CICI:UCSS: Confidential Computing in Reproducible Collaborative Workflows

Keke Chen[1] Zeno Franco[2] Zeyun Yu[3]

University of Maryland, Baltimore County[1], Medical College of Wisconsin[2], University of Wisconsin at Milwaukee[3]

Award #: 2517121

Project Motivation & Overview

Scientists' Needs: confidential processing of their sensitive assets
- Protect intellectual property,
- Share sensitive data,
- Prevent data (or algorithm) sharing before publications
- Conform to legal requirements

Most practical solution:
Confidential computing with CPU-enabled trusted execution environments (TEE)

Challenges

- Developing TEE-based algorithms is not trivial for domain scientist users
  - TEEs may have specific APIs and application development and deployment frameworks
  - Challenging to write secure code to protect from side-channel attacks
- In collaborative workflows, the interplay between private components and other critical system components and additional side channels for adversaries demand a study on
  - New attacks
  - New defenses

Intellectual Merit

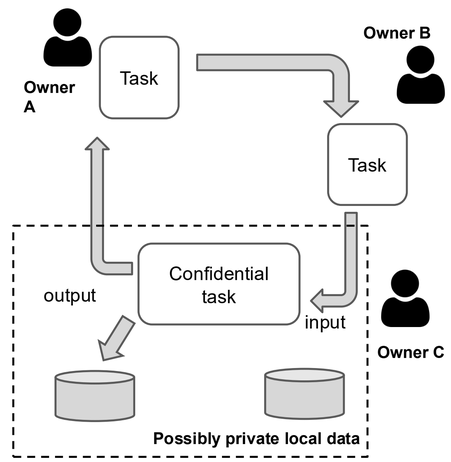
- A scientist-friendly oblivious program development framework for TEE programs
- Different protection and usability solutions for domain scientists to trade-off
- A holistic approach to studying possible security and privacy threats in a collaborative workflow with confidential component
- Science-driven, motivated and validated by collaborative research projects in science data analytics

Broader Impacts

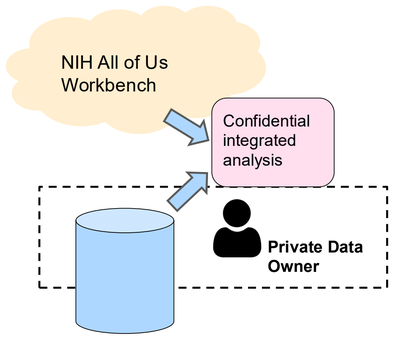
- Open-source library for scientist-friendly oblivious programming
- Secure methods for integrating TEEs into reproducible workflows.
- Boost the idea of open, collaborative science without the concern of confidentiality and privacy breaches.
- Applications in Internet-of-Things (IoT), clouds, and edges.
- Educational and outreach activities

Technical Approach

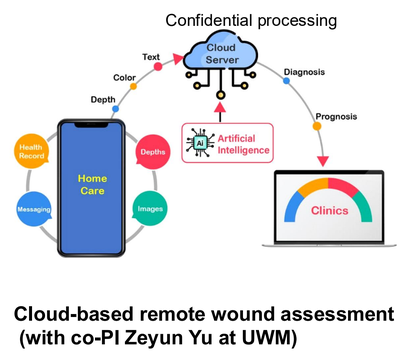
1. Scientist-friendly TEE development: usability and protection
   - Oblivious program development framework against side-channel attacks
   - Usability-first no side-channel-protection approaches
2. Study attacks on confidential components in a reproducible collaborative workflow
   - curious participants
   - dishonest owners
   - issues in reproducibility verification
3. Study defense measures
4. Validate with/apply to scientific workflows



Sample Scientific Applications



Integrating private data and All of Us data (with co-PI Zeno Franco at MCW)



Cloud-based remote wound assessment (with co-PI Zeyun Yu at UWM)

Related Recent Publications

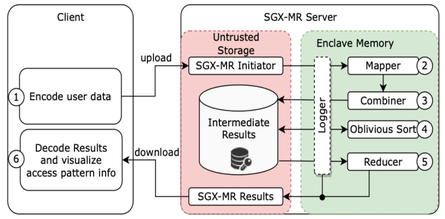
- Yuechun Gu, Jiajie He, and Keke Chen, "Adaptive Domain Inference Attack with Concept Hierarchy", in Proceedings of ACM SIGKDD conference, Toronto, 2025
- Mubashwir Alam and Keke Chen, "TEE-MR: Developer-friendly data oblivious programming for trusted execution environments," in Computers and Security Journal, Vol 148, 2024
- Mubashwir Alam and Keke Chen, "TEE-Graph: efficient privacy and ownership protection for cloud-based graph spectral analysis", in Frontiers in Big Data, 2023
- Keke Chen, Yuechun Gu, and Sagar Sharma, "DisguisedNets: Secure Image Outsourcing for Confidential Model Training in Clouds", ACM Transactions on Internet Technology, volume 23, issue 3, 2023
- Mubashwir Alam and Keke Chen, "Making Your Program Oblivious: a Comparative Study for Side-channel-safe Confidential Computing", IEEE Conference on Cloud Computing (CLOUD), Chicago, 2023
- Mubashwir Alam, Justin Boyce, and Keke Chen, "Demo: SGX-MR-Prot: Efficient and Developer-Friendly Access-Pattern Protection in Trusted Execution Environments, IEEE Conference on Distributed Computing Systems (ICDCS), Hong Kong, China, 2023
- Yuechun Gu and Keke Chen, "GAN-Based Domain Inference Attack," AAAI 2023
- Keke Chen, "Confidential High-Performance Computing in the Public Cloud," IEEE Internet Computing, Vol 27, no 1, 2023

Recent Accomplishments

TEE-MR: Scientist-Friendly Oblivious Program Development

Problem:
Difficult to develop data oblivious Programs (resilient to side-channel attacks) for domain scientists

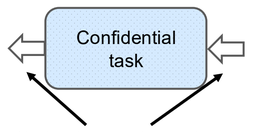
Approach:
Use an application framework (e.g., mapreduce pattern) to regulate dataflow and hide details of protection;
simplify developers' tasks.



Domain Inference Attacks

Problem:
Study attacks on a confidential component in a collaborative Environment.

Model the confidential task as a function y= f(x);
no knowledge about the function;
Knows only input and output
-- What can an attacker learn?



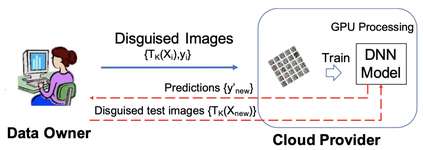
Approach:
- Use confidential deep-learning models as the target
- Develop "domain inference attacks" to infer

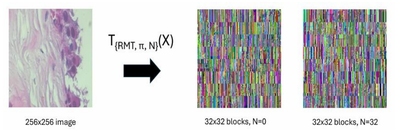
the training data information

Image Disguising Methods

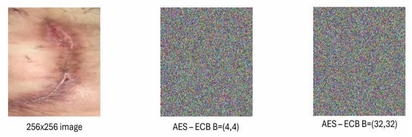
Problem:
Alternative methods for confidential model learning on untrusted platforms where **TEEs are unavailable**

Approach:
- Image Disguising methods for training DNN in the cloud
- Practical for confidential medical image processing





Breast cancer images: encoded learnable Images for classification



Wound images: learnable images for segmentation

Research in Progress

- Scientist-friendly usability-first TEE deployment framework
- Federated learning as an example of collaborative workflow – how TEEs can help and what are the unique problems
- AI-driven methods for critical access-pattern detection and automated oblivious transformation
- Attacks and defenses on multi-agent collaborative AI environments