Dataset link:

https://www.kaggle.com/datasets/dhoogla/cicids2017?utm_source=chatgpt.com

🔹 PRE-THESIS 1 REPORT

Chapter 1: Introduction

1.1 Background

Unmanned Aerial Vehicles (UAVs) are increasingly deployed in Internet of Things (IoT) and edge computing environments for applications such as surveillance, traffic monitoring, disaster response, and smart cities. These UAV networks rely on wireless communication and distributed control, which makes them highly vulnerable to cyber attacks. Attacks such as Distributed Denial of Service (DDoS), brute-force attacks, and botnet-based intrusions can disrupt UAV operations and compromise sensitive data.

Intrusion Detection Systems (IDS) are commonly used to monitor network traffic and detect malicious activities. However, traditional centralized IDS approaches require collecting data from all nodes at a central server, which raises serious concerns related to privacy, scalability, and communication overhead. Federated Learning (FL) has emerged as a promising solution that enables collaborative model training without sharing raw data, making it suitable for UAV-based IoT systems.

1.2 Rationale of the Study / Motivation

The motivation for this study arises from three key challenges. First, UAV networks generate highly distributed and heterogeneous data, which makes centralized learning impractical. Second, data collected by UAVs is often non-independent and non-identically distributed (non-IID), degrading the performance of conventional machine learning models. Third, federated learning systems themselves are vulnerable to malicious clients that may send poisoned or manipulated model updates.

This research is motivated by the need to design a **robust, privacy-preserving, and scalable intrusion detection framework** that can operate effectively under non-IID conditions and resist adversarial behaviors in UAV-based IoT environments.

1.3 Problem Statement

Existing intrusion detection systems for UAV networks suffer from several limitations: centralized data dependency, privacy leakage risks, poor performance under non-IID data distributions, and vulnerability to poisoning attacks in federated learning settings. There is a lack of robust federated IDS frameworks that simultaneously address data heterogeneity, class imbalance, and malicious client behavior.

1.4 Objectives

The main objectives of this research are:

1. To design a federated learning–based intrusion detection system for UAV networks.
2. To evaluate the impact of non-IID data distribution on federated IDS performance.
3. To incorporate robust aggregation techniques to mitigate malicious client updates.
4. To compare centralized and federated intrusion detection approaches.
5. To analyze system performance using standard evaluation metrics.

1.5 Methodology (Brief)

The proposed methodology uses the CIC-IDS2017 dataset to train a lightweight convolutional neural network. Data is distributed among multiple simulated UAV clients using a Dirichlet-based non-IID partitioning scheme. Federated learning is implemented using FedAvg and robust trimmed-mean aggregation. Performance is evaluated through accuracy, precision, recall, F1-score, and ROC-AUC metrics. Poisoning attacks are simulated to test robustness.

1.6 Scope and Challenges

This research focuses on binary intrusion detection using flow-level network features. The scope includes federated learning, robustness analysis, and performance evaluation. Challenges include handling extreme data imbalance, ensuring model convergence under non-IID conditions, and mitigating adversarial client behavior without excessive communication overhead.

Chapter 2: Literature Review

2.1 Preliminaries

This section introduces key concepts such as intrusion detection systems, federated learning, non-IID data distributions, and robust aggregation techniques. It also discusses common network attack types relevant to UAV networks.

2.2 Review of Existing Research

Prior research has explored centralized IDS using deep learning models such as CNNs, LSTMs, and autoencoders. Recent studies have applied federated learning to network security, demonstrating privacy benefits. However, many existing works assume IID data and ignore robustness against poisoning attacks. Robust aggregation methods such as trimmed mean and median have been proposed, but their application in UAV-based IDS remains limited.

2.3 Summary of Key Findings

The literature indicates that federated learning is effective for privacy-preserving IDS, but performance degrades under non-IID data. Robust aggregation improves resilience but introduces trade-offs between precision and recall. There is a clear research gap in designing and evaluating robust FL-based IDS frameworks for UAV networks.

🔹 PRE-THESIS 2 REPORT

Chapter 4: System Design

4.1 Design Process or Methodology Overview

The system design follows a modular approach consisting of data preprocessing, local model training at UAV clients, federated aggregation at the edge server, and global model evaluation. Each module is designed to operate independently while supporting seamless integration within the federated learning framework.

4.2 Preliminary Design or Design (Model) Specification

A lightweight one-dimensional convolutional neural network is selected to ensure computational efficiency on resource-constrained UAVs. The model processes normalized flow-level features and outputs binary intrusion predictions. The federated server aggregates local updates using either FedAvg or trimmed-mean aggregation.

🔹 FINAL THESIS REPORT

Chapter 5: Performance Evaluation and Analysis

5.1 Performance Evaluation

The proposed system is evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Results show that federated learning with class-weighted loss achieves performance close to centralized training while preserving data privacy.

5.2 Analysis of Design Solutions

Comparative analysis reveals that FedAvg provides higher overall accuracy, whereas robust trimmed-mean aggregation improves attack recall under adversarial conditions. This demonstrates a trade-off between detection sensitivity and false alarm rates.

5.3 Final Design Adjustment

Based on experimental results, class-weighted loss and gradient clipping are incorporated to stabilize training. The trim ratio is optimized to balance robustness and accuracy.

5.4 Statistical Analysis

Statistical comparisons across multiple training rounds confirm the consistency of federated learning performance. Variance analysis shows stable convergence behavior under non-IID data distributions.

5.5 Comparisons and Relationships

Centralized and federated approaches are compared to analyze performance gaps. Relationships between data heterogeneity, robustness, and detection accuracy are established.

5.6 Discussion

The results demonstrate that federated learning is a viable alternative to centralized IDS for UAV networks. Robust aggregation improves resilience but introduces performance trade-offs that must be carefully managed.

5.7 Economic Analysis

Federated learning reduces communication costs by eliminating raw data transfer. The proposed lightweight model minimizes computational overhead, making the system economically viable for large-scale UAV deployments.

5.8 Ethical Issues

The proposed framework preserves user privacy by keeping raw network data local to UAVs. Ethical concerns related to surveillance and data misuse are mitigated through privacy-preserving learning and limited data exposure.

Chapter 3.3.4 Ethical Issues *(as required separately)*

The ethical considerations of this project include responsible data usage, avoidance of personal data exposure, and ensuring that intrusion detection mechanisms are not misused for unauthorized surveillance. Federated learning inherently supports ethical AI principles by minimizing data sharing.