SEONGMIN LEE

University of California, Los Angeles (UCLA)
Engineering VI, Room 474
404 Westwood Plaza
Los Angeles, CA 90095
United States

📞 +1 (310) 806 7239   ✉ seongminlee@sigsoft.org   🏛 [Google Scholar](#)

Research Summary

The overarching objective of my research is to *ensure software correctness in complex and critical modern software systems* through **scalable program analysis** and **reliable software testing**. To achieve this, I **leverage interdisciplinary statistical methods**, such as *causal inference, biostatistics, and machine learning*, to analyze the dynamic behavior of software in operational environments. My work has been published in top-tier software engineering venues, including ICSE, FSE, and JSS, and I have served as a program committee member for leading conferences, such as ICSE, FSE, and ASE.

Employment

University of California, Los Angeles	United States
Postdoctoral Researcher, Software Evolution and Analysis Laboratory (SEAL)	Jun. 2025 – Present
Group head: Prof. Miryung Kim	
Max Planck Institute for Security and Privacy	Germany
Postdoctoral Researcher, Software Security Research group	Sep. 2022 – Jun. 2025
Group head: Dr. Marcel Böhme	

Education

Korea Advanced Institute of Science and Technology	Republic of Korea
Doctor of Philosophy, School of Computing	Sep. 2016 – Aug. 2022
Advisor: Prof. Shin Yoo	
Bachelor of Science, School of Computing	Feb. 2012 – Aug. 2016
Bachelor of Science, Department of Mathematical Sciences	

Publications *Each marker text contains the hyperlink to the corresponding paper.

Refereed Journal Articles

SCP'25 Seongmin Lee, Dave Binkley, Robert Feldt, Nicolas Gold, and Shin Yoo. Causal program dependence analysis. *Science of Computer Programming*, 240:103208, February 2025

JSS'21 Seongmin Lee, David Binkley, Robert Feldt, Nicolas Gold, and Shin Yoo. Observation-based approximate dependency modeling and its use for program slicing. *Journal of Systems and Software*, 179:110988, September 2021

JSS'20 Seongmin Lee, David Binkley, Nicolas Gold, Syed Islam, Jens Krinke, and Shin Yoo. Evaluating lexical approximation of program dependence. *Journal of Systems and Software*, 160:110459, February 2020

Refereed Conference Publications

ICLR'25 Seongmin Lee and Marcel Böhme. How Much is Unseen Depends Chiefly on Information About the Seen. In *Proceedings of the 13th International Conference on Learning Representations (ICLR'25)*, May 2025
Spotlight

ICSE'25 Seongmin Lee, Shreyas Minocha, and Marcel Böhme. Accounting for Missing Events in Statistical Information Leakage Analysis. In *Proceedings of the IEEE/ACM 47th International Conference on Software Engineering (ICSE'25)*. Association for Computing Machinery, June 2025

ICSE'24 Danushka Liyanage, Seongmin Lee, Chakkrit Tantithamthavorn, and Marcel Böhme. Extrapolating Coverage Rate in Greybox Fuzzing. In *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering (ICSE'24)*, ICSE '24, pages 1–12, New York, NY, USA, April 2024. Association for Computing Machinery
(*Co-first authors with equal contribution*)

FSE'23 Seongmin Lee and Marcel Böhme. Statistical Reachability Analysis. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE'23)*, ESEC/FSE 2023, pages 326–337, New York, NY, USA, November 2023. Association for Computing Machinery

SCAM'19 Seongmin Lee, David Binkley, Robert Feldt, Nicolas Gold, and Shin Yoo. MOAD: Modeling Observation-Based Approximate Dependency. In *2019 19th International Working Conference on Source Code Analysis and Manipulation (SCAM'19)*, pages 12–22, September 2019

ICST'19 Seongmin Lee, Shin Hong, Jungbae Yi, Taeksu Kim, Chul-Joo Kim, and Shin Yoo. Classifying False Positive Static Checker Alarms in Continuous Integration Using Convolutional Neural Networks. In *2019 12th IEEE Conference on Software Testing, Validation and Verification (ICST'19)*, pages 391–401, April 2019

Preprints

PP'25 Jing Liu, Seongmin Lee, Eleonora Losiouk, and Marcel Böhme. Can LLM Generate Regression Tests for Software Commits?, 2025
(**Co-first authors with equal contribution*)

Invited Articles

GI'20 William B. Langdon, Westley Weimer, Justyna Petke, Erik Fredericks, Seongmin Lee, Emily Winter, Michail Basios, Myra B. Cohen, Aymeric Blot, Markus Wagner, Bobby R. Bruce, Shin Yoo, Simos Gerasimou, Oliver Krauss, Yu Huang, and Michael Gerten. Genetic Improvement @ ICSE 2020. *SIGSOFT Softw. Eng. Notes*, 45(4):24–30, October 2020

Refereed Workshop Publications

ICST'21 Saeyoon Oh, Seongmin Lee, and Shin Yoo. Effectively Sampling Higher Order Mutants Using Causal Effect. In *2021 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 19–24, April 2021

ICSE'20 Seongmin Lee. Scalable and approximate program dependence analysis. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Companion Proceedings*, ICSE '20, pages 162–165, New York, NY, USA, October 2020. Association for Computing Machinery

ICSE'18 Seongmin Lee, David Binkley, Nicolas Gold, Syed Islam, Jens Krinke, and Shin Yoo. MOBS: Multi-operator observation-based slicing using lexical approximation of program dependence. In *Proceedings of the 40th International Conference on Software Engineering: Companion Proceeedings*, ICSE '18, pages 302–303, New York, NY, USA, May 2018. Association for Computing Machinery

KCC'17 Gabin An, Jinhan Kim, Seongmin Lee, and Shin Yoo. PYGGI: Python General Framework for Genetic Improvement. *Journal of Korean Institute of Information Scientists and Engineers*, pages 536–538, December 2017

SBSE'17 Seongmin Lee and Shin Yoo. Hyperheuristic Observation Based Slicing of Guava. In Tim Menzies and Justyna Petke, editors, *Search Based Software Engineering*, pages 175–180, Cham, 2017. Springer International Publishing

SBSE'16 Jeongju Sohn, Seongmin Lee, and Shin Yoo. Amortised Deep Parameter Optimisation of GPGPU Work Group Size for OpenCV. In Federica Sarro and Kalyanmoy Deb, editors, *Search Based Software Engineering*, pages 211–217, Cham, 2016. Springer International Publishing

Grants and Fellowships

- Title: *Statistical Security Analysis for Large, Evolving Software*
  Funding Agency: CASA - Cyber Security in the Age of Large-Scale Adversaries
  Grant ID: DFG under Germany's Excellence Strategy - **EXC 2092 CASA - 390781972**
  Amount: Salary according to the remuneration group E 14 TV-L (full time, $\sim$ €136,000)
  Duration: 2024.01.01 – 2025.12.31
  Role: Sole Principal Investigator (PI)

Awards and Honors

- **Spotlight Paper Award**, 13th International Conference on Learning Representations (ICLR'25), 2025
  - *Title of Paper: How Much is Unseen Depends Chiefly on Information About the Seen*

- **Distinguished Artifact Reviewer Award**, 33rd USENIX Security Symposium, 2024

- **PhD Dissertation Award**, School of Computing, KAIST, 2022
  - *Title of Dissertation: Statistical Program Dependence Approximation*

- **2021 Naver Ph.D. Fellowship Award**: Awarded by NAVER Corp. to Ph.D. candidates who have published an outstanding research paper or have excellent publication performance, 2021

- Government-sponsored Scholarship, Ministry of Science and ICT of Korea, 2016 - 2022

- Government-sponsored Scholarship, Ministry of Science and ICT of Korea, 2012 - 2016

Service

Academic Service
- Program committee: (Main Track) FSE '26, ICSE'26, QUATIC'25, FUZZING'25, LLanMER'25, CauSE'25, ASE'24, ISSTA'24, FUZZING'24, SCAM'24, ASE'23 / (Artifact Evaluation Track) ISSTA'24, ECOOP'24, USENIX Security'24, ICSE'24, ISSTA'23, ICSME'22, ICSME'21 / (Student Research Competition Track) FSE'24 / (Tool Demonstration Track) ISSTA'25, ASE'24

- Reviewer: TOSEM'25, TOSEM'24, TSE'24, IST'24, ASE'24, TOSEM'22, JSS'21, JSS'20 / (External) ICSE'24, FSE'24, ECOOP'24, ICSE'23, ISSTA'23

Institutional Service
- Early Career Researcher (ECR) Board, CASA - Cyber Security in the Age of Large-Scale Adversaries   2024 – 2025

  The ECR Board decides all funding allocations for the CASA Graduate School (CGS). The board has its own budget and the autonomy to dynamically adapt the programs to the changing needs of ECR members. Representing ECRs in CASA governance, the board organizes activities and meets quarterly to address priorities.

- Open Science Ambassador, Max Planck Institute for Security and Privacy (MPI-SP)   2023 – 2025
  As representatives of the institute, Open Science Ambassadors raise awareness and provide valuable information to their colleagues about Open Science practices. They meet annually, in person or online, to discuss strategies, participate in workshops, and evaluate the status of Open Science within and beyond the Society.

Advising Experience

Ph.D. Students
- **Danushka Liyanage** (Monash University, moved on to University of Sydney Postdoc, 2024)   Nov. 2022 – Dec. 2023
  Co-advised with Dr. Marcel Böhme on extrapolating the coverage rate of the Greybox Fuzzing using the statistical model. A **full conference paper** where Danushka is the first author has been accepted to ICSE'24.

Undergraduate Students
- **Jing Liu** (Shanghai University, moved on to MPI-SP Ph.D. student, 2025)   Apr. 2024 – Present
  Co-advised with Dr. Marcel Böhme on the LLM-based regression test generation for the software testing. A **full conference paper** where Jing is the first author has been submitted to ICSE'26.

- **Shreyas Minocha** (Rice University, moved on to Georgia Tech Ph.D. student, 2024)   Feb. 2023 – Jul. 2024
  Primary advisor on the statistical information leakage analysis to suggest the accurate and safe information leakage estimator. A **full conference paper** has been accepted to ICSE'25.

- **Saeyoon Oh** (KAIST, moved on to FuriosaAI)   Jul. 2020 – Apr. 2021
  Co-advised with Dr. Shin Yoo on the effective sampling of higher-order mutants for mutation testing. A **workshop paper** where Saeyoon is the first author has been accepted to ICST'21.

Teaching Experience

Invited Talks
- **Invited Tutorial: "The Magic of Statistics for Software Testing: How to Foresee the Unseen,"**
  18th International Workshop on Search-Based and Fuzz Testing (SBFT'25)   Apr. 2025

- *Open Peer Review in ICLR'25*, Open Science Ambassador Meeting, Max Planck Society   Feb. 2025

- *Guarantees in Software Security – 2nd part: Extrapolating Software Testing*,
  Fuzzing and Software Security Summer School @ National University of Singapore (NUS)   May 2024

- *Magic of Statistics for Software Testing: How to Foreseen the Unseens*, Korea University   Jan. 2025
- *Statistical Program Analysis*, Korea Advanced Institute of Science and Technology (KAIST)   Jan. 2024
- *Statistical Program Analysis*, Ulsan National Institute of Science and Technology (UNIST)   Jan. 2024
- *Causal Program Dependence Analysis*, Sheffield Causality and Testing Workshop   Sep. 2023
- *Statistical program dependence analysis*, Handong Global University   Aug. 2022
- *Observation-based approximate dependency modeling and its use for program slicing*,
  Korea Conference on Software Engineering   Jan. 2022
- *MOBS: Multi-Operator Observation-Based Slicing using Lexical Approximation of Program Dependence*,
  59th CREST Open Workshop – Multi-language Software Analysis   Mar. 2018

Lecturer
- Advanced Automatic Testing (Co-Lecturing with Dr. Flavio Toffalini and Dr. Yannic Noller), Faculty of Computer Science, Ruhr University Bochum   Spring 2025
- MPI Reading Group on Software Security (Co-Lecturing with Dr. Marcel Böhme and Ardi Madadi), Faculty of Computer Science, Ruhr University Bochum   Spring 2025

Teaching Assistant
- Automated Software Testing (CS453), School of Computing (SoC), KAIST   Spring 2019
- Artificial Intelligence Based Software Engineering (CS454), SoC, KAIST   Fall 2018
- Introduction to Logic for Computer Science (CS402), SoC, KAIST   Spring 2018
- Artificial Intelligence Based Software Engineering (CS454), SoC, KAIST   Fall 2017
- Introduction to Logic for Computer Science (CS402), SoC, KAIST   Spring 2017
- Special Topics in Computer Science ⟨Search Based Software Engineering⟩ (CS492), SoC, KAIST   Fall 2016

Research Experience

Software Security Group, MPI-SP   **Sep. 2022 – Present**
Postdoc   *Bochum, Germany*
- Working on structure-aware residual risk estimation to achieve more accurate and efficient software security testing – *Structure-aware Residual Risk*
- Investigating the reliability of coverage- and bug-based fuzzer benchmarking, focusing on the internal consistency of results and agreement across fuzzers/metrics – *Fuzzer Benchmarking Internal Consistency*
- Worked on unbiased estimation of the missing mass, probability, or expected number of newly discovered classes in an unknown multinomial distribution
- Worked on LLM-based regression test generation for software security testing – *Cleverest*
- Worked on applying biostatistics to information leakage analysis to suggest an accurate and safe information leakage estimator – *Statistical Information Leakage Analysis*
- Worked on extrapolating the coverage rate of Greybox Fuzzing using a statistical model – *Greybox Fuzzing Extrapolation*
- Worked on applying statistical methods to program analysis to overcome the scalability issues of static analysis – *Statistical Reachability Analysis*

Computational Intelligence for Software Engineering Laboratory (COINSE), KAIST   **Sep. 2016 – Aug. 2022**
Ph.D. Student   *Daejeon, Republic of Korea*
- Worked on approximating the degree of dependence between program elements using causal inference – *CPDA*
- Worked on applying statistical models to observational data to approximate program dependence – *MOAD*
- Worked on inferring type information in binary executables using RNNs in collaboration with the National Security Research Institute
- Worked on classifying false-positive alarms from a static checker in a continuous integration pipeline using CNNs in collaboration with Samsung Research
- Worked on program dependence approximation using a lexical model on source code – *MOBS*

Computational Intelligence for Software Engineering Laboratory (COINSE), KAIST   **Mar. 2016 – Aug. 2016**
Undergraduate Research Intern   *Daejeon, Republic of Korea*
- Worked on the amortized deep parameter optimization of GPGPU workgroup sizes for OpenCV.
- Accelerated the scalablility of Observation based slicing (ORBS) by applying a code distance metric during the slicing.

Programming Language Research Group (PLRG) Lab, KAIST   **Jul. 2015 – Feb. 2016**
Undergraduate Research Intern   *Daejeon, Republic of Korea*
- Developed a source code translator from C# to C++ with F#.
- Developed a frontend of Scalable Analysis Framework for ECMAScript (SAFE), a Javascript static analysis tool.