Yifei Pang

Pittsburgh, PA 15213
yifeip@andrew.cmu.edu
https://2020pyfcrawl.github.io/

EDUCATION

M.S.: Carnegie Mellon University (CMU), Pittsburgh, USA Aug. 2023 – Present (Expected May 2025)
Major: Information Security
CQPA: 4.0/4.0

B.Eng.: Zhejiang University (ZJU), Hangzhou, China Sept. 2019 - Jun. 2023
Major: Computer Science and Technology
Overall GPA: 3.80/4.00 **The last two years GPA**: 3.94/4.00

RESEARCH INTERESTS

I am broadly interested in **identifying vulnerabilities of machine learning (ML) models** and **designing practical solutions for them.** Specifically, I am deeply interested in exploring **the security and privacy of ML models** and advancing **empirical methods for auditing privacy and security of ML applications**. I am also interested in ML for security.

PUBLICATONS

Yifei Pang, Sreenidhi Ganachari, Yuan Yuan, Steven Wu, Xiaojing Dong, Jin Xu, Zhenyu Yan. "A New Approach to Generate Individual Level Data of Walled Garden Platforms: Linear Programming Reconstruction" *NeurIPS 2024 Workshop on Behavioral ML*.
He, Anxiao, Kai Bu, Jiongrui Huang, **Yifei Pang**, Qianping Gu, and Kui Ren. "SwiftParade: Anti-burst Multipath Validation." *IEEE Transactions on Dependable and Secure Computing (2023)*.
Yifei Pang, Anxiao He, Wenjie Hou, Yunyi Teng, Kai Bu, and Kui Ren. "SwiftOracle: Orthogonality-driven Private Multipath Validation", ready to submit to *IEEE Transactions on Dependable and Secure Computing*.

RESEARCH EXPERIENCE

A New Approach to Generate Individual Level Data of Walled Garden Platforms: Linear Programming Reconstruction | Research intern | CMU Apr. 2024 - Oct. 2024
Advisor: Steven Wu, Assistant Prof., School of Computer Science, CMU; Xiaojing Dong, Associate Prof., Leavey School of Business, SCU; Yuan Yuan, Austin Xu, Adobe.

- This project aims to bridge the gap between aggregate statistics released under privacy regulations and individual-level data needed for machine learning applications.
- I proposed and implemented a novel Linear Programming-based three-step algorithm to reconstruct individual-level data from aggregated statistics of Walled Garden Platforms, and evaluated its effectiveness using the Markov attribution model.
- The reconstruction achieved less than 10% error in most cases; the paper was accepted by **NeurIPS 2024 Workshop on Behavioral ML** [paper, code, poster].

SwiftOracle: Orthogonality-driven Private Multipath Validation | Research Assistant | ZJU Feb. 2023 - Present
Advisor: Kai Bu, Associate Prof., College of Computer Science and Technology; Member, Institute of Cyberspace Research, ZJU

- This project focuses on designing a private multipath validation protocol that efficiently enforce and verify packet transmission paths while preserving path privacy.
- I proposed a novel orthogonality-driven approach, leveraging linearly independent orthogonal vectors to construct a privacy-preserving multipath validation algorithm, and implemented it in DPDK using C++ for real-world performance evaluation.
- The approach demonstrated over a 20-fold throughput improvement and applicability for larger-scale networks compared to state-of-the-art methods. The paper is ready to submit to **IEEE Transactions on Dependable and Secure Computing** (TDSC).

SwiftParade: Anti-burst Multipath Validation | Research Assistant | ZJU Jul. 2022 - Jun. 2023

Advisor: Kai Bu, Associate Prof., College of Computer Science and Technology; Member, Institute of Cyberspace Research, ZJU

- This project focuses on designing an efficient multipath validation protocol to process burst traffic, using a noncommutative homomorphic asymmetric encryption scheme with constant proof size and group-wise proof generation and verification.
- I refined the algorithm by identifying and fixing vulnerabilities, pruning dispensable parts, enabling group-wise computation to improve efficiency, and implemented the algorithm in DPDK using C++ for multi-core performance evaluation.
- The algorithm speeds up packet processing by 2.5×~8.3× and increases communication throughput by 2.8×~10.2× compared to state-of-art approach. The paper was accepted by **IEEE Transactions on Dependable and Secure Computing** (TDSC) in Sep. 2023 [paper].

PROJECTS

ML For Mobile-App Fingerprinting | Course Project in CMU 14742 (Security in Networked Systems) Feb. 2024 - Apr. 2024

- This project focuses on analyzing and improving the **FlowPrint** model, which generates and classifies mobile app fingerprints based on encrypted network traffic.
- I refined the model by filtering frequent flows, such as common DNS interactions with public DNS resolvers, to improve its accuracy and evaluate the performance to recent app traffic analysis.
- The refined model shows slight improvements in accuracy, and demonstrated effectiveness in analyzing recent mobile app traffic with low latency.

Efficient Hyperparameter Tuning For GPT-2 Tiny Model with 30M Parameters

| Course Project in CMU 10605 (Machine Learning with Large Datasets) Nov. 2024

- This project focuses on finding optimal hyperparameters (here learning rate) for GPT-2 Tiny Model with a limited budget.
- I Designed and implemented the **Heuristic Model-Scaling Projection** method, by first tuning the smallest model (3M parameters) using adaptive search with early stopping, then guided the search for two larger models (7M and 11M parameters) based on the initial optimal value, ultimately projecting these results to determine parameters for the full 30M parameter model.
- The approach achieved a learning rate of 0.004625 for GPT-2 Tiny model, approximating the true optimal value of 0.005, while only consuming the budget of training the 30M model once. I was invited to present my methodology in class.

TEACHING

14740 Fundamentals of Telecommunications Networks, CMU, Teaching Assistant 2024 Fall

EXTRACURRICULAR ACTIVITIES

Mental Health Association Union | Member | ZJU Oct. 2020 - Jun. 2021

- Participated in organizing college psychological association communication activities in Hangzhou
- Organized activities about mental health in school, and did volunteer work

SKILLS

Programming Languages: C/C++, Python, Java, Go