MINGYU GUAN

(+1)470-334-6144 · myguan1@gmail.com

EDUCATION

Georgia Institute of Technology Aug. 2019 - May. 2024
Ph.D in Computer Science (GPA:4.0) **Atlanta, GA**

The Chinese University of Hong Kong(CUHK) Aug. 2017 - May. 2019
B.S. in Computer Science with Honours, First Class (GPA:3.5) **Hong Kong, China**

Sun Yat-Sen University(SYSU) Sep. 2015 - Jun. 2017
B.S. in Electronic Information Science (2+2*) (GPA:4.0) **Guangzhou, China**
*A joint program offered by SYSU and CUHK.

PUBLICATION

Mingyu Guan, Anand Padmanabha Iyer, and Taesoo Kim. DynaGraph: Dynamic Graph Neural Networks at Scale. *In Proceedings of the 5th ACM SIGMOD Joint International Workshop on Graph Data Management Experiences & Systems and Network Data Analytics (GRADES-NDA)*, Philadelphia, PA, June 2022.

WORK EXPERIENCE

RESEARCH INTERN May. 2022 - Aug. 2022
Microsoft Research *Redmond, Seattle*

· Designed and implemented a novel heterogeneous Graph Neural Network (GNN) for compromised email detection, which encodes heterogeneity of graphs efficiently by considering both path and hop information;
· Outperformed state-of-the-arts solutions in terms of accuracy and scalability;
· Cooperated with a research team and a product team to construct heterogeneous graphs from a large-scale noisy enterprise email data set and built an automatic system for detecting compromised email accounts.

RESEARCH PROJECTS

System for Dynamic Graph Neural Networks at Scale May. 2021 - Apr. 2022

· Supported efficient dynamic GNN training in both single machine and distributed settings;
· Leveraged the computational structure in the GNN-RNN approach to propose cross-layer optimizations;
· Used time-step fusion to reduce the underutilization of GPUs and cached message passing to eliminate redundant neighborhood aggregations;
· Incorporated the above techniques on DGL and PyG, two widely popular GNN processing frameworks.
· Outperformed existing state-of-the-art GNN frameworks by up to 2.7x on a number of dynamic GNN architectures and workloads.

Processing Billion-scale Dynamic Graphs on a Single Machine Jan. 2020 - Jul. 2021

· Introduced the design of cell abstraction, allowing a significant reduction in overall storage space as well as enabling a simple, yet effective load-balancing strategy;
· Proposed an API and execution model tailored for streaming graphs by incorporating a hybrid edge- and vertex-centric API coupled with the *edgeChanged* API to allow a timely reaction to graph changes;
· Designed a technique for concurrent analytics on streaming graphs, which fully exploits the similarities in data access among concurrent graph processing jobs.

Automating Massively Parallel Heterogeneous Computing Jan. 2020 - May. 2021

· Modeled input program as a hierarchical data flow graph (HDFG) to perform a set of graph-based operations and transformations for automatic optimization and parallelization;
· Performed purity checking automatically by traversing abstract syntax tree(AST) module;
· Inferred types of variables and objects automatically with both static analysis and dynamic analysis.