

# Woosuk Lee

Post-doctoral Researcher  
Department of Computer and Information Science  
University of Pennsylvania  
Room L575, Levine Hall  
3330 Walnut St, Philadelphia, PA 19104, USA  
✉ : woosuk@cis.upenn.edu    🌐 : <http://ropas.snu.ac.kr/~wslee>

## Education

---

**Computer Science and Engineering, Seoul National University**      Mar 2009 – Feb 2016  
Doctor of Philosophy in Computer Science and Engineering  
Thesis: Improving the Usability of Static Analyzers  
Advisor: Prof. Kwangkeun Yi

**Computer Science and Engineering, Seoul National University**      Mar 2005 – Feb 2009  
Bachelor of Science in Computer Science

## Research Interests

---

My research interest spans all aspects of programming systems with the goal of improving software quality and programmer productivity. In particular,

- ▶ **Sound, Scalable, & User-friendly Static Analysis**  
I have addressed the three major usability issues: false-alarms, limited scalability, and copy-right concerns [2]. All the techniques have been implemented on top of Sparrow\*, a realistic static analyzer for C programs.
- ▶ **Scalable Program Synthesis guided by Machine Learning**  
I have developed techniques to improve the scalability of program synthesis tools. I have proposed a new approach to guide program synthesis by leveraging a probabilistic program model learnt from a corpus of existing programs. Also, I have developed an interactive synthesis approach that systematically asks the user minimal amount of information to converge to their desired program.

## Experience

---

**University of Pennsylvania**      Jul 2017 – present  
Post-doctoral Researcher  
Advisor: Prof. Mayur Naik

**University of California, Berkeley**      Jul 2012 – Nov 2012  
Visiting Student  
Advisor: Prof. Dawn Song

**Fasoo.com**      Jul 2010 – Aug 2010  
Intern

---

\*Sparrow Analyzer, <http://ropas.snu.ac.kr/sparrow>

Advisor: Dr. Kyujin Cho

## Research Project

---

- ▶ **Guiding Program Synthesis Using a Learnt Probabilistic Model** 2017 – present  
I have designed and implemented a program synthesis tool called Euphony [6]. By leveraging statistical regularity in programs, Euphony outperforms existing synthesis tools on the benchmarks of 2016 syntax-guided synthesis (SyGuS)<sup>†</sup> competition in terms of synthesis time and the number of synthesized programs.
  
- ▶ **Synthesizing Datalog Programs via Active Learning** 2016 – present  
I have designed and implemented a program synthesis tool called Alps [5]. Alps is an interactive synthesis tool that finds a program written in Datalog, a declarative logic programming language, which is widely used to express deductive rule-based programs. A growing range of usage scenarios motivates the need to synthesize such programs from examples that are labeled positive or negative. Alps systematically asks the user to label a minimal number of examples in order to converge to their desired program. To minimize the number of labels, we use an active learning technique called query-by-committee to choose the example that is most controversial among programs in the current candidate space.
  
- ▶ **Static Analysis on Encrypted Programs** 2014 – 2015  
I have designed and implemented a static analysis on encrypted programs. The motivation was from the users' copyright concerns in our *static-analysis-as-a-service*<sup>‡</sup> system. The technique allows to compute encrypted pointer analysis results from encrypted programs without decryption with the aid of a state-of-the-art homomorphic encryption scheme [7].
  
- ▶ **Sparrow: a static analyzer for C program** 2011 – present  
I have been a core developer of Sparrow, a state-of-the-art static analyzer that aims to verify the absence of fatal bugs in C source. In particular, I have been developing techniques for *sound non-statistical alarm clustering* [8, 9] that reduce the number of false alarms to be inspected, *sparse relational analysis* [4] that eliminates redundant computations during static analyses without loss of precision, and static analysis progress estimation [10]. I have also contributed to the theoretical sparse analysis framework [3] that provides a general method for achieving global static analyses that are precise, sound, yet also scalable. In the project, I participated in designing a pre-analysis for the octagon analysis and implementing the sparse octagon analysis. (<http://ropas.snu.ac.kr/sparseanalysis>). Sparrow is publicly available on GitHub. (<http://www.github.com/ropas/sparrow>)
  
- ▶ **Static Analysis for Detecting Android Malicious Apps** 2012 – 2012  
I have implemented static analyzers for the purpose of detecting malicious behaviours including privacy leaks and sneaking phone calls. The implementation was rewarded with shares of En-sighta Security Inc. founded by Prof. Dawn Song in UC Berkeley. The company was acquired by FireEye.
  
- ▶ **Shovel: A SAT-based Tool for Information Flow Alarm Classification** 2015 – 2015  
I have implemented and designed a tool called Shovel [1], that assists the user to quickly classify, as true or false, alarms reported by information flow analyzers. Specifically, SHOVEL helps the user by finding a shortest function call-return path from the source to the sink that satisfies a

---

<sup>†</sup><http://www.syguus.org>

<sup>‡</sup><http://rosaec.snu.ac.kr/clinic>

given user constraint. We empirically showed that our approach is very effective by classifying 351 alarms for 42 open-source C programs and identifying 48 true alarms with crash bugs, three of which were assigned CVE numbers.

## Publications

- [1] Jong-Gwon Kim, **Woosuk Lee**, Jaeseung Choi, Chung-Kil Hur, and Kwangkeun Yi. Shovel: A sat-based tool for information flow alarm classification. Technical memorandum, Research On Software Analysis for Error-free Computing Center, Seoul National University, March 2015.
- [2] Woosuk Lee. *Improving the Usability of Static Analyzers*. PhD thesis, Seoul National University, February 2016.
- [3] Hakjoo Oh, Kihong Heo, Wonchan Lee, **Woosuk Lee**, Daejun Park, Jeehoon Kang, and Kwangkeun Yi. Global sparse analysis framework. *ACM Trans. Program. Lang. Syst.*, 36(3):8:1–8:44, September 2014.
- [4] Hakjoo Oh, Kihong Heo, Wonchan Lee, **Woosuk Lee**, and Kwangkeun Yi. Design and implementation of sparse global analyses for c-like languages. In *Proceedings of the 33rd ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '12*, pages 229–238, New York, NY, USA, 2012. ACM.
- [5] Xujie Si\*, **Woosuk Lee\***, Richard Zhang, Aws Albarghouthi, Paraschos Koutris, and Mayur Naik. Synthesizing datalog programs via active learning. submitted to POPL'18 (\*contributed equally).
- [6] **Woosuk Lee**, Rajeev Alur, and Mayur Naik. Guiding search-based program synthesis using a learnt probabilistic model. submitted to POPL'18.
- [7] **Woosuk Lee**, Hyunsook Hong, Kwangkeun Yi, and Jung Hee Chun. Static analysis with set-closure in secrecy. In *Proceedings of the 22nd International Static Analysis Symposium, SAS'15*. Springer International Publishing, 2015.
- [8] **Woosuk Lee**, Wonchan Lee, Dongok Kang, Kihong Heo, Hakjoo Oh, and Kwangkeun Yi. Sound non-statistical clustering of static analysis alarms. *ACM Trans. Program. Lang. Syst.*, 39(4):16:1–16:35, August 2017.
- [9] **Woosuk Lee**, Wonchan Lee, and Kwangkeun Yi. Sound non-statistical clustering of static analysis alarms. In *Proceedings of the 13th International Conference on Verification, Model Checking, and Abstract Interpretation, VMCAI'12*, pages 299–314, Berlin, Heidelberg, 2012. Springer-Verlag.
- [10] **Woosuk Lee**, Hakjoo Oh, and Kwangkeun Yi. A progress bar for static analyzers. In Markus Muller-Olm and Helmut Seidl, editors, *Proceedings of the 21st International Static Analysis Symposium, SAS'14*, pages 184–200. Springer International Publishing, 2014.

## Software

---

I have contributed to the following open-source software:

- ▶ Sparrow: a static analyzer for C program  
<http://www.github.com/ropas/sparrow>

- ▶ EUSolver: a program synthesis tool based on divide-and-conquer enumeration  
<https://bitbucket.org/abhishekudupa/eusolver/>
- ▶ Alps: an interactive synthesis tool for Datalog  
<https://github.com/XujieSi/SpeedyFOIL>

## Patents

---

- ▶ Copyright information inserting system and method (Registration No. (date): 1010971040000 (12/15/2011))
- ▶ Electronic apparatus for determining whether program comprises malicious code and method for controlling thereof (Application No. (date): 10-2015-0055481 (04/20/2015))

## Reviewer

---

- ▶ ACM TOPLAS, IEEE TSE, ACM PLDI 2016, ACM SAC 2013, ACM POPL 2013, SAS 2012, SPLASH 2010, PADL 2012, PEPM 2012

## Talks

---

- ▶ Structure-guided Synthesis.  
Weekly Group Seminar. University of Pennsylvania. 19 April 2017.
- ▶ Interactive Synthesis of Logic Programs.  
Invited talk. Korea University. 9 February 2017.
- ▶ Static analysis with set-closure in secrecy.  
Paper presentation. In 22nd International Static Analysis Symposium (SAS 2015). Saint-Malo. 9-11 September 2015.  
[http://ropas.snu.ac.kr/~wslee/sas15\\_talk.pdf](http://ropas.snu.ac.kr/~wslee/sas15_talk.pdf)
- ▶ A Progress Bar for Static Analyzers.  
Paper presentation. In 21st International Static Analysis Symposium (SAS 2014). Munich. 11-13 September 2014.  
[http://ropas.snu.ac.kr/~wslee/sas14\\_talk.pdf](http://ropas.snu.ac.kr/~wslee/sas14_talk.pdf)
- ▶ Sound Non-Statistical Clustering of Static Analysis Alarms.  
Paper presentation. In 13th International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI 2012). Philadelphia. 25-27 January 2012.  
[http://ropas.snu.ac.kr/~wslee/vmcai12\\_talk.pdf](http://ropas.snu.ac.kr/~wslee/vmcai12_talk.pdf)

## Teaching Experience

---

- ▶ Teaching Assistant, SNU 4541.664A Program Analysis: Theories and Practices    Spring 2014

- ▶ Teaching Assistant, SNU 4541.664A Program Analysis: Theories and Practices      Fall 2010
- ▶ Teaching Assistant, SNU 4190.210 Principle of Programming      Fall 2009

## Vulnerabilities Reported

---

- ▶ CVE-2015-8106 (Latex2rtf format string vulnerability)
- ▶ CVE-2015-8107 (Assigned)

## References

---

**Kwangkeun Yi**

Professor  
Dept. of Computer Science and Engineering  
Seoul National University  
Email: kwang@ropas.snu.ac.kr

**Rajeev Alur**

Professor  
Dept. of Computer and Information Science  
University of Pennsylvania  
Email: alur@cis.upenn.edu

**Hakjoo Oh**

Assitant Professor  
Dept. of Computer Science and Engineering  
Korea University  
Email: hakjoo\_oh@korea.ac.kr

**Mayur Naik**

Associate Professor  
Dept. of Computer and Information Science  
University of Pennsylvania  
Email: mhnai@cis.upenn.edu

**Jung Hee Cheon**

Professor  
Department of Mathematics  
Seoul National University  
Email: jhcheon@snu.ac.kr

Last updated: September 2, 2017  
<http://ropas.snu.ac.kr/~wslee>