Research Statement

My research interest lies on privacy-preserving machine learning. Currently there're three main problems I want to solve, which are listed below.

Learning private representation.

Data releasing contains risk of privacy leakage. It's useful to know if there's a obfuscation mechanism that could obfuscate the private features while maintains data's utility. Inspired by recent advancement of generative adversarial networks, learning private representation could be formed as a two-player privacy game: a defender that obfuscates data to minimize the risk of private attributes being inferred without losing too much utility, while an active adversary keeps trying to infer sensitive information from the obfuscated data released by defender. The defender and adversary are both assumed neural networks here.

There're already some works on this kind of problems, like Generated Adversarial Privacy(GAP) which uses GAN model to perturbate private information, and applications like DeepPrivacy, which directly blurs human faces for identity. However, the real problenm is to know a fast way to converge the defender's neural network to the optimal setting, or whether there exists such optimal point. Since this is a non-linear optimization problem, we need to define a proper local optimal, which could be determined in a data-driven way.

Privacy-Preserving Federated Learning

Federated learning is proposed by Google to train machine learning models without users uploading the raw data. The server and users train same model and server side receives only gradient updates from user side. However, researchers have proved that this is not safe enough. Gradient itself could release privacy, sometimes even the raw training data. There're three tools that have the potential to prevent this kind of privacy leakage: secure multi-party computation, differential privacy and homomorphic encryption.