STEVENS
INSTITUTE OF TECHNOLOGY
1870

Ph.D. DISSERTATION DEFENSE

Candidate:	Xiaoxue Han
Degree:	Doctor of Philosophy
School/Department:	Charles V. Schaefer, Jr. School of Engineering and Science / Computer Science
Date:	Tuesday, April 21st , 2026
Time/Location:	1:00 PM, Gateway North Building, Room 303
Title:	Towards Open-World Graph Learning: Dynamics, Robustness, and Interpretability
Chairperson:	Dr. Yue Ning, Department of Computer Science, School of Engineering & Sciences
Committee Members:	Dr. Wendy Hui Wang, Department of Computer Science, School of Engineering & Sciences Dr. Tian Han, Department of Computer Science, School of Engineering & Sciences Dr. Zhuo Feng, Department of Electrical and Computer Engineering, School of Engineering & Sciences

ABSTRACT

Graph Neural Networks (GNNs) have emerged as the de facto model for representation learning on structured data and have demonstrated strong performance across a wide range of applications. However, while these models excel in closed-world settings—where data is static, and distributions are fixed—they face significant challenges when deployed in unpredictable, open-world environments. In the open world, graph data is inherently dynamic, with continuous temporal and structural evolution. Moreover, models must be robust to maintain reliability under distribution shifts and interpretable to enable transparent reasoning in high-stakes decision-making.

In this dissertation, we address these fundamental challenges by developing a cohesive framework for trustworthy graph learning. First, we tackle the complexity of data dynamics by moving beyond the limitations of the Markov assumption. We propose a text-enhanced multi-granularity temporal graph learning model (MTG) that integrates long-term and short-term dependencies with auxiliary textual signals to capture the evolving nature of real-world events. To further handle structural evolution in streaming settings, we introduce TACO, a topology-aware graph coarsening framework that enables continual learning by preserving essential topological "memory" while maintaining computational efficiency as graphs grow.

Second, to ensure model reliability under environmental changes, we address the challenge of out-of-distribution (OOD) generalization. We develop DeCaf, a causal decoupling framework grounded in Structural Causal Models, which independently learns unbiased feature–label and structure–label mappings. This design mitigates spurious correlations and enhances robustness under distribution shifts. Finally, to bridge the gap between complex model predictions and human understanding, we propose the Graph Concept Bottleneck (GCB). This paradigm shifts GNN interpretability from opaque subgraph-based explanations to

a human-understandable concept space, encouraging models to “reason” through meaningful semantic concepts before making predictions.

Together, these contributions advance graph learning from static pattern matching toward a more robust, interpretable, and dynamic-aware paradigm, enabling reliable deployment in open-world environments.