Federated vs Centralized Malware Detectors Under Adversarial Poisoning Attacks

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the robustness of federated learning-based malware detectors against adversarial poisoning attacks compare to centralized models in terms of precision degradation. This work investigates the possibilities enabled by federated learning concerning IoT malware detection and studies security issues inherent to this new learning paradigm. In this context, a framework that uses federated learning to detect malware affecting IoT devices is. 13 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Federated Learning for Malware Detection in IoT Devices. Research question: How does the robustness of federated learning-based malware detectors against adversarial poisoning attacks compare to centralized models in terms of precision degradation?.

2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.2/10.

3 Results

13 papers retrieved. 13 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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
In Federated Learning (FL), algorithm training is performed in a decentralized manner by different nodes using local dat	×	0.09
In FL, decentralized nodes share model parameters instead of raw data with the rest of the network.	×	0.05
The exchange and aggregation of model parameters in FL can be performed through a central server or via a peer-to-peer a	×	0.04
After several iterations in FL, each client possesses a global model obtained as an aggregation of individual models fro	×	0.07
Previous works on FL for intrusion detection lack the use of realistic datasets in the FL context.	×	0.04
Previous works on FL for intrusion detection lack analysis on adversarial impact.	×	0.04
Previous works on FL for intrusion detection lack discussion of deployment in B5G scenarios.	×	0.04
The paper presents a use case involving a B5G scenario with Non-IID (Independent and Identically Distributed) data and n	×	0.12
The proposed security framework covers both anomaly detection and classification approaches.	×	0.06
The dataset split described in Table (p6) allocates 79% of data for training, 20% for known device testing, and 1% as un	×	0.04
An alternative dataset split described in Table (p6) allocates 39.5% for training, 39.5% for threshold selection, 20% fo	×	0.03
Table (p12) reports a centralized model performance metric of 95%.	×	0.06
Table (p12) contains data points indicating 7.8% and 7% related to benign or enigmatic classifications.	×	0.02

References

- http://arxiv.org/abs/2605.25937v1

- http://arxiv.org/abs/2104.09994v3
- http://arxiv.org/abs/2602.16480v1