Ph.D. in Information Technology
Thesis Defense

August 29th, 2025
at 16.00 pm
Online Meeting

Hassan Nazeer CHAUDHRY – XXXII Cycle

Efficient Processing of Graph-based Data Streams

Supervisor: Prof. Matteo Giovanni Rossi
Co-supervisor: Prof. Alessandro Margara

Abstract:

Graph data structures model relations between entities in many different application domains. With the emergence of the Internet of Things (IoT) and Big Data, graphs have become very large. Modern-day graph processing systems partition graphs across a distributed system and process them in parallel, combining the final result on a centralised system. Although Graph processing systems enable scalable distributed computations over large graphs, they are limited to static scenarios in which the graph's structure does not change. Most modern-day applications are dynamic, resulting in graphs that continuously evolve. Understanding the evolution of graphs is key to enabling timely reactions when necessary. This Ph.D. addresses the problem by proposing a new model to express temporal patterns over graph data structures. The model seamlessly integrates computations over graphs to extract relevant values and temporal operators that define patterns of interest in the evolution of the graph. During the research, the syntax and semantics of this model were developed, and its concrete implementation was in a framework called FlowGraph, a middleware for temporal pattern recognition in large-scale graphs. The performance and scalability of FlowGraph are thoroughly evaluated using various workloads and use cases. FlowGraph presents a level of performance that is comparable to any state-of-the-art graph processing tool which processes static graphs. In the presence of temporal patterns, it can further optimise processing by avoiding complex graph computations until strictly necessary for pattern evaluation.

PhD Committee

Prof. Gianpaolo Cugola**, Politecnico di Milano**

Prof. Daniele Dell'Aglio, **Allborg University**

Prof. Davide Frey, **INRIA**