

# Kai-Siang Wang

Website: [kwang.cc](http://kwang.cc)

Github: [github.com/kwang1012](https://github.com/kwang1012)

Email: [kw37@illinois.edu](mailto:kw37@illinois.edu)

Mobile: +1-447-902-0906

## RESEARCH INTERESTS

Systems for ML, Distributed Systems, Internet of Things, Edge Computing, Cloud Computing, Algorithmic optimization

## EDUCATION

- **University of Illinois Urbana-Champaign (UIUC)** Urbana, IL, USA  
*Ph.D. in Computer Science*  
*Aug 2023 – Present*  
*Advised by Professor Indranil (Indy) Gupta*
- **National Tsing Hua University** Hsinchu, Taiwan  
*Master's degree - CS*  
*Sep 2021 - Jan 2023*  
*Advised by Professor Jerry Chou*
- **National Tsing Hua University** Hsinchu, Taiwan  
*Bachelor's degree - EECS*  
*Sep 2017 - Jun 2021*  
*Advised by Professor Jang-Ping Sheu*

## PUBLICATIONS

- **Kai-Siang Wang, Jerry Chou. "A Reservation-Based List Scheduling for Embedded Systems with Memory Constraints."**: In *PDCAT*, 2022.
- **Kai-Siang Wang, Cheng-Han Hsieh and Jerry Chou. "Optimal Static Bidding Strategy for Running Batch Jobs with Hard Deadline Constraints on Spot Instances."**: In *CLOSER*, 2023
- **Chen-Chun Chen, Kai-Siang Wang, Yu-Tung Hsiao, Jerry Chou. "ALBERT: An automatic learning based execution and resource management system for optimizing Hadoop workload in clouds."**: in *Journal of Parallel and Distributed Computing*, vol. 168, pp. 45-56, Oct 2022

## RESEARCH EXPERIENCE

- **DPRG Research Group, University of Illinois Urbana-Champaign** Urbana, IL, USA  
*Graduate Research Assistant*  
*Aug 2023 – Present*
  - **Smart Home**: Design a new abstraction for smart home environments to provide better observability
  - **Fault-tolerant Pipeline Parallelism**: Analyze churns that occur during long-term training with pipeline parallelism.
- **LSA Lab, National Tsing Hua University** Hsinchu, Taiwan  
*Graduate Research Assistant*  
*Mar 2021 - Jan 2023*
  - **ML task scheduling**: Design and implement a reservation-based scheduling algorithm to reduce the makespan of a schedule under memory constraints
  - **Cloud provisioning**: Analyzed Spot Instances price history and proposed an optimal bidding strategy to reduce the cost.
- **HSCC Lab, National Tsing Hua University** Hsinchu, Taiwan  
*Undergraduate Research Assistant*  
*Sep 2020 - Jan 2021*
  - **CNN Parallelization**: Work on the project, *Parallelize CNN inference on multiple Raspberry Pi devices*. We build a system that provides end-to-end and automatic framework for parallel image inferencing.

## INDUSTRY EXPERIENCES

- **Skymizer Inc.** Taipei, Taiwan  
*Software engineer Part-time*  
*May 2021 - March 2023*
  - **Forest Runtime - Runtime that can execute executables provided by ONNC**: Participated in the runtime implementation in C and provided Python API.
  - **DL Pipeline Implementation**: Based on Pipedream(proposed by Microsoft), implement an asynchronous pipeline for DNN inference with ONNX.
- **Industrial Technology Research Institute** Hsinchu, Taiwan  
*System Architect Intern*  
*Jul 2018 - Dec 2018*
  - **Implementation**: Implement NVDLA using C++.
  - **Version Control**: Work with collaborators and use Git for version control.
- **Twister5 Inc.** Taipei, Taiwan  
*Full-stack developer Part-time*  
*Dec 2021 - May 2022*
  - **Admin Panel**: Use Vuejs to build the admin panel monitor the status of servers
  - **Backend Development**: Build a REST api with NEST.js and mongoDB.

TEACHING EXPERIENCE

---

- **Computer Science Department, University of Illinois Urbana-Champaign** Urbana, IL, USA  
*Graduate Teaching Assistant* Fall 2024
  - CS425 Distributed Systems
- **Computer Science Department, National Tsing Hua University** Hsinchu, Taiwan  
*Graduate Teaching Assistant* Fall 2022
  - CS542200 Parallel Programming

SKILLS

---

- **Languages:** Python, C++, SQL, Java, JavaScript
- **Frameworks:** TensorFlow, PyTorch, NodeJS
- **Tools:** Kubernetes, Docker, Git, PostgreSQL, MongoDB
- **Platforms:** Linux, Web, Windows, AWS, GCP