Minguk Choi

mgchoi@dankook.ac.kr | min-guk.github.io | linkedin.com/in/mingukchoi

Research Interest

Past: Linux Kernel, Key-Value Store
Present: Learned Data Structures
Future: *ML for Systems* - Learned Cache, Learned Query Optimizer
Systems for ML - Scheduler/Pipeline for ML Cluster

On-going Project

Semi-ordered Parallel Updatable Learned Index: Enhances performance in multi-threaded environments by reducing the critical section and data sorting overhead while maintaining lookup performance and size.

Adaptive Sample Training Algorithm for Learned Index: An algorithm for dynamically training with the optimal sampling interval by considering the data distribution and error-bound.

Accelerating Learned Index via SIMD: Accelerating the unique training and operations (e.g., model-biased insert) of learned index through the parallelism of internal ML models, which has not yet been actively studied.

Education

Dankook University — Yongin, South Korea
M.S. in AI-based Convergence — Mar 2023 – Aug 2024

Dankook University — Yongin, South Korea
B.S. in Software Science , **Salutatorian** (4.1/4.5) — Mar 2017 – Feb 2023

Publications

International Conference

(SIGMOD 2024) Can Learned Indexes be Built Efficiently? A Deep Dive into Sampling Trade-Offs
Minguk Choi, Seehwan Yoo, Jongmoo Choi

International Journal

(Electronics 2023) An Empirical Study of Segmented Linear Regression Search in LevelDB
Ramadhan Agung Rahmat, Minguk Choi, Yoojin Chung, Jongmoo Choi.

Domestic Conference

(KCC 2024) Analysis of RMI Using CPU-Optimized Search Algorithms ***Best Paper Award***
Yeojin Oh, Minguk Choi, Boseung Kim, Yongjie Zhu, Seehwan Yoo, Jongmoo Choi

(KCC 2024) Accelerating RMI Training with SIMD
Boseung Kim, Minguk Choi, Yeojin Oh, Yongjie Zhu, Seehwan Yoo, Jongmoo Choi

(KCC 2024) Performance Analysis of Batch Prediction Using SIMD in RMI
Yongjie Zhu, Minguk Choi, Yeojin Oh, Boseung Kim, Seehwan Yoo, Jongmoo Choi

(KCC 2024) Breakdown Internal Operations in Updatable Learned Index
Suhwan Shin, Minguk Choi, Nakyeong Kim, Seehwan Yoo, Jongmoo Choi

(KCC 2024) Analysis of Updatable Learned Indexes with Index Size Perspective
Nakyeong Kim, Minguk Choi, Suhwan Shin, Seehwan Yoo, Jongmoo Choi

(KSC 2022) Bloom Filter Optimization in LevelDB based on Hit-Ratio
Hansu Kim, Minguk Choi, Seehwan Yoo, Jongmoo Choi

(KSC 2022) Read performance analysis according to Compaction Trigger
Zhao Guangxun, Sangwoo Kang, Minguk Choi, Seehwan Yoo, Jongmoo Choi

(KSC 2022) LevelDB Cache Structure and Performance Analysis
Subin Hong, Minguk Choi, Seehwan Yoo, Jongmoo Choi

(KSC 2022) Per Key-Value Checksum Analysis on RocksDB
Suhwan Shin, Seyeon Park, Minguk Choi, Seehwan Yoo, Jongmoo Choi

Experience

Student Researcher — Dankook University
System Software Laboratory — Jun 2021 – Present
- Linux Kernel (CPU Scheduler, Memory Allocator, Block I/O Layer)
- Key-Value Storage (LevelDB, RocksDB)
- Learned Index Structure (Read-only/Updatable, Sampled Learning, Semi-ordered, SIMD)

Teaching Assitant — Dankook University
Operating Systems — Spring 2024
- Assignment 1: CPU scheduler simulator (8-type scheduler, various workloads, auto-graded through gtest)
- Assignment 2: Concurrent data structure (Queue/BST, various workloads, auto-graded through gtest)
- Assignment 3: Ext2 file forensics

Index Structure Study — Winter 2024
- Lecture: Introduction on (Learned) Index Structures
- Research Guidance: 5 domestic conference papers, 1 best paper award

Operating System Practice — Summer 2023
- Lecture: 40 hours
- Set and grade final-exam questions
- Practice: Sloppy Counter, Dining Philosophers, Readers-Writers, and myShell

System Programming — Fall 2023
- Assignments: myLs, myCreate, myCat, myCopy, and myShell

LevelDB Study — Summer 2022
- Lecture: Introduction on LevelDB (Key-Value Store)
- Research Guidance: 4 domestic conference papers, tuning contest (YCSB), open-source doucument

Compulsory Military Service — Korean Air Force
Honorable Discharge — Jan 2019 – Oct 2020

Opensource Contribution

BASIL — Owner
Benchmark of sampling applied learned indexes (To be released)

Leveldb Study/Wiki — Owner
Analysis document about LevelDB Key-Value Store

YCSB-CPP — Contributer
Yahoo! Cloud Serving Benchmark(YCSB) written in C++

Uftrace — Wiki Contributer
Function graph tracer for C/C++/Rust/Python

PGM-Index — Bug Report
State-of-the-art learned data structure

Research Projects

Future Vehicle Security System, KEA — May 2023 - Present
Key-Value Store with Predictable Latency Support, NRF — Mar 2022 - Present
(SW StarLab) Key-Value DB for Unstructured Big Data , IITP — Jun 2021 – Present

Language and Technical Skills

Languages: Native korean, Englsih (TOEIC 970/990, OPIc IH)
Programming Languages: c/c++, python
Developer Tools: gcc/g++, gdb, uftrace, git, matplotlib

Honors and Awards

Best Assitant Award, Dankook University — Jul 2024
Best Paper Award, KCC (Korea Computuer Congress) 2024 — Jun 2024
Certificate of Appreciation, KCC (Korea Computuer Congress) 2024 — Jun 2024
Complimentary Professional Membership, ACM (Association for Computing Machinery) — Jun 2024

Personal interests

Fitness: Running, Weight lifting
Cooking: Korean Cuisine, Japanese Cuisine, Grilling

References

Prof. Jongmoo Choi choijm@dankook.ac.kr — Software Science, Dankook University
Prof. Seehwan Yoo seehwan.yoo@dankook.ac.kr — Mobile Systems Engineering, Dankook University