PRIVACY PRESERVING FEDERATED LEARNING DETECTING INTRUSION

A PREPRINT

December 3, 2023

Problem Statement

Cybersecurity remains a critical concern as evolving attack patterns continue to challenge conventional defense strategies. The unpredictable nature of recent cyberattacks necessitates innovative approaches, and machine learning is emerging as a potent tool for intrusion detection. This project delves into federated learning, a decentralized paradigm that ensures client privacy while aggregating insights from diverse sources.

Dataset Overview

The dataset (1) forms the cornerstone of this study. Capturing activities over a 5-day period, it encapsulates normal human behaviors and simulates various attacks—Brute Force, DoS, DDoS, Heartbleed, Web Attack, Infiltration, Botnet, and DDoS. The dataset comprises 9 CSV files, each representing a client in the federated learning experiments. Initially It does not contain any information about IP addresses, I added private IP addresses synthetically.

Methodology

Technologies and Algorithms

To implement this project, I used Python programming languages and the following libraries summarised in Table 1.

Library	Version
pandas	2.0.3
matplotlib	3.7.3
flwr	0.15.0
torch	2.1.0
torchvision	0.16.0
numpy	1.24.4
scikit-learn	1.3.2
phe	1.5.0

Table 1: List of Libraries and Versions

To train individual client models, four machine learning algorithms were initially selected, each offering unique strengths:

1. **Logistic Regression:** A linear model for binary classification, providing probabilities and assigning the class with the highest probability.
2. **Support Vector Machine (SVM):** A powerful classification algorithm effective for linear and non-linear data, aiming to find optimal hyperplanes.
3. **Multi-layer Perceptron (MLP):** MLP is a type of artificial neural network that consists of multiple layers of interconnected nodes, or neurons. It is a feedforward neural network architecture, meaning that information

flows through the network in one direction—from the input layer to the output layer. The typical architecture includes an input layer, one or more hidden layers, and an output layer.

Algorithm's Working Process

The server model incorporates a custom aggregation strategy called FedSum. The FedSum strategy extends the default FedAvg strategy and overrides the aggregate fit method to implement a weighted average based on the number of examples each client contributes.

The client loads a preprocessed dataset, standardizes it, and initializes a machine learning model. During training, different strategies applied for different models based on the necessity.

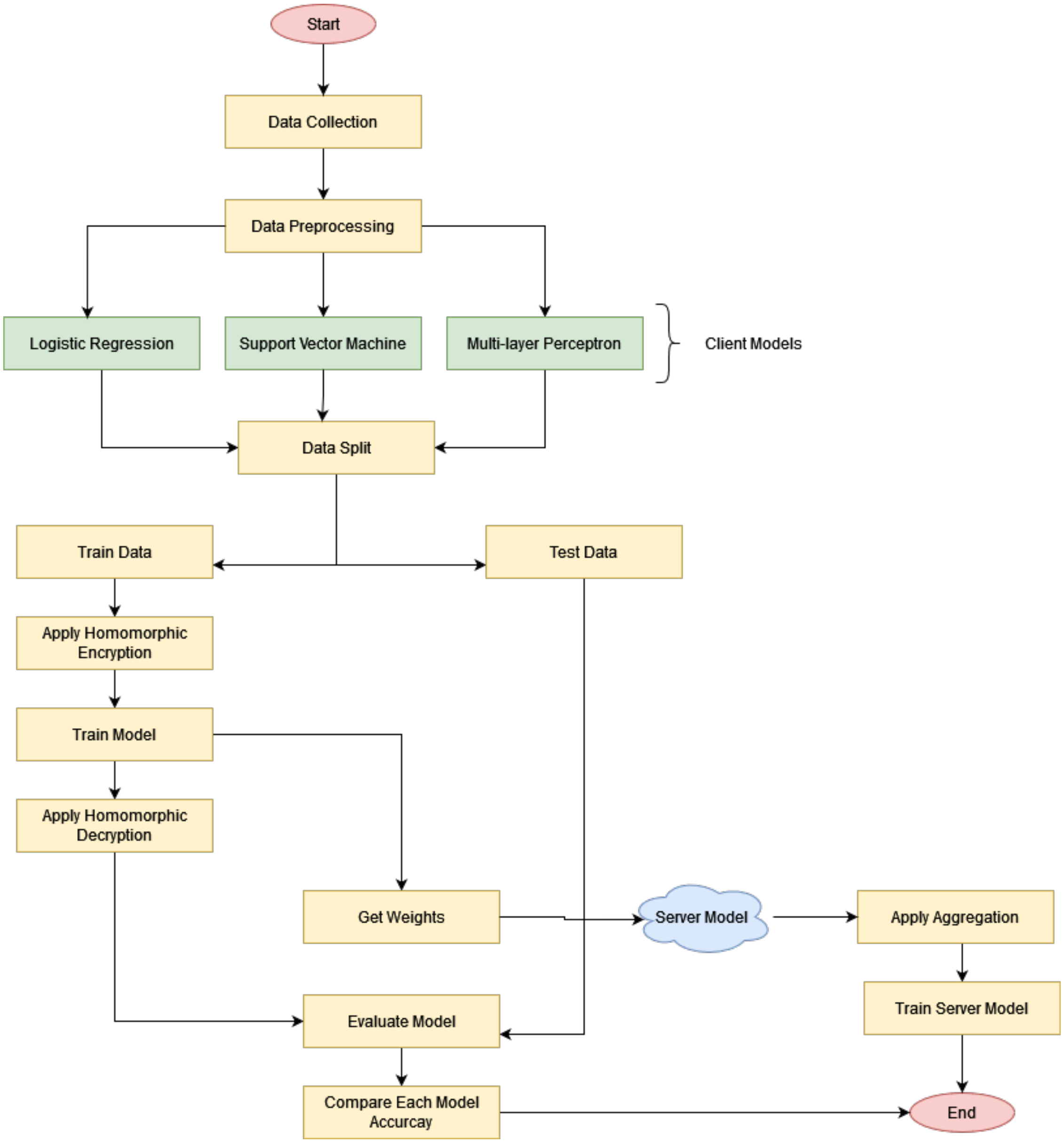

Figure 1: Process Overview

In addition, the code also includes functionality for testing the performance of each client's model. The fit method is utilized for training the model on the locally held dataset, updating the model parameters according to the received encrypted weights. The number of local epochs for training is specified by the local epochs parameter. After training, the evaluate method is employed to assess the model's accuracy on the locally held dataset against the provided parameters. Figure 1 summarised the whole process.

Encryption and Decryption Process

To ensure the privacy of the model updates during federated learning, each client employs homomorphic encryption, specifically Paillier encryption. In the encryption process, the encrypt method is responsible for encrypting the weights of the models before transmitting them to the server. If the nohomo flag is not set, each layer of the model's weights undergoes encryption. On the server side,it aggregates and update the global model using contributions from different clients while preserving the privacy and confidentiality of individual model updates. The choice of paillier encryption ensures secure communication and collaboration between clients and the server in the federated learning framework.

Conclusion

In conclusion, this project unfolds a comprehensive methodology for applying federated learning to enhance intrusion detection. Leveraging the strengths of individual client models and a collaboratively trained global model, the approach aims to strike a balance between model performance and privacy preservation. By scrutinizing the diverse facets of model evaluation, privacy preservation, and addressing inherent challenges, this study contributes to the evolving landscape of federated learning in cybersecurity.

Acknowledgments

We acknowledge the invaluable insights gained from the dataset (1) and express gratitude to the research community for fostering an environment of collaboration and knowledge sharing.

References

[1] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *4th International Conference on Information Systems Security and Privacy (ICISSP)*, (Portugal), January 2018.