical device for the detection of several analytes.
The reported activities have been connected within a EU project, a Regional project, a collaboration with companies, and an internal project. Both label- free and label-based approaches have been explored. With label-free assay, the biotransducer has to convert the biological event without the support of any other kind of additional elements or processes. An optical device based on integrated microring resonators, whose resonance wavelength is affected by the presence of analytes over the waveguide has been conceived, realized, and tested. The platform showed important results, for example by detecting Ovalbumin protein (a Ricin A toxin simulant) down to a concentration of 400 pM. Detection of DNA was also achieved, focusing in particular on the functionalization of the microring resonator with a novel method based on Micro-Contact Printing.

We decided to move to label-based approach, and in particular to exploit a combination of optical and magnetic approaches, in collaboration with Nanomagnetism for Biology and Spintronics Group (NaBiS) of prof. R. Bertacco. In our research, we combined the use of magnetic beads as la- bels with our optical device, exploiting the invasive impact that these beads have on the optical mode. The first results showed an enhancement of the detection limit of almost three orders of magnitude when magnetic beads were used. Following this path, we designed a novel platform that increases the interaction of optical detection and magnetic properties. By the use of an electromagnet, we were able to actually move the beads. The oscillation of the beads induces a variation of the evanescent optical field and hence a fluctuation of the optical phase. If a biological molecule links the beads to the optical device, the oscillation applied to the beads is transformed into a stretch of the molecule itself and can be detected from the output optical power. This opto-magnetic platform has some advantages, the main one is the possibility of an on-off detection, since the signal is revealed just in presence of molecules, magnetic beads, and magnetic field. We tested this platform by using DNA strand to link the optical waveguide to the beads. The results show that the platform works correctly, as the signals matched with the expected values, in particular the stretch constant of the DNA matched with the observed displacement of the beads. A patent was filled to cover this innovative concept.

The second path was to improve the resolution in the detection of resonance wavelength shift. In the first experiments, detection of wavelength shifts was performed by spanning a tunable laser source over a certain wavelength range. With this method, achieving a resolution under 10 pm is very critical. To overcome these limitations, we developed an electronic-photonic platform together with STMicroelectronics (STm). The main concept is the locking of a tunable laser to a microring resonator through an electronic feedback loop. In particular, we designed an add-drop microring resonator, whose powers at the output ports are balanced by the tuning of a laser wavelength. In this scheme, biological events will induce a shift of the microring resonance that is tracked by the electrical signal that drive the laser. The platform required a lay activity in collaboration with STm in order to be designed and tested. As results, we were able to distinguish step changes in the wavelength shift down to 9 femtometers, three orders better than state-of-the-art.

In conclusion, with this Ph.D. work, we developed optical platforms for biosensing. The obtained results are beyond the state-of-the-art concerning optical biosensors. The close relationship with other research groups and industries have been essential for this path, allowing to generate innovative ideas and concepts. The experimental side of the Ph.D. has been wide and essential, in particular in the Photonic Device Laboratory and in Polifab. The future steps will focus on the integration of the two platforms in a single solution, capable of joining the main features of the both.

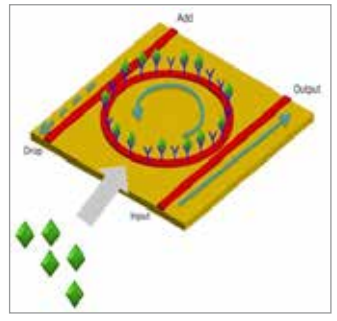

Fig. 1

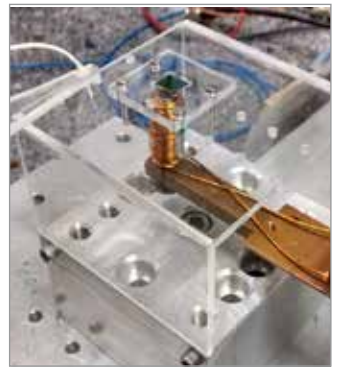

Fig. 2

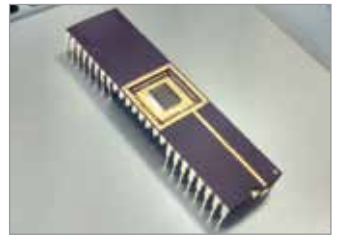

Fig. 3

ENERGY EFFICIENCY IN LARGE DATA-CENTERS USING PERFORMANCE EVALUATION TECHNIQUES

Riccardo Pinciroli - Supervisor: Prof. Marco Gribaudo

Data-centers power consumption is one of the greatest issues that IT organizations must face. Indeed, to enable future expansions, it is necessary to take under control the energy related costs. The U.S. data-centers energy consumption will be 73 billion kWh in 2020. Although it is estimated to increase only by 4% with respect to 2014, it is still an important problem, if considering that the worldwide amount of energy consumed for data-centers in 2014 has been estimated to be 270 TWh. Moreover, even if large data-centers have made great improvements in energy efficiency, small and medium data-centers – that are generally used in medical, retail, office, and education sectors – are difficult to monitor and may affect the global energy consumption trend due to their wide diffusion.
Besides being a huge cost for every IT company, the energy high demand of data-centers has also an environmental impact due to the emission of large quantities of greenhouse gases. Indeed, according to the literature, data-centers are estimated to be responsible for 2% to 10% of global CO2 emissions and the U.S. Environmental Protection Agency estimated that 67.9 million metric tons of CO2 have been released in the atmosphere in 2011 by IT industry.
For these reasons, the reduction of data-centers energy consumption is crucial when new infrastructures and applications are developed. Reducing energy consumption allows IT companies to increase their revenues in several different ways: lowering the electricity bill, thus letting the company save several money; reducing the carbon tax the organization must pay when the CO2 emitted by their infrastructure exceeds some thresholds; the governments may also apply some credit and cap-and-trade systems in order to limit the data-centers CO2 production. For instance, when a business exceeds allowed carbon dioxide emission levels, it must purchase either credits or carbon offsets from the market usually, less polluting companies to be permitted to keep on working and producing CO2.
Due to its relevance for IT organizations, the data-centers energy consumption problem has been largely studied by scientists from industry and academia. Many different techniques have been proposed, such as frameworks and tools to directly reduce the energy consumption of a server, or scheduling strategies to better allocate the incoming requests, thus being able to turn off unused resources. Nonetheless, several efforts are still required in order to further improve the available techniques and introduce new approaches. The main problems studied in this thesis are described in the following.
Power models. Over the years, several techniques and strategies have been proposed in literature in order to improve the data-centers energy efficiency. However, most of the proposals require a suitable power model that can estimate the power consumption of the system starting from simple system parameters, such as the utilization of its resources. Especially when the techniques considered are based on results that exhibit a non-linear behavior, the accuracy of the power consumption model is of paramount importance to correctly identify the optimal system configurations that can achieve the target performance with the lowest possible power budget. The typical use of analytic expressions for the power computation is the inclusion of energy characterization in models of the system defined using suitable formalisms, i.e., queuing networks, Petri nets, Markov chains. Taking into account widespread commercial power reduction techniques, like dynamic voltage/frequency scaling (DVFS) and multi-threading (SMT), we can provide a more detailed power consumption expressions that can be used to obtain more accurate estimates. Unfortunately, the available power models either do not consider these power-saving features, or do not account for both of them at the same time.
In this thesis, a model to estimate the power consumption of a multi-core CPU, when both DVFS and SMT are enabled, is provided. It needs some parameters that may be easily derived from operational and machine characteristics and some calibrating coefficients that must be derived from the experiments.
Multi-class workloads. The available techniques and metrics proposed to reduce power consumption through the application of efficient load control strategies rarely take into consideration the multi-class nature of system workload, i.e., jobs with different characteristics and behavior. However, this type of workloads may be composed by a mix of jobs that saturates different resources, thus it is an important system's feature which may be exploited in order to increase the utilization of each resource and decreasing the idle period of a server. Indeed, a scheduler that accounts for the mix of requests in execution could make the system work with an optimal mix. This way, the utilization of each server increases and the system's energy consumption decreases, since a strong relationship between a server power consumption and its CPU utilization has been proven in literature. However, the available power metrics analyze the power consumption of a data-center assuming it is processing only a single class of job. This assumption may decrease the accuracy of the power/energy metrics and, in the worst case, also make the user take a sub-optimal decision about the strategy that should be adopted.
To address this problem, a new energy metric and Pool depletion systems paradigm have been proposed. The former is exploited to evaluate the performance of energy saving strategies in a multi-class environment, the latter lets the users consider a system made of a pool of heterogeneous tasks that must be depleted in the shortest time by machines with finite capacity.
Epistemic uncertainty. When a model is adopted to study a physical system, its input parameters are usually estimated from observations. The value of the parameters is not the exact one since it is derived from a finite number of samples. For this reason, the input parameters of the model are assessed with uncertainty and are stochastic. Those input parameters become input random variables with a probability density function. Thus, the model does not return an exact value, but some stochastic results with a confidence interval. Differently from aleatory uncertainty that has been largely considered in literature and that is due to the natural variations of the physical phenomenon modeled epistemic uncertainty is introduced into the model by a lack of knowledge (i.e., finite number of observations) and needs to be propagated to the output. Albeit in the former case the uncertainty is reduced improving the model itself, in the latter one it may be curtailed by collecting a larger amount of samples for a more accurate input parameters estimation. In this thesis, epistemic uncertainty propagation is studied for M/M/1 queues and power consumption models.
Other systems and applications. Energy consumption has become such an important problem also due to the wide diffusion of portable devices (e.g., smartphones, sensors, etc.). Indeed, they are usually powered by a battery with a short lifetime, thus it is important to efficiently manage power and energy requirements in order to extend their life. In other words, reducing energy consumption is fundamental when dealing with mobile devices to increase the lifetime of their battery, differently from the data-centers case where energy efficiency is crucial to increase the companies revenues. For these reasons, analyzing energy consumption in small devices to make their lifetime longer is another important feature that may be enabled by power models. In particular, the case of Mobile CrowdSensing a contribution based paradigm involving mobiles in pervasive application deployment and operation is studied in this thesis.

MODELLING AND CONTROL OF AIRCRAFT ENVIRONMENTAL CONTROL SYSTEMS

Alexander Pollok - Supervisor: Prof. Francesco Casella

Unseen by the passengers, aircraft environmental control systems are complex thermodynamic systems, requiring a large quantity of power. The tasks of designing, modelling, optimising and controlling all these systems leave many degrees of freedom to the respective expert, and typically require many design loops to arrive at satisfactory results. This thesis contributes to multiple aspects of this process.
In aircraft Environmental Control Systems (ECS), Limit Cycle Oscillations (LCO) can occur. Those are problematic since the life expectancy of the ECS is affected. Using an Equation-based object-oriented modelling language (EOOML), a complete, detailed and dynamic simulation model of an ECS. This model includes the Engine Bleed-Air system (EBAS), the air conditioning pack, the cabin and ducting dynamics as well as the recirculation system. Using simulations, it is shown that LCO occurring in ECS cannot be explained by Helmholtz resonance effects.
To further investigate the cause of the LCO electropneumatic valves - as used in EBAS - are modelled in more detail, using the Lund-Grenoble friction model. Using this model, LCO in aircraft ECS are predicted for the first time.
Several control strategies are devised, implemented and evaluated against this model. A strategy based on a combination of feed-forward control, feed-back control and online tuning of the integral action outperforms all other candidates. A 46% reduction of the developed objective function is achieved when compared to the state of the art.

Current architectures for aircraft cabin climatisation only allow for a small number of temperature zones. Differences in heat load, generated for instance by nonconforming seating class layouts, cannot be compensated by the control system.
A new architecture is presented that allows for an infinite number of temperature control zones - at the cost of a more involved control system.
Suitable control strategies, as well as failure management strategies are demonstrated.

For optimisation studies in the context of simulation models, a class of controllers is found, based on Boundary Layer Sliding Mode Control These controllers do not require any tuning effort and show good performance for many systems during simulations. The high sensitivity to noise is unproblematic, as the system is purely virtual at this stage. These features make them suitable for modelling experts at development stages where the architecture design is not yet finalised.

On the most basic level, usability aspects of EOOML are explored. It is found that the use of inheritance can severely retard the understanding of simulation models. Results also suggest graphical representations to be superior to block diagrams with equation-based and algorithm-based representations taking the middle spot.

Fig. 1 - Design of typical Environmental Control SystemS

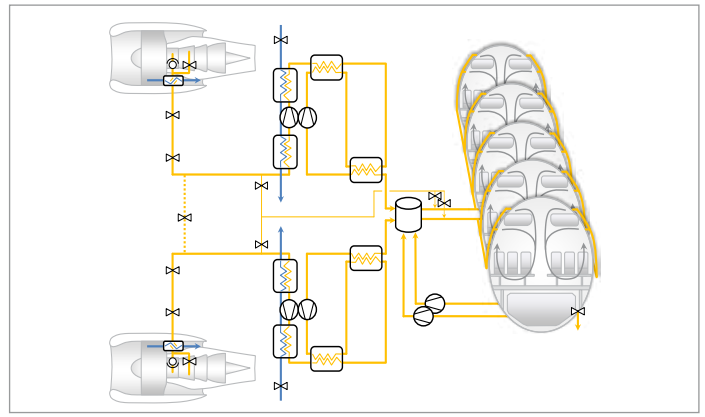

Fig. 2 - simplified structure of a state of the art Environmental Control System

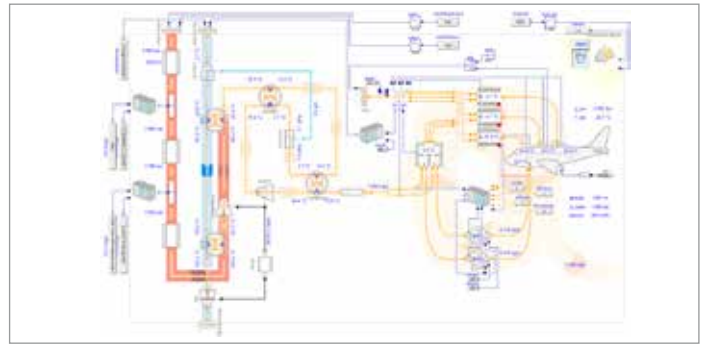

Fig. 3 - ECS model in the Modelica modelling language

ALGORITHMS FOR SEQUENCE-AWARE RECOMMENDER SYSTEMS

Massimo Quadrana - Supervisor: Prof. Paolo Cremonesi

Recommender Systems are one of the most successful applications of data mining and machine learning technology in practice and significant technological advances have been made over the last two decades. Academic research in the field in the recent past was strongly fueled by the increasing availability of large datasets containing user-item rating matrices. Many of these works were therefore based on a problem abstraction where only one single user-item interaction is considered in the recommendation process. The recommendation problem is therefore framed as matrix-completion, in which the missing entries in the user-interaction matrix must be predicted.

In many application domains, however, multiple user-item interactions of different types can be recorded over time. Most algorithms that are optimized for this particular problem setting cannot make use of the rich information that is hidden in the sequentially-ordered user interaction logs which are often available in practical applications. In addition, there are application domains, in which the items must be recommended in a certain order. Such situations are typically not covered as well in research setups that rely on a user-item rating matrix.

To address this problem, in the recent years researchers have developed a new breed of algorithms named *sequence-aware recommender systems* (SARS). Such algorithms can handle the information in user interaction logs by design without resorting on abstractions such as the user-item matrix.

This thesis focuses on the study of novel algorithms for sequence-aware recommender systems and their applications. We first provide a characterization of the problem; we highlight the relations and differences with respect to other related recommendation problems, namely recommendation based on matrix-completion, and with respect to context-aware and time-aware recommender systems. We provide an in-depth review of the state of the art, a categorization of the existing approaches and evaluation methodologies.
We then focus on the problems of *session-based* and *session-aware* recommendation. These problems have gained attention recently, given their proximity with many real-world recommendation scenarios.
We first validate the usefulness of personalized sequence-aware recommendations in session-based scenarios through a user study run in the hotel booking domain.
We then present novel sequence-aware algorithms for session-based and session-aware recommendation. In such a setting, we are given the sequence of the most recent actions of a user and the problem is to find items that are relevant in the context of the session and, when historical information on the user is available, match the user's general taste.
We investigate on models based on *Recurrent Neural Networks* (RNNs), the neural network configuration of choice for processing sequentially-ordered data. We show the effectiveness of sequence-aware recommenders based on RNNs in several real-life scenarios, namely session-based recommendation with rich product descriptors, personalized session-based recommendation for returning users, modeling of musical taste evolution, and automated playlist generation.
We devised an advanced Parallel-RNN mechanism trained with an alternated variant of the back-propagation algorithm that jointly exploits item features and identifiers to infer the intent of the user within the session and to provide more accurate recommendations accordingly (see Figure 1).
To address personalized session-based recommendation, we present a flexible architecture based on Hierarchical Recurrent Neural Networks that transfers information across sessions to personalize the user experience across every new user session (see Figure 2).
In our experimental evaluation, we empirically evaluate the proposed models on large datasets from several domains, namely video, classified advertisement, hotel, job and music. The results show that our sequence-aware models significantly outperform the state-of-the-art, showing that our models are indeed effective in several session-based recommendation scenarios in terms of recommendation accuracy.

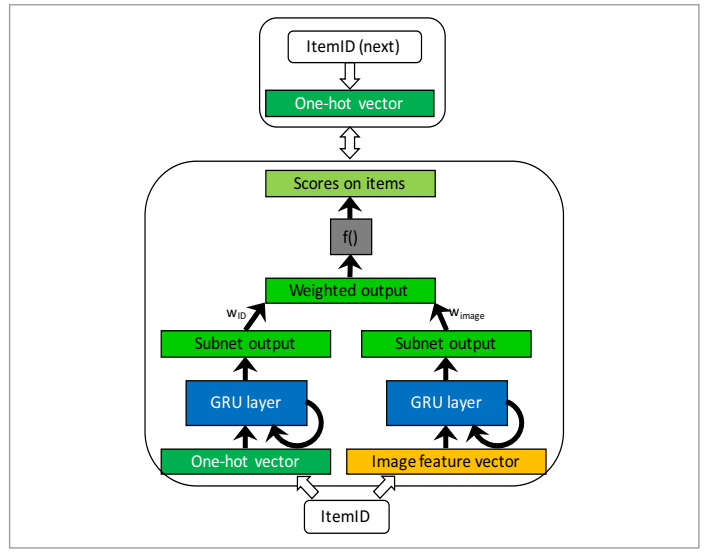

Fig. 1

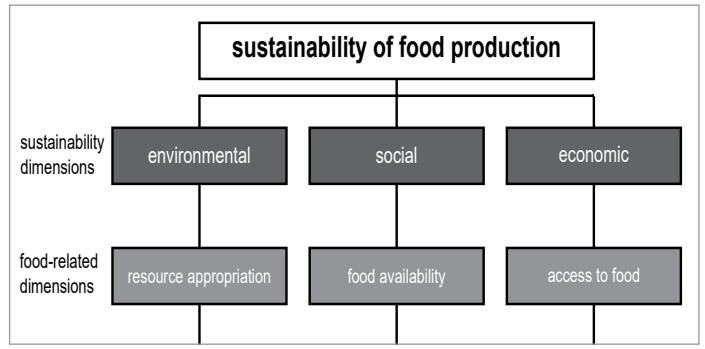

Fig. 2

EMBEDDED SYSTEM SECURITY: ATTACKS, IMPACTS & DEFENSES

Davide Quarta - Supervisor: Prof. Stefano Zanero
Co-Supervisor: Dr. Federico Maggi

Embedded systems ar are nowadays ubiquitous: smartphones, home appliances, medical devices, Industrial Control Systems (ICS). These device, often display a common subset of characteristics: limited processing power, limited available power, physical exposure, remoteness and unmanned operation, and network connectivity. Such characteristics create venues for peculiar attacks: we focus on software attacks, and in particular those stemming from the interconnection of a particular kind of device, Industrial Robot. Regarding defenses, we focus instead on mitigations to the aforementioned attacks, and strategies to strive for a more secure future of the industrial robotics ecosystem.

In the context of the Industry 4.0 revolution, these devices are interconnected, with other industrial robots, and more in general with other \Industrial Internet-of-Things devices (e.g., sensors, actuators, tools), and collaborate with human to shorten production times, and reduce costs. This revolution is bringing industrial robots closer to the forefront. The improvements in the way industrial robots work and communicate unlock new venues for attacks.

On the one hand industrial devices are designed according to strict safety standards in order to work in rough conditions, and to guarantee the safety of the human operator. On the other hand industrial robots are also designed to provide extreme flexibility and adaptability to different necessities, this is required both by the Industry 4.0 concept, but also by the very same definition of an industrial robot which needs to be a multipurpose manipulator.

Embedded systems have grown in complexity over time, and likewise, their interconnections and communication primitives, the ubiquitousness of these devices calls for more attention to their security, and to the security problems arising from their interconnection.

The Mirai botnet incident is an excellent example of what happens when IoT systems, once devised to be not remotely accessible, and not interconnected, are exposed to external networks. The Mirai botnet was composed mostly of IoT devices, and grew up to include a whopping 600k of infected devices. It's nevertheless easy to find many instances of security incidents in the embedded system world, in particular, the ICS environment in which embedded devices (e.g., controllers, sensors, actuators) are displaying peculiar challenges in securing them have seen an increased attention by researchers in the last few years.

The dissertation explores also defensive aspects: in particular we focused on the software security of a specific class of embedded systems, Industrial Robots, by exploring attacks and the extent of the impact of such attacks, moving later to to discussing mitigations and the role of standards in the security Cyber-Physical Systems (CPS). Moreover we propose a system to obtain transparent execution of arbitrary code in a TEE environment while achieving code confidentiality and being easy to use.

To the best of our knowledge, there is no systematic analysis of the attack surface and of the impact of attacks against industrial robots enabled by software vulnerabilities and architectural flaws.
In this chapter, we systematically analyze the feasibility of attacking a modern industrial robot by exploring concrete attack vectors that, when exploited, can subvert the interaction between a robot and the surrounding environment, thereby violating its basic requirements. In other words, we wonder to what extent, starting from the exploitation of the "cyber" components of a robot, an attacker can affect the physical environment.
To answer this question, we propose a domain-specific attacker model, discuss how certain combinations of software vulnerabilities enable classes of attacks unique to industrial robots (e.g., circumventing safety measures, impairing the precision of movements), and evaluate their potential impact. More specifically, we enumerate five classes of attacks, based on the observation that a robot working under normal circumstances should at least be able to read accurately from sensors, execute its control logic, perform precise movements, and not harm humans.

To show the feasibility of the attacks, we present a case study on a modern industrial robot sold by a major manufacturer. Due to standards and architectural commonalities among most modern industrial robots, the robot we chose is representative of a large class of industrial robots. Guided by our attacker model, we discover various vulnerabilities that allow an attacker to completely and remotely compromise the robot-controlling computers, and show how we used them to implement some of the proposed classes of attacks.

In summary, the first part of the work presents the following contributions: We define an attacker model for industrial robots, describing the capabilities of an adversary to successfully develop and convey attacks; Starting from the ``core" architectural features of a generic industrial robot, we describe concrete, robot-specific attack classes; We conduct an experimental security assessment on a de-facto standard robot;
We analyze how, concretely, an attack can leverage software vulnerabilities to carry out robot-specific attacks, evaluating their impact, and discussing the future security challenges. The second part of the thesis further explores novel aspects of the security of industrial robots, focusing on attacks and vulnerabilities originating from user interaction (i.e., programming the robot) and from the interconnection and integration of end effectors and IIoT devices to the controller: We systematize the existing knowledge about the attack surface of an industrial robot controller, and further expand it focusing on attack vectors arising from the connection with the external world, the expressiveness of robotic programming languages, and on use of the device's capabilities;
We analyze peculiar issues arising from the use of robotic Domain-Specific Languages (DSL) such as ABB's RAPID, and Universal Robot URScript; We highlight dangerous patterns and issues stemming from untrusted input processing and permission circumvention issues

In the last part of the thesis we critically analyze the findings discussed before, focusing on mitigations and countermeasures: We discuss mitigations for the attack we developed, and discuss the challenges involved in securing industrial robots; We discuss the existing standards for the cyber-security in the manufacturing ecosystem, and the specific challenges we must face in an effort to reduce the attack surface;

Last, we propose Tarnhelm: the first system leveraging a Trusted Execution Environment (TEE), specifically ARM TrustZone, for protecting code confidentiality while providing transparency for arbitrary parts of an application, moreover the tool is easy to use and provides transparent execution for a process running partially in the normal operating system and partially in secure operating system: We propose two new powerful and flexible primitives to achieve confidentiality as well as transparency; We design and implement these primitives in \tool, which allows for the transparent execution of arbitrary parts of an unmodified application in different isolated execution environments, thus facilitating its adoption by developers; Our evaluation shows that Tarnhelm presents a low performance overhead and modest additions to the Trusted Computing Base (TCB).

FAST AND FINE-GRAINED ELASTIC RESOURCE PROVISIONING FOR MODERN SOFTWARE SYSTEMS

Giovanni Quattrocchi - Supervisor: Prof. Luciano Baresi

In the last decade the increasing number of connected devices, the rise of cloud computing and the explosion of the size and velocity of data changed and shaped the requirements, the development process and the architectures of software systems. Nowadays, for example, webapplications are deployed and executed into the cloud, and single monolithic deployments are split in many modular services to individually develop and manage functionalities. Moreover specialpurpose frameworks for parallel batch processing are emerging to cope with large amounts of data. One of the challenges to be addressed on these software systems is how to provision and optimize resources to meet a varying demand since fluctuating workloads, unpredictable peaks of traffic, and unexpected changes are increasingly common phenomena. Without taking care of these contingencies, service providers struggle in satisfying functional and non-functional requirements, usually defined in SLAs (Service Level Agreements). To avoid resource saturation and unresponsiveness, users dissatisfaction and unnecessary costs, the provisioning of resources must be elastic, that means being capable of automatically adapt to changes in the execution environment that could affect the quality of service perceived by the users. Therefore resources allocated to a system should match as closely as possible the demand. One of the key metrics that captures the quality of a service of a system is the speed of answer that is often the subject of non-functional requirements defined in SLAs as a constraints over the time to answer one or more requests. Moreover the elastic provisioning of resources must be automatic to be less error prone and to reduce the complexity of managing the services at runtime. For this reason, elasticity can be materialized in autonomic systems whose behavior can change automatically according to the state of the system itself and of the environment. State of the art solutions focus on the control of software systems deployed in the cloud by changing the number of allocated virtual machines using mainly heavyweight techniques such as optimization problems or delegating to the system administrators part of the adaptation process (i.e., rule-based approaches). Moreover virtual machines are relatively slow to be provisioned (around six minutes on average) and only available in fixed configurations limiting how fast and precise the adaptation could be. This thesis aims to study, analyze and evaluate novel technologies and models that enable the creation of lightweight, autonomous, fine-grained and fast elastic resource provisioning for modern software systems. Its main contribution is the technique used to plan and enact the control actions. It mixes containers, a lightweight virtualization technology that enables the fast and fine-grained elasticity, and control theory that provides a lightweight theoretical foundation for controlling these systems. Containers can be provisioned in seconds and re-configured in milliseconds while control theory enables the computation of next allocations in constant amount of time. In the solution we used gray-box controllers: they capture just the main characteristics of the controlled system abstracting away many implementation details. We applied this methodology to two real-world scenarios: web and big-data applications. For the first area we present ECoWare, an autonomic system that allows containerized web-applications to scale their resources both at the VM level and at the container level. Furthermore, applications can combine this infrastructural adaptation with higher level (such as middleware or application level) adaptation actions. The experiments show that our planner outperforms Amazon's AutoScaling by 78% on average without containers; and that the introduction of containers allows us to improve by yet another 46% on average. Instead, xSpark is dedicated to the control of big-data batch applications. This type of systems manipulate huge quantities of data and users are often interested in quantifying and constraining the execution times (deadlines) for completing single runs. xSpark is a novel container-based extension to the Spark framework that exploits container technology to provide extremely fine-grained resource allocation. While vanilla Spark does not allow users to constrain deadlines, xSpark allows that and allocates required resources accordingly at runtime. The evaluation on four well known benchmark applications witnesses that xSpark was able to use less resources than native Spark and complete the executions with a less than 2% error in terms of set deadlines. Since this work mainly focuses on CPU allocation, future works include the support in the adaptation process of more resource types such as memory, storage and networking. Moreover the approach could be extended to handle the simultaneous control of web and big-data applications in heterogeneous cluster. Finally we plan to apply a similar methodology to new areas such as the Internet of Things and Fog computing.

ROBUSTNESS IN DATA-DRIVEN CONTROL: THEORY AND AUTOMOTIVE APPLICATIONS

Gianmarco Rallo - Supervisor: Prof. Sergio Matteo Savaresi

Model-based design represents a standard practice in industrial control. According to this concept, a model of the plant is identified from data or developed from first principles and then used to design a feedback controller satisfying some closed-loop requirements. Unfortunately, the resulting feedback controller is not necessarily optimal when connected to the plant, as it is designed using information about the model of the plant to control. In particular, the closed-loop performance is limited by modeling errors, which are typically due to incomplete knowledge of parameters or inappropriate "a-priori" model structure selection. In fact, a model suitable for control design needs to be both simple and reliable and the best trade-off is not easy to find. Moreover, a good model describing the main dynamics of a system may not be the best for control.

Standard model-based design is in fact an indirect data-driven procedure, in that the controller is derived from a model, which in turn is usually derived or refined using experimental data. An alternative is represented by "data-to-controller" algorithms. According to these approaches, only the final control cost is taken into account and the controller is directly obtained from experimental data without first deriving a model of the system. This feature may lead to a significant performance improvement in many practical situations. Several direct data-driven controller tuning techniques have been proposed, both iterative and noniterative. The advantages provided by data-driven control strategies are evident. The control design phase is much faster and the problem of undermodeling is avoided since no model of the plant is needed. For the same reason, direct approaches are suitable when the knowledge of the physics of the plant is poor or if the modeling cost is high due to the complexity of the system to control, as for automotive applications. Moreover, the structure of the controller designed according to direct strategies does not depend on the structure of the model. Therefore, the parameters of the controller structure are directly optimized from data and the reduction phase is unnecessary.

There are also drawbacks when dealing with data-driven control approaches. First, without the model of the plant it is not possible to check closed-loop performance before implementing the control law. Moreover (as is the case with indirect model-based strategies) direct methods are limited by real-world uncertainties that act at different stages of the design process and could limit the applicability of these techniques. Strictly speaking, a problem of robustness – in a broad sense – arises in this field. The following list helps to tidily provide an insight into the possible causes of direct approaches effectiveness reduction.

- **Robust controller design.** When linear controllers are designed for real-world plants, a prescribed level of robustness is necessarily required. Although this problem has been thoroughly addressed in the model-based literature, it currently represents a challenge for data-driven control approaches. In fact, the *classical* concept of robustness, *i.e.,* the one meant for indirect methods, can be seen as the capability of handling model uncertainties or model parameters variations throughout the controller design procedure and, for this reason, it strongly depends on the employed system representation. On the other hand, data-driven strategies exclusively rely on experimental data and there are not model uncertainties to deal with. However, the robustness property is here directly related to the plant under control (*e.g.*, cope with system dynamics variations or changing operating condition) rather than on its mathematical description. This difference contributes to highlight the potentiality of direct strategies but research efforts are needed in order to explore such a new and promising perspective.
- **Robust closed-loop reference matching.** If the selected fixed-order controller structure does not allow to match the reference model closed-loop behaviour, the data-driven parameters optimization provides a biased result. This problem can be tackled within the controller design phase by means of tuning a suitable prefilter. However, a non-optimal controller structure selection may emphasize direct approaches robustness problems.
- **Data uncertainty.** Experimental data are inevitably affected by measurement noise, which is one of the main causes of uncertainty for direct methods. The detrimental effect of the noise must be addressed during the controller design phase but this task is not trivial without exploiting a model of the system to control. Moreover, a closed-loop stability certificate can be derived from the same batch of data used for the controller tuning. This assessment is extremely useful for safety-critical applications. However, existent criteria need to be conservative for dealing with the effect of the noise but when the data-driven stability condition is not satisfied the designer has to lower tracking performance requirements. Lastly, a robustness problem arises also for what concerns closed-loop stability assessment procedures, which must consider system variations.
- **Experiment design.** The design of the experiment (*i.e.*, determining the input sequence to use for feeding the plant and gather data) is crucial for both system identification and direct controller tuning frameworks. While several contributions can be found in the "identification for control" literature, it is still necessary to explore experiment design for data-driven approaches.

The overall aim of this work is to extend the effectiveness of noniterative direct methodologies for control system design by proposing solutions for counteracting part of the above listed limitations. Addressing these critical points allows to fully compare data-driven techniques to indirect approaches, therefore it could be possible for the control engineer to be free of choosing the most suitable method for a given application.
The present study aims attention at time-domain approaches and it mainly concerns the linear framework. The focus on noniterative data-driven strategies, which allow to design a ready-to-use controller from a single batch of data, is motivated by their lower experimental cost combined with a null on-line computational effort that make these approaches more appealing/easy-to-use with respect to the iterative ones. Furthermore, assessing or improving the robustness for the iterative techniques requires additional experiments devoted to gather information at each iteration. Among all the noniterative techniques, the Virtual Reference Feedback Tuning (VRFT) method represents the main strategy exploited for the analysis proposed in this work (this technique was developed at Politecnico di Milano). Within this framework, linear-in-the-parameters controllers, like PIDs, can be rapidly designed and possibly recalibrated with new experiments.
Automotive applications can serve as an important benchmark for direct techniques, given the complexity of the systems usually involved. In this thesis, the novel robust VRFT technique is successfully applied to a vehicle lateral dynamic control problem. Lastly, this work proposes also an open-loop data-driven control strategy, motivated by a particular automotive case-study that involves Electric Parking Brake systems, which extends the effectiveness of direct approaches to feedforward control problems.

SUSTAINABLE DESIGN AND MANAGEMENT OF AGROECOSYSTEMS - INTEGRATING ECOLOGICAL MODELS AND OPTIMIZATION TECHNIQUES TO SUPPORT DECISION-MAKING

Francesca Recanati - Supervisor: Prof. Paco Melià

Sustainability in the agri-food sector is key in the global development context defined by the Agenda 2030 and the 17 Sustainable Development Goals. The presence of nutritional, economic, and environmental goals, which are all directly or indirectly interlinked and are often conflicting, makes food systems complex and further challenges our capacity to achieve overall sustainability standards. Moreover, promising approaches to increase agri-food sustainability, such as agroecology and multi-functional agriculture, highly differ from the currently widespread intensive agricultural practices based on monoculture, both for planning and management choices. Appropriate mathematical tools and quantitative assessment methods are needed to explore the complexity of agricultural systems, and ultimately support actors in the food sector in the transition towards more sustainable agri-food practices via informed decisions. This being the context, the goal of this thesis is the design of quantitative tools to support farmers, farmers' associations, NGOs, and other actors and decision makers in the agri-food sector to develop more sustainable agricultural systems, with a focus on multi-species agroecosystems. The work carried out during this research has been divided into three stages, ranging from knowledge acquisition to the optimized design of agroecosystems. The first step is two-fold: it includes (i) understanding the predominant biophysical components and processes of an agroecosystem, and (ii) developing a mathematical model of the system. The second step consists in the investigation of the sustainability concept, and of how it can be put into operation to assess agroecosystems. Finally, the last step aims at integrating models and sustainability assessments into an optimization problem to provide a decision support tool to help develop 'more' sustainable agroecosystems.

The first outcome of this thesis is a crop dynamic model that simulates agroecosystems, including mechanisms of competition for limited resources, and it is flexible enough to describe both monoculture and multi-species systems (Figure 1). The development process behind the model and the model itself allowed to understand the functioning of agroecosystems and to highlight the main processes and components involved. It ultimately aims to support the comparison of alternative agroecosystems, characterized by different species composition and biodiversity levels (e.g., monoculture vs. agro-forestry) from an ecosystem services perspective. Secondly, we develop a comprehensive and multi-spatial scale sustainability assessment framework for agricultural systems (Figure 2), including a set of quantitative indicators covering the three dimensions of sustainability (i.e., environmental, economic, and social one, the latter being focused on nutrition), and we test it onto the assessment of domestic food production in the Gaza Strip. The framework definition and its application supported the understanding of how the sustainability concept can be put into operation to quantitatively assess food systems. Focusing on the environmental sphere of sustainability, we investigated how the consumption and depletion of natural resources, the impacts on the environment and on the functioning of ecosystems due to food production can be assessed through the Life Cycle Assessment (LCA) and Ecosystem Services (ESs) Assessment methods. On the one hand, the application of LCA to study food supply chains allowed to highlight the potential importance of the agricultural phase within the environmental profile of food products. On the other hand, it emerged the high variability of environmental impacts brought by alternative agroecosystems. From the analysis of material and energy balances at the base of LCA we move then to the ESs assessment, which emphasizes the agroecosystems composition and functions, and their link with ESs, and allows us to better understand the relation between agroecosystem functioning and environmental sustainability and to increase the comprehensiveness of our assessment. In this direction, we propose and test two methodologies for the assessment of climate and water regulation services of agroecosystems, and apply the dynamic crop model previously developed to assess and compare monoculture against multi-species agroecosystems in terms of major ecosystem services, e.g., climate and nutrient regulation. The last outcome of this thesis is a decision support tool for the design of sustainable multi-species agroecosystems, obtained from the integration of agroecosystem simulation models and sustainability assessments into an optimization problem. Its formulation, implementation, and application allowed (i) to understand the link between species composition of agroecosystems and sustainability performances, (ii) to investigate conflicts and synergies between the considered (sustainability) objectives, and (iii) to highlight the main methodological challenges in the development of such tools linked to lack of proper data, time horizon, and objective function selection. The work included in this thesis contributed to put the sustainability concept into operation within the agri-food sector through both the definition of general and widely applicable indicators and specific mathematical models and frameworks, and the application of all the developed tools to real-world case studies in the Amazon basin (e.g., Peru and Brazil) and in the Gaza Strip demonstrated how comprehensive and quantitative assessments can effectively support informed development of sustainable agroecosystems.

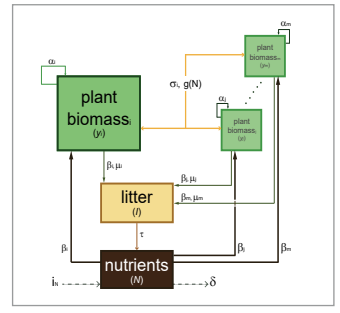

Fig. 1 - Flowchart of the agroecosystem model. The state variables are the biomass of each involved plant species (yi), the organic nitrogen in litter (l) and the inorganic nitrogen in the soil (N). The arrows indicate the relation between those state variables: the maximum potential intrinsic capacity of plant to grow (α_i), the competition for resources (for light through the σ_i function, for nutrients through the function g(N)), and the litter formation (β_i and μ_i).

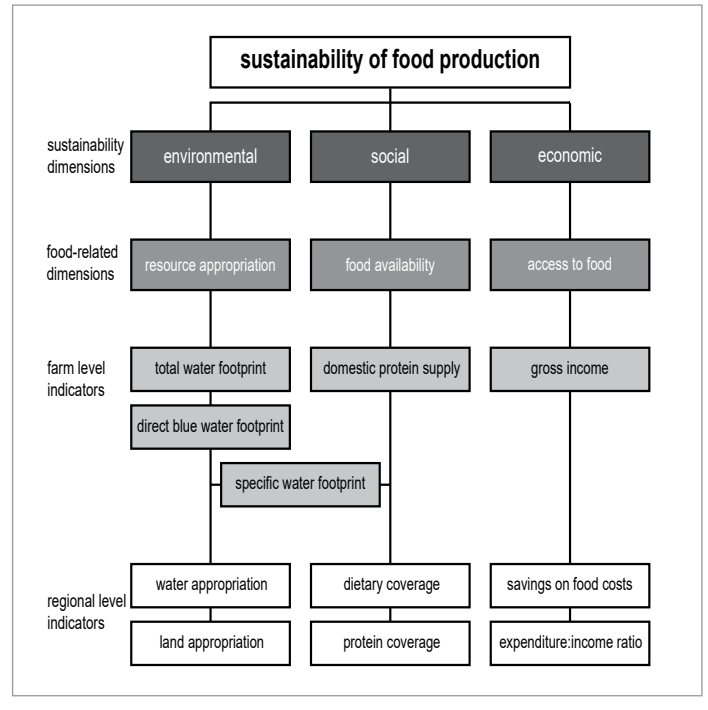

Fig. 2 - Framework for the sustainability assessment of food production scenarios. The achievement of the general objective, i.e., sustainable food production, is measured along three basic dimensions: environmental, social, and economic.

PHYSICAL MODELING OF NANOSCALE NAND FLASH MEMORY RELIABILITY

Davide Resnati - Supervisor: Prof. Christian Monzio Compagnoni

The research activity carried out during the three years of Ph.D. program was focused on the reliability issues related to V_T instabilities in NAND Flash memory arrays and their modeling, with the aim of developing predictive tools able to reproduce the cell behavior in presence of these phenomena and to assess their impact on the array operation.

First, a comprehensive experimental investigation of the V_T instabilities coming from cycling-induced charge trapping/detrapping in planar NAND Flash memory arrays was performed, allowing to extract the main features of the phenomenon, such as its variability along a memory page and its dependences on the main array operating conditions. These included the dependence on retention temperature, which highlighted a thermal activation of the detrapping process around 1.1eV, confirming previous observations, and the impact of the cycling dose, which revealed a non-negligible amount of charge detrapping occurring even in fresh, i.e. uncycled, arrays.
Moreover, results regarding the impact of cell V_T were found to be in contrast with the current microscopic pictures for charge detrapping, demanding a review of the physical mechanisms assumed to be involved in the process.

A new microscopic picture for charge trapping/detrapping was then conceived, based on a multi-step view of the process which was assumed to involve a structural relaxation of negatively-charged defects in the tunnel-oxide followed by the exchange of charge carriers with the substrate. This picture was implemented in a statistical model that simulates the evolution of the single defects inside a NAND cell. When validated against experimental data coming from large ensembles of cells, the model proved to be able to explain all the dependences observed experimentally, including the complete statistical distribution of V_T shift caused by charge detrapping during retention. Furthermore, the model was shown to reproduce with satisfying accuracy the results coming from complex experimental schemes that are part of typical data retention stress tests of NAND arrays.

Together with charge trapping/detrapping, RTN fluctuations represent a strong reliability constraint, posing limitations to tight V_T placement in NAND arrays and hindering the capability of storing more bits in a single Flash cell.
Beside the widely studied impact of RTN fluctuations on array V_T distributions, this work highlighted a RTN-induced sensitivity of the data retention characteristics of NAND cells depending on their position in the programmed V_T distributions. The signature of the phenomenon is a more positive V_T shift during retention for cells placed in the lower tail of the distribution, while the opposite holds for the cells in the upper tail. This peculiar result was explained in terms of different occupation of RTN traps along the programmed V_T distribution, when this distribution is placed using a program-and-verify algorithm. Specifically, cells below the verify level will have predominantly empty RTN traps that will lead them to more positive V_T shift during the retention period, while cells well above the verify level will have the opposite behavior.
A Monte Carlo simulation was then shown to reproduce this effect, where the simulated cell V_T was determined by a simple program-and-verify algorithm together with employing established models for RTN dynamics.

Finally, the attention was focused on the impact of temperature on the main parameters controlling the operation of NAND arrays, with a comparison between planar NAND arrays and modern three-dimensional (3-D) NAND arrays, representing the state-of-the-art of Flash technology.
The temperature effects were addressed by investigating the temperature dependence of three main parameters, namely cell V_T, the string saturation current and the amplitude of RTN fluctuations. The differences between the temperature behavior of these parameters in the two technologies are particularly interesting considering that 3-D NAND strings are realized with polycrystalline silicon, in place of the conventional monocrystalline silicon substrate of planar NAND strings.
The work showed that the string saturation current in 3-D strings has a positive activation energy E_a, as opposite to the negative E_a of the planar case. The positive E_a in 3-D NAND strings is dictated by the presence of potential barriers formed by charge carriers trapped in the highly defective grain boundaries delimiting crystalline domains.
The statistical distribution of RTN-induced V_T shifts in 3-D NAND arrays was also found to be markedly dependent on temperature, with a dominant role played by the change in the amplitude of the V_T shift induced by single RTN traps. This was explained in terms of non-uniformities of inversion charge density between the grain centers and the grain boundaries, which increase at lower temperatures and make the grain boundaries a stronger bottleneck for current conduction.

The practical impact of the V_T shifts introduced by all the discussed mechanisms consists in the increased width of the V_T distributions, resulting in the insurgence of read errors. The work highlighted how charge detrapping is more relevant on high V_T levels, implying that in MLC/TLC architectures the most critical read operation is the one at the highest read voltage. The highest V_T distribution, as a matter of fact, will have the most pronounced detrapping-induced negative tail after a data retention period. At a controller level, the resulting increase of the bit error rate can be alleviated by optimally shifting the read levels to compensate the shifts of the V_T distribution tails.
Conversely, RTN was shown to have no relevant V_T level dependence and consequently it affects all the V_T distributions in the same fashion. Moreover, RTN-induced V_T shifts are relevant in determining the width of the time-0 V_T distribution, that is the distribution immediately after the program operation. Despite the reduced magnitude of RTN-induced V_T shift in 3-D NAND cells with respect to planar NAND cells, RTN is still an important source of read errors, especially in architectures storing 3 or more bits per cell, where very tight time-0 V_T distributions are required. Considering the temperature dependence of RTN observed in 3-D NAND, this represents a stronger constraint at lower operating temperature, leading to a worse time-0 V_T placement and to increased bit error rates.

In conclusion, the research work provided an in-depth analysis of the reliability issues of modern NAND Flash technologies coming from phenomena impacting the stability of cell V_T, addressing both planar and 3-D architectures. The experimental and modeling efforts which focused on cycling-induced charge trapping/detrapping and RTN fluctuations in planar NAND arrays represent a great improvement in the understanding of the phenomenon and in the ability to predict its impact on memory arrays in realistic operating conditions.

VEHICLE DYNAMICS PLANNING AND CONTROL FOR SAFETY AND COMFORT IN AUTONOMOUS CARS

Federico Roselli - Supervisor: Prof. Sergio Matteo Savaresi

The automotive industry is rapidly evolving and the advent of fully autonomous cars seems every day closer. The problem of autonomous driving has been widely discussed in the academic environment in the last decades and it covers many fields of the Information Technology from the sensing to the control. This research work is focused on the main aspects related to the control of vehicle dynamics. The main control systems developed so far in the automotive field aimed to improve the driving experience or help the diver in the most challenging maneuvers. With the advent of fully autonomous cars these features will lose their value, but two factors will gain significant importance: comfort and safety. Since several studies conclude that the most of car accidents are caused by human mistakes, safety has become the main driver for autonomous drive development. This research aims to improve these two important requirements by controlling the vehicle dynamics in case of emergency situations and in very uncomfortable scenarios. High speed driving and emergency manoeuvres mainly excites the longitudinal and lateral vehicle dynamics. Lateral dynamics control is the most challenging problem when dealing with quick maneuvers like an obstacle avoidance. The trajectory preview plays an important role in this case and most of the works in the literature exploit this feature using the Model Predictive Control (or other optimization based techniques). While these techniques give the possibility of considering the future reference trajectory and the nonlinearities of the system, it is not suitable for sudden obstacle avoidance due to the high computational time, especially when cheap hardware has to be used. In this work two path tracking methods for accurate tracking of high speed and emergency maneuvers are developed. The proposed algorithms exploit the preview without asking an excessive computational burden and without the need of lateral speed measure that is often unavailable. Another important issue in emergency manoeuvres is the planning of the avoidance trajectories: the most of the approaches in the literature don't consider the path tracking control dynamics that becomes fundamental in case of quick maneuvers. Moreover the planning must be completed as fast as possible not to risk to make the avoidance unfeasible. The approach presented in this work is based on the characteristics of the best performing controller between the two developed to guarantee feasibility and optimality in the path generation. The standard presence of "preview sensors" like cameras in self-driving cars can become useful also in case of driving on particularly uncomfortable scenarios, for example in presence of potholes and bumps. A vehicle equipped with stereo cameras is used to experimentally explore the possibilities of having preview information about the presence of a bump/obstacle. Depending on the available information two comfort oriented control strategies are designed. The experimental analysis and validation was performed with a vehicle made available by the European Honda Research Department.

Lateral Control

Two control techniques have been developed for high-speed/highly-dynamic driving exploiting the preview. To deal with non-linearity introduced by the variation of longitudinal speed, the system is considered as a LPV system, where the speed is a low varying parameter.

- Look-ahead based LPV control: it consists in a self-scheduled H-infinity controller. The objective is to control 2 variables: the lateral error with respect to the reference trajectory at the CoG of the vehicle and the first derivative of a look-ahead error defined as the distance between a virtual point on the same longitudinal axis of the vehicle and the trajectory. The look-ahead distance is a tuning parameter that can be used to anticipate the steering when approaching to a corner.
- Curvature preview based LPV control: the control system is composed of two controllers. The first controller is a yaw rate controller: the reference yaw rate depends on the track curvature, moreover a pre-filter is used to invert the system dynamics and anticipate the reference yaw rate. Then a lateral error control loop is added to the yaw rate controlled system in order to avoid the drifting of the lateral error and to guarantee robustness.

The control system has been validated on two different platforms:

- A Dodge Dart provided by Magneti Marelli: the car is equipped with a electric motors connected at the steering wheel and brake pedal to allow autonomous drive. Differential GPS is used to have accurate position feedback. The controllers were experimentally tested on the Rivalta track circuit performing obstacle avoidance and double lane change maneuvers: results show that very accurate path tracking is achieved guaranteeing errors lower than 40 cm .
- A 1/10th scale RC Car called BARC developed at UC Berkeley. The car is equipped with a processor with Linux, an Arduino, encoders, a IMU and a Camera. ROS is used to develop the control software and communicate with the different sensors. The path following is performed only using on-board sensors. The camera is used both to perform path planning and to estimate the vehicle position: image processing is used to detect the lanes and obtain information about the future trajectory. The Curvature preview LPV controller was able to make the car follow obstacle avoidance maneuvers with errors lower than 10 cm.

Path Planning for emergency maneuvers

Due to the importance of vehicle dynamics in emergency maneuvers a specific planner for this circumstance has been developed. The algorithm assumes to drive in a structured environment and only one dynamic obstacle is considered; The generated path aims to shift the car form the current unsafe lane to another obstacle free lane according to the obstacle position and speed. These maneuvers are particularly critical for two reasons: they involve operating close to the handling limits and they require a fast lateral motion of the vehicle.
The planner exploit the capability of the curvature preview based controller of tracking a reference path within a specific frequency bandwidth to generate a control feasible trajectory based on sinusoidal curvature sections. The algorithm is tested in simulation on different scenarios assuming the obstacle is seen when the vehicle is very close to it.

Vertical Control:

In this work two cameras are used to detect the presence of isolated impulsive irregularities such as potholes and bumps. This information is then used by a control software to act on the electronic semi-active dampers. The capabilities of the available vision algorithm are quite limited: only some characteristics of the future road are available, such as the distance of the bump/pothole and its height/depth, moreover both these estimates have significant uncertainties.
Due to this limitation the designed control logics will be simple in order to be robust to misdetections and inaccurate estimations.

1) Adaptive control with preview: the information about the presence of the bump/pothole is used to set the best fixed damping to improve the comfort before reaching the obstacle. The main advantage of this logic is that no accelerometers or stroke sensors are needed.
2) Switch control with preview: this logic is based on some considerations that can be deduced by the adaptive control analysis and aims to switch between two levels of damping according to the wheel position on the bump/pothole.

The simulations conducted on CarSim and the experimental results showed that even a simple adaptive logic can produce significant improvements w.r.t. standard semi-active controls if the preview is exploited. The main advantages are seen especially in the case of wide potholes and high bump where the effect of the suspension stops is crucial and the preview can predict their hit.

OPTIMAL KINODYNAMIC PLANNING FOR AUTONOMOUS VEHICLES

Başak Sakcak - Supervisor: Prof. Gianni Ferretti

Motion planning is one of the most important factors affecting the functioning of an autonomous vehicle and aims to find a collision-free motion guiding the vehicle from an initial configuration to a final one. In particular, kinodynamic motion planning addresses this very issue, taking also into account the dynamics of the system so that the resulting solution would be executable by the vehicle. For most applications, the solution to the trajectory planning problem has to satisfy some desirable properties; e.g., reaching the goal in minimum time by minimizing the energy consumption or by maximizing the safety. These requirements shift the focus from finding *any feasible trajectory* to finding the *optimal trajectory* for a given objective function. This thesis addresses the optimal motion planning problem with a strong emphasis on inclusion of the underlying realistic dynamic system as an object of the planning. The proposed approaches consider arbitrary dynamics without imposing strong restrictions on their characteristics. Furthermore, the motion planning problem is also widened to encompass some particular aspects such as the role of the topology on the generated trajectories and constraints imposed by the landmarks.

The problem of optimal kinodynamic motion planning is addressed, using important classes of motion planning algorithms: the exact approach and the sampling based approach. The former looks for a solution in the continuous state space, while the latter samples this space redefining it as a graph where nodes are connected via edges representing local trajectories between sampled states.

The optimality for a sampling based planner is achieved once all the nodes are connected to the underlying graph representation optimally, possibly rewiring the graph by testing connections within a neighborhood. The same strategy applies to kinodynamic planning as well, with the additional difficulty that when optimality is required, connecting two nodes involves solving a *two point boundary value problem* (TPBVP), which is computationally challenging especially when dealing with complex dynamics, such as for non-holonomic systems, in presence of actuation constraints. To this end, the first contribution of this thesis is a resolution optimal motion planner, RRT* with Motion Primitives (RRT*_MP) that samples the state space in a grid representation and generates a tree of trajectories using a database of motion primitives. Therefore, the computationally intensive part of solving for a TPBVP is carried to the preliminary phase of database generation by alleviating the computational load during planning. This represents a promising result for online applications, especially in dynamic environments where the planner has to generate a new trajectory in response to changes in the environment. Furthermore, it allows to compute the true cost-to-go in order to guide the expansion of the tree, improving the convergence properties. The algorithm is proven to be asymptotically optimal as the grid resolution goes to zero and number of nodes goes to infinity. The (sub)optimality caused by the gridding can be tuned to compromise between the size of the database and this performance degradation.

Planning with topology constraints has gained an increasing attention in the recent years. Other than obtaining a collision free trajectory that minimizes a selected objective, one might be interested in how the resulting trajectory avoids the obstacles. This interest a number of fields in the robotics and autonomous vehicle applications, ranging from avoiding dangerous zones, coordinating multiple agents for better exploration, following rules on how to avoid obstacles. Furthermore, other than imposing some rules about how to avoid the obstacles, it is also possible to use the topology information for obtaining a number of distinct good trajectories or for decomposing the otherwise complex trajectory generation problem.

Motivated with these examples, the second part of the thesis focuses on planning approaches that take into account also the specific topology of the environment. In particular, it elaborates on the decoupled approach of systematically generating homotopy class constraints and imposing these constraints during the planning phase to obtain a set of distinct local optimal trajectories belonging to different homotopy classes. The contributions in this case is twofold: first, an exact optimal planner based on homotopy class constraints and optimal control is proposed to decompose the overall trajectory generation problem into simpler sub-problems and, second, RRT*_MP is enhanced with homotopy awareness that allows generation of distinct good trajectories.

With the continual increase in the number of tasks that are handled by the help of autonomous vehicles, the role of the planner has been changing from returning optimal trajectories that are collision-free and dynamically feasible to considering advanced tasks that gradually exclude the human operator from the navigation loop. Another such problem has also been addressed in this thesis is the generation of motion plans that consider the connection between the autonomous vehicle and a signal emitting base. To illustrate, for autonomous vehicles used in surveillance or search and rescue missions, it is important that the robot is always in communication with the bases to transmit data or to receive tactical information. Furthermore, for mobile robots or autonomous vehicles that work in Global Positioning System (GPS) denied environments, such as indoor robots or autonomous underwater vehicles, it is essential for vehicles to maintain connection with a signal emitting beacon for improved localization. Without loss of generality, this definition of the problem can be applied to any artificial landmark that can be perceived from a limited range.

Finally, the proposed algorithms are applied to a particular problem of landmark-aware motion planning. A continuous connectivity problem, where the vehicle has to be connected with a landmark at all times, and a relaxed problem, where the vehicle can travel without coverage for a predefined distance are proposed. Both problems are reformulated as multi-objective problems and are tested in an omni-comprehensive scenario. The main contribution in that case is application of the proposed approaches to this particular problem, which allows us to address a more complex problem without explicitly interpolating a set of selected ones.

DATA MIGRATION AND REPLICATION ACROSS HETEROGENEOUS NOSQL DATABASES

Marco Scavuzzo - Supervisor: Prof. Elisabetta Di Nitto

During the last years we have been witnessing a radical change in the IT world. Specifically, thanks to social-networks, cloud computing and IoT, we have entered in the era of, the so called, Big Data: i.e., an increasing amount of data continuously produced by applications, sensors, smartphones, etc.. The need of coping with these new challenges has paved the way for new information management approaches. In particular, in the data management field, new types of databases like, e.g., NoSQLs, have been created. NoSQL stands for *Not-Only SQL* and it identifies those databases that, differently from traditional Relational Database Management Systems (RDBMSs), can handle the above mentioned Big Data. The birth of NoSQLs was, mostly, spontaneous and it first took place at big Internet companies (i.e., Google and Amazon). Such origins, together with the lack of common and shared guidelines to build such systems, has recently brought to the proliferation in the market of hundreds of NoSQL databases, each one having different characteristics and interfaces from one another. These heterogeneities can be found in different aspects, like:

- different type of deployments, i.e., on-premise or offered as-a-service by some vendor;
- different data models, i.e., Key-value, Document, Column-family, Graph; • different schema management capabilities, i.e., none at all, or very limited;
- different querying languages or APIs;
- different data access guarantees in terms of consistency and concurrency.

In fact, in order to build databases that could manage vast amounts of semi-structured, or unstructured, data produced at very high rates, the different NoSQLs vendors have adopted reduced functionalities for their databases, with respect to RDBMSs, delegating the burden to develop such functionalities (if needed) to applications developers, within the applications using these databases. Historically, in software engineering practice, designing applications that used RDBMSs (hence, using SQL as *lingua franca*) implied a decoupling between the application logic and the data management layer. These practices have allowed for the creation of easily maintainable applications that could also be easily ported, with none or very limited changes, from one RDBMS product to another. Notwithstanding, the usage of NoSQL databases, as replacement of RDBMSs, because of the less tight properties and of the absence of some characteristics typical of RDBMSs, and finally, because of the heterogeneities discussed above, have forced developers to take into account factors, hitherto negligible, like data storage and management. All of this has brought to the conception of approaches and libraries allowing developers to abstract the application logic from the data management and storage aspects when using NoSQL databases. Indeed, these approaches can reduce applications re-engineering efforts in case the applications need to be ported to a different NoSQL database, thus minimizing vendor lock-in. Notwithstanding, these approaches only focus on the interfacing between the applications and the databases and do not consider the problem of migrating the data from one NoSQL database to another, thus, if used alone, these libraries do not offer a complete solution to the vendor lock-in problem. In fact, additional efforts would be needed to transfer data to the new database, before such applications could actually work correctly again. Moreover, if applications should keep working during data transfers, the databases should be kept synchronized through a data replication process. For this reason, my research focuses on data migration and replication techniques across heterogeneous NoSQL databases. We define data migration as the process that typically occurs offline and that transfers all the data from a source database to a target one, preserving those properties that are expressed both at the data schema and at the data model levels. Data replication is, in this context, an online process that keeps different databases aligned and consistent even in the presence of an ongoing data migration. The data migration and replication system I have developed is called Hegira4Cloud. This system is basically a middleware capable of:

- interfacing with different NoSQL databases in order to migrate their data to other databases;
- intercepting user operations, thus acting like a query mediator, with the purpose of correctly replicating them, in accordance with ongoing data migration processes.

Hegira4Cloud is, thus, capable of performing data migrations and replications while allowing applications to keep operating on the database hitherto used. Once both processes are concluded, and the source and target databases are aligned, the application can operate seamlessly on the target database. Hegira4Cloud adopts an original Common Intermediate Data Model (CIDM) that is generic enough to abstract and properly represent all the characteristics of the data stored in a source NoSQL database, thus preserving those characteristics also in the database chosen as the target for a given data migration. Additionally, the CIDM makes it easier to add support for new databases; in fact, it would suffice to create two translators – one translating the data from the source database data model into the CIDM and another one able to translate from the CIDM to the target data model representation – to support a new database. The system currently supports Key-Value and Column-family based NoSQLs by relying on a CIDM instance specific for these classes of databases. This instance relies on the fact that NoSQLs typically store a schema along with the data they contain. This schema contains the necessary information about the data types, secondary indexes (when supported) and, occasionally, about data partitioning. As previously explained, not all NoSQLs maintain a schema of the data contained within the database; in fact, they delegate the schema management to the applications using the database. Having the schema distributed across many applications prevents a third-party application – that has no knowledge about other applications schemas – from correctly interpreting the data stored in the database and, moreover, it makes schema maintenance difficult. As regards the data migration, if Hegira4Cloud cannot derive the schema from the database, it cannot be able to properly interpret the data so as to preserve any information expressed in the schema. Hence, Hegira4Cloud integrates a Schema Registry for managing the schemas that all applications, interacting with the source database, use to process the data. The storage of these schemas within a common, shared registry allows the migration system to correctly process data in the source database and to preserve its characteristics even when the source database does not maintain any schema. Given that data migration processes typically require the transfer of huge quantities of data and that, as a consequence, these processes can last several days, Hegira4Cloud has been built to tolerate faults and interruptions both of the databases, or of any of the internal system's components. In fact, by means of a highly replicated migration log it is possible for a data migration process to be consistently recovered from the exact point where it stopped, thus reducing the data migration time in case of faults. Finally, since there may be several applications using the same database subject to data migration, Hegira4Cloud has been carefully designed to scale horizontally, so as to be able to cope with increasing applications workloads. Thorough evaluations of Hegira4Cloud have been carried out in order to demonstrate that this system is able to, consistently and correctly, migrate and replicate big volumes of data across many different NoSQL databases even in presence of faults during the process.

FEATURE-BASED ANALYSIS AND SYNTHESIS OF VIOLIN TIMBRE

Francesco Setragno - Supervisor: Prof. Augusto Sarti

The timbre of musical instruments is one of the most complex and ambiguous case of study in music research. The lack of a mathematical formulation, the subjectivity of timbre description and the dependency on the data make this property far from being exhaustively understood. Nevertheless, timbre is a very important aspect of music and it is of great interest for manufacturers, musicologists and researchers to be able to analyse and control the sound properties of musical instruments.
Among the others, the sound of the violin received particular interest for decades, due to its complex behaviour and the aura of legend that surrounds the masterpieces of the ancient Cremonese masters - Stradivari, Guarneri, Amati. The violin is a complex instrument made of tens of different pieces and materials and involving non-linear interactions between its components. Researchers correlate acoustical properties, materials and structural behaviour to timbre perception, but many aspects are still unclear.
In this thesis, timbral analysis techniques are studied and applied to the specific case of violins. The choice of this class of instruments depends on different reasons: the aforementioned complexity of the bowed instruments family; the great interest in the manufacturing, music and research communities; and the availability of the remarkable collection of the Violin Museum in Cremona, that includes historical and top-quality contemporary violins.
The musical timbre depends both on the physics of sound (low-level perspective) and human perception (high-level perspective). For this reason, this study is conducted at different levels of abstraction. The fist level concerns the measurable aspects of the audio signal. The second level is related to how the listener perceives and describes the sound.
This study requires a proper data collection stage. In particular, this stage includes audio recordings, interviews with violin makers and a listening test where the subjects evaluate the timbre of the recorded violins.
First we recorded 50 violins in Cremona including historical instruments from the Violin Museum, top-quality contemporary instruments from the Triennale competition and low-quality violins from the violin making school Istituto Antonio Stradivari. We obtained a considerably large, diverse and unique dataset which is the basis for this study. The recordings include different performances: open strings, single notes, a major scale and six musical excerpts.
The interviews have been conducted with the purpose of collecting the words that are typically used to describe the timbre of violins. We interviewed 18 violin makers in Cremona and we collected several descriptors, their meaning and the relations among them.
As for the listening test, the purpose was to annotate each violin in the dataset with a set of numbers describing its timbre. The descriptors collected during the interviews were arranged in pairs (e.g. Dark/Bright) and, for each violin recording and each pair, the subjects were asked to provide a number from 1 to 10.
For what concerns the low-level analysis, the purpose is to investigate measurable properties of the audio signal that timbre relies on. A typical way of analyzing such properties is called feature-based analysis. This approach is often used in Music Information Retrieval (MIR) and it is based on the extraction of many objective quantities from the raw audio (low-level features). These quantities have a low degree of abstraction and are obtained by means of signal processing algorithms.
We define and extract a set of low-level features in order to capture various aspects of violin timbre. Since the literature suggests that the timbral relations between instruments depends on the pitch, we investigate this dependency. The results suggest that the differences between violins change with the played note. This phenomenon is evident especially for the A string and the E string.
Another important aspect that we take into account is the impact of materials on the final sound. Many studies investigate the role of the varnishes used in the violin making process. We address this problem by analysing a violin during the finishing steps. We consider the violin untreated, the violin after the application of a casein-based ground coat and the varnished violin. For each step we recorded a performance using the violin under study and we extracted low-level features from the recordings. The results show that the application of the ground coat has a significant impact on the produced sound, especially for what concerns the emitted energy. The application of the varnish results to have a minor impact.
Finally, we investigate and characterize the differences between historical and contemporary violins, which is a topic of great interest for violin enthusiasts and researchers. We compare the feature extracted from the recordings and we study the differences between the two classes of instruments. By means of feature selection algorithms we retrieve the features that best allow to discern the two categories. We show that the performance of a classifier trained to recognize an historical violin using the selected features are prominent. The results point out that there are objective cues that allow to discern historical and contemporary violins.
For what concerns the high-level perspective, we employ knowledge management tools to address the problem of the timbre description semantics. Human listeners make use of words from the natural language to describe the sound quality. Such words lack a mathematical formulation and their semantics is not formally defined. In order to make a machine able to process this type of information, a formal representation is needed. A solution is the implementation of an ontology, a knowledge management tool used in many fields of information science, particularly in the Semantic Web.
A first version of the ontology is implemented using the descriptors that we collected during the interviews and the relations among them. Then we validate and extend the ontology with concept analysis techniques, proving its semantic validity.
The core implementation of the ontology is finally enriched with other information related to the perceived timbre, like materials and low-level features.
The relation between low-level properties of violin sound and perceived timbre is still unclear. For this reason, we employ machine learning techniques to model sound descriptors based on low-level features and bridge the gap between the objective and subjective description of violin timbre.
We test both methods based on hand-crafted features and methods based on unsupervised feature learning using neural networks. In particular, we employ a Deep Belief Network to build a representation of the audio signal that is used as input for the regressors. We test different regressors: Linear, Ridge, SVR, AdaBoost, GradientBoost.
The regression performance for the two methods are evaluated using the RMSE and the R^2 metrics. The method based on hand-crafted features exhibits better performance, probably because deep learning methods require a higher amount of data.
We also investigate the relation between low-level features and high-level descriptors, using the Pearson Correlation Index and the Distance correlation index.
The results allow to discover the relation between the two faces of timbre - objective and subjective - and show that it is possible to train a machine to produce a human-like sound description with a good degree of accuracy.
The techniques employed in this thesis can be generalized to other musical sources and used to implement tools that efficiently describe and model the timbral properties of musical instruments and the timbral differences between different sources.
This work is part of a wider project conducted with the collaboration of the Violin Museum and the University of Pavia.

DESIGN AND IMPLEMENTATION OF MICROWAVE AND TERAHERTZ MATERIAL CHARACTERIZATION METHODS

Saleem Shahid - Supervisor: Prof. Gian Guido Gentili

1. Introduction

1.1. Dielectric Measurement Systems

Dielectric constant measurement has been playing a vital role in various medical, security and non-destructive testing applications from the last few decades. Different type of measurement setups has been developed and improved by enhancing the antenna focusing performance, positioning systems etc. in order to achieve ultra-wideband (UWB) reflected waves or precisely the scattering fields from the object or dielectrics. Microwave and millimeter frequency signals are widely used to measure the s-parameters. The measured data is further processed to estimate the electrical properties of materials, particularly dielectrics. The sub-sections are organized below to briefly introduce the concept of dielectric measurement systems, mainly, background, reported characterization techniques and measurement systems used up till now.

1.1.1. Background

The dielectric material characterization has been divided into various types of measurements which depends on the nature of characterization i-e, (a) measurement in narrow or broad band, (b) the material under test (MUT) is low loss or high loss medium, (c) whether to measure the electrical and/or magnetic properties, etc. [1–2]. The free-space dielectric measurement method have been used widely over the years at microwave and millimeter frequency bands due to easy extraction of the transmission and/or reflection scattering parameters from MUT which is placed between transmitter and receiver antennas [3–5]. The free-space measurement method is mostly used for imaging and material characterization applications which particularly prefers non-destructive evaluation technique in order to avoid physical contact with a sample and reduce extra machining care needed for an MUT [2].

1.1.2. Material Characterization Methods

The conventional free-space measurement method for dielectric material characterization have been first experimented at the MIT Radiation Laboratory [6]. Later, a low-loss dielectric measurement based on the method in [6] at millimeter frequency bands has also been demonstrated by Breeden [7] which provided a comprehensive uncertainty analysis for the characterization of the materials. Similarly, another method is developed for the dielectric permittivity estimation using Brewster's angle and the magnitude of transmission and reflection coefficients are measured from reference metal and MUT [8]. Furthermore, in free-space measurement methods for microwave, millimeter or sub-millimeter frequencies, the dielectric constant of low loss and thinner materials has been estimated from envelopes of measured transmission and/or reflection spectra instead of the whole scattering parameters unlike in visible light or X-rays [9–10]. In most of the studies, we have noticed the use of only normal incidence wave interaction with materials to build the maximal and/or minimal envelopes. Many other interesting discussions on analytical analysis and derivations for the envelope method has been reported and their utilization in the dielectric measurement systems has been studied [11-12].

Another method which is gaining the reputation in material characterization is Bayesian inversion [13]. This method is proven to be efficient when it comes to the reduction of uncertainties in electrical or magnetic parameters estimation. Bayesian inversion can be useful to get the detailed information about the model parameters in terms of a probability density. Later, this probability density can be used to measure the uncertainty in the prior model parameters to build the confidence on observed parameters. In our dielectric measurement work, we have been able to find the uncertainties in the MUT thickness and its positioning. The uncertainty quantification (UQ) may be defined as: the process of quantifying uncertainties associated with extraction of scattering parameters and quantifying the uncertainty contributions of all other sources in data space [14].

1.1.3. Measurement Systems

The dielectric measurements encounter two type of problems; the error in positioning of MUT and data extraction method, and inefficient post-processing techniques to accurately estimate the dielectric properties. It is convenient to implement the measurement system using off-the-self performance instruments such as, commercial vector network analyzers (VNAs) and positioning apparatus [15-17]. The VNA uses the step frequency technique in order to perform the measurement in usually wide frequency range and it transmit each signal at different step frequency which illuminates the MUT individually, and the scattering parameters (transmitted and/or reflection) are received back [18]. We have used frequency domain scattering parameters for dielectric properties estimation as well as we have inverted it to time domain response using Fourier transform and Bayesian inversion. Despite that, the time domain experimental systems are also available where a pulse generator and real time oscilloscope is used instead of VNA [19]. Usually the pulse generator is used to generate the required pulse which illuminates the MUT and the time response is recorded at the oscilloscope [19-20].

1.2. Motivation and Contribution

To our understanding, the antenna systems for dielectric measurement can be further improved which will help to extract more accurate scattering parameters. There is a scope in the designing of UWB microstrip patch antennas as well as horn antennas with lens correction. The lens correction is usually done to achieve the better spot focusing. The antennas with wide bandwidth and directive radiation patterns are needed to get high resolution data from dielectric measurement systems. The lens corrected horn also limits the use of measurement setup at only single distance from transmitter to MUT (where distance is focal length of lens). Avoiding the lens correction will allow us to measure the scattering parameters at variable distance in near field. Furthermore, the measurement setups normally hold inaccuracy in the measured data due to different factors including positioning tolerance and physical limitation in setup, which makes the dielectric properties estimation more challenging. We have found the scope of improvement in the time domain data inversion techniques which can in cooperate all the challenges. In this thesis, we have developed couple of time domain data inversion techniques which are fast in processing and more accurate in dielectric properties estimation making this research work interesting and useful for dielectric measurement and imaging applications.

The contributions of this thesis are divided into two parts; the antenna designing and data inversion methods. The first part presents the designing of microstrip patch and horn antennas for bandwidth and radiation performance enhancement. We have proposed novel designs of microstrip patch antennas i-e; opposite slots and stacked patches, whereas the proposed horn antenna designs allowed us to measure the MUT at variable near field distances which was not possible with lens corrected horns. We have designed two different spot focused horn antennas; one for frequency domain analysis and another for time domain analysis for better removal of all the spurious elements after TRL calibration. In the second part, the research focus is shifted to the characterization of electrical properties of MUT such as, dielectric constant, loss tangent and MUT thickness for single and multi-layered materials at different frequency bands of 26-40GHz, 75-110GHz

and 915-925GHz. The free space measurement method is used for data acquisition using commercially available VNA's and custom built positioning system. In order to extend the comprehensive study, the measurement are carried out at PoliMI and UPC Barcelona. The frequency domain analysis is done for 915-925GHz band while the different techniques in time domain analysis are used for 26-40GHz and 75-110GHz bands. The time domain Fourier inversion and Time domain Bayesian inversion techniques are developed with the focus on improvement of the results accuracy as well as reduction of computational power and time. We have successfully achieved convincible accuracy in dielectric properties estimation as compared to the reported techniques whereas the computation power and time is significantly reduced when compared to commercial softwares like Ansys HFSS and CAD FEKO.

1.3. Applications

This research work can be useful for the applications requiring the estimation of physical and/or electrical properties particularly dielectric constant, loss tangent and material thickness. The possible list of applications can be;

1) Non-destructive test for structural defects detection in small objects as well as buildings.
2) Product quality control e-g, food, drugs and commercial grade items etc.
3) Medical imaging for early cancer diagnoses and detection.
4) Security scanning of the concealed objects at airports and sensitive places

HIERARCHICAL AND MULTILAYER CONTROL STRUCTURES BASED ON MPC FOR LARGE-SCALE SYSTEMS

Xinglong Zhang - Supervisor: Prof. Riccardo Scattolini

Over the last decades, the complexity of systems is continuously increasing due to economic reasons and technological advances. It is known that the centralized Model Predictive Control (MPC) solutions for such large-scale systems might result in unacceptable control performance due to various factors, such as high dimension of the system, computation efficiency and communication bandwidth.

Moreover, centralized controllers are not scalable and difficult to maintain. For these reasons, in the last twenty years, decentralized and distributed MPC algorithms have been developed with a number of local problems solved in parallel to achieve global or local objectives.

An alternative to decentralized and distributed control consists in the use of hierarchical control structures based on MPC. This approach is very powerful especially for control of systems with separable fast and slow dynamics, for the coordination of subsystems and when it is required to consider different objectives in the long term and regulation problems in the short term.

This Thesis addresses the theoretical development of hierarchical and multilayer control algorithms based on MPC for large-scale systems.

In Chapter 2, we develop a two-layer control structure for the coordination of independent linear dynamic systems with input and joint output constraints. At the higher layer, a reduced order dynamic model of the system's components is used to state and solve an economic MPC algorithm in a long time scale. The outcomes of this layer are the components of the control variables to be held constant over the long sampling periods. At the lower layer, decentralized MPC controllers, one for each subsystem, are implemented in a shorter time scale and according to a shrinking horizon strategy to compensate for the model inaccuracies at the high level. The overall convergence, recursive feasibility, as well as the fulfillment of the joint constraints, are obtained under mild assumptions.

A fully scalable hierarchical control scheme for coordination of similar independent systems with joint output and input constraints is presented in Chapter 3. Differently from Chapter 2, a scalable low-dimensional model mapping the common input to the collective output is used at the high layer, this model is easily determined from the impulse responses of the subsystems. The outcome of the high layer is the value of the common input to be held constant and to be distributed among the subsystems based on a specific weight associated with each subsystem. This approach allows to modify the system configuration with time varying weights, in terms of the contribution provided by any subsystem to the overall system performance, and to implement plug-and-play operations. The recursive feasibility is guaranteed also during plug-in and plug-out operations, and the overall convergence of the system output to the set-point is proven.

Finally, in Chapter 4, we extend the hierarchical control structure to large-scale interconnected systems. At the higher layer, a robust centralized MPC algorithm based on a reduced order dynamic model optimizes a long-term performance index, while at the lower layer local MPC regulators, are designed for the full order models of the subsystems to refine the control action computed at the higher layer. The recursive feasibility and robustness of the two layer algorithm are guaranteed and the overall convergence of the state to the steady state is fully discussed.

Several simulation examples are reported to show the effectiveness of all the proposed algorithms.