

# Minguk Choi

mgchoi@dankook.ac.kr | Homepage | LinkedIn | Github

## RESEARCH INTEREST

---

**Past:** Linux Kernel, Key-Value Store

**Present:** Learned Index Structures

**Future:** *ML for Systems* - Learned Data Structures, Learned Query Optimizer, Learned Cache  
*Systems for ML* - Scheduler/Pipeline for ML, Flash(CXL)-LLM

## 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**  
M.S. in AI-based Convergence

Yongin, South Korea  
Mar 2023 – Aug 2024

**Dankook University**  
B.S. in Software Science , **Salutatorian** (4.1/4.5)

Yongin, South Korea  
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**)  
Yejin 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

System Software Laboratory

June 2021 – Present  
Yongin, Korea

- Linux Kernel (CPU Scheduler, Memory Allocator, Block I/O Layer)
- Key-Value Storage (LevelDB, RocksDB)
- Learned Index (Read-only/Updatable, Sampled Learning, Semi-ordered, SIMD)

### Teaching Assistant

Operating System Practice

Dankook University  
Sep 2023

- Conduct 40 hours of lectures on system programming and operating systems
- Create a various practice (Sloppy Counter, Dining Philosophers Problem, Readers-Writers problem, myShell)
- Setting and grading final-exam questions

Operating Systems

Mar 2024 – Jul 2024

- Create an CPU scheduler simulator assignment (8-type scheduler, various workloads, auto-graded through gtest)
- Create a concurrent data structure assignment (Queue/BST, various workloads, auto-graded through gtest)
- Give a ext2 file forensics assignment

System Programming

Sep 2023 – Dec 2023

- Give a myls, myCreate, myCat, myCopy, and myShell assignments

Index Structure Study

Jan 2024 – Feb 2024

- Conduct an introductory lecture on (Learned) Index Structures
- Guidance for student research (5 domestic conference papers, 1 best paper award)

LevelDB Study

Jul 2022 – Aug 2022

- Conduct an introductory lecture on (Learned) Index Structures
- Hosting a LevelDB tuning contest using YCSB
- Guidance for student research (4 domestic conference papers, open-source analysis document)

### Compulsory Military Service

Honorable Discharge

Jan 2019 – Oct 2020  
Korean Air Force

## OPENSOURCE CONTRIBUTION

---

### BASIL

Benchmark of sampling applied learned indexes (To be released)

Owner

### Leveldb Study/Wiki

Analysis document about LevelDB Key-Value Store

Owner

### YCSB-CPP

Yahoo! Cloud Serving Benchmark(YCSB) written in C++

Contributer

### Uftrace

Function graph tracer for C/C++/Rust/Python

Wiki Contributor

### PGM-Index

State-of-the-art learned data structure

Bug Report

## PROJECTS

---

Key-Value DB for Unstructured Big Data (SW StarLab)

IITP  
June 2021 – Present

Key-Value Store with Predictable Latency Support

NRF  
Mar 2022 - Present

Future Vehicle Security System

KEA  
May 2023 - Present

## HONORS AND AWARDS

---

Best Paper Award, Korea Computer Congress

## LANGUAGE AND TECHNICAL SKILLS

---

**Languages:** Native korean, English (TOEIC 970/990, OPIc IH)

**Programming Languages:** c/c++, python

**Developer Tools:** gcc/g++, gdb, uftrace, git, matplotlib

## PERSONAL INTERESTS

---

**Fitness:** Running, Weight lifting

**Cooking:** Korean Cuisine, Japanese Cuisine, Grilling

## REFERENCES

---

Prof. Jongmoo Choi [choijm@dankook.ac.kr](mailto:choijm@dankook.ac.kr)

Software Science, Dankook University

Prof. Seehwan Yoo [seehwan.yoo@dankook.ac.kr](mailto:seehwan.yoo@dankook.ac.kr)

Mobile Systems Engineering, Dankook University