University of Pittsburgh | School of Computing and Information

Dissertation Defense

Doctor of Philosophy in Computer Science

"**Towards Efficient and Robust of Graph Deep Learning**" by **Yue Dai**

Date: May 30, 2025
Time: 10:00 a.m. – 12:00 p.m.
Place: 6329 Sennott Square, 210 S. Bouquet Street, Pittsburgh, PA 15260

Committee:

- Youtao Zhang, Professor, Department of Computer Science, School of Computing and Information
- Xulong Tang, Assistant Professor, Department of Computer Science, School of Computing and Information
- Stephen Lee, Assistant Professor, Department of Computer Science, School of Computing and Information
- Jun Yang, Professor, Department of Electrical and Computer Engineering, Swanson School of Engineering

Abstract:
Inspired by the success of Graph Neural Networks (GNNs), recent graph deep learning studies have introduced GNN-based models like Graph Matching Networks (GMNs) and Temporal Graph Neural Networks (TGNNs) for diverse tasks in various domains such as social media, chemistry, and cybersecurity. Despite these advances, deploying such models efficiently and robustly in real-world settings remains challenging. Three core issues impede their broader adoption: (1) suboptimal inference latencies, which fail to meet real-world responsiveness needs; (2) limited training efficiency and scalability, hindering rapid model development for targeted applications; and (3) fragile robustness against adversarial attacks, posing serious security and privacy concerns.

In this defense, will present my research on full-stack optimizations for GNN-based models. First, I will introduce Cascade, a dependency-aware TGNN training framework that boosts training parallelism without compromising vital graph dependencies, resulting in faster training while preserving model accuracy. Next, I will detail CEGMA, a software-hardware co-design accelerator that eliminates redundant computations and memory accesses in GMNs, which leads to faster inference. Finally, I will introduce Memfreezing, a novel memory-targeting adversarial attack that investigating the adversarial vulnerabilities of TGNNs alongside robust defenses. In addition, I will outline future directions that optimizes efficiency of emerging GNN–LLM hybrid models. Through this holistic approach, I aim to enable efficient, scalable, and secure GNN-based solutions across a wide range of real-world applications.