Hao Jiang, Ph.D.

Postdoctoral Fellow
Harvard John A. Paulson
School Of Engineering And Applied Sciences
150 Western Ave, Boston, MA 02134

(315)267-6271
hajiang@seas.harvard.edu
https://harperjiang.github.io

As a Postdoctoral Researcher at Harvard University, my current research focuses on building the next-generation low-latency blockchain. During my Ph.D. study at the University of Chicago, my research focuses on Databases, Distributed Systems, and Machine Learning has resulted in numerous publications at leading academic conferences. With six years of experience as a Software Engineer, I am proficient in Java/C++/Python. I have contributed to the development of innovative software solutions for a range of industries.

Research and Work Experience

- **Harvard University** — Boston, MA
  Postdoctoral Researcher with Prof. Stratos Idreos — *Oct 2021 – Now*
  - Design and build the next generation high-throughput low-latency blockchain system that adapts to workloads

- **Facebook Inc.** — Menlo Park, CA
  Software Engineer Intern — *July. 2018 – Sept. 2018*
  - Design, implement and evaluate a load balance algorithm for PHP requests

- **The University of Chicago** — Chicago, IL
  Research Assistant under Prof. Aaron J. Elmore — *Sept. 2015 – Aug. 2021*
  - PIDS: Unsupervised Pattern Inference and Compression for String Attributes
  - SBoost: Use SIMD to speed up scanning on compressed data without decompression
  - CodecDB: Use Neural Network to optimize Lightweight Encoding for Columnar Database.
  - Design a new algorithms for Stream Partitioning of Large-Scale Graph
  - Design a sampling based method to classify whether a Large-Scale Graph satisfy power-law distribution

- **Clarkson University** — Potsdam, NY
  Research Assistant under Prof. Jeanna N. Matthews — *Sept. 2012 – Jul. 2015*
  - Design, implement and experiment several heuristic based partitioning methods on Internet Topology. Experiment on building a Internet Topology Platform support thirdparty data analysis programs to access Internet structure with ease.
  - Participate in GreenDataCenter (GDC) Project. Design and implement a simulation environment to study the feasibility of using pure green energy, such as solar and wind to power distribute data center and provide service with high availability.

- **Baidu Inc.** — Shanghai, China
  System Architect — *Sept. 2011 – Aug. 2012*
  - Design and implement a customizable MySQL replication framework. This framework intercepts message sent by a master database during MySQL master-slave replication, rewrite it with user provided function and send it to slave database. This framework allows an easy customization of replication behaviors.

- BigData Analysis System using Hadoop. Lead a 4-developer team to design and develop a Hadoop-based BigData analysis system. The system processes a daily data volume of over 10 terabytes.

- **OOCL Co. Ltd.** Shanghai, China
  Senior Software Engineer *Sept. 2005 – Sept. 2011*
  - Production Server JVM Performance Tuning. Design and implement a log analysis system for production environment JVM resource leak and memory leak tracing and tuning.
  - Design and optimization of an accounting system containing billions of rows and processing millions of queries hourly based on Oracle DBMS.

Publications

1. Chunwei Liu, **Hao Jiang**, John Paparrizos, Aaron J. Elmore *Decomposed Bounded Floats for Fast Compression and Queries VLDB 2021*
2. **Hao Jiang**, Chunwei Liu, John Paparrizos, Andrew A. Chien, Jihong Ma, Aaron J. Elmore, *Good to the last bit: Data-Driven Encoding with CodecDB*, *SIGMOD 2021*
3. **Hao Jiang**, Chunwei Liu, Qi Jin, John Paparrizos, Aaron J. Elmore, *PIDS: Attribute Decomposition for Improved Compressionand Query Performance in Columnar Storage*, *VLDB 2020*
4. Chunwei Liu, McKade Umbenhower, **Hao Jiang**, Aaron J. Elmore, *Mostly Order Preserving Dictionaries*, *ICDE 2019*
5. **Hao Jiang**, Aaron J. Elmore, *Boosting Data Filtering on Columnar Encoding with SIMD*, *DaMon 2018*
6. Dixin Tang, **Hao Jiang**, Aaron J. Elmore, *Adaptive Concurrency Control: Despite the Looking Glass, One Concurrency Control Does Not Fit All*, *CIDR 2017*
7. **Hao Jiang**, Yaoqing Liu, Jeanna N. Matthews, *IP Geolocation Estimation using Neural Networks with Stable Landmarks*, *IEEE INFOCOM Workshop GI 2016*
8. Wenjin Hu, Long Zhang, **Hao Jiang**, Jeanna N. Matthews, *A Quantitative Study of Virtual Machine Live Migration*, *CAC 2013*

Education

- **The University of Chicago** Chicago, IL
  Ph.D. Computer Science *2015 – 2021*
  - Advisor: Aaron J. Elmore
- **Clarkson University** Potsdam, NY
  M.Sc. Computer Science *2012 – 2015*
  - Advisors: Jeanna N. Matthews
- **Fudan University** Shanghai, China
  B.Sc. Computer Science *2001 – 2005*

– Advisor: Liang Zhang

Skills

- **Computer Science:** Database, Distributed Systems, Parallel Programming, Machine Learning
- **Tech Stack:** Java, Scala, C/C++, Python, Golang, Matlab, Javascript, HTML/CSS
- **Teamworking:** Experience of leading small development teams