POLITECNICO DI TORINO

UNIVERSITÀ DEGLI STUDI DI TORINO

integrazione tematiche del 19/06/2017

Dottorato/PhD in Matematica Pura e Applicata / Pure and Applied Mathematics

(in convenzione con Università degli Studi di Torino e il Politecnico di Torino)

XXXIII ciclo

Elenco delle tematiche per specifiche borse di Dottorato / List of research topics bound to PhD scholarships

1. **HOMOLOGICAL SUMMARIZATION OF HIGH DIMENSIONAL STATIC AND DYNAMIC SIMPLICIAL DATA -** *(Politecnico di Torino)*
2. **DATA-DRIVEN MACHINE LEARNING METHODS FOR PHYSICALLY-BASED SIMULATIONS -** *(Politecnico di Torino)*
3. **RESILIENT CONTROL, INFORMATION FLOWS, AND INCENTIVE MECHANISMS FOR INTELLIGENT TRANSPORTATION SYSTEMS** *(Politecnico di Torino)*

POLITECNICO DI TORINO

UNIVERSITÀ DEGLI STUDI DI TORINO

Titolo del progetto di Ricerca/Research Title:

HOMOLOGICAL SUMMARIZATION OF HIGH DIMENSIONAL STATIC AND DYNAMIC SIMPLICIAL DATA

FINANZIATO DA / FUNDED BY: Politecnico di Torino

SUPERVISOR: Prof. Francesco Vaccarino *(francesco.vaccarino@polito.it)*
Prof. Paolo Garza (paolo.garza@polito.it)
Dott. Francesco Bonchi (francesco.bonchi@isi.it)

CONTATTO / CONTACT: http://bigdata.polito.it

CONTESTO DELL'ATTIVITÀ DI RICERCA / CONTEXT OF THE RESEARCH ACTIVITY:

While network approaches can describe the fabric of relations between agents in a systems or molecules in an organism, they are constrained in their descriptive power to pairwise interactions (i.e. edges), which might not always be justified when focusing on phenomena that involve group dynamics (e.g. scientific collaboration, genetic pathways) or higher-order descriptions (e.g. viral evolution, molecule folding). Against this background, a set of new techniques for data analysis, based on a set-theoretic (hence topological and geometry-independent) formalism, has been gaining traction over the last decade and has come to be collectively referred to as Topological Data Analysis (TDA). The novelty of TDA is that it studies the shape of topological spaces at the mesoscopic scale by going beyond the standard measures defined on data points' pairs.

This is done by moving from networks to simplicial complexes. The latter are obtained from elementary objects, called simplices, built from such simple polyhedral as points, line segments, triangles, tetrahedra, and their higher dimensional analogues glued together along their faces.

TDA is still very much developing as a branch of data science. While it provides a new paradigm, based in algebraic topology, to how we think about data, and has obtained its first successes, there still are many challenges to be met to fully exploit its potential. The most pressing one is the computational scalability of persistent homology which still prevents large-scale applications. Classic algorithms for topological features extraction had a memory and time complexity which scales exponentially on the number of simplices. On the other hand, large-scale static and dynamic graphs can be challenging to process and store, due to their size and the continuous change of communication patterns between nodes. The problem of summarizing large-scale dynamic graphs, maintaining the evolution of their structure and the communication patterns has been successfully addressed by F.Bonchi and cooperators. Their approach was based on grouping the nodes of the graph in supernodes according to their connectivity and communication patterns. The resulting summary graph preserves the information about the evolution of the graph within a time window. They also proposed two online, distributed, and tunable algorithms for summarizing this type of graphs.

POLITECNICO DI TORINO

UNIVERSITÀ DEGLI STUDI DI TORINO

In this context, the research activities of the candidate will focus on exploring to which extent it will be possible to adapt to the TDA setting techniques developed in studying large-scale static and dynamic graphs.

The research activity fits in the SmartData@PoliTo interdepartmental center, that brings together competences from different fields, ranging from modeling to computer programming, from communications to statistics. The candidate will join this interdisciplinary team of experts and collaborate with them.

The research activity will be co-advised by. Francesco Bonchi (ISI Fundation), and in the context of the joint PhD between Politecnico di Torino and University of Turin.

OBIETTIVI/OBJECTIVES: The goal of the research activity is threefold:

1. To find new efficient ways to represents in a compressed way simplicial complexes losing in a controllable way the topological information stored thereby;
2. To define new algorithms to find "canonical bases" of the homology groups, i.e. the harmonics, which, in the one dimensional case are nothing but the (polygonal) cycles of the underlying graph.
3. Apply 1 and 2 to real data

At the end, the candidate is expected to maturate competences in applied mathematics and computer science. They will gain a strong expertise in topological data analysis applied to big data, design of algorithms for graph and simplicial complexes, competences in big data handling and processing.

CAPACITÀ E COMPETENZE RICHIESTE PER LO SVOLGIMENTO DELL'ATTIVITÀ DI RICERCA/SKILLS AND COMPETENCIES FOR THE DEVELOPMENT OF THE ACTIVITY: The candidate is required to have very good competences in linear algebra, topology, geometry and discrete mathematics, experience in algorithm design/analysis and good programming skills.

POLITECNICO DI TORINO

UNIVERSITÀ DEGLI STUDI DI TORINO

Titolo del progetto di Ricerca/Research Title:

DATA-DRIVEN MACHINE LEARNING METHODS FOR PHYSICALLY-BASED SIMULATIONS

FINANZIATO DA / FUNDED BY: Politecnico di Torino

SUPERVISOR: Prof. Francesco Vaccarino *(francesco.vaccarino@polito.it)*
Prof. Stefano Berrone (stefano.berrone@polito.it)

CONTATTO / CONTACT: http://bigdata.polito.it

CONTESTO DELL'ATTIVITÀ DI RICERCA / CONTEXT OF THE RESEARCH ACTIVITY:

Machine learning (ML) is one of the most successful technology used in current days to gather information out of big data. Methodologies developed thereby, in particular deep neural networks with their various training settings, were proven to be incredibly powerful in gaining new knowledge overcoming very often human experts like e.g. in breast cancer prediction or in performing automated tasks, e.g. Tesla self-driving car, and in complicated games like Google Alpha Go victory against the world Go champion.

One of the problem with machine learning is that as exciting as their performance gains have been, though, nobody knows quite how they work. And that means no one can predict when they might fail.

This pose several questions, on the business side, where data and their accompanying analytics bring real value if and only if they are "actionable", but also and even more seriously, on the scientific side, especially the medical one.

On the other hand, there is another, older, aspect of the use of computers in science and business, namely simulation. Natural and human-generated systems such as weather, biological processes, supply chains, or computers, can be represented by mathematical models and computer software. Such models are widely used today to better understand and predict the behavior of such systems by means of simulation.

Here we face a sort of dual situation with respect to ML, namely, we know how to produce data and they might form a quite big ensemble, which has all the Vs characterizing big data.

The candidate will focus his activity on studying the interplay between machine learning (but not only), computer simulations and statistical models: by analyzing using ML techniques the configuration space coming out from a knows mathematical/statistical model we will try to identify the relevant parameters and to refine/simplify the model. After this they will use the knowledge acquired to infer models form big data, whose configuration space approximate the given data and to use this simplified model to transform correlation in causation to make our information finally "actionable".

We will use also the developed framework to increase the population of high quality-low quantity dataset and to work on model reduction, assessment and validation in the area of FEM.
The research activity fits in the SmartData@PoliTo interdepartmental center, that brings together competences from different fields, ranging from modeling to computer programming, from communications to statistics. The candidate will join this interdisciplinary team of experts and collaborate with them, and in the context of the joint PhD between Politecnico di Torino and University of Turin.

OBIETTIVI/OBJECTIVES: To develop a mathematical and computational framework addressing the questions raised above.

CAPACITÀ E COMPETENZE RICHIESTE PER LO SVOLGIMENTO DELL'ATTIVITÀ DI RICERCA/SKILLS AND COMPETENCIES FOR THE DEVELOPMENT OF THE ACTIVITY: The candidate is required to have very good competences in basic machine learning, topology/geometry, numerical analysis, experience in algorithm design/analysis and good programming skills.

Titolo del progetto di Ricerca/Research Title:

RESILIENT CONTROL, INFORMATION FLOWS, AND INCENTIVE MECHANISMS FOR INTELLIGENT TRANSPORTATION SYSTEMS

FUNDED BY: Politecnico di Torino

SUPERVISOR: Giacomo Como - Lagrange Department of Mathematical Sciences - Politecnico di Torino - giacomo.como@polito.it

CONTACT: http://www.disma.polito.it/personale/scheda/(nominativo)/giacomo.como
giacomo.como@polito.it

CONTEXT OF THE RESEARCH ACTIVITY

Efficient and resilient operation of transportation systems are among today's most pressing societal challenges. Fast-increasing demand combined with slower and very costly growth in physical infrastructure capacity has led to widespread traffic congestion on several urban andhighway networks across the world. Recent technological advancements in terms of smart sensing, communication, and computation devices havecreated unprecedented possibilities for increasing the transportation system's performance without expensive infrastructure investments. The resulting cyber-physical architecture is often referred to as Intelligent Transportation System (ITS).
On the other hand, the pervasive use of smart technologies has increased the complexity of the ITS and amplifying its inherent systemic risks. In fact, even when designed to perform well under normal operation conditions, ITSs tend to exhibit critical fragilities in response to unexpected disruptions that, started as small local perturbations, have the potential to build up through cascading mechanisms. A key challenge in mitigating such fragilities and increasing the efficiency and resilience of ITSs comes from the fact that the system involves human decision makers
with competing interests and heterogeneous behaviors. The challenges above necessarily require novel interdisciplinary approaches. This Joint Research Project with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology will build novel mathematical methodology leveraging on the underlying problem structure and combining tools from different fields such as control systems, convex optimization, game theory, and mechanism design.

SKILLS AND COMPETENCIES FOR THE DEVELOPMENT OF THE ACTIVITY

The candidate should have earned an undergraduate degree in the mathematical, physical, or engineering sciences. She/he is expected to have a solid mathematical background and to be eager to apply it to the solution of current societal challenges. The student will spend 18-24 months of her/his PhD at the Laboratory for Information and Decision. Systems at the Massachusetts Institute of Technology.