Supervised and Unsupervised Federated Learning for Zero-Day Malware Detection in IoT Deployments

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How do supervised versus unsupervised federated learning approaches compare in terms of model accuracy trade-offs when detecting zero-day malware variants in cross-device IoT deployments. 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. 9 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.1/10. 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 do supervised versus unsupervised federated learning approaches compare in terms of model accuracy trade-offs when detecting zero-day malware variants in cross-device IoT deployments?.

2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.1/10.

3 Results

8 papers retrieved. 9 claims extracted; 1 independently verified. Quality review score: 5.1/10.

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 (FL) enables data privacy by design, as data is not shared with any external identity.	×	0.11
Previous works dealing with FL for intrusion detection lack the use of realistic datasets in the FL context, the analysi	×	0.04
The proposed security framework uses FL to detect, in a privacy-preserving fashion, cyberattacks affecting IoT devices.	✓	0.15
The proposed framework covers both anomaly detection and classification approaches using multi-client FL.	×	0.12
The use case presents a B5G scenario where there is a necessity of detecting cyberattacks affecting IoT devices, managin	×	0.07
The model aggregation in FL can be performed through a central entity, called server, or following a peer-to-peer approa	×	0.08
The dataset is split into 79% for training, 1% unused, and 20% for known device test.	×	0.03
The dataset is split into 39.5% for training, 39.5% for threshold selection, 1% unused, and 20% for known device test.	×	0.02
The centralized model achieves 95% accuracy with a 7.8% benign rate and 7% enigmatic rate.	×	0.04

References

- http://arxiv.org/abs/2104.09994v3
- http://arxiv.org/abs/2204.07772v1
- http://arxiv.org/abs/2404.10012v1