SOVEREIGN: Generalizing GNNs with Tokenized Mixture of Experts

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 27, 2026

Abstract

Deployed graph neural networks (GNNs) are frozen at deployment yet must fit clean data, generalize under distribution shifts, and remain stable to perturbations. We show that static inference induces a fundamental tradeoff: improving stability requires reducing reliance on shift-sensitive features, leaving an irreducible worst-case generalization floor. Instance-conditional routing can break this ceiling, but is fragile because shifts can mislead routing and perturbations can make routing fluctuate. We capture these effects via two decompositions separating coverage vs selection, and base sens

1 Introduction

Analysis of: Generalizing GNNs with Tokenized Mixture of Experts. Research goal: How robust is SMoES to cross-domain generalization and modality imbalance in multimodal benchmarks like MMBench and SEED-Bench, compared to hard-routing MoE baselines, under controlled perturbations to image-text alignment?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

10 papers retrieved. 3 claims extracted, 3 verified. Tribunal: 7.5/10 $\rightarrow$ APPROVE (revision_round=0). Policy: AUTO_APPROVE.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
STEM-GNN achieves a stronger three-way balance, improving robustness to degree/homophily shifts and to feature/edge corr	✓	0.39
Static inference induces a fundamental tradeoff: improving stability requires reducing reliance on shift-sensitive featu	✓	0.37
Instance-conditional routing can break this ceiling, but is fragile because shifts can mislead routing and perturbations	✓	0.32

References

- http://arxiv.org/abs/2604.21478v1
- http://arxiv.org/abs/2604.23996v1
- http://arxiv.org/abs/2602.09258v1