

Wangsong Yin (尹望嵩)

Ph.D. Candidate (4th Year)

Peking University, China

Email: yws@stu.pku.edu.cn

Education

09/2022 – present

Ph.D. in Computer Science, PKU

- Advisor: Gang Huang, Xuanzhe Liu
- Remote Advisor: Mengwei Xu
- Research Topic: Mobile AI / AIoT

09/2018 – 07/2022

B.E. in Software Engineering, NWPU

- GPA Ranking: 1/258
- National Scholarship: Each year (2019/2020/2021)

Publications (* = equal contributions)

(full list at <https://scholar.google.com/citations?user=9hp2UsAAAAJ&hl=zh-CN>)

[MobiCom'25] “Elastic On-Device LLM Service”

Wangsong Yin, Rongjie Yi, Daliang Xu, Gang Huang, Mengwei Xu, Xuanzhe Liu

[SenSys'24] “PieBridge: Fast and Parameter-Efficient On-Device Training via Proxy Networks”

Wangsong Yin, Daliang Xu, Gang Huang, Ying Zhang, Shiyun Wei, Mengwei Xu, Xuanzhe Liu

[arXiv] “ShadowNPU: Dynamic Sparse Attention on Mobile SoCs”

Wangsong Yin, Daliang Xu, Mengwei Xu, Gang Huang, Xuanzhe Liu

[SenSys'26, To Appear] “LLM as a System Service on Mobile Devices”

Wangsong Yin, Mengwei Xu, Yuanchun Li, Xuanzhe Liu

[TMC'24] “EdgeLLM: Fast On-Device LLM Inference with Speculative Decoding”

Daliang Xu, **Wangsong Yin**, Hao Zhang, Xin Jin, Ying Zhang, Shiyun Wei, Mengwei Xu, Xuanzhe Liu

[ACM Computing Surveys'25] “Resource-Efficient Algorithms and Systems of Foundation Models: A Survey”

Mengwei Xu*, Dongqi Cai*, **Wangsong Yin***, Shangguang Wang, Xin Jin, Xuanzhe Liu

Teaching / Service Experience

- Teaching Assistant, Introduction to Computing (C++), 2022, PKU
- Reviewer, IEEE Transactions on Mobile Computing (TMC), 2025

Skills

- Programming Languages: Python/C/C++/Shell/SQL/Java
- Frameworks/Libraries: Pytorch, MLLM, QNN, MNN, HF Transformers, ARM NEON