Full-Graph vs. Mini-Batch Training Robustness in Adversarial Graph Neural Networks

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: Does the choice between full-graph and mini-batch training pipelines affect the robustness of Graph Neural Networks against adversarial perturbations in control flow graphs used for security analysis. Malware remains a big threat to cyber security, calling for machine learning based malware detection. While promising, such detectors are known to be vulnerable to evasion attacks. 12 claims were extracted from source literature; 6 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.6/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Adversarial Deep Ensemble: Evasion Attacks and Defenses for Malware Detection. Research question: Does the choice between full-graph and mini-batch training pipelines affect the robustness of Graph Neural Networks against adversarial perturbations in control flow graphs used for security analysis?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.6/10.

3 Results

14 papers retrieved. 12 claims extracted; 6 independently verified. Quality review score: 6.6/10.

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
246,002,762 new malware variants emerged in 2018	×	0.09
Kaspersky detected 5,321,142 malicious Android packages in 2018	×	0.02
Ensemble methods significantly enhance the robustness of deep neural networks against a wide range of attacks	✓	0.24
Ensemble methods promote robustness when base classifiers are robust enough	✓	0.24
Ensemble attacks can evade the enhanced malware detectors effectively, even notably downgrading the VirusTotal service	✓	0.33
A few manipulations are enough to perturb a malware example into an adversarial one, by which the perturbed malware exam	×	0.11
Researchers have proposed enhancing the robustness of classifiers using ensemble methods such as adversarial training of	✓	0.15
Attackers can leverage ensemble methods to promote attack effectiveness as well such as by evading several classifiers o	✓	0.17
The proposed mixture of attacks enables attackers to leverage multiple generative methods and multiple manipulation sets	✓	0.21
The adapted 'max' attack and its iteration in a greedy manner boost attack effectiveness	×	0.06
Salt and pepper noises attack and pointwise attack are adapted for effective attacks when gradients of loss function suf	×	0.03
Adversarial training is instantiated using a mixture of attacks and a manipulation set with the cardinality as la	×	0.09

References

- http://arxiv.org/abs/2006.16545v1
- http://arxiv.org/abs/2601.22678v3
- http://arxiv.org/abs/2004.07919v3