

JIHEON CHOI

Contact Information

Ajou University | Depts. of Artificial Intelligence, @ Distributed and Parallel Computing Lab (Lab. <https://wise.ajou.ac.kr>)
206, Word cup-ro, Yeongtong-Gu, Suwon, Gyeonggi-do, 16499 Republic of Korea | E-mail: unidev@ajou.ac.kr

Summary

A security-focused researcher and software engineer with 5+ years of experience spanning academic research and industry applications. Specializes in optimizing high-performance computing systems, developing secure distributed machine learning architectures, and implementing scalable cloud-based solutions. My research interests include, but not limited to the following areas.

- **High-Performance & Distributed Computing:** HPC Systems Optimization and Job Scheduling
- **AI & Machine Learning:** Distributed ML Training, Federated Learning and Privacy-Preserving ML Techniques
- **Security & Privacy:** Differential Privacy and Homomorphic Encryption, Automatic Security Vulnerability Assessment, Secure ML Systems Design

Professional Experience

■ Graduate Research Assistant

@ Ajou University, Dept. of Artificial Intelligence (Sep. 2023 – now) *Undergraduate RA. Mar. 2022 – Aug. 2023
Suwon, Republic of Korea

Distributed and Parallel Computing Lab. (formerly WISE Lab)

- Laboratory Manager: Oversaw overall laboratory operations and management. Mentored and managed team members with handling research planning and human resource allocation. Established and maintained relationships with external partner organizations.

■ Freelance Security Engineer, FIG Inc. (Oct 2023 ~ Mar 2025)

- Performed personal information processing agent inspection for a marketing agency, implementing all extensive enterprise security requirements while building new AWS architecture and migrating services
- Leveraged enterprise client's transit gateway connections to meet security requirements without expensive solutions—achieving significant cost savings beyond what development teams could accomplish alone

■ Lead Software and Security Research Engineer

MONTHLY HACKING Inc., Seoul, Republic of Korea (Aug. 2020 – Dec. 2023)

- Co-founded the company with a non-technical partner and directed technical R&D division, pioneering South Korea's first subscription-based penetration testing SaaS platform, securing investment through successful hacking demonstrations in real-world scenarios.
- Engineered a REST API for an automated penetration testing SaaS platform that transformed manual security processes into scalable solutions, enabling protection of enterprise-level systems with startup-sized security teams
- Delivered high-impact security consulting to well-known domestic startups (approximately 50+), uncovering critical vulnerabilities through penetration testing and providing rapid incident response during security breaches

■ Software Engineer

WAPAS SYSTEMS Inc., Incheon, Republic of Korea (Oct. 2019 – Aug. 2020)

- Implemented a web application for a Philippine water shop that included both a customer-facing ordering platform using Node.js and React framework on the AWS platform
- Led the stakeholder communication to clarify business requirements during the service operation that streamlined their water delivery service

Education

■ M.S. & Ph.D. in Artificial Intelligence

Ajou University, Suwon, Republic of Korea (Sep. 2023 – now)

- Advisor: Prof. Sangyoon Oh (Ajou University)
- Dissertation: TBD

■ B.Sc. in Software and Computer Engineering

Ajou University, Suwon, Republic of Korea (Mar. 2019 – Aug. 2023)

Research Project

- A Study on Graph Parallel Processing Algorithm Based on GPU (June 2024 – now)
funded by KISTI (Korea Institute of Science and Technology Information), NST (National Research Council of Science and Technology)
 - Optimized GPU-based breadth-first search algorithms through level-by-level execution techniques. Researched sparse matrix multiplication algorithms specifically tailored for GPU memory architecture.
 - Developed efficient graph partitioning methods to enhance data movement in GPU-based GNN implementations, effectively addressing memory bottlenecks in neural network training.
- Development of multi-tiered, multi-purpose autotuning framework for exa-scale supercomputers (Nov 2023 – now) *funded by Ministry of Science and ICT*
 - Collaborated with KISTI Supercomputing Center to analyze job logs from 5th generation supercomputers.
 - Implemented ML-based prediction models for scheduling algorithms to optimize resource utilization.
 - Enhanced execution time through parameter autotuning and applied active learning to predict queued job applications with minimal data requirements.
- Development of High Efficiency HPC Job Scheduling Algorithm (Jan 2023 – Feb 2025)
funded by Samsung Display Inc.
 - Analyzed topology-level computing resources to identify critical scheduling factors, and leveraged machine learning and data mining on simulation job logs to determine key scheduling parameters
 - Developed an efficient HPC scheduling algorithm that improved data center resource utilization from 72% to 93%. Led migration from IBM LSF to SLURM with custom web interface and API implementation
 - Validated on 16-node cluster before successful deployment to 300-node production environment
- Development of Safety of the intended functionality from insufficiency of perception and decision making (Sep 2022 – now) *funded by Korea Automotive Technology Institute*
 - Designed and implemented a Node.js-based server architecture for processing – 100TB of autonomous driving sensor data, with an efficient compression interface that reliably delivers high-volume data by numerous small files into manageable packages within the data processing workflow
- Development of Web Application Vulnerability Assessment Software (Apr 2022 ~ Dec 2022)
funded by Ministry of SMEs and Startups
 - Led R&D proposal writing and project management from concept through product design and project launch.
 - Successfully passed performance evaluation by authorized testing institution to verify achievement of target metrics specified in the proposal.
 - Transformed automated vulnerability checker tools (e.g., automated Log4J vulnerability detection) into SaaS offerings by implementing Python Flask API connected to Spring Boot backend services.

Selected Publications and Patents

If you would like to review full publications (12+) and patents (4+), please see the personal homepage: <https://jiheon.dev>

- [1] **Jiheon Choi**, Jaehyun Lee, Minsol Choo, Taeyoung Yoon, Oh-Kyoung Kwon, Sangyoon Oh. When HPC Scheduling Meets Active Learning: Maximizing The Performance with Minimal Data, *High Performance Computing in Asia-Pacific Region (HPC Asia)*, 2025.
- [2] Miri Yu, **Jiheon Choi**, Jaehyun Lee, Sangyoon Oh. Staleness Aware Semi-asynchronous Federated Learning, *Journal of Parallel and Distributed Computing*, IF: 3.4, JCR Q1. (2024)
- [3] **Jiheon Choi** and Sangyoon Oh “Method and Apparatus for Classifying Workload Using Artificial Intelligence Model” South Korea 10-2025-0029398 (Applied)
- [4] **Jiheon Choi** and Sangyoon Oh “Apparatus and Method for Hotspot Prevention Based on Bloom Filter in Distributed Database Management System” South Korea 10-2024-0162896 (Applied)

Skills

- Programming Language – C, Java, JavaScript, Python
- Framework & Libraries – Node.js, React, Next.js, Vue.js, Nest.js, Spring Boot, Python Flask, PyTorch, Tensorflow
- Web Servers – Nginx, Apache Tomcat, Docker, docker-compose
- Databases – MySQL, MongoDB, AWS DynamoDB, SQLite, NoSQL