Federated Malware Detection F1-Score Transfer Across IoT Datasets Under Differential Privacy

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How does the F1-score of federated malware detection models trained on N-BaIoT transfer to unseen IoT network traffic datasets (e.g., BoT-IoT, CIC-IoT-2021) under varying differential privacy noise. 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. 12 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 F1-score of federated malware detection models trained on N-BaIoT transfer to unseen IoT network traffic datasets (e.g., BoT-IoT, CIC-IoT-2021) under varying differential privacy noise levels (ε=0.5, ε=2, ε=5)?.

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 3.2/10.

3 Results

8 papers retrieved. 12 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	×	0.10
In FL, decentralized nodes share model parameters instead of raw data with the rest of the network.	×	0.04
The aggregation of model parameters to create a global model can be performed through a central server or a peer-to-peer	×	0.08
Previous works on FL for intrusion detection lack the use of realistic datasets in the FL context.	×	0.06
Previous works on FL for intrusion detection lack analysis on adversarial impact.	×	0.05
Previous works on FL for intrusion detection lack discussion of their deployment in B5G scenarios.	×	0.05
The paper presents a use case involving a B5G scenario with Non-IID data and non-trusted stakeholders.	×	0.02
The proposed security framework covers both anomaly detection and classification approaches.	×	0.06
The experimental setup includes a data split where 79% is used for training, 20% for known device testing, and 1% is unu	×	0.04
An alternative data split configuration uses 39.5% for training, 39.5% for threshold selection, 20% for known device tes	×	0.02
A centralized model configuration achieved a result of 95% in the reported metrics.	×	0.04
A specific configuration reported a benign classification rate of 50% and a value of 7.8.	×	0.02

References

- http://arxiv.org/abs/2104.09994v3
- http://arxiv.org/abs/2404.19114v1
- http://arxiv.org/abs/2207.02337v1