

# Wei Wang

✉ wweione@outlook.com    🌐 <https://wwei.one>    🔄 wweione

## Education

---

- 2024 – now    **Ph.D. in Computer Science @ University of Maryland**
- 2023 – 2024    **Ph.D. in Computer Engineering @ Boston University**
- 2017 — 2021    **B.E. in Computer Science @ Northeastern University, China. GPA: 3.98**

## Research Interests

---

Interested in the interdisciplinary studies in **Networking Systems** and **Machine Learning Systems**.

## Professional Experience

---

- Sep 2023 – Present | **Research Assistant @ Boston University, University of Maryland**; Advisor: **Prof. Alan Zaoxing Liu**  
R2CC: Reliable & Resilient Collective Communication (NCCL/RCCL drop-in)  
• Designed a failure-tolerant CCL that reroutes around NIC/link faults to keep jobs running **without restarts**. • Prototype on H100/InfiniBand; training overhead ~1.5% (7B@512 GPUs), inference latency impact ~3%.
- Jul 2022 – Jul 2023 | **AI Framework Engineer, Full Time @ Intel Corporation**  
🔄 Intel Extension for TensorFlow: An Open Source TensorFlow Backend for Intel CPU & GPU  
• Improved hardware utilization and peak performance by optimizing TensorFlow computation graph.  
• Designed an int8 quantization solution for framework, which increased general inference speed by 2-4 times.
- Jul 2020 – Feb 2022 | **Research Assistant @ Peking University**; Advisor: **Prof. Qun Huang**  
AutoSketch: Automatic Sketch-Oriented Compiler for Query-driven Network Telemetry  
A telemetry system combing sketch with high-level query interfaces to meet telemetry intent with minimal resources.  
• Designed algorithms to bridge the gap between user accuracy intent and complex sketch parameters.
- Dec 2018 – Dec 2021 | **Research Assistant @ Northeastern University**; Advisor: **Prof. Fuliang Li**  
TalentSketch: LSTM-based Sketch for Adaptive and High-precision Network Measurement  
• Proposed a deep learning based adaptive and high-precision measurement framework to support typical sketch to recognize poorly estimated flows and improve accuracy by 12%~23%.

## Honors and Awards

---

- Jul 2021 | Excellent Undergraduate Thesis, Northeastern University, **15/491**
- 2019 | The 2019 **ACM-ICPC, Asia Regional, Gold Medal**  
The 2019 China Collegiate Programming Contest, **Silver Medal**
- 2018 | The 2018 **ACM-ICPC, Asia Regional, Xian, Xuzhou, Bronze Medal**
- 2017 - 2019 | Outstanding Student Award, Northeastern University, **21/491**  
Northeastern University Scholarship, Northeastern University, **21/491**

## Publications

---

- 1 Reliable and Resilient Collective Communication Library for LLM Training and Serving  
**Wei Wang**, Nengneng Yu, Sixian Xiong, Zaoxing Liu. **under review**, arXiv.
- 2 AutoSketch: Automatic Sketch-Oriented Compiler for Query-driven Network Telemetry.  
Haifeng Sun, Qun Huang, Jinbo Sun, **Wei Wang**, Fuliang Li, Yungang Bao. In **NSDI 2024**
- 3 TalentSketch: LSTM-based Sketch for Adaptive and High-precision Network Measurement.  
Yangsheng Yan, Fuliang Li, **Wei Wang**, Xingwei Wang. In **ICNP 2022**