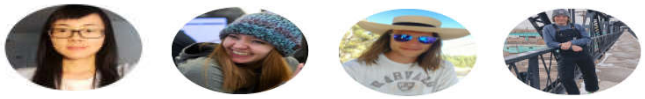
# Privacy Inference under Side Channel Power Attacks

Summer 2021

## Team Members



**Hui Hu:** I am a graduate researcher in CEDAR lab with interests in privacy-preserving machine learning, side-channel attacks. I am leading this project.

**Shaya Wolf:** I am a graduate researcher in CEDAR lab with interests in distributed systems, swarm communications, and encryption mechanisms.

**Rafer Cooley:** I am a graduate researcher in CEDAR lab with interests in bio-inspired algorithms, complex adaptive systems, and network protocols.

**Jayden Parker Vap:** I am an undergraduate researcher in CEDAR lab with interests in hardware security and machine learning.

## Problem Statement

In this work, we study privacy inference under side channel power attacks, as Figure 1 shows.

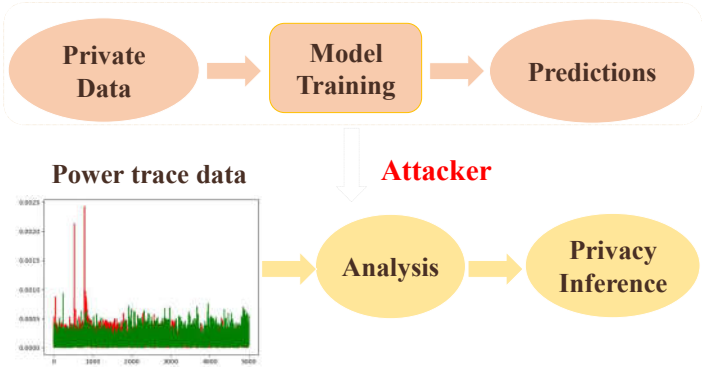



*Figure 1. Privacy inference under side channel power attacks.*

## Experiment Results

Figure 3 shows the power traces for two random samples in two different groups (1000 dimensions vs. 5000 dimensions ). Table 1 shows inference accuracy.

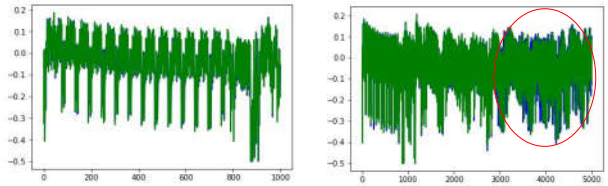


*Figure 3. Power traces of two random samples in different groups.*

*Table 1. Privacy inference accuracy*

| Classifier | k value | 1000 dimensions | 5000 dimensions |
|---|---|---|---|
| k-NN | k=3 | 0.5976 | 0.7734 |
| k-NN | k=6 | **0.6445** | 0.8008 |
| k-NN | k=10 | 0.6328 | **0.8379** |

## Introduction

Privacy has emerged as a big concern in machine learning as witnessed by increasing interests among researchers in privacy-preserving AI in recent years. While many privacy-preserving machine learning techniques have been proposed[1,2], most of them still focus on algorithm or framework designs and ignore the characteristics of hardware in the modeling process may leak privacy. To fill this gap, in this study, we investigate privacy leakage under side channel power attacks.

Side channel attack techniques have been shown to be powerful for key extraction from a cryptographic algorithm[3,4]. Motivated by this finding, in this work, we first investigate privacy inference in machine learning under side channel power attacks, including model privacy inference and data privacy inference. Further, we will propose an efficient privacy-preserving mechanism under side channel power attacks when needed.

## Methods

Our method includes the following four steps (as Figure 2 shows):

**Step 1:** Split data into different groups according to the values of sensitive feature $X \rightarrow \{G_1, \ldots, G_i\}$.

**Step 2:** Input each training and inference sample into a three-layer neural network to learn weights.

**Step 3:** Collect the power trace of each training and inference sample.

**Step 4:** Predict the labels of power traces of all inference samples using k-NN classifier.

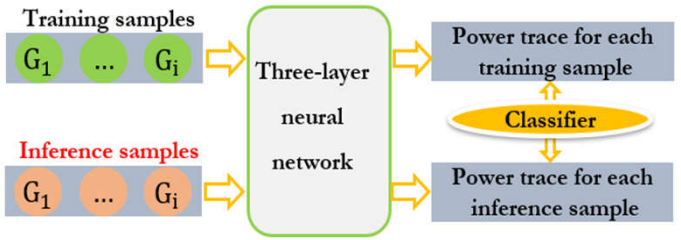



*Figure 2. The method for privacy inference*

## Results Analysis

We have the following three observations:

(1) The power traces with 5000 dimensions are more differentiable than 1000 dimensions, as the red circled part shows in Figure 3.

(2) The inference accuracy is higher on the power traces with 5000 dimensions than on the power traces with 1000 dimensions (0.8379 vs. 0.6445). This observation implies privacy inference accuracy will be higher when the power traces are more differentiable.

(3) As Table 1 shows, the best inference accuracy is close to 0.84 on the COMPAS data set. This means data privacy is leaked severely under side-channel power attacks.

Therefore, how to protect privacy more efficiently under side channel power attacks is our next work.


**Advisor**: Dr. Mike Borowczak

**Group Members**:

- Hui Hu (hhu1@uwyo.edu)
- Shaya Wolf (swolf4@uwyo.edu)
- Rafer Cooley (rcooley2@uwyo.edu)
- Jayden Parker Vap (jvap2@uwyo.edu)



**References:**

[1] Al-Rubaie, M., & Chang, J. M. (2019). Privacy-preserving machine learning: Threats and solutions. *IEEE Security & Privacy*, *17*(2), 49-58.

[2] Liu, B., Ding, M., Shaham, et al. (2021). When machine learning meets privacy: A survey and outlook. *ACM Computing Surveys (CSUR)*, *54*(2), 1-36.

[3] Banerjee, U., Ho, L., & Koppula, S. (2015). Power-based side-channel attack for aes key extraction on the atmega328 microcontroller. *Computer Systems Security*.

[4] Gamaarachchi, H., & Ganegoda, H. (2018). Power analysis based side channel attack. *arXiv preprint arXiv:1801.00932*.