Student: Wouter Jehee
Responsible professor: Kaitai Liang
Supervisor: Rui Wang

# Improving privacy in FL-GANs using Intel SGX

TU Delft

## 1. Background

**Generative Adversarial Networks (GANs) [1]**
**Federated Learning (FL) [2]**
**Intel Software Guard Extensions:** a set of CPU instructions that allow for the creation of **Trusted Execution environments (TEEs)** in which secure remote computations can be performed. [3]

## 2. Research questions

How can Intel SGX improve the privacy of clients in a FL-GAN?

- What is the attack surface in a FL-GAN (not using SGX) and what can malicious actors achieve?
- How can the data be protected using secure enclaves?
- What is the impact on the performance of the training algorithm?

## 3. Threat model

The following threats are considered

- Inference attacks by the server
- Inference attacks by clients
- Poisoning attacks by clients

## 4. Architecture

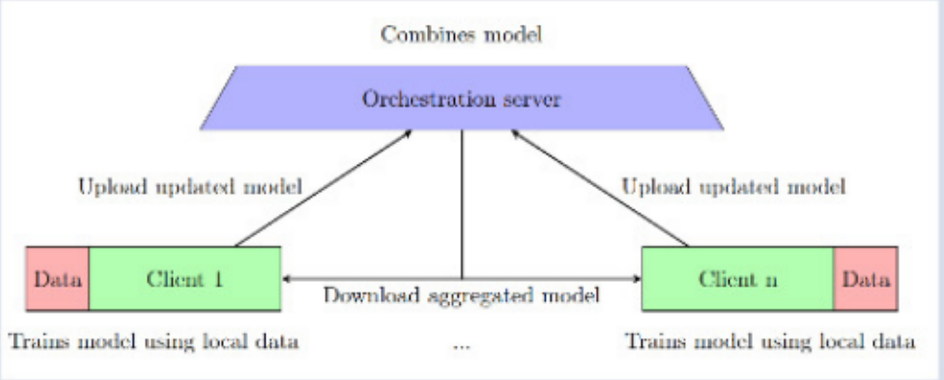


Figure 1: FL overview

Client and server verify quote provided by an external caching service to prove **Confidentiality & Integrity** through **Attestation**.

## 5. Results

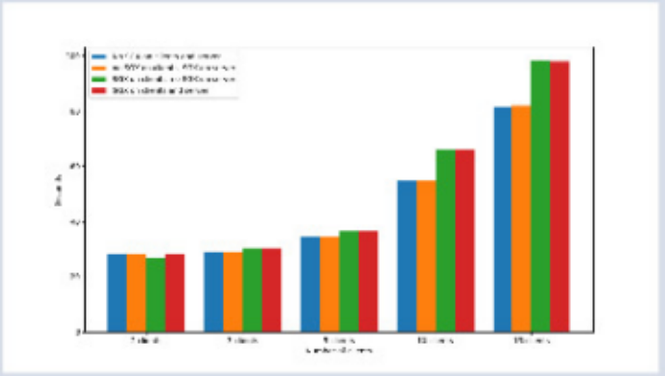

Figure 2: performance parabola

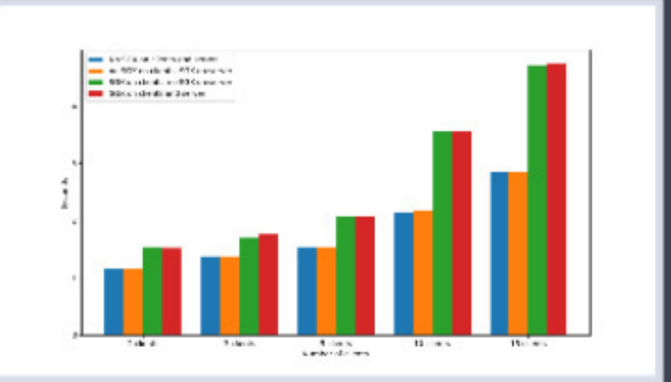

Figure 3: performance iris data set

## 6. Conclusion

**Key takeaways:**

- Using SGX on the orchestration server has low overhead and is worthwhile considering.
- SGX on the client side has more overhead, especially as the number of features in the dataset increase.

**Future work:**

- Compare training with GPU's vs CPU's with SGX
- Using SGX in a fully decentralized setting (without an orchestration server)

[1] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath,
"Generative adversarial networks: An overview," IEEE Signal Processing Magazine, vol. 35,
no. 1, pp. 53–65, 2018. doi: 10.1109/MSP.2017.2765202
[2] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods,
and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, 2020. doi:
10.1109/MSP.2020.2975749
[3] V. Costan and S. Devadas, "Intel sgx explained." Cryptology ePrint Archive, 2016.