From Data To Models And Back - DataMod 2016

Vienna, 8th July 2016

Co-Located with STAF 2016

Program of the day

9:00 - 9:20 – Welcome

9:20 - 10:10 - Invited Talk - Mirco Musolesi: “Mining Big (and Small) Mobile Data for Social Good”

Abstract: An increasing amount of data describing people’s behaviour is collected by means of applications running on smartphones or directly by mobile operators through their cellular infrastructure. This information is extremely valuable for marketing applications, but it has also an incredible potential to be beneficial for society as a whole, thanks to applications in a variety of fields, from healthcare to transportation, from geodemographics to national security. In particular, mobile data can be extremely valuable for developing and evaluating quantitative models of human behaviour, which can be used as a basis for the development of intelligent mobile systems. In this talk I will analyze the challenges and opportunities in using big (and small) data for applications of high societal and commercial impact discussing the current work of my lab in the area of mobile data mining and anticipatory mobile computing. The scope of my talk will be broad, encompassing both modelling and systems-oriented issues.

Bio: Mirco Musolesi is a Reader in Data Science at the Department of Geography at University College London and a Faculty Fellow at the Alan Turing Institute, the UK National Institute for Data Science. He received a PhD in Computer Science from University College London and a Master in Electronic Engineering from the University of Bologna. He held research and teaching positions at Dartmouth College, Cambridge, St Andrews and Birmingham. He is a computer scientist with a strong interest in sensing, modelling, understanding and predicting human behaviour and dynamics in space and time, at different scales, using the "digital traces" we generate daily in our online and offline lives. He is interested in developing mathematical and computational models as well as implementing real-world systems based on them. This work has applications in a variety of domains, such as intelligent systems, ubiquitous computing, networked systems, healthcare, security & privacy, and data analytics for "social good”. More details about his research profile can be found at: http://www.ucl.ac.uk/~ucfamus/

10:10 – 10:30 Questions and Panel Discussion

10:30 – 11:00 Coffee Break

11:00 - 11:30 - Riccardo Guidotti, Giulio Rossetti and Dino Pedreschi. “Audio Ergo Sum: A Personal Data Model For Musical Preferences”

11:30 - 12:00 - Mohamed Aymen Ben Hajkacem, Chiheb Eddine Ben N'Cir and Nadia Essoussi. “An Accelerated MapReduce-based K-prototypes for Big Data”

12:00 - 12:30 - Antonio Cerone. “Refinement Mining: Using Data to Sift Plausible Models”

12:30 - 13:00 - Daniel Reijsbergen. “Probabilistic Modelling of Station Locations in Bicycle-Sharing Systems”

13:00 – 14:00 - Lunch

14:00 - 14:50 - Invited Talk - Emanuela Merelli “The Topological Field Theory of Data: a program towards a novel strategy for data mining through data language”

Abstract: We aim to challenge the current thinking in IT for the “Big Data” question, proposing a program aiming to construct an innovative methodology to perform data analytics in a way that returns an automaton as a recognizer of the data language: a Field Theory of Data. We suggest to build, directly out of probing data space, a theoretical framework enabling us to extract the manifold hidden relations (patterns) that exist among data, as correlations depending on the semantics generated by the mining context. The program, that is grounded in the recent innovative ways of integrating data into a topological setting, proposes the realization of a Topological Field Theory of Data, transferring and generalizing to the space of data notions inspired by physical (topological) field theories and harnesses the theory of formal languages to define the potential semantics necessary to understand the emerging patterns.

Bio: Emanuela Merelli is full professor of Computer Science at the University of Camerino. She founded the BioShape laboratory, a multidisciplinary environment where young researchers and students work for conceiving new formal and computational methods for “decode” the behavior of complex biology systems. She has been the coordinator of the TOPDRIM FP7-FET project and the partner of LITBIO (Laboratory for Interdisciplinary Technologies in Bioinformatics) Italian FIRB project. Currently, she is working for launching a new program on topological field theory of data with the aim to pave a way for synthesizing automata as recognizers of languages derived from topological data space and for understanding the complexity of some computing paradigms. She published many papers in refereed international journals and she is continuously involved in the organization of events with interdisciplinary character.

14:50 – 15:10 Questions and Panel Discussion

15:10 - 15:30 - Nieves Atienza, Rocio Gonzalez-Diaz and Matteo Rucco. “Separating Topological Noise from Features using Persistent Entropy”

15:30 – 16:00 Coffee Break

16:00 - 16:30 - Giovanni Pardini and Paolo Milazzo.” A High-Level Model Checking Language with Compile-time Pruning of Local Variables”

16:30 - 16:50 - Martyn Ellison, Radu Calinescu and Richard Paige. “Towards Platform Independent Database Modelling in Enterprise Systems”

16:50 - 17:30 – Panel Discussion and Closing