Scalable and Sustainable Methodologies for ML-Driven Security Operations: Energy-Aware Scheduling for Secure Model Training at the Edge

Project Overview

This project explores innovative strategies for reducing the carbon footprint and energy consumption associated with secure, distributed training of machine learning (ML) models at the network edge. Leveraging fully distributed learning paradigms—such as federated learning—this project targets both scalability and enhanced security, especially in edge computing environments where resources are constrained and privacy is paramount.

Project Goals

- **Intelligent Client Selection:** Develop methods for selecting which edge devices (clients) participate in each training round, using both security criteria and real-time energy metrics.
- **Energy-Aware Scheduling:** Design protocols for when and what to broadcast to other nodes, with the objective of minimizing energy usage during model training, while maintaining strict accuracy and security requirements.
- **Systematic Trade-Offs:** Analyze how energy availability, device efficiency, and security risks interact to affect overall system performance.
- **Scalability Analysis:** Evaluate the scalability of proposed methods in large, heterogeneous edge networks.

Ultimately, the project aims to deliver practical methodologies that make ML training at the edge both more secure and environmentally sustainable.

Motivation and Significance

Distributed learning—especially federated learning—has emerged as a promising approach for training ML models across large-scale, decentralized systems, including cloud, edge, and IoT networks. The advantages are multifold:

- **Privacy and Security:** By keeping data localized and sharing only model updates, distributed learning ensures compliance with privacy regulations and reduces the risk of data leakage.
- **Resource Optimization:** Edge devices can collectively train models without centralizing all data or computations, saving bandwidth and reducing latency.
- **Energy Efficiency:** As the environmental impact of AI grows, optimizing energy consumption becomes a societal imperative. Energy-aware scheduling can balance training performance with sustainability goals.

Research Challenge: Current distributed ML frameworks often ignore the energy profile and availability of participating devices, as well as real-time security risks. This project will investigate algorithms and protocols that dynamically adjust client participation and communication, ensuring that training remains both accurate and sustainable, even as network conditions and device states evolve.

Learning Outcomes

Students will:

- Gain hands-on experience with federated learning and edge computing.
- Explore the interplay between security, energy management, and ML system design.
- Develop analytical and programming skills in designing, simulating, and evaluating energy-aware protocols.
- Evaluate real-world challenges at the intersection of AI, sustainability, and cyber-physical systems.

Background Required

- Understanding of computer networks.
- Programming (e.g., Python, or another suitable language).
- (Optional but helpful) Prior coursework in machine learning.

Expected Outcomes

- A simulation of an energy-aware scheduling algorithm for federated learning at the edge.
- A study comparing baseline vs. energy-aware protocols, with metrics including energy savings, model accuracy, and security robustness.
- A written report and a presentation summarizing findings, trade-offs, and future directions.