Haibin Lai

Shenzhen | 12211612@mail.sustech.edu.cn | | 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 Distributed Computing, GPU Computing, and High Performance Computing at the SUSTech HPC-Lab.

Education

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

- GPA: 3.70/4.0 (sustech.edu.cn)
- **Coursework:** Operating System(H), Computer Organization(H), Computer Network, Machine Learning(H)

Experience

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

- Regularly participate in the operation and maintenance of SUSTech Qiming and Taiyi Supercomputer. Assist engineers in deploying multiple new nodes and resolving various network interruption issues. Conduct HPL/HPCG benchmark testing on the supercomputing cluster.

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.

Current Research

GPU-accelerated Dynamic Graph Pattern Mining Computing Framework **May 2024 - Present**
Parallel Framework for Continuous Subgraph Matching

Graph databases are being widely applied in industrial fields such as financial fraud detection, leading to various data stream 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, real-time data streams like finance and networking poses challenges that traditional static graph pattern mining algorithms cannot meet. 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

Viritio-gpu support for Asterinas OS Repo: **asterinas** , **Report**

- In this project, we implement virtio-gpu driver in Asterinas Operating System on Qemu. The driver works with page buffer scheme and allow user library like mesa communicates with qemu virtio using specfic syscall.
- Tools Used: Rust

Bayesian Optimization-Based HPL Parameter Tuning System Repo: **HPLAutoTune**

- We applied Bayesian optimization (BO) to tune HPL benchmark on SUSTech Qiming Supercomputer to achieve peak performance.
- Tools Used: Python

CrashSimGen: Generating Safety-Critical Scenarios with Diffusion Models Repo: **CrashSimGen**

- **First prize** project of CS329 Machine Learning(H). CrashSimGen is a project that generates dangerous road scenarios using diffusion models for autonomous driving risk assessment
- Tools Used: Pytorch, Tensorflow

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 SpringBoot and Docker. 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. Gained hands-on experience in deploying seismic stations in the field, data retrieval, and software like SAC for natural earthquake data analysis, along with Linux command usage and Fortran,MATLAB programming.
- Tools Used: Python, MATLAB, Fortran

Awards

7th APAC HPC-AI Student Competition: APAC HPC-AI Dec 2024

- Third Prize
- Leader of HPC Team
- Responsible for parallel optimization on Hoomd-blue HPC software, using HPC-X communication library, and a better scheduling strategy for domain decompostion.

2024 Outstanding Student Award Nov 2024

ASC Student Supercomputer Challenge: ASC24 Apr 2024

- First Prize
- Group Competition Award
- Responsible for LINPACK benchmark and parallel optimization of materials science calculations in the team, as well as Linux and network 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 Jun 2024

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

Georgia Institute of Technology ASP Program Jul 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.

Techniques

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

Tools: PostgreSQL, Docker, Hadoop, Linux