CHAPTER I
INTRODUCTION

1.1 Background

Internet of Things (IoT) has grown significantly and is used in a variety of fields, such as industry, healthcare [1], transportation [2], and smart home [3]. While this technology makes automation and data integration easier, IoT devices have limited processing performance and storage capacity. This makes IoT devices more vulnerable to cyberattacks, making it difficult to use sophisticated security measures [4]. Therefore, a security system that can detect attacks efficiently while maintaining the efficiency of the IoT device system is needed.

An approach commonly used in IoT to treat security issues is the implementation of Intrusion Detection Systems (IDS). IDS functions to monitor and recognize systems related to cyberattack activity. However, conventional IDS deployment is based on the reliance on centralized systems, which can be a point of failure and increase the risk of data leakage [5]. Therefore, an efficient and decentralized approach is required to detect attacks collaboratively, while also preserving data privacy and maintaining system efficiency.

To address these limitations in conventional IDS, Federated Learning (FL) has emerged as a promising alternative. FL enables distributed training of attack detection models on each device without sending raw data to a central server. In this way, each device can still contribute to the model training process without compromising its data privacy [6]. McMahan et al. [7] introduced FedAvg as an early aggregation method that showed effectiveness in distributed training, but had limitations in handling non-IID data and unevenness of system resources. Despite its advantages, the challenge in FL is selecting an aggregation method, particularly under non-IID data conditions. In response to these challenges, Li et al. [8] developed FedProx, a federated optimization framework that introduces the proximal term to handle statistical and systems heterogeneity directly. The research proved that FedProx improves convergence stability compared to FedAvg, and even improves test accuracy in highly heterogeneous environments. Therefore, selecting the aggregation method is crucial, as it directly affects model convergence, detection accuracy, and overall system robustness, especially in the presence of non-IID data distributions common in real-world IoT environments.

Several research have been conducted to evaluate aggregation methods in FL to improve the accuracy of IDS systems on IoT using the CICIoT2023 dataset. Parineeta et al. [9] implemented the Deep Neural Network model and FedAvg aggregation to detect Mirai botnet attacks in Covariate Shift and Concept Drift scenarios, with accuracy reaching 93%. However, this research did not focus on a specialized aggregation solution designed to overcome the heterogeneity of data between clients. Following up on this, Prasad et al. [10] explored intrusion detection in Industrial IoT networks using a Recurrent Neural Network model that is effective in handling sequential data and compared FedAvg aggregation which achieved the highest accuracy of 88.97%. FedProx which is more adaptive to data heterogeneity and produced the highest accuracy of 90.8%. However, this research did not explicitly explain the data distribution scheme using IID or non-IID distributions. To overcome the previous limitation, Chen et al. [11] evaluated the FedAvg and FedProx aggregation methods using Artificial Neural Network model and applied a straggler scheme in a non-IID data scenario. The results showed that FedAvg only achieved 79% accuracy, while FedProx achieved 84%. However, this research does not explain the hyperparameter values used in FedProx, such as the penalty parameter (μ), which has the potential to affect model performance significantly.

Although several research have been conducted to improve the effectiveness of attack detection in IoT networks using FL, there are still some important aspects that have received less attention. Most research focuses on using aggregator methods such as FedAvg and FedProx without deeply evaluating the effect of regularization strength on FedProx under non-IID data distribution conditions that are common in IoT environments. Therefore, this thesis is conducted to explicitly explore the effect of the regularization value on FedProx aggregation method, which has not been widely discussed in previous research. In addition, this thesis systematically implements and compares the performance of FedAvg and FedProx in two data distribution scenarios (IID and non-IID) to improve accuracy using Deep Neural Network (DNN) architecture trained using the CICIoT2023 dataset. This approach is expected to provide a more in-depth analysis of the effectiveness of aggregation methods in FL for IoT-based IDS.

1.2 Problem Identification

Based on previous research in the background, the following concerns will be addressed in this thesis.

- How does the choice of aggregation method affect the performance of FL in

detecting cyberattacks?

- What are the performance differences between FedAvg and FedProx in IID and non-IID data distributions?
- How does the regularization strength in the FedProx method affect model convergence and detection accuracy under IID and non-IID conditions?

1.3 Objective and Contributions

The objectives of this thesis are as follows.

- This thesis implements a FL-based IDS using a Deep Neural Network (DNN) model trained on the CICIoT2023 dataset.
- This thesis compares FedAvg and FedProx under IID and non-IID data distributions to measure the effectiveness and robustness of each method in real-world scenarios.
- This thesis explicitly investigates the impact of different regularization strengths in FedProx aggregation method.
- This thesis evaluates the model using several performance metrics such as loss, accuracy, precision, recall, and F1-score.

1.4 Scope of Work

This thesis will focus on the following areas within the field of FL.

- Only two aggregation methods are used, namely FedAvg and FedProx.
- The CICIoT2023 dataset is sampled down to 10% of its original size.
- The classification model is a fixed-architecture DNN.
- The research does not address security threats to FL, such as model poisoning or inversion attacks.

1.5 Expected Results

The expected result of this thesis is that the developed model can accurately detect and classify various types of cyberattacks on IoT networks. This thesis consists of two main stages, namely the implementation of FL using two aggregation methods FedAvg and FedProx, and the evaluation of their performance under IID and non-IID data distribution scenarios. The model training process uses the DNN algorithm, while the FL process is implemented using the Flower framework based on TensorFlow. The evaluation process uses loss, accuracy, precision, recall, and F1-score as the main performance metrics. From the analysis conducted based on the experimental results, it is expected that the research can demonstrate the effectiveness of FedProx in handling heterogeneous data conditions and its superiority over FedAvg in non-IID scenarios. This thesis is also expected to contribute valuable insights to the development of privacy-preserving IDS, particularly in the field of IoT security using collaborative artificial intelligence approaches.

1.6 Research Methodology

To achieve success in conducting this thesis, the methodology used is as follows:

- **WP 1: Literature Review**
  This work package involves reviewing previous research on IDS, FL, and aggregation methods such as FedAvg and FedProx, particularly in IoT environments. Key research gaps are identified, including issues with centralized IDS architecture, challenges under non-IID data distributions, and the limited exploration of FedProx regularization strength. This stage provides the theoretical foundation and justification for conducting a comparative analysis between FedAvg and FedProx on IID and non-IID datasets.

- **WP 2: System Design**
  In this work package, the design of the proposed system is developed using a simulation-based approach. The system comprises a FL architecture with two client nodes and one server node built using the Flower framework. The CICIoT2023 dataset is selected due to its rich attack variations across 105 IoT devices. The system design also defines how the IID and non-IID data distributions are simulated, and how the FedAvg and FedProx algorithms are implemented and parameterized.

- **WP 3: Data Preprocessing and Model Development**
  This work package includes merging, cleaning, and sampling 10% of the CICIoT2023 dataset. Preprocessing includes removing duplicates, handling missing and infinite values, encoding attack labels using one-hot encoding, and splitting data into training and testing sets. A DNN classifier is constructed using TensorFlow with three hidden layers and trained using the Adam optimizer and the categorical cross-entropy loss function.

- **WP 4: Federated Learning Simulation**
  This work package involves the execution of the FL simulation. Each client trains the local model with local data, and model parameters are aggregated by the server using FedAvg or FedProx. Two scenarios are tested: IID (uniform label distribution across clients) and non-IID (label skew distribution).

- **WP 5: System Analysis and Evaluation**
  The final work package consists of analyzing performance based on loss, accuracy, precision, recall, and F1-score. Results are compared across the aggregator configurations (FedAvg and FedProx) in both IID and non-IID settings.

1.7 Research Plan and Action Point

The research plan and action points outlined in this thesis aim to guide the systematic execution of the research. The plan details the key steps to be taken and the specific actions required to achieve the research objectives. A comprehensive overview of these steps can be found in Table 1.1, which provides a clear timeline and structure for the successful completion of the research.

Month	Action Plan
October 2024	Conduct literature review on FL, IDS, FedAvg, and FedProx
November 2024	Write a literature review emphasizing the gaps in current research and the potential areas for contribution.
December 2024	Perform data preprocessing on CICIoT2023 dataset including cleaning, sampling, and encoding.
Januari 2025	Develop and validate a baseline DNN model for intrusion detection.
Februari 2025	Implement FL simulation using Flower framework with FedAvg and FedProx
April 2025	Conduct experiments under IID and non-IID scenario using both FedAvg and FedProx aggregators.
May 2025	Evaluate and compare all experimental results.

Table 1.1 Research plan and action point