Sooyoung Lim

https://sooyounglim.github.io

Email: slim@keti.re.kr
Mobile: +82-10-9389-7541

Research Interests

Storage Systems, Operating Systems, Distributed Systems, Systems for ML

Education

- **Sookmyung Women's University** — Seoul, Republic of Korea
  Master of Engineering in Computer Science (GPA: 4.44/4.5) — *Mar 2022 – Feb 2024*
- **Sookmyung Women's University** — Seoul, Republic of Korea
  Bachelor of Engineering in Software Convergence (GPA: 3.92/4.5) — *Mar 2017 – Feb 2022*

Publications

International Journal

- Automatic Internal Parallelism Reconfiguration on Heterogeneous Low-Power Hadoop Clusters *[preprint]*
  Sooyoung Lim and Dongchul Park
  Future Generation Computer Systems, 2025 (*under review*)
- Improving Hadoop MapReduce performance on heterogeneous single board computer clusters *[doi]*
  Sooyoung Lim and Dongchul Park
  Future Generation Computer Systems, 2024
- Efficient Stack Distance Approximation Based on Workload Characteristics *[doi]*
  Sooyoung Lim and Dongchul Park
  IEEE Access, 2022

International Conference

- Toward Heterogeneity-Aware Striping in Lustre
  Sooyoung Lim, Jaegi Son and Dongmin Kim
  IEEE ICTC, Jeju Island, Republic of Korea, 2025 (*accepted*)

Work Experience

- **Korea Electronics Technology Institute (KETI)** — Seong-nam, Gyeong-gi, Republic of Korea
  Researcher — *May 2025 – Present*
  - **I/O Optimization for ML Frameworks on Object Storage Systems**
    Offloading multimodal data preprocessing operations to storage engines in DAOS
  - **I/O Optimization for Distributed File Systems on Heterogeneous Storage Systems**
    Implementing a resource-aware striping mechanism in Lustre

Research Experience

- **Chung-Ang University** — Seoul, Republic of Korea
  Research Assistant — *Jun 2024 – Feb 2025*
  - **Auto-Tuning for Resource Scheduling on Heterogeneous Low-Power Clusters**
    Developed a novel Hadoop YARN by implementing a resource-aware tuning method based on the node-level parallelism
- **Sookmyung Women's University** — Seoul, Republic of Korea
  Research Assistant — *Jan 2022 – Feb 2024*
  - **Big Data Processing on Heterogeneous Low-Power Clusters**
    Developed a novel Hadoop YARN by implementing two dual-mode scheduling strategies and a MapReduce task placement policy
  - **Data Access Pattern Profiling for Cache Simulation**
    Designed a workload-aware stack distance approximation algorithm for efficient cache behavior simulation

Teaching Experience

- **Sookmyung Women's University** Seoul, Republic of Korea
  Teaching Assistant
  - Linux System *Fall 2022, Fall 2023*
  - Data Structures *Spring 2023*
  - Introduction to Programming *Spring 2023*
  - Big Data Processing *Fall 2022*

Programming Projects (selected)

- **Chatbot for Clothing Recommendation** *[github]* *Sep 2020 – Mar 2021*
  Implemented a CNN-based recommendation model on the customized web-based chatbot service and integrated YOLACT for CD/CI pipelines
- **Application for Discovering Nearby Discounted Expiring Foods** *[github]* *Nov 2019 – Jan 2020*
  Developed an Android application that alerts users to nearby discounted food items nearing expiration, integrating real-time backend updates and location-based filtering via RESTful API communication
- **Application for Voting on Member of Parliament** *[github]* *Jun 2018 – Jul 2018*
  Built Android frontend interfaces by converting design mockups into XML layouts and custom view components

Patents

- Method and Device for Allocating MapReduce Task in Heterogeneous Cluster Environment
  Dongchul Park and Sooyoung Lim
  Korea Patent, 10-2025-0008411, 2025

Honors and Awards

- **Excellent Alumna Scholarship**, Sookmyung Women's University (*full scholarship*) *2022 – 2023*
- **Research Support Scholarship**, Sookmyung Women's University *2022*
- **2nd prize, Sookmyung Hackathon powered by AWS**, Sookmyung Women's University *2019*
- **Scholarship**, Sookmyung Women's University *2017, 2019, 2020, 2021*

Technical Skills

- **Proficient** C/C++, Python, Java, UNIX Shell, LaTeX, Git, SQL, R
- **Familiar** QEMU, Go, Kotlin

Language Proficiency

- **TOEFL iBT** Total 104 (Reading 28, Listening 22, Speaking 26, Writing 28)