

Shahrooz Pouryousef

Santa Monica, CA | +1-413-404-2650 | email | github.com

Work authorization: U.S. Permanent Resident (Green Card)

Summary

Networking researcher focused on DCN transport and traffic engineering and the AI \times Networking seam. Built event-driven simulators and ran measurement studies; Now applying networking methods to decentralized/distributed training (DDP/micro-batching, tail-latency profiling) at the AI \times Networking seam.

Skills

- **Networking & Data-Center:** Congestion control (TCP/DCTCP/CUBIC concepts), ECN/AQM, load balancing (ECMP/flowlets), routing/TE, Clos/BCube/leaf-spine
- **AI Systems (applied):** PyTorch, Hugging Face Transformers; offline evaluation (Precision/Recall@K, NDCG@K); basic distributed training concepts (DDP, micro-batching, input pipeline throughput); latency/memory profiling; *model-efficiency basics from coursework*: pruning and post-training quantization.
- **Simulation & Experimentation:** ns-3, OMNeT++; event-driven simulation (SimPy/asyncio).
- **Programming & Data:** Python (pandas/NumPy/SciPy, NetworkX), Jupyter; Git/Linux; L^AT_EX; optimization tooling (Gurobi, IBM CPLEX).
- **Course-derived competencies (selected):** Systems for Deep Learning (COMPSCI 690AB) \rightarrow DDP fundamentals, profiling, pruning/quantization basics; Neural Networks (682) \rightarrow backprop, regularization; Distributed & Operating Systems (677) \rightarrow concurrency, scheduling; Computer Networks (653) \rightarrow congestion/ECN/queueing; Advanced Algorithms (611) \rightarrow complexity/approximation intuition; AI (683) \rightarrow search/intro RL; Security (560) \rightarrow secure systems mindset.

Experience

Research Scientist (Quantum Systems) — Cisco *Santa Monica, CA* 2025–Present

- Networking for distributed systems: topology-aware benchmarking, event-driven simulation, and scheduler-centric performance analysis for communication-bound workloads (done in quantum-DCN context).

Ph.D. Researcher (Networks/Systems) — University of Massachusetts Amherst *Amherst, MA* 2019–2025

- Developed **event-driven simulators** and measurement pipelines; ran congestion-control and TE experiments (ns-3/OMNeT, Python);
- Published in NSDI/SIGCOMM/INFOCOM/QCE/TQE/QCNC; mentored undergraduate researchers

Research Intern (Traffic Engineering) — ByteDance *Remote* 2021

- Prototyped **RL-based traffic engineering** to improve **link utilization** and flow completion times relative to an **ECMP** baseline in a Python simulation environment.

Research Intern (Quantum Networking) — Cisco *San Jose, CA* 2024

- Performed **performance analysis of quantum networking protocols** using a **SimPy event-driven simulator**; modeled EPR generation success/latency, switching delay, and queueing at constrained resources.
- Ran parameter sweeps (link loss, attempt rate, path length) and reported throughput/tail latency and EPR consumption

Selected Projects

- **Benchmarking DCN Topologies (Cisco):** Studied how leaf-spine variants, oversubscription, and latency paths affect step-time sensitivity and end-to-end throughput.
- **TCP Congestion-Control Simulation (Simplex/ns-3):** Event-driven experiments logging CWND, queue occupancy, and latency; CDF/time-series analysis under ECN/AQM settings
- **Topology Generator & Scheduling Experiments: Layered/Clos topology generator** with per-switch capacity limits; explored path diversity vs. blocking and simple scheduling sensitivities under varying loads.

- **Do LLMs Understand Collaborative Signals?:** Analyzed LLM limitations on collaborative-filtering structure; explored contrastive “repair” strategies and compared with MF/NCF baselines for Top- N recommendation. [arXiv]
- **Tradeoffs: LLM Accuracy vs. Context Length vs. Processing Time:** Measured how prompt length impacts quality and decoding latency/throughput using public `lightning.ai` code; connects token budgeting to tail latency in distributed serving. [link]
- **LLM from Scratch (PyTorch):** Minimal transformer (positional encodings, scaled dot-product attention, causal masking, basic generation) with quick latency/memory checks [GitHub]

Publications (selected)

- S. Pouryousef, L. Gao, A. Venkataramani, “Towards Logically Centralized Interdomain Routing,” *NSDI*, 2020. [link](#)
- **Shahrooz Pouryousef**, et al. Do LLMs Understand Collaborative Signals? Diagnosis and Repair. In *(CIKM)*, 2025.
- S. Pouryousef, L. Gao, D. Towsley, “Robust Path Selection in Software-defined WANs using Deep Reinforcement Learning,” 2022 (preprint). [arXiv:2212.11155](#)
- Shahrooz Pouryousef, Muhammad Daniyal Dar, Suleman Ahmad, Phillipa Gill, Rishab Nithyanand, Extortion or Expansion? An Investigation into the Costs and Consequences of ICANN’s gTLD Experiments — Passive and Active Measurement (PAM) (2020)
- **Shahrooz. Pouryousef**, et al. A Quantum Overlay Network for Efficient Entanglement Distribution INFO-COM 2023.

Education

Ph.D., Computer Science — University of Massachusetts Amherst 2025

Dissertation: *Resource Allocation in Quantum Networks* (scheduling/orchestration and evaluation under probabilistic constraints)

Advisor: Prof. Don Towsley

M.S., Computer Science — University of Massachusetts Amherst 2020

M.S., Computer Engineering — Sharif University of Technology 2015

B.S., Information Technology — University of Tabriz 2013

Service & Teaching

Instructor for two undergraduate CS courses; mentored six undergraduates across networking/ML projects. Reviewer: IEEE TQE/ToN conferences/journals; student travel grants at major networking venues.