

Huanchen Zhang

Tsinghua University
Institute for Interdisciplinary Information Sciences
4-6014 FIT Building, Tsinghua University, Beijing, China 100084

Email: huanchen@tsinghua.edu.cn
Web: <http://people.iis.tsinghua.edu.cn/~huanchen/>
Twitter: @huanchenzhang

Education

Carnegie Mellon University Sept. 2013 - Feb. 2020
Pittsburgh, PA USA
Ph.D. in Computer Science Department
Thesis: “Memory-Efficient Search Trees for Database Management Systems”
Advisor: David G. Andersen

University of Wisconsin - Madison June 2009 - May 2013
Madison, WI USA
B.S. in Computer Engineering, with *Distinctive Scholastic Achievement*
2nd and 3rd Major: Computer Sciences and Mathematics
Advisor: Remzi Arpaci-Dusseau

Professional Experience

Assistant Professor Jan. 2021 - Present
Beijing, China
Tsinghua University, Institute for Interdisciplinary Information Sciences (IIIS)

Postdoctoral Research Fellow April 2020 - Dec. 2020
San Mateo, CA USA
Snowflake Inc.

Graduate Research Assistant Sept. 2013 - Feb. 2020
Pittsburgh, PA USA
Carnegie Mellon University

Research Intern May 2016 - August 2016
Palo Alto, CA USA
Hewlett Packard Labs

Awards & Honors

- **2022 ACM China Rising Star Award Honorable Mention**
- 2022 ACM SIGMOD China Rising Star Award
- 2021 World Artificial Intelligence Conference (WAIC) Yunfan Award
- 2021 Excellent Young Scholar (Overseas), National Natural Science Foundation of China
- **2021 ACM SIGMOD Jim Gray Dissertation Award**
- 2020 Communications of the ACM (CACM) Research Highlights
- 2019 ACM SIGMOD Research Highlight Award
- **Best Paper Award, ACM SIGMOD 2018**
- Student Travel Grant, ACM SOSP 2015, USENIX NSDI 2014
- UW-Madison Graduate with Distinctive Scholastic Achievement, 2013
- Edgar H. and Laverne R. Krainer Memorial Scholarship, 2011 & 2012
- UW-Madison Computer Sciences Summer Research Assistant Award, 2012
- The John and Elizabeth Moore Award for Excellence in General Chemistry, 2011

Publications

- [1] Jiaoyi Zhang, Kai Su, and **Huanchen Zhang**. “Making In-Memory Learned Indexes Efficient on Disk”. In: *Proceedings of the ACM on Management of Data (SIGMOD’24)*, 2.3: Article 151, 26 pages.
- [2] Hengrui Wang, Te Guo, Junzhao Yang, and **Huanchen Zhang**. “GRF: A Global Range Filter for LSM-Trees with Shape Encoding”. In: *Proceedings of the ACM on Management of Data (SIGMOD’24)*, 2.3: Article 141, 27 pages.
- [3] Shuangyu Cai, Boyu Tian, **Huanchen Zhang**, and Mingyu Gao. “PimPam: Efficient Graph Pattern Matching on Real Processing-in-Memory Hardware”. In: *Proceedings of the ACM on Management of Data (SIGMOD’24)*, 2.3: Article 161, 25 pages.
- [4] Yihao Liu, Xinyu Zeng, and **Huanchen Zhang**. “LeCo: Lightweight Compression via Learning Serial Correlations”. In: *Proceedings of the ACM on Management of Data (SIGMOD’24)*, 2.1: Article 65, 28 pages.
- [5] Jiake Ge, **Huanchen Zhang**, Boyu Shi, Yuanhui Luo, Yunda Guo, Yunpeng Chai, Yuxing Chen, and Anqun Pan. “SALI: A Scalable Adaptive Learned Index Framework based on Probability Models”. In: *Proceedings of the ACM on Management of Data (SIGMOD’24)*, 1.4: Article 258, 25 pages.
- [6] Prashant Pandey, Martín Farach-Colton, Niv Dayan, and **Huanchen Zhang**. “Beyond Bloom: A Tutorial on Future Feature-Rich Filters”. In: *Proceedings of the ACM on Management of Data (SIGMOD’24)* Tutorial, June 2024.
- [7] **Huanchen Zhang**, Yihao Liu, and Jiaqi Yan. “Cost-Intelligent Data Analytics in the Cloud”. In: *Proceedings of the 2024 Conference on Innovative Data Systems Research (CIDR’24)*, January 2024.
- [8] Xinyu Zeng, Yulong Hui, Jiahong Shen, Andrew Pavlo, Wes McKinney, and **Huanchen Zhang**. “An Empirical Evaluation of Columnar Storage Formats”. In: *Proceedings of the VLDB Endowment (VLDB’24)*, 17.2: 148-161.
- [9] Zhiyu Mei, Wei Fu, Guangju Wang, **Huanchen Zhang**, and Yi Wu. “SRL: Scaling Distributed Reinforcement Learning to Over Ten Thousand Cores”. In: *Proceedings of the Twelfth International Conference on Learning Representations (ICLR’24)*, May 2024.
- [10] Leon Windheuser, Christoph Anneser, **Huanchen Zhang**, Thomas Neumann, and Alfons Kemper. “AdaCom: Adaptive Compression For Databases”. In: *Proceedings of the 27th International Conference on Extending Database Technology (EDBT’24)*, March 2024.
- [11] Jiesong Liu, Feng Zhang, Lv Lu, Chang Qi, Xiaoguang Guo, Dong Deng, Guoliang Li, **Huanchen Zhang**, Jidong Zhai, Hechen Zhang, Yuxing Chen, Anqun Pan, Xiaoyong Du. “G-Learned Index: Enabling Efficient Learned Index on GPU”. In: *IEEE Transactions on Parallel and Distributed Systems (TPDS’24)*, 35.6: 795-812.
- [12] Weitao Wan, Feng Zhang Zhang, Chenyang Zhang, Mingde Zhang, Jidong Zhai, Chai, **Huanchen Zhang**, Wei Lu, Yuxing Chen, Haixiang Li, Anqun Pan, and Xiaoyong Du. “Compressed Data Direct Computing for Databases”. In: *IEEE Transactions on Knowledge and Data Engineering (TKDE’23)*, 36.5: 1902-1918.
- [13] Junyi Zhao, **Huanchen Zhang**, and Yihan Gao. “Efficient Query Re-optimization with Judicious Subquery Selections”. In: *Proceedings of the ACM on Management of Data (SIGMOD’23)*, 1.2: Article 158, 26 pages.
- [14] Ke Wang, Guanqun Yang, Yiwei Li, **Huanchen Zhang**, and Mingyu Gao. “When Tree Meets Hash: Reducing Random Reads for Index Structures on Persistent Memories”. In: *Proceedings of the ACM on Management of Data (SIGMOD’23)*, 1.1: Article 105, 26 pages.
- [15] Zheng Chen, Feng Zhang, Jiawei Guan, Jidong Zhai, Xipeng Shen, **Huanchen Zhang**, Wentong Shu, and Xiaoyong Du. “CompressGraph: Efficient Parallel Graph Analytics with Rule-Based Compression”. In: *Proceedings of the ACM on Management of Data (SIGMOD’23)*, 1.1: Article 4, 31 pages.
- [16] Hokeun Cha, Xiangpeng Hao, Tianzheng Wang, **Huanchen Zhang**, Aditya Akella, and Xiangyao Yu. “Blink-hash: An Adaptive Hybrid Index for In-Memory Time-Series Databases”. In: *Proceedings of the VLDB Endowment (VLDB’23)*, 16.6: 1235-1248.
- [17] Ziwei Wang, Zheng Zhong, Jiarui Guo, Yuhan Wu, Haoyu Li, Tong Yang, Yaofeng Tu, **Huanchen Zhang**, and Bin Cui. “REncoder: A Space-Time Efficient Range Filter with Local Encoder”. In: *Proceedings of the 39th IEEE International Conference on Data Engineering (ICDE’23)*, April 2023, pp. 2036-2049.

- [18] Christoph Anneser, Andreas Kipf, **Huanchen Zhang**, Thomas Neumann, and Alfons Kemper. “Adaptive Hybrid Indexes”. In: *Proceedings of the 2022 ACM International Conference on Management of Data (SIGMOD’22)*. June 2022, pp. 1626-1639.
- [19] Eric R. Knorr, Baptiste Lemaire, Andrew Lim, Siqiang Luo, **Huanchen Zhang**, Stratos Idreos, Michael Mitzenmacher. “Proteus: A Self-Designing Range Filter”. In: *Proceedings of the 2022 ACM International Conference on Management of Data (SIGMOD’22)*. June 2022, pp. 1670-1684.
- [20] Ling Zhang, Matthew Butrovich, Tianyu Li, Yash Nannapanai, Andrew Pavlo, John Rollinson, **Huanchen Zhang**. “Everything is a Transaction: Unifying Logical Concurrency Control and Physical Data Structure Maintenance in Database Management Systems”. *Conference on Innovative Data Systems Research (CIDR’21)*. Jan. 2021.
- [21] **Huanchen Zhang**. “Memory-Efficient Search Trees for Database Management Systems”. *Ph.D. Thesis*. **ACM SIGMOD Jim Gray Dissertation Award**
- [22] **Huanchen Zhang**, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Succinct Range Filters”. *Communications of the ACM (CACM)*. 4 (2021): 166-173.
- [23] **Huanchen Zhang**, Lily Liu, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Order-Preserving Key Compression for In-Memory Search Trees”. In: *Proceedings of the 2020 ACM International Conference on Management of Data (SIGMOD’20)*. June 2020, pp. 1601-1615.
- [24] **Huanchen Zhang**, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Succinct Range Filters”. *ACM Transactions on Database Systems (TODS)*. 45.2 (2020): 1-31.
- [25] **Huanchen Zhang**, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Succinct Range Filters”. *ACM SIGMOD Record*, 48.1 (2019): 78-85.
- [26] **Huanchen Zhang**, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “SuRF: Practical Range Query Filtering with Fast Succinct Tries”. In: *Proceedings of the 2018 ACM International Conference on Management of Data (SIGMOD’18)*. June 2018, pp. 323-336. **Best Paper Award (1 out of 90 accepted papers)**
- [27] Ziqi Wang, Andrew Pavlo, Hyeontaek Lim, Viktor Leis, **Huanchen Zhang**, Michael Kaminsky, and David G. Andersen. “Building a Bw-Tree Takes More Than Just Buzz Words”. In: *Proceedings of the 2018 ACM International Conference on Management of Data (SIGMOD’18)*. June 2018, pp. 473-488.
- [28] **Huanchen Zhang**, David G. Andersen, Andrew Pavlo, Michael Kaminsky, Lin Ma, and Rui Shen. “Reducing the Storage Overhead of Main-Memory OLTP Databases with Hybrid Indexes”. In: *Proceedings of the 2016 International Conference on Management of Data (SIGMOD’16)*. June 2016, pp. 1567-1581.

Non Peer-Reviewed

- [29] **Huanchen Zhang**. “The End of the x86 Dominance in Databases?” Abstract. In: *Conference on Innovative Data Systems Research (CIDR’19)*. Jan. 2019.

Patents

- [30] **Huanchen Zhang** and Kimberly Keeton. “Data Storage over Immutable and Mutable Data Stages”. Filed Sept. 2017, Granted Sept. 2019.
- [31] **Huanchen Zhang** and Kimberly Keeton. “Changing Concurrency Control Modes”. Filed May 2017. Patent Pending.

Service

- SIGMOD Program Committee – 2023, 2022, 2021, 2020
- VLDB Program Committee – 2025, 2024, 2023
- ICDE Program Committee – 2022
- Jim Gray Award Committee – 2024, 2023, 2022
- DaMoN Program Committee – 2024
- AIDB Program Committee – 2021
- Journal Reviewer – TODS’24, SPE’24, TKDE’24, TKDE’23, TKDE’21, DISC’21, KAIS’18

Teaching

- Instructor – Database Systems (Tsinghua 40470414) – Spring 2024, Spring 2023, Fall 2021
- Instructor – Introduction to Programming (C/C++) (Tsinghua 30470332) – Fall 2023, Spring 2022, Fall 2022
- Instructor – Data Structures in the Real World (Tsinghua 20470102) – Summer 2021
- Head TA & Guest Lecturer – Advanced OS and Distributed Systems (CMU 15-712) – Fall 2017
- Head TA – Database Applications (CMU 15-415/615) – Fall 2016

Students

- Jiaoyi Zhang (Ph.D.)
- Junyi Zhao (Ph.D.)
- Yihao Liu (Ph.D.)
- Hengrui Wang (Ph.D.)
- Xinyu Zeng (Ph.D.)
- Jiansheng Qiu (Ph.D.)
- Yiming Qiao (Ph.D.)
- Yipeng Liu (Ph.D.)
- Yulong Hui (Ph.D.)
- Kai Su (Ph.D.)
- Jiahong Shen (Ph.D. incoming)
- Weitao Wan (M.S. w/ Mingyu Gao)
- Fangzhou Yuan (M.S.)

Selected Talks

- “Cost-Intelligent Data Analytics in the Cloud.”
 - CIDR Conference Talk, Jan. 2024
 - WAIC General Artificial Intelligence Forum, July 2023
 - SiftDB, July 2023
 - DaMoN @ SIGMOD, June 2023
 - Great Bay Area Digital Tech Workshop, June 2023
 - Data Technology Carnival, April 2023
 - Northeastern University (China), February 2023
- “Cloud Data Warehouses: Snowflake and Beyond.”
 - ByteDance, April 2021
 - Alibaba Cloud, April 2021
 - CCF Forum, January 2021
- “Order-Preserving Key Compression for In-Memory Search Trees.”
 - SIGMOD’20 Conference Talk, June 2020
 - University of Chicago Database Group Seminar, June 2020
 - Snowflake Inc, May 2020
- “Memory-Efficient Search Trees for Database Management Systems.”
 - SIGMOD’21 Jim Gray Award Talk, June 2021
 - Snowflake Inc, January 2020
 - Salesforce, January 2020

- Tsinghua University IIS, October 2019
- “SuRF: Practical Range Query Filtering with Fast Succinct Tries.”
 - CMU Network Group Seminar, November 2018
 - Parallel Data Lab Annual Retreat, October 2018
 - SIGMOD’18 Conference Talk, June 2018
 - CMU Parallel Data Lab Seminar, May 2018
 - CMU Database Group Seminar, May 2018
- “Succinct Trie Indexes Made Practical.”
 - CMU Database Group Seminar, February 2017
- “Distributed Metadata Store for RePO”
 - HP Labs, August 2016
- “Reducing the Storage Overhead of Main-Memory OLTP Databases with Hybrid Indexes.”
 - Parallel Data Lab Annual Retreat, October 2016
 - SIGMOD’16 Conference Talk, June 2016
 - UC-Berkeley Database Group Seminar, June 2016
 - HP Labs, May 2016
 - Parallel Data Lab Annual Retreat, October 2015