

# Unsupervised Federated Learning for Zero-Shot Malware Detection in IoT Networks

Assignee Research

May 30, 2026

## Abstract

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: To what extent do unsupervised federated learning approaches maintain detection precision-recall performance when adapting to new malware families not present in the original N-BaIoT training set. 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 4.5/10. 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: To what extent do unsupervised federated learning approaches maintain detection precision-recall performance when adapting to new malware families not present in the original N-BaIoT training set?.

## 2 Methodology

Systematic literature search across multiple databases yielded 10 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.5/10.

## 3 Results

10 papers retrieved. 9 claims extracted; 1 independently verified. Quality review score: 4.5/10.

## 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
FL approaches lack the use of realistic datasets in the FL context, the analysis on adversarial impact, or the discussio	×	0.04
The proposed security framework uses FL to detect, in a privacy-preserving fashion, cyberattacks affecting IoT devices.	✓	0.18
The framework covers both anomaly detection and classification approaches using multi-client FL.	×	0.10
The use case presents a B5G scenario with Non-IID data and non-trusted stakeholders or clients.	×	0.03
The model aggregation can be performed through a central entity or following a peer-to-peer approach.	×	0.09
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.	×	0.03
The centralized approach achieves 95% accuracy with a 7.8% benign rate and 7% malicious rate.	×	0.03

## References

- <http://arxiv.org/abs/2104.09994v3>
- <http://arxiv.org/abs/2405.00839v1>
- <http://arxiv.org/abs/2207.08187v1>