Federated Deep Learning in IoT Cybersecurity: Client Heterogeneity Impacts on Inference Efficiency and Accuracy

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 6 peer-reviewed papers addressing the following research question: What is the effect of varying client heterogeneity on the inference efficiency and accuracy of federated deep learning systems in IoT cybersecurity applications. Federated Learning (FL) enables decentralized training of machine learning models on distributed data while preserving privacy. However, in real-world FL settings, client data is often non-identically distributed and imbalanced, resulting in statistical data heterogeneity which. 8 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: FedDiverse: Tackling Data Heterogeneity in Federated Learning with Diversity-Driven Client Selection. Research question: What is the effect of varying client heterogeneity on the inference efficiency and accuracy of federated deep learning systems in IoT cybersecurity applications?.

2 Methodology

Systematic literature search across multiple databases yielded 6 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.8/10.

3 Results

6 papers retrieved. 8 claims extracted; 0 independently verified. Quality review score: 3.8/10.

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 standard FEDAVG, parameter aggregation is performed by computing the weighted mean of client parameters, where weight	×	0.03
Statistical data heterogeneity emerges when there is a subpopulation shift, meaning the representation of subpopulations	×	0.12
In the context of this paper, subpopulations are defined by the combination of target labels and attributes (Y $\times$ A).	×	0.01
Class Imbalance (CI) is defined as a condition where the distribution of target labels differs between training and test	×	0.04
Class Imbalance can yield a biased classifier that performs poorly on samples from the minority class.	×	0.04
Attribute Imbalance (AI) occurs when the probability of a certain attribute in the training set is much smaller than oth	×	0.03
Attribute Imbalance can produce a classifier biased towards the majority attribute.	×	0.05
Spurious Correlation (SC) is characterized by a statistical dependency between variables.	×	0.04

References

- http://arxiv.org/abs/2007.04806v1

- http://arxiv.org/abs/2504.11216v2
- http://arxiv.org/abs/2506.02887v2