Haibin Lai

Shenzhen | 12211612@mail.sustech.edu.cn | | my blog | github.com/HaibinLai

About Me

I am a Junior Turing Class student at SUSTech, majoring in Computer Science. Problem-solving is my lifelong delight. I am currently participating in research in the fields of High Performance Computing, GPU Computing, and Databases at the SUSTech HPC-Lab.

Education

Southern University of Science and Technology, BS in Computer Science Sept 2022 – Present

- GPA: 3.67/4.0 (sustech.edu.cn)
- **Coursework:** Computer Organization(H), Principle of Database(H), Artificial Intelligence(H)

  I am a member of the Turing class, which is designated for elite CS students at SUSTech.

Experience

Student Assistant, SUSTech Center for Computational Science and Engineering Feb 2024 – Aug 2024

- Regularly participate in the operation and maintenance of infrastructure in the center's machine room. Assist engineers in deploying multiple new nodes and resolving various network interruption issues. Conduct HPL/HPCG benchmark testing on the supercomputing cluster, demonstrating strong technical skills and effective communication awareness.

Deputy Class Monitor, Turing Class, Computer Science Department, SUSTech Sep 2022 – Present

- Assist class advisor in managing daily affairs, organize class activities such as discussions with Turing Award winner Professor Joseph Sifakis, special reports, and class tournaments. Demonstrated strong collective awareness along with effective English communication and writing skills.

Current Research

GPU-accelerated Dynamic Graph Pattern Mining Computing Framework **May 2024 - Present**

With the rapid development of big data, graph databases are being widely applied in industrial fields such as financial fraud detection, leading to various data mining algorithms like subgraph matching and frequent subgraph mining. Among these, Graph Pattern Mining (GPM) algorithms have become increasingly important in recent years, with a growing demand for their efficient deployment in distributed and heterogeneous systems. However, in real-time data environments like finance and networking, the timeliness of data streams poses challenges that traditional static graph pattern mining algorithms cannot meet. Therefore, this research aims to design a GPU-accelerated dynamic graph pattern mining system to enhance the responsiveness of GPM algorithms to dynamic graphs.
Tools Used: CUDA, C++

Projects

Bayesian Optimization-Based HPL Parameter Tuning System HaibinLai/HPLAutoTune

- In this project, we applied Bayesian optimization (BO) to tune HPL benchmark running on a CPU node inside a cluster system. We implement Bayesian optimization, a fast and rapid optimization method on tuning the hyperparameter, to achieve peak performance.
- Tools Used: Python

RISC-V Pipeline CPU with interruption CS214-Project-CPU

- Course project of Computer Organization. It contains several features: 5 stage pipeline, interruption dealing, UART communication.
- Tools Used: FPGA Verilog, Assembly

Experience

Online Intern, Zhejiang University DataEarth Lab Aug 2024 - Present

- Building Geochemistry Pi, a Python ML framework. Geochemistry Pi is an open-sourced highly automated machine learning Python framework for data-driven geochemistry discovery on tabular data.
- Geochemistrypi
- agupubs.onlinelibrary.wiley.com/doi/10.1029/2023GC011324

Intern, Beijing Sunway World Technology Co., Ltd. Aug 2023 - Sep 2023

- Studied the Laboratory Information Management System (LIMS) solutions, gaining insights into software workflows and architecture alongside the technical department. Learned to set up applications using SpringBoot and Docker container deployment. Collaborated with the sales team to engage with clients, understanding the full process from market research and client interactions to closing deals and contract signing. Acquired skills in market demand analysis tools and techniques, as well as in drafting client need analyses and summary reports.

SUSTech Geophysics Field Internship Aug 2023

- Accompanied the SUSTech Earth and Space Science Department to visit the Yunnan Earthquake Bureau, Dali Seismological Station, and Lijiang Astronomical Observatory for practical training. Applied Python deep learning techniques to analyze seismic wave data and compute source locations. Gained hands-on experience in deploying seismic stations in the field, data retrieval, and software processing like SAC software for natural earthquake data processing and MATLAB for data analysis, along with Linux command usage and Fortran programming.
- Tools Used: Python, MATLAB, Fortran

Awards

ASC Student Supercomputer Challenge: ASC24 April 2024

- First Prize
- Group Competition Award
- Responsible for parallel optimization of high-performance computing tasks such as LINPACK, AI LLM, and materials science calculations in the team, as well as Linux system debugging and cluster monitoring.

Contemporary Undergraduate Mathematical Contest in Modeling (CUMCM) Jan 2024

- Provincial 2nd Award

2023 Outstanding Student Award Nov 2023
2022 President's Special Scholarship Sep 2022

English Skills & Communication

CET6 June 2024

- Ministry of Education, China
- Score: 649/710

Georgia Institute of Technology ASP Program July 2023

- Participated in summer courses such as Business Case Study and Leadership Across Cultures at Georgia Tech, enhancing international communication skills and gaining insights into American culture.

Technologies

Languages: C++, C, CUDA, Java, python, Rust, SQL

Technologies: PostgreSQL, Docker, Hadoop, Linux