Postgraduate Seminar Series

Venue: Graduate Seminar Room
Date & Time: August 02, 2025 at 2:30 PM

Speaker Information

Israt Jahan Mouri (Std No. 0421054001) is a full time Ph.D. student in the department of CSE, BUET. She completed her undergraduate studies from Military Institute of Science and Technology (MIST) in 2013 and M.Sc. from BUET in 2020. Her research interest lies in the field(s) of Federated Learning, Cloud Computing, Security of Distributed Learning, and Poisoning Attacks on Machine Learning. She is currently doing her doctoral thesis under the supervision of Prof. Dr. Muhammad Abdullah Adnan. She will be speaking about her ongoing PhD research in this talk.

XGDFed: Exposing Vulnerability in Byzantine Robust Federated Binary Classifiers with Novel Aggregation-Agnostic Model Poisoning Attack

Federated Learning (FL) is a machine learning method in which multiple edge devices collaborate to train an ML model without exchanging data. Instead, each device sends its local update to the central server, which then aggregates these updates using a Byzantine-robust aggregation rule to prevent Byzantine failures of edge devices. Recent studies have shown that some carefully crafted model poisoning attacks can evade Byzantine robust aggregations. However, these state-of-the-art Byzantine robust attacks rely on knowing the central server's aggregation algorithm or other benign model updates from all the edge devices, which is impractical. To address this, we have proposed a novel model poisoning attack called *XGDFed* which effectively targets the decision function of binary classifiers regardless of the server's aggregation method. Our experiments indicate that *XGDFed* outperforms state-of-the-art attacks using the CIFAR-10, FashionMNIST, MNIST, IJCNN1, and Acoustic datasets. With only 10% of attackers in the Fashion-MNIST dataset, *XGDFed* reduced the accuracy of the federated binary classifier from 83% to 51%. This research is the first empirical evaluation of the robustness of Byzantine-robust aggregations against state-of-the-art model poisoning attacks on binary classifiers in a federated environment. Although binary classifiers are often overlooked in the research literature, they can be relevant in different applications of federated learning.