Mining Large Multi-Aspect Data: Algorithms and Applications

Evangelos (Vagelis) Papalexakis
Carnegie Mellon University

10/20/2015 at 02:00 pm
Core A (Room 301)

Abstract

Given a Knowledge Base that records millions of relations of the form Barack Obama is the president of USA, how can we automatically learn new synonyms and enhance the Knowledge Base? Imagine now measuring the brain activity of a person while reading words that appear in this Knowledge Base; how can we relate information processing in the brain, and information found on the World Wide Web? Can we use both pieces of data in order to enhance knowledge extraction in both scenarios? On a third, seemingly unrelated, application, consider having different views of a social network, e.g. observing who is calling whom, who sends e-mails to whom, and who texts whom; can we use this rich information towards community and anomaly detection? What if we also have demographic information about the people of the network? Can we further enhance our analysis? The key underlying theme behind all the above applications is the multi-aspect nature of the data, with the ultimate question being: how can we take advantage of all different aspects? And if so, can we analyze sets of multi-aspect data jointly? Finally, can we automatically, and in a mostly unsupervised setting, filter out aspects of the data which are redundant or not beneficial for the task at hand, and at the same time, being able to assess the quality of our analysis despite the lack of ground truth data? In this talk, we discuss fast, scalable, and interpretable algorithms (with specific emphasis on Tensor Analysis), and their application to a wide variety of multi-aspect data problems.

Bio

Evangelos (Vagelis) Papalexakis is a PhD candidate at the School of Computer Science at Carnegie Mellon University (CMU), under the supervision of Prof. Christos Faloutsos since August 2011. Prior to joining CMU, he obtained his Diploma and MSc in Electronic & Computer Engineering at the Technical University of Crete, in Greece. Broadly, his research interests span the fields of Data Mining, Machine Learning, and Signal Processing. His research involves designing scalable algorithms for mining large multi-aspect

datasets, with specific emphasis on tensor factorization models, and applying those algorithms to a variety of real world multi-aspect data problems. His work has appeared in KDD, ICDM, SDM, ECML-PKDD, WWW, PAKDD, ICDE, ICASSP, IEEE Transactions of Signal Processing, and ACM TKDD. He has a best student paper award at PAKDD'14, finalist best papers for SDM'14 and ASONAM'13 and he was a finalist for the Microsoft PhD Fellowship and the Facebook PhD Fellowship. Besides his academic experience at CMU, he has industrial research experience working at Microsoft Research Silicon Valley during the summers of 2013 and 2014 and Google Research during the summer of 2015.

Faculty Host: Tina Eliassi-Rad