PhD in Computer and Control Engineering

Research Title: Urban intelligence

Funded by: DAUIN + “SmartData@Polito” Center

Supervisor:
- Silvia Chiusano (DIST) - silvia.chiusano@polito.it
- Danilo Giordano (DAUIN) – danilo.giordano@polito.it

Contact:
- http://smartdata.polito.it
- http://dbdmg.polito.it/
- www.tlc-networks.polito.it/

Context of the research activity:
Due to the rapid development of cities and the increase in urban population, past decades have seen many urban issues arise, such as traffic congestion, energy shortages and air pollution. At the same time, an increasing volume of a variety of data is generated and collected with the help of new technologies.

Urban intelligence entails the acquisition, integration, and analysis of big and heterogeneous data collections generated by a diversity of sources in urban spaces to profile the different facets and issues of the urban environment.

Digging deep urban data collections can unearth a rich spectrum of knowledge valuable to characterize citizen behaviours, identify weaknesses and strengths of the services provided in the urban area as well as improve the quality of these services or even devise new ones. However, data analytics on these data collections is still a daunting task, because they are generally too big and heterogeneous to be processed through data analysis techniques currently available. Consequently, from a data science perspective, data emerging from today's urban environment give rise to a lot of challenges that constitute a new inter-disciplinary field of research.

The research activity fits in the SmartData@Polito interdepartmental centre, that brings together competences from different fields, ranging from modelling to computer programming, from communications to statistics. The candidate will join this interdisciplinary team of experts and collaborate with them.
The PhD student will work on the study, design and development of proper data models and novel solutions and for the acquisition, integration, storage, management and analysis of big volumes of heterogeneous urban data.

The research activity involves multidisciplinary knowledge and skills including database, machine learning techniques, and advanced programming.

Different case studies in urban scenarios such as citizen-centric contexts, urban mobility, and healthy city will be considered to conduct the research activity. The objectives of the research activity consist in identifying the peculiar characteristics and challenges of each considered application domain and devise novel solutions for the management and analysis of urban data for each domain. More urban scenarios will be considered with the aim of exploring the different facets of urban data and evaluating how the proposed solutions perform on different data collections.

More in detail, the following challenges in the context of data analytics will be addressed during the PhD:

(i) Suitable data fusion techniques and data representation paradigms should be devised to integrate the heterogeneous collected data into a unified representation describing all facets of the targeted domain. For example, urban data are usually acquired by means of different sensor networks deployed in the city. Since these data are often collected with different spatial and temporal granularities, suitable data fusion techniques should be devised to support the data integration phase, and provide a spatio-temporal alignment of collected data.

(ii) Adoption of proper data models. The storage of heterogeneous urban data collections requires the use of alternative data representations to the relational model such as NoSQL databases (e.g., MongoDB).

(iii) Design and development of algorithms for big data analytics. Urban data is usually characterized by spatio-temporal coordinates describing when and where data has been acquired. Spatio-temporal data has unique properties, consisting of spatial distance, spatial hierarchy, temporal smoothness, period and trend, which entails the design of suitable data analytics methods. Moreover, huge volume of data demands the definition of novel data analytics strategies also exploiting recent analysis paradigms and cloud based platforms as Hadoop and Spark.
Proper strategies will be also devised for data and knowledge visualization.

The PhD work plan can be summarized as follows:

(a) During the 1st Year the PhD student will review the recent literature on urban computing to identify the up-to-date research directions and the most relevant open issues in the urban scenario. Based on the outcome of this preliminary explorative analysis, an application domain (such as urban air pollution) will be selected as a first reference case study. This domain will be investigated to identify the most relevant data analysis perspectives for gaining useful insights and to assess of main data analysis issues. The student will perform an exploratory evaluation of state-of-the-art technologies and methods on the considered domain, and will present a preliminary proposal for the optimization techniques of these approaches.

(b) Based on the results of the 1st year activity, during the 2nd and 3rd Year the PhD student will design and develop a suitable data analysis framework including innovative analytics solutions to efficiently extract useful knowledge in the considered domain, aimed at overcoming weaknesses of state-of-the-art methods. Moreover, the PhD student will progressively consider a larger spectrum of application domains in the urban scenario. The student will evaluate if and how his/her proposed solutions can be applied to the new considered domains as well as he/she will propose novel analytics solutions.

During the PhD, the student will have the opportunity to cooperate in the development of solutions applied to research projects on smart cities. The student will participate to conferences presenting the results of his/her research activity.

Skills and competencies for the development of the activity

The candidate must have excellent knowledge of machine learning techniques, data analysis methodologies, and business intelligence methodologies.

The candidate should have excellent programming skills, with also knowledge of Big Data platforms, such as Hadoop, Spark and MongoDB.