An Introduction to Computational Urban Science

Konstantinos Pelechrinis (University of Pittsburgh) & Daniele Quercia (Bell Labs)

Currently more than 50% of the world's population lives in cities. This percentage is projected to increase to 70% by 2050 based on a recent report by the United Nations. These levels of urbanization can be partially explained from the long-lived thought of cities as the paramount instrument for innovation and wealth creation. However, the unprecedented levels of urban population has lead the cities to become the main source of crimes, diseases and pollution. Social and urban scientists have been always interested on how people interact within the cities - both with other dwellers and with the city infrastructure as well. However, until recently many of the theories could not be tested at a large scale and/or were based on theoretical assumptions. With the advancements in information technology we are now able to collect data from a large spectrum of urban life facets - ranging from our interaction with the public transportation network to the amount of trash each dweller generates - that can consequently allow us to validate and/or update existing theories, quantify the interactions between dwellers and the environment, as well as inform public policy decisions.

In this tutorial, we will begin by covering urban and social theories developed in the past 50 years including the gravity model, psychological maps and the notion of aesthetic capital. We will further presents an application of each theory and ways that we can utilize Web technologies to validate these theories and further exploit them for applications such as navigation algorithms that target happy pedestrian routes.

We will then move on to describe data analytics methods that can be used in the context of urban informatics. In particular, we will first discuss quasi-experimental techniques (e.g., difference-in-differences, regression discontinuity) that can be used to analyze and reach conclusions from observational data in a way that mimics an experimental design. Furthermore, we will cover basic tensor-based techniques for analyzing the highly heterogenous urban data that originate from a variety of urban data sources. The focus will be on matrix and tensor factorization techniques that can provide information for latent urban activity patterns.

We will present these analytical methods through examples from the urban space (e.g., transportation, urban economy, public policy etc.). Throughout our tutorial we will also provide pointers to the relevant methodological literature.

Target audience: The target audience for this tutorials includes academic and industry researchers as well as government agencies. The audience is expected to have *basic familiarity* with research methods and statistics. While pointers to advanced materials will be provided, advanced knowledge of statistics and linear algebra is not necessary to follow the tutorial.

Bios

Konstantinos Pelechrinis Dr. Pelechrinis is interested in all aspects of the information cycle (collection, transfer and analysis) and his ultimate goal is to deliver information-centric solutions for the social good. He is an assistant professor at the School of Information Sciences at the University of Pittsburgh, where he is mainly focused on the analysis of social network/media data with an application on urban informatics. He is involved in understanding the social, spatial, temporal and network dynamics of the behavior of people, developing models and algorithms for intelligent urban services as well as studying the effect of location-based social media on local businesses. In the past, he has been involved in research in computer networking and, in particular, wireless and mobile networks - protocol design, real world experimentation, and performance analysis for wireless network systems. He is extremely interested in the design and implementation of practical systems, based on analytical frameworks. Konstantinos received his PhD from the Computer Science department of University of California, Riverside, in 2010. Previously he obtained his MSc degree from the Computer Science department of University of California, Riverside in 2008 and the diploma of Electrical and Computer Engineering from the National Technical University of Athens, Greece, in 2006. He has also held research intern positions at LANL, Thomson Research Labs Paris and MSR Cambridge. He was a visiting researcher at the University of Thessaly during Fall 2008. He is a recepient of the Army Research Office Young Investigator (2015) for his work on composite networks.

Daniele Quercia Dr. Quercia leads the Social Dynamics team at Bell Labs in Cambridge (UK). He is interested in the relationship between online and offline worlds and his work has been focusing in the areas of urban informatics. His research has been published in leading venues including ICSE, Ubicomp, ICDM, CSCW, RecSys, WSDM, and WWW, received a best paper award from ACM Ubicomp and from AAAI ICWSM, and an honorable mention from AAAI ICWSM, and has been featured on more than 80 international news outlets. He spoke at TED, wrote for BBC, and has been named one of Fortune magazine's 2014 Data All-Stars. He was Postdoctoral Associate at the Massachusetts Institute of Technology where he worked on social networks in a city context, and his PhD thesis at UC London was nominated for BCS Best British PhD dissertation in Computer Science. During his PhD, he was a Microsoft Research PhD Scholar and MBA Technology Fellow of London Business School, and he also interned at the National Research Council in Barcelona and at National Institute of Informatics in Tokyo. He studied at Politecnico di Torino (Italy), Karlsruhe Institute of Technology (Germany), and University of Illinois (USA).