Federated Learning for Cybersecurity

Mehul Mathur | Yash Shrivastava | Velidindi Kaarthika Dhruv | Dr. Dilipkumar S | School of Computer Science and Engineering

Introduction

The project **Federated Learning in Cybersecurity** investigates integrating decentralized machine learning with privacy-fencing mechanisms in order to provide solutions to the cybersecurity issues in the modern age.
Introduced by Google (McMahan et al., 017), Federated Learning enables decentralized model training on edge devices without transferring raw data to a central server. The Federated Averaging (FedAvg) algorithm aggregates model updates from multiple clients, combining local stochastic gradient descent with weighted averaging. Dwork et al. (2006) introduced Differential Privacy as a formal privacy-preserving technique. In the FL context, it involves adding calibrated noise to model updates, ensuring individual data points cannot be inferred from shared gradients. Issues like data heterogeneity, communication bottlenecks, and security threats (e.g., model poisoning) are challenges that are threat to FL based systems

Motivation

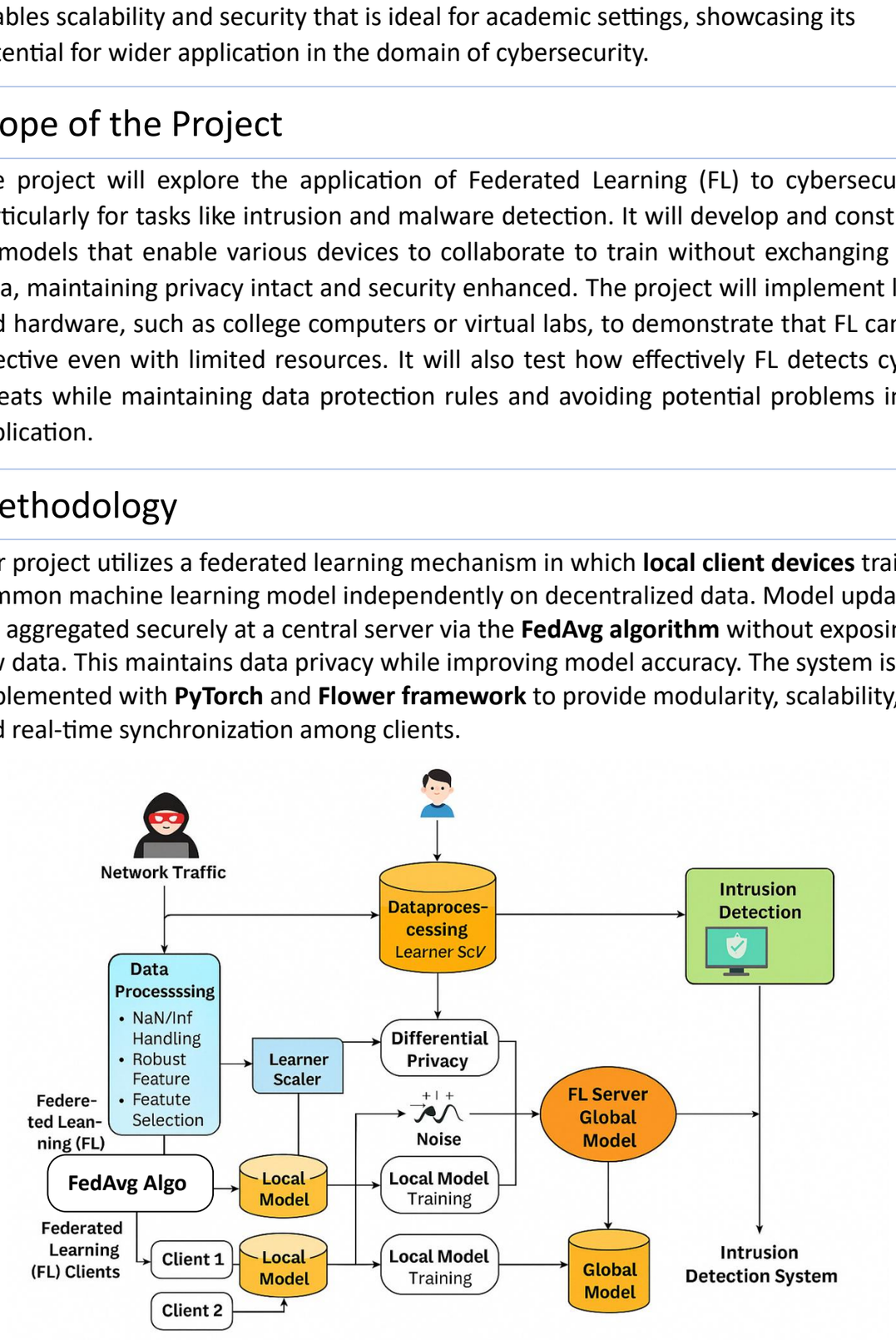
With a focus on simplicity of implementation, the project will implement low-specification, affordable hardware, like university laptops or virtualized systems, to establish the feasibility of Federated Learning in practical scenarios. Not only does the decentralized aspect of Federated Learning ease data privacy concerns, but it also enables scalability and security that is ideal for academic settings, showcasing its potential for wider application in the domain of cybersecurity.

Scope of the Project

The project will explore the application of Federated Learning (FL) to cybersecurity, particularly for tasks like intrusion and malware detection. It will develop and construct FL models that enable various devices to collaborate to train without exchanging raw data, maintaining privacy intact and security enhanced. The project will implement low-end hardware, such as college computers or virtual labs, to demonstrate that FL can be effective even with limited resources. It will also test how effectively FL detects cyber threats while maintaining data protection rules and avoiding potential problems in its application.

Methodology

Our project utilizes a federated learning mechanism in which **local client devices** train a common machine learning model independently on decentralized data. Model updates are aggregated securely at a central server via the **FedAvg algorithm** without exposing raw data. This maintains data privacy while improving model accuracy. The system is implemented with **PyTorch** and **Flower framework** to provide modularity, scalability, and real-time synchronization among clients.



Architecture Diagram for FL Model – Depicting connection with clints and server

The central server performs Federated Averaging (FedAvg) across the noisy local models to generate a global model.
FedAvg Operation:
Given weights $\theta 1, \theta 2, \ldots, \theta_n$ from n clients:

$$\theta' = \theta + N(0, \sigma^2)$$

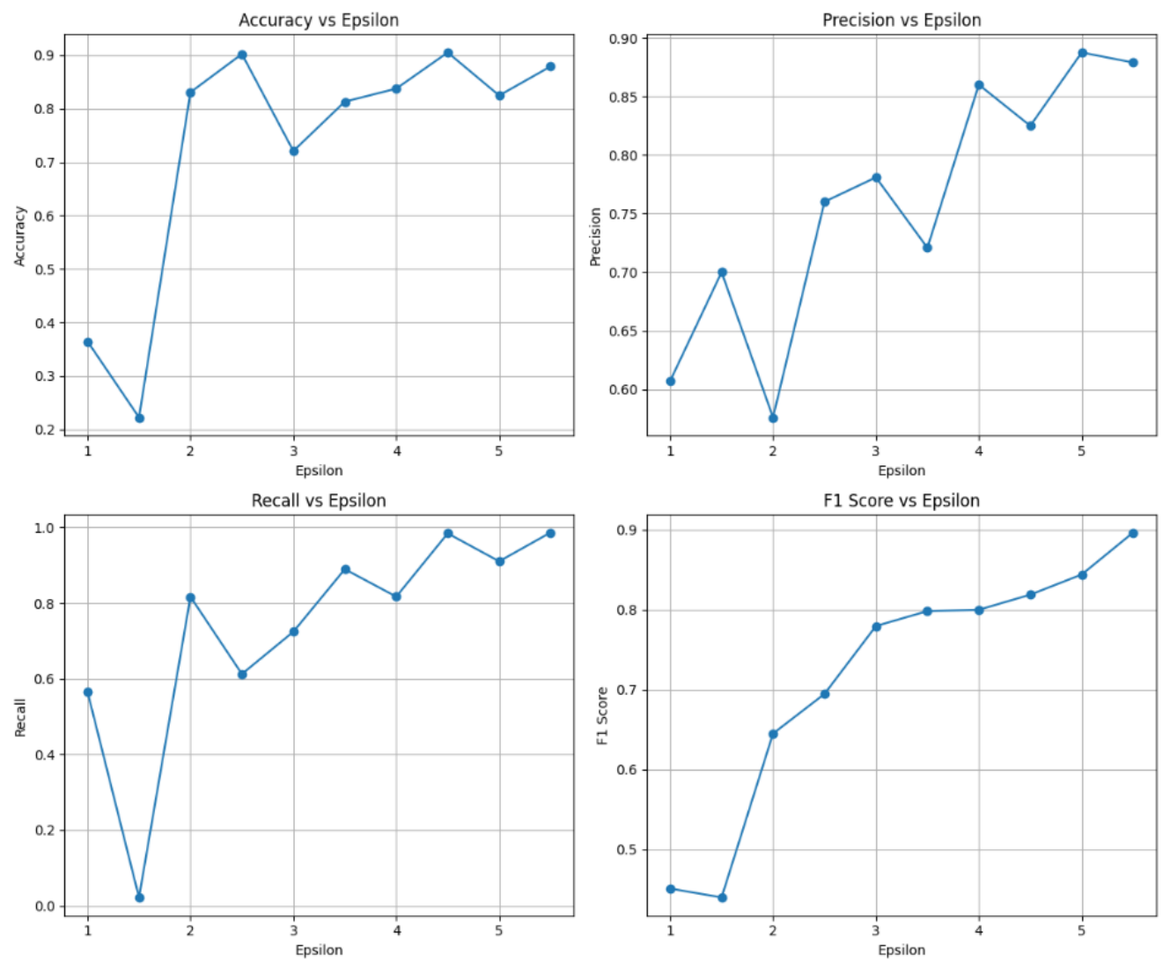
$$\theta_{global} = \frac{1}{n}\sum_{i=1}^{n}(\theta_i)$$

The first equation adds Gaussian noise $N(0,\sigma^2)$ to model parameters θ for **differential privacy**, ensuring data security. The second equation performs federated averaging, where the global model θ_{global} is computed by averaging client models θ_i enabling secure, collaborative training without sharing raw data.

Results

Global Model (Post Aggregation):

- Accuracy: 0.916
- F1-Score: 0.90
- Recall: 0.886



The set of graphs shows how performance metrics vary with different values of privacy parameter epsilon (ε). As ε increases (implying weaker privacy), all four metrics—accuracy, precision, recall, and F1 score—generally improve. This highlights the inverse relationship between privacy and model performance. Notably, very low ε values degrade results significantly due to the heavy noise added for stronger privacy. A balanced value of ε must be chosen to optimize privacy while maintaining satisfactory prediction quality.

Metric Pair	Trade-off	Explanation
Privacy (ε) vs Accuracy	Inverse	Stronger privacy (lower ε) introduces more noise → reduced accuracy.
Accuracy vs Recall	Balanced	Too high a threshold increases precision but decreases recall.
Latency vs Throughput	Not measured but relevant	Local model training increases latency, but allows parallelism across clients.
Feature Dimensionality vs Performance	Decreases	Fewer features improve training speed but may reduce model expressiveness.
Data Imbalance vs Precision	Not measured but relevant	Minority class (attacks) underrepresented → may hurt precision.

Conclusion

The findings show a high correlation between larger epsilon values and better model performance—accuracy, precision, recall, and F1 score—up to a threshold. This shows that increasing the privacy budget (ε) reduces noise and enhances detection capacity with a good privacy level.
The idea that Federated Learning with Differential Privacy can help with effective and secure intrusion detection is valid. The evidence shows that it is possible to perform well while adhering to privacy rules, especially when the privacy noise is set correctly.
Future work could investigate making the deployment functional for larger and more complex networks, utilizing actual cybersecurity information, and enhancing defences against privacy attacks. The project incorporates encryption techniques, enhances how fast it can learn and maintain privacy rigorous, and enables communication to become more efficient. This is relevant in sensitive applications such as health and finance, balancing functionality and innovation.

References

1. Sen, J. (2024). *Privacy in Federated Learning*.
2. Hasan, J. (2023). *Security and Privacy Issues of Federated Learning*.
3. Zhang, Y., & Wang, X. (2025). *Enhancing AI Cyber Security with Privacy-Preserving Federated Learning*. SSRN.
4. Zhang, L., et al. (2021). *Trustworthy Federated Learning: Privacy, Security, and Beyond*.
5. Harder, T., & Lutz, C. (2021). *Privacy-first health research with federated learning*. Nature Digital Medicine, 4(1), 1-10.
6. Liu, Y., & Chen, H. (2023). *Federated Learning for Cybersecurity: A Survey*. IEEE Transactions on Information Forensics and Security.