NJIT

Department of Computer Science

A New Generation of Query Processing for Advanced Data Analytics

Nikolaos Tziavelis
Northeastern University

Hosted by Cristian Borcea

Date: Wednesday, March 20, 2024
Coffee: 10:45 AM – 11:00 AM
Time: 11:00 AM – 12:00 PM
Location: GITC 4402 (4th floor Seminar Lecture Hall)
WebEx link: https://njit.webex.com/njit/j.php?MTID=m32a65e8538951a9235a4610ec62ead32

Abstract:
While both the amount of data and our need to analyze them are growing, today's data systems are limited in the analytical tasks they can handle efficiently. When presented with a complex analytical task, they often become unresponsive because they resort to naive query-processing algorithms. A major culprit is a join operation in the middle of the query-processing pipeline: that join can often produce unnecessarily large intermediate results even though the final output is small. Overcoming this challenge by "pushing" different kinds of operators behind joins is not always simple and often requires advancements in algorithms. In this talk, I will present my work on improving the algorithmic capabilities of data systems, focused on the case where users have preferences over the query answers and are interested in retrieving only a select few of them. I show that for a large class of these queries, it is possible to avoid materializing intermediate joins, resulting in non-trivial complexity guarantees and efficient implementations that outperform existing systems by orders of magnitude. Besides addressing fundamental questions regarding the limits of query processing, this work opens up unexplored possibilities for system design. The long-term goal is a new generation of scalable systems for data analytics that can efficiently handle queries and data that are considered infeasible today.

Bio:
Nikolaos (Nikos) Tziavelis is a PhD candidate at Northeastern University, advised by Mirek Riedewald and Wolfgang Gatterbauer, and before that, he received a Diploma in Electrical and Computer Engineering from the National Technical University of Athens. His research interests lie in the design and implementation of novel practical algorithms for query processing in data systems. His work has been awarded with a Google PhD fellowship, a PODS 2021 "Best of" recognition, and a 2023 PhD Workshop Best Paper Award.
https://ntzia.github.io/