STEVENS
INSTITUTE OF TECHNOLOGY
1870

Ph.D. DISSERTATION DEFENSE

Candidate:	Zhuosheng Zhang
Degree:	Doctor of Philosophy
School/Department:	Department of Electrical and Computer Engineering
Date:	Tuesday, December 12, 2023
Time/Location:	11:00 am / https://stevens.zoom.us/j/96005852449
Title:	Improving Security of Privacy-Preserving Federated Learning
Chairperson:	Dr, Shucheng Yu, Department of Electrical and Computer Engineering, Charles V. Schaefer, Jr. School of Engineering and Science
Committee Members:	Min Song, Department of Electrical and Computer Engineering, Charles V. Schaefer, Jr. School of Engineering and Science Koduvayur Subbalakshmi, Department of Electrical and Computer Engineering, Charles V. Schaefer, Jr. School of Engineering and Science Jie Shen, Department of Electrical and Computer Engineering, Charles V. Schaefer, Jr. School of Engineering and Science

ABSTRACT

The emergence of next-generation communications and computing paradigms such as edge computing has made distributed machine learning technologies like federated learning increasingly popular in emerging applications. Unlike centralized learning, which trains models at the central server, federated learning enhances data privacy by allowing participants (i.e., users) to train models locally and only send the trained model parameters or gradients to the remote server without revealing their private datasets. However, recent research has shown that attackers can infer training examples held by users by disclosing local model parameters or gradient updates. To protect local models against disclosure, only the user should know their own local model, while the global model will be disclosed to all users. This variant of federated learning, which protects users' private data and models, is known as Privacy-Preserving Federated Learning (PPFL). However, privacy comes at a cost, and two major challenges in PPFL are efficiency and security. In this research, the author approaching the PPFL in following aspects: (1) Enable backdoor attack detectability in PPFL; (2) Improving computational and communicational efficiency in PPFL and its variant: Federated Distillation; (3) Exploring the vulnerability of PPFL trained black-box model.