Suyoung Lee

suyoung.lee@kaist.ac.kr | leeswimming

KAIST E3-1 #4431, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

https://leeswimming.com

1 Summary

Suyoung Lee is a postdoctoral researcher in the WSP Lab at KAIST whose research lies at the intersection of artificial intelligence (AI) and computer security. His work focuses on securing modern AI-integrated applications and developing AI-based frameworks for identifying software vulnerabilities. His research has uncovered over 50 critical flaws in AI-based systems, web browsers, and web applications, resulting in publications in top venues such as USENIX Security, NDSS, WWW, ICML, NeurIPS, and IEEE TDSC. He was selected as one of Most Valuable Security Researchers by Microsoft Security Response Center (MSRC). He has also served as a program committee member for The ACM Web Conference (WWW) and a reviewer for ACM Transactions on Software Engineering and Methodology (TOSEM).

2 Employment History

Mar. 2026 – Current **Postdoctoral Researcher**
WSP Lab, KAIST

3 Education

Sept. 2019 – Feb. 2026 **Ph.D., Graduate School of Information Security**
Korea Advanced Institute of Science and Technology (KAIST)
Advisor: Sooel Son

Sept. 2017 – Aug. 2019 **M.S., Graduate School of Information Security**
Korea Advanced Institute of Science and Technology (KAIST)
Advisor: Sooel Son

Mar. 2013 – Aug. 2017 **B.S., Computer Engineering**
Sungkyunkwan University

4 Research Interests

Security, software engineering, and machine learning.

5 Honors and Awards

2025	Cybersecurity Paper Competition by KACS, 2nd place
2022	Best Paper Award (*KCC* 2022)
2019	Cybersecurity Research Competition by KIISC, 3rd & 4th place
2019	MSRC's 2018-2019 Most Valuable Security Researchers
2015—2016	Kwanjeong Educational Foundation Scholarship
2014	Distinguished Freshman Award

6 Publications

(*: equal contribution)

6.1 Conference Papers

[1] Dongwon Shin, **Suyoung Lee**, Sanghyun Hong, and Sooel Son. You Only Perturb Once: Bypassing (Robust) Ad-Blockers Using Universal Adversarial Perturbations. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC)*, pages 190—206, 2024

[2] Byungjoo Kim, **Suyoung Lee**, Seanie Lee, Sooel Son, and Sung Ju Hwang. Margin-based Neural Network Watermarking. In *Proceedings of the International Conference on Machine Learning (**ICML**),* pages 16696—16711, 2023

[3] Dongwon Shin*, **Suyoung Lee***, and Sooel Son. RICC: Robust Collective Classification of Sybil Accounts. In *Proceedings of the ACM Web Conference (**WWW**),* pages 2329—2339, 2023

[4] Hoyong Jeong, **Suyoung Lee**, Sung Ju Hwang, and Sooel Son. Learning to Generate Inversion-Resistant Model Explanations. In *Proceedings of the Advances in Neural Information Processing Systems (**NeurIPS**)*, pages 17717—17729, 2022

[5] Gyumin Lim, Gihyuk Ko, **Suyoung Lee**, and Sooel Son. Adversarial Activation based Neural Network Pruning Revision. In *Proceedings of the Korea Computer Congress (KCC)*, 2022

[6] **Suyoung Lee**, HyungSeok Han, Sang Kil Cha, and Sooel Son. Montage: A Neural Network Language Model-Guided JavaScript Engine Fuzzer. In *Proceedings of the USENIX Security Symposium (**USENIX Security**)*, pages 2613—2630, 2020

[7] Taekjin Lee, Seongil Wi, **Suyoung Lee**, and Sooel Son. FUSE: Finding File Upload Bugs via Penetration Testing. In *Proceedings of the Network and Distributed System Security Symposium (**NDSS**)*, 2020

6.2 Journal Papers

[1] **Suyoung Lee**, Wonho Song, Suman Jana, Meeyoung Cha, and Sooel Son. Evaluating the Robustness of Trigger Set-Based Watermarks Embedded in Deep Neural Networks. *IEEE Transactions on Dependable and Secure (**TDSC**)*, 20(4):3434—3448, 2023

[2] Gyumin Lim, Gihyuk Ko, **Suyoung Lee**, and Sooel Son. Pruning Deep Neural Networks Neurons for Improved Robustness against Adversarial Examples. *Journal of KIISE,* 50(7):588—597, 2023.

7 Patents

[1] Sooel Son, Dongwon Shin, and **Suyoung Lee**. Universal Adversarial Attack Method, Apparatus, and Computer Program for Bypassing Machine Learning-based Ad Blocking Systems. Korea Patent 10-2025-0155816, 2025

[2] Sooel Son, Dongwon Shin, and **Suyoung Lee**. Random Sampling-based Collective Classification Method and System for Sybil Account Detection. Korea Patent 10-2875522, 2025

[3] Sooel Son, Sung Ju Hwang, Hoyong Jeong, and **Suyoung Lee**. Method and Apparatus for Generating Inversion-resistant Model Explanations. Korea Patent 10-2024-0181066, 2024

[4] Sooel Son, Dongwon Shin, and **Suyoung Lee**. Random Sampling-based Collective Classification Method and System for Sybil Account Detection. US Patent 18491570, 2023

[5] Sooel Son, Gyumin Lim, Gihyuk Ko, and **Suyoung Lee**. Method and Apparatus of Revising a Deep Neural Network for Adversarial Examples. Korea Patent 10-2601761, 2023

[6] Sooel Son and **Suyoung Lee**. Evaluating Method of Evaluating Robustness of Artificial Neural Network Watermarking against Model Stealing Attacks. US Patent 17361994, 2021

[7] Sooel Son and Suyoung Lee. Evaluating Method for the Robustness of Watermarks Embedded in Neural Networks against Model Stealing Attacks. Korea Patent 10-2301295, 2021

[8] Sooel Son, Sang Kil Cha, **Suyoung Lee**, Insung Kim, and Taekyu Kim. Method and Apparatus for Testing JavaScript Interpretation Engine using Machine Learning. Korea Patent 10-2132450, 2020

8 Software Artifacts

[1] YOPO, A Universal Adversarial Attack Framework against Machine Learning-based Ad-Blockers, 2024
https://github.com/WSP-LAB/YOPO

[2] RICC, A Collective Classification Framework for Finding Sybil Accounts, 2023
https://github.com/WSP-LAB/RICC

[3] GNIME, A Defense Framework against Model Inversion Attacks, 2022
https://github.com/WSP-LAB/GNIME

[4] Montage, A Neural Network Language Model-based JavaScript Engine Fuzzer, 2020
https://github.com/WSP-LAB/Montage

[5] FUSE, A Penetration Testing Tool for Finding File Upload Bugs, 2020
https://github.com/WSP-LAB/FUSE

9 Reported Security Vulnerabilities

CVE-2019-8594	Arbitrary code execution in JavaScriptCore of Safari
CVE-2019-0923	Memory corruption in ChakraCore of Edge
CVE-2019-0860	Arbitrary code execution in ChakraCore of Edge **($5,000 reward)**
ID #474359	Stored XSS in WordPress **($600 reward)**
XEVE-2019-001	Unrestricted file upload in XE
CVE-2018-19419	Unrestricted file upload in CMSMadeSimple
CVE-2018-19422	Unrestricted file upload in Subrion
CVE-2018-19421	Unrestricted file upload in GetSimpleCMS
CVE-2018-19420	Unrestricted file upload in GetSimpleCMS
CVE-2018-19172	Unrestricted file upload in Elgg
CVE-2018-19146	Unrestricted file upload in Concrete5
CVE-2018-19062	Unrestricted file upload in CMSSimple
CVE-2018-18966	Unrestricted file upload in OsCommerce2
CVE-2018-18965	Unrestricted file upload in OsCommerce2
CVE-2018-18964	Unrestricted file upload in OsCommerce2
CVE-2018-18694	Unrestricted file upload in Monstra
CVE-2018-18637	Unrestricted file upload in ECCube3
CVE-2018-18574	Unrestricted file upload in CMSMadeSimple
CVE-2018-18572	Unrestricted file upload in OsCommerce2
CVE-2018-6383	Unrestricted file upload in Monstra

10 Professional Services

10.1 Conference Reviewer

2026	WWW (Social Networks and Social Media Track)

10.2 Journal Reviewer

2023	ACM TOSEM

11 Teaching

11.1 Teaching Assistant

Spring 2021	IS542 – Web Service Security and Privacy	KAIST
Fall 2019	CS492 – Machine Learning Application Trends in Information Security	KAIST
Spring 2019	IS542 – Web Service Security and Privacy	KAIST
Fall 2018	IS593 – Machine Learning Application Trends in Information Security	KAIST