

Taming Large Intermediate Results for Joins over Graph-Structured Relations Using Factorization

Amine Mhedhbi, Ph.D. student, University of Waterloo.

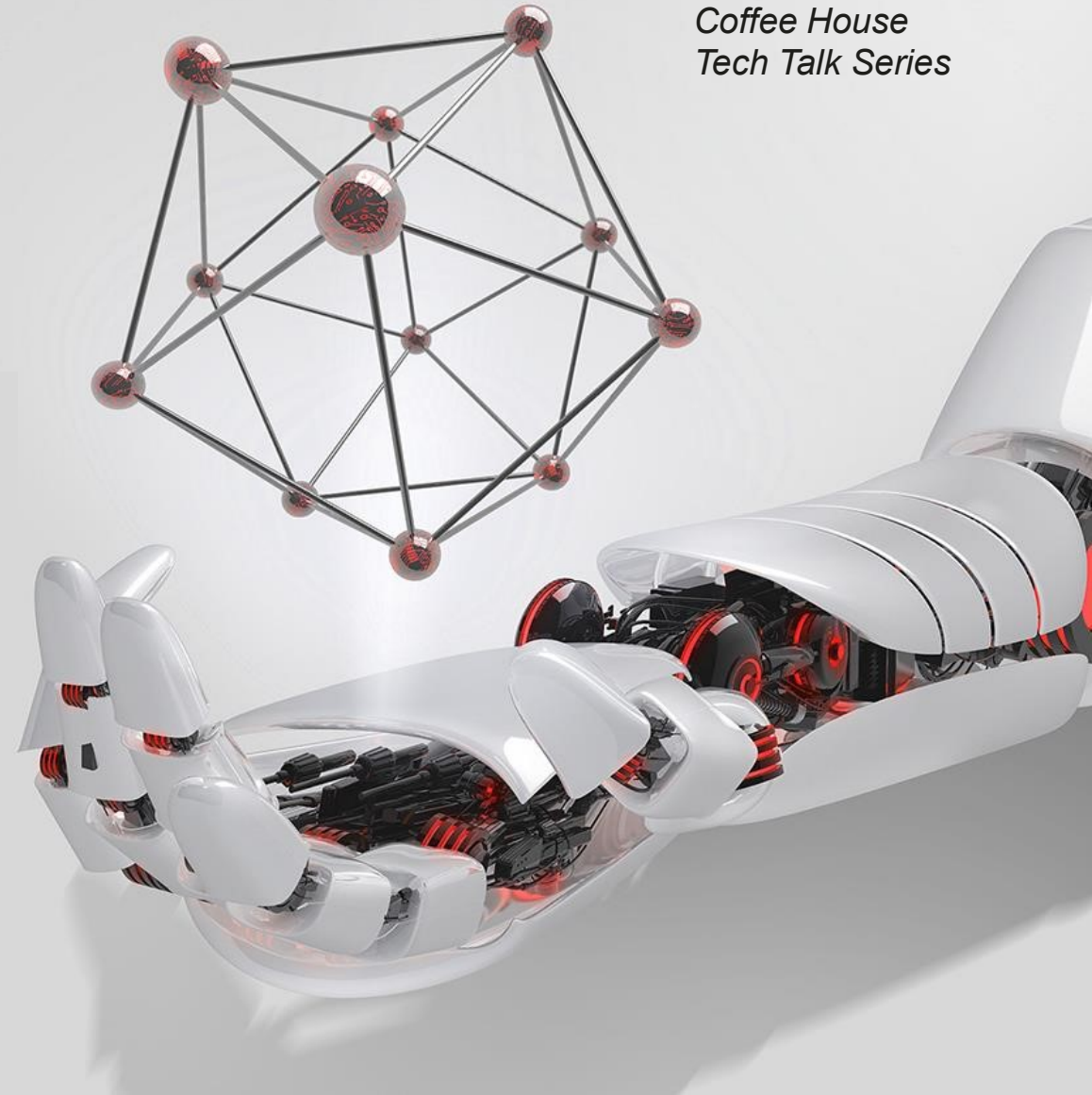
Abstract

Select-project-join (SPJ) queries over graph-structured relations i.e., primarily over many-to-many relations lead to an explosion in the size of intermediate results. We introduce SPJ query evaluation using factorized representations. Factorized representations are succinct representations for intermediate relations that are asymptotically smaller than flat representations and have been shown to lead to asymptotically faster query run times. Existing implementations, however, incur a lot of overhead as they rely on full materialization and therefore abandon pipelining, among other classical optimizations. Furthermore, they use more memory resources than necessary. Other implementations are not generalizable and rely on specific aspects of factorization to target certain queries. All of these implementations are harder to adopt by database management systems thus requiring a complete system rewrite.

In this talk, I will present a comprehensive design and implementation of a query processor that relies on factorized representations that is practical for relational or graph analytical database management systems. The system primarily relies on: i) pipelining while maintaining a factorized representation that postpones or completely avoids Cartesian Products; and ii) late materialization of intermediate relations that we cache and reuse. The query processor design, combined with a new dynamic programming optimizer, provides significant speedups when queries benefit from factorization with no penalty otherwise. The approach also enables robust query processing.

About the Speaker

Amine Mhedhbi is a Ph.D. student at the University of Waterloo. His work focuses on query processing, optimization, and storage techniques for large-scale graph-structured data management. As part of his research, he designs and implements the GraphflowDB system. He received the VLDB Best Paper Award in 2018 and the Microsoft Ph.D. Research Fellowship 2020-2022.



Date & Time: Tuesday 26 April 2022, 14:00pm

WeLink: <https://welink.zhumu.com/j/184647198>

Meeting ID : 0184647198