

Metapath Sampling Strategies and Their Impact on HGNN Convergence and Multi-Task Accuracy

Assignee Research

June 2, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: What is the impact of different metapath sampling strategies (random vs. heuristic-based) on the convergence speed and final accuracy of HGNNs in multi-task learning settings (e.g., node classification and link prediction) on citation graphs? 10 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: MECCH: Metapath Context Convolution-based Heterogeneous Graph Neural Networks. Research question: What is the impact of different metapath sampling strategies (random vs. heuristic-based) on the convergence speed and final accuracy of HGNNs in multi-task learning settings (e.g., node classification and link prediction) on citation graphs?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.2/10.

3 Results

16 papers retrieved. 10 claims extracted; 0 independently verified. Quality review score: 2.2/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
MECCH is evaluated on five heterogeneous graph datasets: IMDB, ACM, DBLP, LastFM, and PubMed.	×	0.11
For node classification, MECCH is evaluated on IMDB, ACM, and DBLP datasets.	×	0.07
For link prediction, MECCH is evaluated on LastFM and PubMed datasets.	×	0.11
The IMDB dataset contains 12,722 nodes of 3 types and 37,288 edges of 4 types.	×	0.03
The ACM dataset contains 8,994 nodes of 3 types and 25,922 edges of 4 types.	×	0.03
The DBLP dataset contains 18,405 nodes of 3 types and 67,946 edges of 4 types.	×	0.03
The LastFM dataset contains 20,612 nodes of 3 types and 201,908 edges of 5 types.	×	0.03
The PubMed dataset contains 63,109 nodes of 4 types and 368,245 edges of 16 types.	×	0.03
MECCH achieves a Macro-F1 score of 0.905 and a Micro-F1 score of 1.727 on the ACM dataset.	×	0.04
MECCH achieves a Macro-F1 score of 4.142 and a Micro-F1 score of 3.645 on the DBLP dataset.	×	0.04

References

- <http://arxiv.org/abs/2211.12792v2>
- <http://arxiv.org/abs/2407.11790v4>
- <http://arxiv.org/abs/2002.01680v2>