Jianwei Hao

Ph.D.
Computer Science
University of Georgia

GitHub
LinkedIn
+1-7064618898
glennglennhowe@gmail.com

Education

•University of Georgia *-2024.5*

Ph.D. in Computer Science

Projects

•IoT Energy Scheduling System

An on-device system with energy prediction models and dynamic scheduler.

- Two-step energy prediction: solar irradiance and energy gain. Time series solar irradiance prediction with algorithms **ARIMA**, **Prophet**, **LSTM**.
- Predicted energy is allocated based on sensors' **priority** and **energy cost** to dynamicly minimize sensing intervals.
- Data is boosted to adapt to Edge Computing systems. Regression and classification are used for energy inference.
- Energy regression models include **SVM**, **MLP**, **XGBoost**, and **Random Forest**.

•DNN Accelerator

*An **on-device** DNN model inference accelerator.*

- Use **concurrency**, **multiple processors**, and batching to accelerate DNN models' inference throughput on Edge Computing devices: Nvidia Jetson GPU and Coral TPU. The throughput increased 2-6.2x.
- Tools: Python, PyTorch, TensorFlow, TF Lite.
- Models: AlexNet, MobileNet, DenseNet, InceptionV3.
- Hardware: Raspberry Pi, Nvidia Jetson GPU, Jetson Nano, Jetson Xavier NX, Coral TPU.

Experience

•Honeywell *2022.5-2022.8*

Software Engineer

- Built up a cloud service cost **dashboard website** for CEOs to monitor the cloud resource (VM, Kubernetes, DB, Serverless) cost of the enterprise. The backend is built with Django platform. **REST APIs** are developed to communicate with the frontend (React JS). **Caching** is implemented to cache the query data and database. The database is refreshed periodically through scraping the cloud provider's API.
- Tools: **Python**, **PostgreSQL**, **Django**, **Redis**, **Azure API**.

•Ericsson *2008.7-2018.8*

Developer

- **Air Interface** feature verification for newly developed features on **3GPP** standard systems **3G/4G/5G**.
- Maintained network performance, identified bottlenecks, and implemented solutions to improve the network.
- Managed network **parameter tuning**, **handover optimization**, **RRC/RAB** KPIs, and frequency allocation to meet KPI requirements.
- Tech leader in managing hardware and software of 4G/3G RAN nodes: **eNodeB**, **RNC**, **NodeB**.
- Mentored junior engineers, providing technical guidance and contributing to knowledge sharing.
- Tools: **3GPP**, **Moshell**, **RAN simulator**, **CSR**, **TR**, **Jira**, **Jenkins**.

Technical Skills

VM, Docker, Kubernetes, Hadoop, Serverless, AWS, GCP, Azure, Python, Java, Pandas, scikit-learn, TensorFlow, PyTorch, Seaborn, Django, Redis, MySQL, PostgreSQL.

Publications

[1] **J. Hao** et al. "AI Multi-Tenancy on Edge: Concurrent Deep Learning Model Executions and Dynamic Model Placements on Edge Devices". In: *IEEE CLOUD* (2021).

[2] **J. Hao** et al. "An Empirical Analysis of VM Startup Times in IaaS Clouds". In: *IEEE CLOUD* (2021).

[3] **J. Hao** et al. "Characterizing Resource Heterogeneity in Edge Devices for Deep Learning Inferences". In: *SNTA* (2020).

[4] **J. Hao** et al. "DynaES: Dynamic Energy Scheduling for Energy Harvesting Environmental Sensors". In: *IEEE IPCCC* (2023).

[5] **J. Hao** et al. "Reaching for the Sky: Maximizing Deep Learning Inference Throughput on Edge Devices with AI Multi-tenancy". In: *ACM Transactions on Internet Technology (TOIT)* (2023).