Wonkyo Choe

wonkyochoe@virginia.edu | +1 434 234 7911 | https://wonkyoc.github.io/

Professional Summary

Ph.D. Candidate in Computer Science specializing in system infrastructure for large-scale machine learning. My dissertation builds efficient, scalable infrastructure for serving generative AI workloads, designing systems that coordinate compute, memory, and storage across heterogeneous resources. I work across the full stack, from operating systems and virtualization to LLM inference, with published systems research spanning memory management, edge inference, and model compression.

Education

University of Virginia, Ph.D. in Computer Science | Advisor: Felix Xiaozhu Lin — 2021 – 2027 (Expected)
Ajou University, M.Eng in Computer Engineering (With Full Scholarship) — 2018 – 2021
Ajou University, B.Eng in Computer Engineering — 2011 – 2018

Industry Experiences

Web Development Intern — Feb 2017 – Feb 2018
Biztech Inc, Irvine, CA

- Engineered end-to-end web applications for clients; managed the full development lifecycle from initial front-end design to backend architecture and post-launch support; Funded by university international scholarship.

Selected Publications

SpMAP: Transparent Sparsity for LLMs, **MobiSys'26** [PDF]
Wonkyo Choe, Felix Xiaozhu Lin. Syscall-driven sparse LLM loading; cut resident memory by 63% on edge devices.

RWKV-Lite: Deeply Compressed RWKV for Resource-Constrained Devices, **arXiv 2025** [PDF]
Wonkyo Choe, Yangfeng Ji, Felix Xiaozhu Lin. SVD-based compression; up to 5× smaller memory footprint.

Proto: A Guided Journey through Modern OS Construction, **SOSP'25** [PDF]
Wonkyo Choe*, Rongxiang Wang*, Afsara Benazir*, Felix Xiaozhu Lin. Educational OS with progressive apps (* equal contribution).

AnA: An Attentive Autonomous Driving System, **ASPLOS'25** [PDF]
Wonkyo Choe, Rongxiang Wang, Felix Xiaozhu Lin. Dual-perception modules; 2× fewer collisions, 43% less compute.

Skills

Languages: C, C++, Python3, Bash, Latex, Markdown
Skills: LLM / Diffusion Inference, Linux Kernel Development, System optimization, Performance Profiling
Technologies: PyTorch, Docker, K8s, Git

Honors & Awards

Departmental Scholarship, Ajou University (2011, 2016, 2017, 2018)
Graduate Outstanding Teaching Award, University of Virginia (2025)

Professional Activities

Peer Reviewer, IEEE Transactions on Mobile Computing (2024)