Privacy-Preserving Secure Aggregation in Federated Learning for Underground Mining Surveillance

Motivation and Introduction

To improve safety and efficiency in subterranean operations, the mining sector is progressively embracing AI-driven automation and monitoring technologies. But for remote and bandwidth-limited underground mines, conventional centralized learning methods are unworkable since they present privacy concerns and demand too much data transfer. Federated Learning (FL) lets distributed cameras train models locally without sharing raw image data, hence providing a privacy-preserving option. FL is nevertheless susceptible to untargeted attacks like sign-flip, additive noise, and model poisoning that can compromise model performance even with its advantages. Homomorphic Encryption (HE) could help to lower these risks by means of secure aggregation methods employing cryptographic tools as Message Authentication Codes (MACs), hence preserving privacy.

This project aims to:

- This project aims to build a real-world federated learning dataset from an experimental mine by taking pictures of miners, helmets, tunnels, passages, and other subterranean features.
- Simulate non-IID and IID FL scenarios to explore real data variance.
- Design secure aggregation methods using cryptography techniques.
- Design algorithms to detect attacks.

Flow of Project

The project is organised into the following main phases:

Phase One: Dataset Development

- Visit the experimental mine to get photos of miners, helmets, tunnels, rooms, and other items.
- For model training, label and classify photos into relevant categories.

Phase 2: Federated Learning Installation

- Reflecting realistic deployment situations by simulating IID and non-IID data distributions.
- Build a federated learning system in which every camera (FL client) train models locally.
- Assess baseline models—e.g., CNNs, Vision Transformers—for picture classification.

Phase 3: Secure Aggregation Employing Cryptographic Techniques

- Use Message Authentication Codes (MACs) to check FL update integrity.
- Protect model modifications using homomorphic encryption (HE) and differential privacy.

Phase 4: Defense Mechanisms and Attack Detection

- Simulate data poisoning attacks in FL, additive noise, and sign-flip.
- Create detection algorithm to detect malicious clients.

Phase 5: Assessment and Enhancement

- Examine FL model performance both with and without secure aggregation.
- Improving privacy-preserving technologies will enable more processing efficiency.
- Evaluate attack detection success rates in various hazardous settings.

Phase Six: Publication and Documentation

- Create open access, technical reports on datasets.
- Forward results to conferences on artificial intelligence security and the mining sector.

Participant Desired Skills

- Machine Learning & Deep Learning, particularly CNNs and Vision Transformers.
- Decentralized artificial intelligence methods and federated learning.
- Cryptographic Security: homomorphic encryption, differential privacy, MACs.
- Computer vision and Image processing.
- Model implementation using Python and PyTorch.

Prior experience with these technologies is beneficial but not required -students will receive mentorship and training throughout the project.

Advantages of the Project

- The creation of a unique underground mining dataset for FL research.
- Improvements in security: Development of privacy-preserving and attack-resistant FL algorithms.
- Journal submissions and conference papers are viable options for research and publication.
- Enhances the safety monitoring and surveillance of mining operations through the use of AI.