Transformer-Based Anomaly Detection Outperforms Graph Neural Networks in Cross-Domain Network Traffic

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: To what extent do transformer-based anomaly detection models improve cross-domain generalization compared to graph neural networks when measured by AUC-PR on heterogeneous network traffic benchmarks. Anomaly detection is defined as discovering patterns that do not conform to the expected behavior. Previously, anomaly detection was mostly conducted using traditional shallow learning techniques, but with little improvement. 13 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.4/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Mul-GAD: a semi-supervised graph anomaly detection framework via aggregating multi-view information. Research question: To what extent do transformer-based anomaly detection models improve cross-domain generalization compared to graph neural networks when measured by AUC-PR on heterogeneous network traffic benchmarks?.

2 Methodology

Systematic literature search across multiple databases yielded 4 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.4/10.

3 Results

4 papers retrieved. 13 claims extracted; 1 independently verified. Quality review score: 4.4/10.

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
Mul-GAD is the first approach to analyze the anomaly detection problem from the perspective of objective functions.	×	0.12
Mul-GAD finds that label-oriented objective functions have more generalized performance compared to other types.	×	0.07
Mul-GAD provides two effective fusion strategies at the view and feature levels.	×	0.14
Both view-level and feature-level fusion strategies boost detection performance.	✓	0.22
Computing the feature similarity matrix plays an important role in boosting detection performance.	×	0.04
Mul-GAD outperforms state-of-the-art methods on detection performance across the majority of datasets.	×	0.13
Mul-GAD outperforms state-of-the-art methods in terms of generalization across the majority of datasets.	×	0.11
Shallow learning methods handle anomaly problems based on spatial density, statistical distribution, and variants of cla	×	0.05
Local Outlier Factor (LOF) acquires anomaly scores by computing the spatial density of each node, where lower density co	×	0.03
K-nearest neighbor (KNN) determines the class of a node by seeking the k closest neighbors and using the majority class.	×	0.01
Shallow learning methods are constrained by inductive bias, making it hard to spot abnormal nodes masquerading as normal	×	0.03
Anomaly detection algorithms can be categorized into shallow learning and graph neural network methods.	×	0.14
Anomaly detection objective functions can be categorized into label-oriented, reconstruction-oriented, and ssl-oriented.	×	0.05

References

- http://arxiv.org/abs/2305.02496v1
- http://arxiv.org/abs/2212.05478v1
- http://arxiv.org/abs/2211.12792v2