

# Hongbin Zhong

☎ +86 18379727287    ✉ [rj986215159@outlook.com](mailto:rj986215159@outlook.com)  
💻 [github:https://github.com/rjzhb](https://github.com/rjzhb)

## INTEREST

---

- Machine Learning Systems(Model Inference, Model Compression, Distributed Training)
- Data-Intensive Systems(Data Analytical Systems, Database Systems)

## EDUCATION

---

<b>Georgia Institute of Technology</b>	Aug 2024 - Jun 2029
Ph.D in Computer Science	Atlanta
<b>Northeastern University</b> 985	Sep 2020 - Jul 2024
Bachelor in Computer Science	China

## PUBLICATIONS

---

[SIGMOD 2024] PECJ: Stream Window Join on Disorder Data Streams with Proactive Error Compensation.

Xianzhi Zeng\*, Shuhao Zhang(Corresponding author), **Hongbin Zhong**, Hao Zhang, Mian Lu, zhao zheng, Yuqiang Chen

[\[code\]](#) camera ready

## RESEARCH EXPERIENCE

---

<b>Columbia University</b>	Jul 2023 - Present
Research Assistant	New York City, NY
Adviser: <a href="#">Eugene Wu</a>	

- *FADE Project - Query Explanation (Answering "Why" Made Fast)*
  - **Using Sparse Matrix:** explore and implement how interventions evaluation can be adopted to work with sparse arrays without compromising performance
  - **Avoiding Reading And Writing To Disk:** Optimized the generation of Interventions, significantly reducing disk I/O and allowing most of the required data for the evaluation phase to be directly accessed within the CPP file.
  - **Apply SIMD, MultiThreading Optimizations:** Apply SIMD and MultiThreading to the Evaluation loops of Sparse Matrix data structure.
  - **Benchmarks And Experiments:** Finished numerous benchmarks and completed a significant number of experiments.

<b>Rutgers University</b>	Jul 2023 - Sep 2023
Research Assistant	New Jersey
Adviser: <a href="#">Dong Deng</a>	

- *Find similar data text segment between some documents*
  - **Implement Baseline:** [<DataPortraits>](#) [\[code\]](#)
  - **Implement Baseline(Python to C++):** [<SlimPajama>](#) [\[code\]](#)
  - **Assisting In Some Running Experiments:** Such as running slimpajama and get results.
  - **Optimizing Parallelization:** Optimize parallelization for group function.

<b>4paradigm Company</b>	Jan 2023 - Jul 2023
Research Assistant	Singapore

Adviser: [Mian Lu](#) & [Shuhao Zhang](#)

- **(ML4SYS)High-Accuracy Low-Latency Stream Window Join with Out-of-Order Data Arrival, Stream Processing System** [\[code\]](#)
  - **Implement Baseline:** [ICDE 2016<MultipleWay Sliding Window Stream Join>](#) [\[code\]](#)
  - **Finding And Organizing Datasets:** using scripts to process the datasets for subsequent experiments
  - **Help implement Bayesian variational inference** (transformer + Mathematical method)
  - **Optimizing Intra-Window Join:** Optimizing Intra-Window Join using adaptive filters.
  - **Assisting In Running Experiments Scripts:** Besides, modifying scripts to enhance the visual representation of experimental data.

## INDUSTRY EXPERIENCE

---

**Meituan(Top-tier internet company in China)**

Apr 2022 - Sep 2022

Software engineer

Beijing

- **Meituan's App Short Video project**
- **(I built the majority of the foundation for the entire app)**
  - **Data Transimission:** Developed log->kafka->hive to remotely report user video watching behavior to Kafka and synchronize data to Hive tables for recommendation algorithm training.
  - **User Algorithm And Data Refresh:** Improved user experience in poor network environments or for first-time visitors by pulling various data source videos using *Crane* scheduled task to refresh the data periodically.
  - **Implement user group-based resource access:** Through distributed caching *Cellar*, requiring migration of data from Hive data warehouse to a distributed cache Key-Value database due to significant user base growth.

## MAIN COURSES

---

Advanced Algebra	Probability and Statistics	Analytic Geometry
Mathematical Analysis	Discrete Mathematics	Artificial Intelligence
Data Structure	Operating System	Computer Networks
Database	Computer Architecture	Machine Learning
C++ Programming	Java Programming	Software
Engineering		
Complication Principle	Algorithm Design and Analysis	Linux System
Assembly Programming	Information Security	Computer Interface

Self-Learning(All experiments are finished):

- CMU 15-445 (Intro to Database by Andy Pavlo)
- MIT 6.S081(Operating system kernel)
- Stanford CS144(implement and optimize TCP protocol) [\[code\]](#)

## OTHER

---

- **Programming language Or Api:** C/C++, CUDA, MPI, OpenMP, Java, Rust, Python
- **Languages:** English(TOEFL 104), Chinese(Native)