

# SEONGMIN LEE

Max Planck Institute for Security and Privacy (MPI-SP)

Universitätsstraße 140

44799 Bochum

Germany

☎ +49 177 783 5480 ✉ [seongmin.lee@mpi-sp.org](mailto:seongmin.lee@mpi-sp.org) 🏠 [Google Scholar](#)

## Research Interests

---

My research interest lies in dynamic program analysis, especially using statistical methods on dynamic information from execution to reason about a program's semantic properties. The goal of my research is to bring program analysis closer to real-world circumstances regarding the scale and complexity of software within the presence of non-experimental or missing data in the analysis.

## Education and Employment

---

Max Planck Institute for Security and Privacy

Germany

**Postdoctoral Researcher, Software Security Research group**

Sep. 2022 – Present

Group head: Dr. Marcel Böhme

Korea Advanced Institute of Science and Technology

Republic of Korea

**Doctor of Philosophy, School of Computing**

Sep. 2016 – Aug. 2022

Advisor: Dr. Shin Yoo

GPA - 4.02/4.3

**Bachelor of Science, School of Computing**

Feb. 2012 – Aug. 2016

**Bachelor of Science, Department of Mathematical Sciences**

GPA - 3.48/4.3

## Publications

---

### Peer-Reviewed Journal Articles

- 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*, page 110988, 2021
- 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, 2020

### Peer-Reviewed Conference and Workshop Papers

- Seongmin Lee and Marcel Böhme. Statistical reachability analysis. In *ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*, December 2023
- Saeyoon Oh, Seongmin Lee, and Shin Yoo. Effectively sampling higher order mutants using causal effect effectively sampling higher order mutants using causal effect. In *The 16th International Workshop on Mutation Analysis*, 2021
- 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, 2020. Association for Computing Machinery
- S. Lee, S. Hong, J. Yi, T. Kim, C. Kim, and S. 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)*, pages 391–401, 2019
- S. Lee, D. Binkley, R. Feldt, N. Gold, and S. Yoo. Moad: Modeling observation-based approximate dependency. In *2019 19th International Working Conference on Source Code Analysis and Manipulation (SCAM)*, pages 12–22, 2019
- 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
- Gabin An, Jinhan Kim, Seongmin Lee, and Shin Yoo. PyGGI: Python General framework for Genetic Improvement. In *Proceedings of Korea Software Congress, KSC 2017*, December 2017

- Jeongju Sohn, Seongmin Lee, and Shin Yoo. *Amortised Deep Parameter Optimisation of GPGPU Work Group Size for OpenCV*, pages 211–217. Springer International Publishing, Cham, 2016

## Posters & Technical Reports

- Seongmin Lee, Dave Binkley, Robert Feldt, Nicolas Gold, and Shin Yoo. Causal program dependence analysis and causal fault localization. Technical Report CS-TR-2021-423, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea 34141, January 2021
- 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 Proceedings*, ICSE '18, pages 302–303, New York, NY, USA, 2018. Association for Computing Machinery
- Seongmin Lee and Shin Yoo. Using source code lexical similarity to improve efficiency of Observation Based Slicing. Technical Report CS-TR-2017-412, School of Computing, KAIST, May 2017

## Software Engineering Notes

- 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

## Research Experience

### Software Security Group, MPI-SP

Sep. 2022 – Present

Postdoc

Bochum, Germany

- Researched applying statistical methods for program analysis to overcome the scalability issue of the static analysis – Statistical Reachability Analysis
- Researched unbiased estimation of the missing mass/probability/expected number of discovering new classes of unknown multinomial distribution

### Computational Intelligence for Software Engineering Laboratory (COINSE), KAIST

Sep. 2016 – Aug. 2022

Ph.D. Student

Daejeon, Republic of Korea

- Researched approximating the degree of dependence between program element using causal inference – CPDA
- Researched applying statistical models on the observation data to approximate the program dependence – MOAD
- Researched inferencing the type information in the binary executables using RNN with National Security Research Institute
- Researched classifying the false positive alarms from static checker in continuous integration pipeline using CNN with Samsung Research
- Researched program dependence approximation using the lexical model on the source code – MOBS

### Computational Intelligence for Software Engineering Laboratory (COINSE), KAIST

Mar. 2016 – Aug. 2016

Undergraduate Research Intern

Daejeon, Republic of Korea

- Researched on the amortised deep parameter optimisation of GPGPU work group size for OpenCV.
- Accelerated the scalability 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.

## Academic Services

- Program committee: ASE 2023 / (Artifact Evaluation Track) USENIX 2024, ICSE 2024, ISSTA 2023, ICSME 2022, 2021
- Reviewer: JSS 2020, JSS 2021, TOSEM 2022 / (External) ICSE 2023, ISSTA 2023, FSE 2024

## Invited Talks

- *Causal Program Dependence Analysis*  
Sheffield Causality and Testing Workshop, 2023
- *Statistical program dependence analysis*  
Handong Global University, 2022
- *Observation-based approximate dependency modeling and its use for program slicing*  
Korea Conference on Software Engineering, 2022
- *MOBS: Multi-Operator Observation-Based Slicing using Lexical Approximation of Program Dependence*  
59th CREST Open Workshop - Multi-language Software Analysis, 2018

## Awards and honors

---

- **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

## Teaching Experience

---

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