Cedar Ren

(757) 279-4582 | renresear.ch | cedar.ren@gmail.com

EDUCATION

University of Virginia Sep 2019 – Jan 2023
Dropped out of PhD in Computer Science | Machine Learning and Computer Architecture | GRE: Verbal:167, Quant:168 *GPA*: 4.0
College of William and Mary August 2016 – May 2019
Bachelors of Science | Double Major in Computer Science and Mathematics *GPA*: 3.8

SKILLS

Programming Languages Rust, C++, Python, C , Zig, SQL, MatLab, Lean
Software Torch, Numpy, Keras, TensorFlow, sklearn, Linux, GNUPLOT
Specializations Machine Learning, Queuing Theory, Statistics, Probability, CPU/GPU Performance Profiling
Interests Security, Hardware Accelerators, Formal Verification, Ursula K LeGuin books, FengShui

WORK EXPERIENCE

MatX – Software Stack for Transformer ASICs Apr 2025 – Current
- Understand performance requirements of recent LLMs (Deepseek, Kimi, HRM) and develop simulator to inform hardware design
- Develop research kernels and compilers to support hardware design.

AMD – Open source ML compiler stack Sep 2023 – Apr 2025
- Lead bring-up of performant serving stack for pre-compiled LLMs by writing the CI setup, implementing prefix-sharing kv caching, and squashing numerous bugs. See list of PRs here.
- Enabled performant compilation of LLMs on Vulcan and ROCm via contributions to torch-mlir, IREE & SHARK-TestSuite

Stealth Mode Crypto Data Company / Consulting role Mar-Sep 2023
- Improved Zig Etherium implementation performance by ~2x by migrating to an 100x faster AVX512 vectorized hash function
- Used Tracy with performance counters + instrumentation to detect performance hotspots

Intel Labs – Architecture Tooling Group | Research Intern Oct 2022 – Jan 2023
- Accelerate SimPoint generation by 200x using hardware performance counters sampling to avoid instrumentation
- Achieve <3% CPI and <10% MPKI estimation accuracy while retaining 1,000,000x benchmarking speedup from SimPoint

University of Virginia – Computer Science Department | PhD Candidate Sep 2019 – Current
- Applied formal verification to ensure that quantized machine learning models remained invulnerable to adversarial attacks using DNNV (https://github.com/dlshriver/dnnv), ONNX, and ReluPlex (https://arxiv.org/abs/1702.01135)
- Discovered 2 critical security flaws that threatened execution integrity and data security in modern x86 processors.
- Mentored 5 undergraduate students on computer architecture and machine learning projects, breaking down large projects into digestible chunks, as well as providing instruction on computer architecture, side-channel attacks, machine learning compilers, and ML models (incl. model specification, feature engineering, parameter tuning, and cross-validation).

NXP Semiconductors – Edge Security | ML Research Intern May 2022 – Aug 2022
- Applied statistical and machine learning algorithms (*incl. logistic regression, perceptrons, time-convolutional neural networks, decision trees, k-nearest neighbors, random forests, support vector regressions*) to monitor CPU performance counters for Spectre and Meltdown type side-channel attacks (*Python, scikit-learn, pandas, statsmodels, NumPy, MLJar*)

SELECTED PROJECTS

ProxyVM – In collaboration with Intel Labs and the Semiconductor Research Corporation Jan 2022 - Current
- Augment profiling tools with differential privacy to enable ML hardware supply chain collaboration without loss of privacy
- Accelerated pre-silicon hardware simulations while maintaining high performance predictability by generating augmented performance traces (basic block vectors augmented with data access pattern vectors).
- Extended existing system to emerging hardware and workloads using LLVM and MLIR as a compatibility layer
- Modify cross-platform machine learning compiler to generate execution traces for benchmarking on CPU, GPU, FPGA, and ASIC

I See Dead Micro-Ops Sep 2019 – Jan 2021
- Analyzed Intel x86 processor design documents to discover potential vulnerability, craft microbenchmarks to reverse undocumented CPU features, and design proof-of-concept exploits for novel vulnerabilities
- Designed micro-architectural benchmarks that characterized undocumented x86 instruction translation mechanisms
- Published novel spectre-type attack in International Symposium on Computer Architecture (15% acceptance rate).
- Published SMT performance-preserving speculative side-channel defenses to protect processors without compromising performance. USENIX security (18% acceptance rate)

Equity AI (*Honors Thesis, Summa Cum Laude)* Dec 2017 – May 2019
- Distributed bayesian hyperparameter optimization over a small (~50 machines) cluster, using idle lab computing.

Int'l Genetically Engineered Machines Contest *(iGEM 2017, Int'l 2nd Place & Best Model Award)* Oct 2016 – Sep 2017
- Won 2nd place overall & best math model in international genetic engineering contest (iGEM 2017)
- Modeled genetic circuits with partial differential models + MCMC, testing predictions against wet-lab experimental results
- Designed plasmid to implement protein-protease gene circuit to demonstrate novel gene expression rate control method