Federated and Centralized Learning for Malware Detection under Obfuscation and Adversarial Attacks

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 9 peer-reviewed papers addressing the following research question: How does the detection accuracy of federated learning models compare to centralized deep neural networks when evaluated on the AndroZoo benchmark with varying levels of code obfuscation and. 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. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.3/10. 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 detection accuracy of federated learning models compare to centralized deep neural networks when evaluated on the AndroZoo benchmark with varying levels of code obfuscation and adversarial perturbations?.

2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.3/10.

3 Results

9 papers retrieved. 7 claims extracted; 0 independently verified. Quality review score: 2.3/10.

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
Federated Learning enables data privacy by design as data is not shared with any external identity.	×	0.08
Previous works dealing with FL for intrusion detection lack the use of realistic datasets in the FL context.	×	0.06
The proposed framework covers both anomaly detection and classification approaches using multi-BENCHMARK TABLES.	×	0.07
The server aggregates updated individual models from clients to create a global model.	×	0.05
Data is split into train (79%), unused (1%), and known test (20%) sets for evaluation.	×	0.02
Threshold selection uses 39.5% of the data while 20% is allocated for known device testing.	×	0.02
The centralized approach achieves 95% accuracy in detecting benign vs. malicious samples.	×	0.04

References

- http://arxiv.org/abs/2104.09994v3
- http://arxiv.org/abs/2502.01654v1
- http://arxiv.org/abs/2008.06767v2