Attack and Defense Mechanisms in Federated Learning Considering Resource-Constrained Edge Devices

Syed Hannan Shah, Rutgers University, hannansshah2004@gmail.com

Faculty mentor: Hadi Amini, PhD Student Mentor: Md Jueal Mia, Florida International University

Introduction

This project focuses on enhancing Federated Learning Security for IoT devices while managing resource constraints such as memory, bandwidth, and energy.

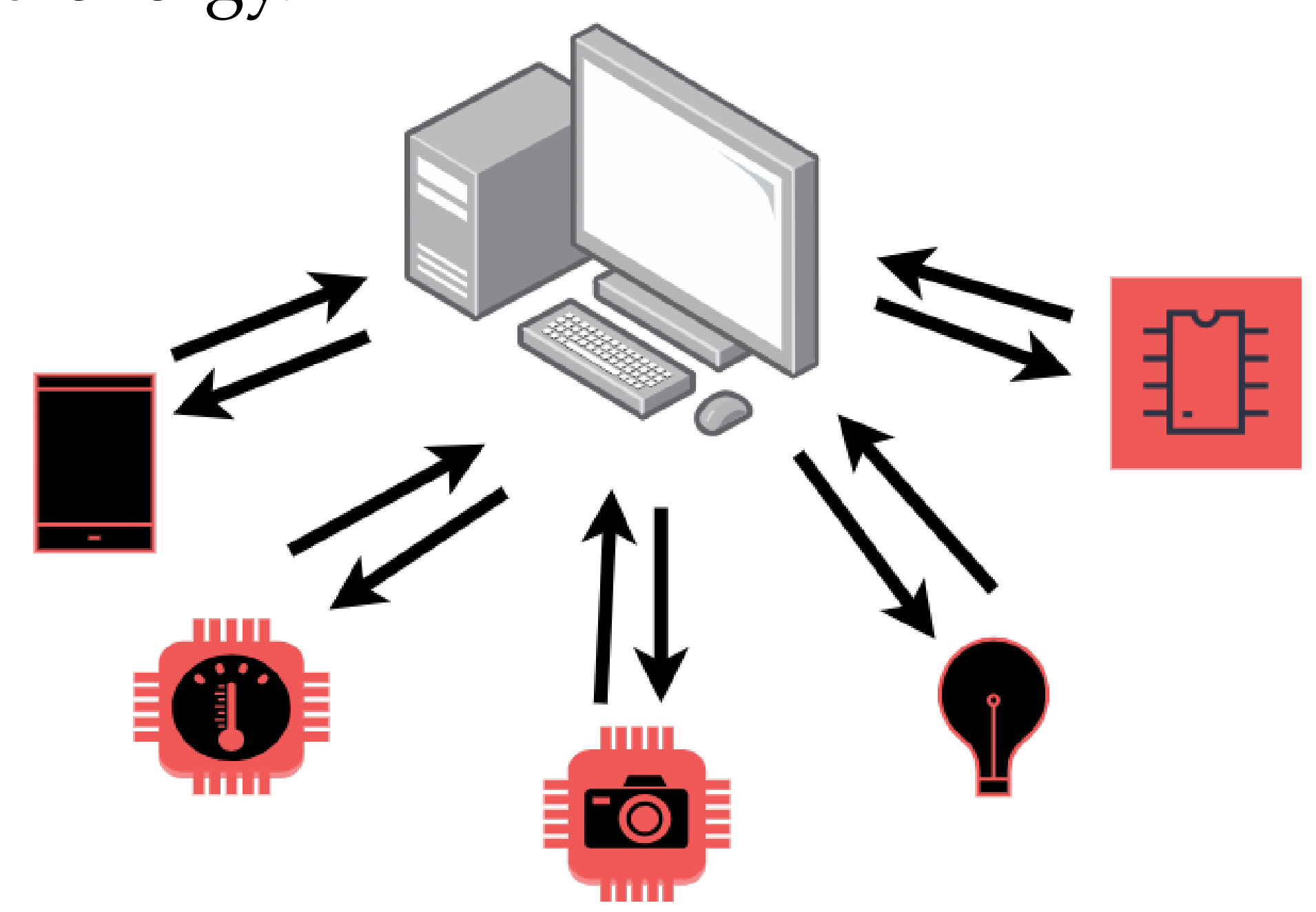

Figure 1 : Traditional Federated Learning Process

Aggregation Techniques

1. Federated Averaging
2. Krum
3. Median
4. Trim Mean
5. FLTrust
6. ROMOA

Malicious Techniques

1. No Byzantine Attack
2. Trim Attack
3. Krum Attack
4. FLTrust Attack
5. Min-Max Attack
6. Min-Sum Attack
7. Scaling Attack

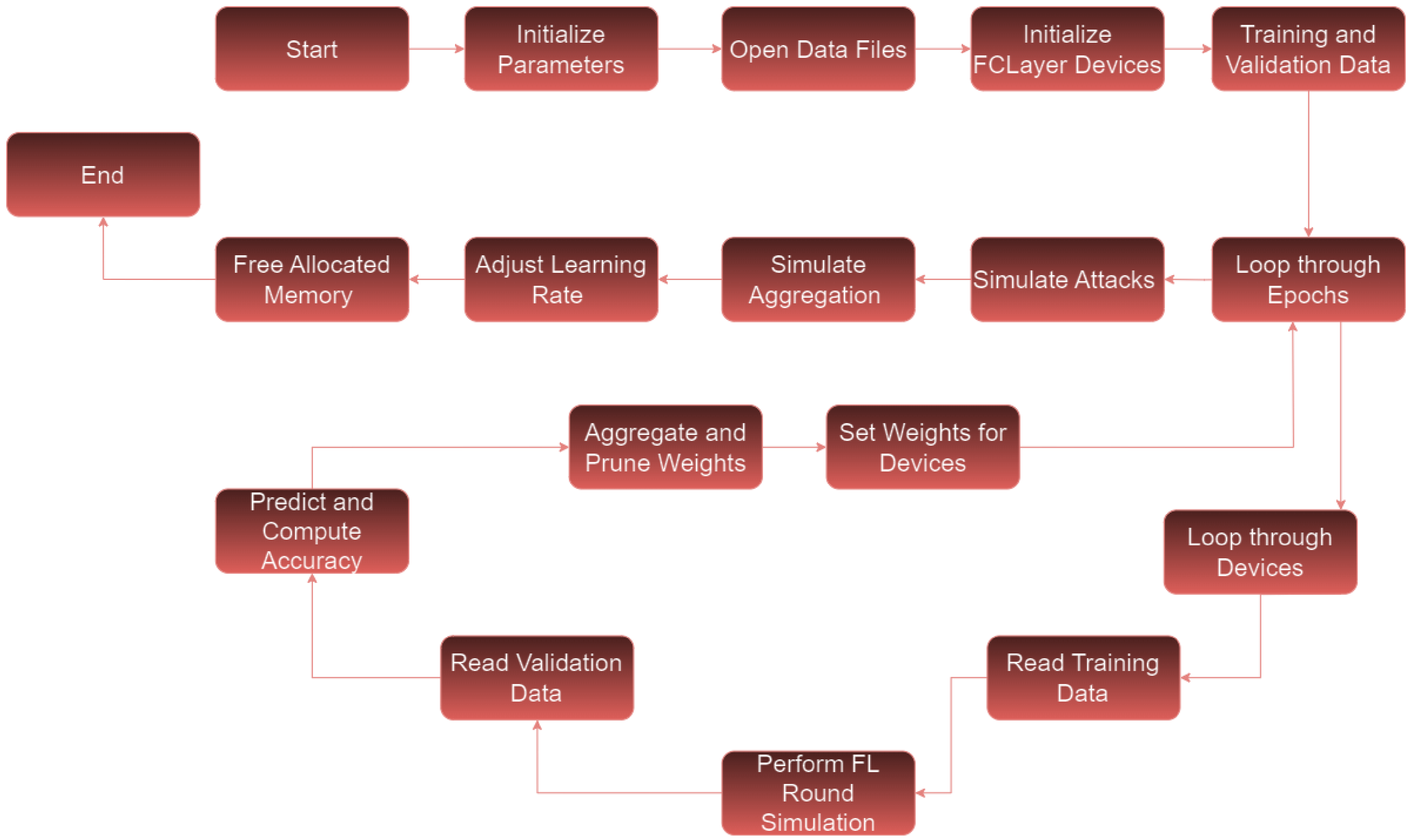
Start → Initialize Parameters → Open Data Files → Initialize FCLayer Devices → Training and Validation Data → Loop through Epochs → Loop through Devices → Read Training Data → Perform FL Round Simulation → Read Validation Data → Predict and Compute Accuracy → Aggregate and Prune Weights → Set Weights for Devices → Loop through Epochs → Simulate Attacks → Simulate Aggregation → Adjust Learning Rate → Free Allocated Memory → End

Figure 2 : Proposed Framework

TinyDistillation Simulation

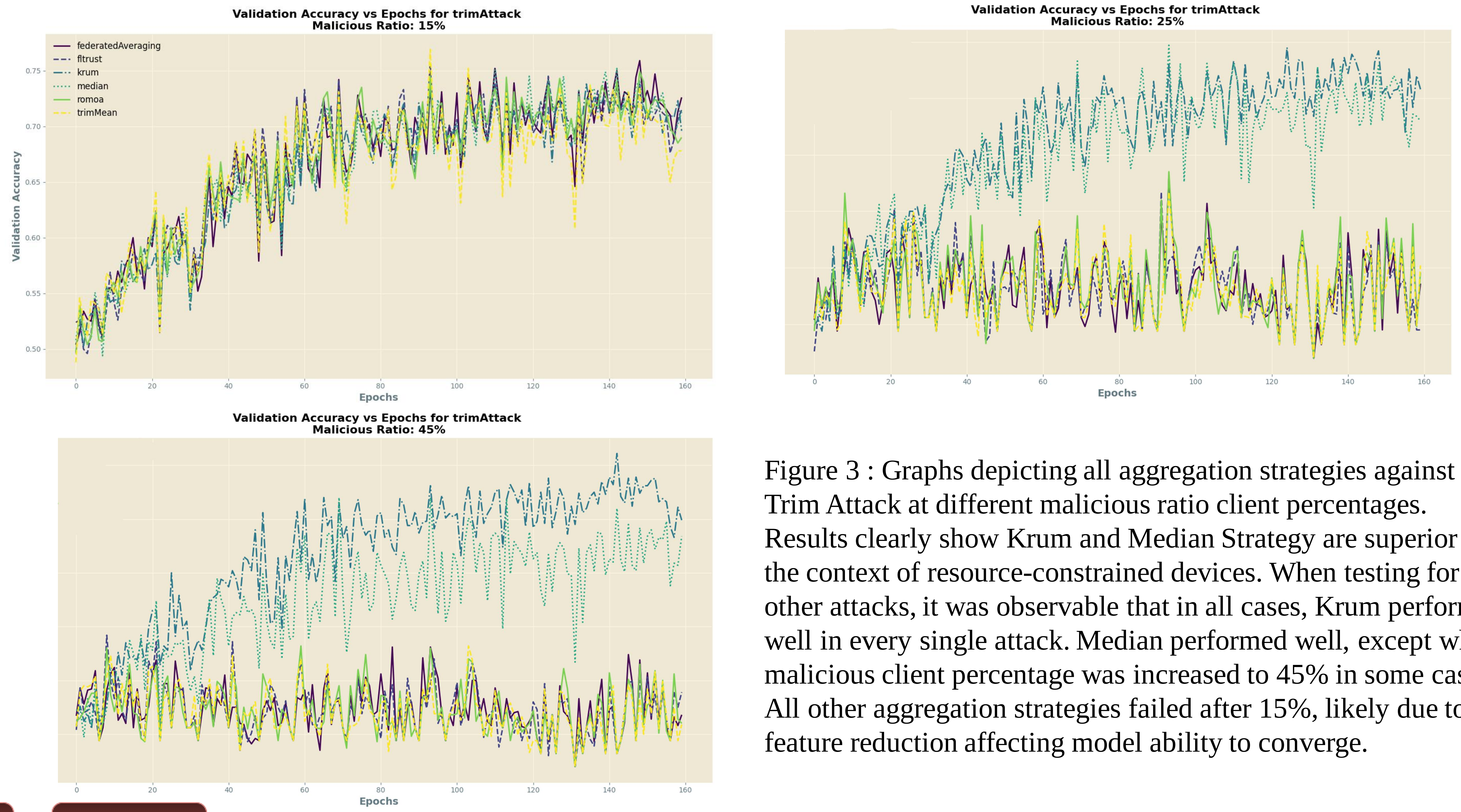

Figure 3 : Graphs depicting all aggregation strategies against Trim Attack at different malicious ratio client percentages. Results clearly show Krum and Median Strategy are superior in the context of resource-constrained devices. When testing for all other attacks, it was observable that in all cases, Krum performed well in every single attack. Median performed well, except when malicious client percentage was increased to 45% in some cases. All other aggregation strategies failed after 15%, likely due to feature reduction affecting model ability to converge.

TinyDistillation Setup

Data Acquisition: Utilized CIFAR-10 dataset for binary classification (horses vs. dogs), resized images to 96x96 pixels, normalized, and extracted features using a pre-trained VGG16 model, going from 3072 features to 256 features.

Model Initialization: Trained a binary classifier (cats vs. dogs) to create initial weights, using a VGG16 base with additional layers, resulting in 512 weights.

Acknowledgement

This work was supported by the U.S. National Science Foundation REU Program.

Conclusion

This research shows the importance of robust aggregation methods like Krum and Median in federated learning for resource-constrained IoT devices. By optimizing resource usage with quantization and pruning, we can maintain model integrity and performance under adversarial conditions. In the future, we should develop more sophisticated defense strategies to enhance FL security further.

References

[1] K. Kopparapu, E. Lin, J. G. Breslin and B. Sudharsan, "TinyFedTL: Federated Transfer Learning on Ubiquitous Tiny IoT Devices," 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Pisa, Italy, 2022

[2] T. Gehlhar, F. Marx, T. Schneider, A. Suresh, T. Wehrle, and H. Yalame, "SAFEFL: MPC-friendly Framework for Private and Robust Federated Learning," Cryptology ePrint Archive, Paper 2023/555, 2023

[3] A. Imteaj, U. Thakker, S. Wang, J. Li and M. H. Amini, "A Survey on Federated Learning for Resource-Constrained IoT Devices," in IEEE Internet of Things Journal, vol. 9, no. 1, pp. 1-24, 1 Jan.1, 2022

Note : List is not conclusive, many more references were used.