SANGHYUN HONG

Assistant Professor | Computer Science | Oregon State University

4013 Kelley Engineering Center,
2500 NW Monroe Ave,
Corvallis, OR 97331

: sanghyun.hong@oregonstate.edu
: sanghyun-hong.com
: +1 (301) 771-1475

RESEARCH INTERESTS

My research concerns the security and dependability of deep learning systems—systems that include deep neural networks (DNNs) as a key component. My recent work studies the computational properties of DNNs that traditional software does not have, which make them *particularly* vulnerable to practical hardware or system-level attacks. My work on identifying such computational properties often exposes distinct internal behaviors of DNNs, *e.g.*, confusion or gradient-level disparity, under adversarial pressure. I built defenses that suppress such malicious internal behaviors.

EMPLOYMENT HISTORY

Oregon State University — Corvallis, OR, USA
Assistant Professor — 09.2021 - **Present**
Department of Electrical Engineering and Computer Science (EECS)

Google Brain — Mountain View, CA, USA
Research Intern in Privacy and Security Team — 10.2020 - 01.2021
Host: Dr. Nicholas Carlini and Dr. Alexey Kurakin

Frame.io — New York, NY, USA
Research Intern in Cloud Security Team — 11.2017 - 05.2018
Host: Dr. Abhinav Srivastava

Openwise Inc. (lab-based mobile solution start-up) — Seoul, South Korea
Co-founder and Chief Technology Officer (CTO) — 12.2011 - 08.2014

MBridge Systems Inc. — Seoul, South Korea
Lead Researcher (R&D Department) — 12.2010 - 12.2013

EDUCATION

University of Maryland, College Park — College Park, MD, USA
Ph.D. in Computer Science — 09.2015 - 08.2021
Academic advisor: Prof. Tudor Dumitraș
Dissertation: *Building Secure and Reliable Deep Learning from A System Security Perspective*

M.S. in Computer Science — 09.2015 - 12.2017
Academic advisor: Prof. Tudor Dumitraș
Scholarly paper: *Characterizing Program Behaviors in a Virtualized Infrastructure without Introspection*

Seoul National University — Seoul, South Korea
B.S. in Electrical Engineering and Computer Science (*magna cum laude*) — 03.2007 - 02.2015
Academic adviser: Prof. Seongsoo Hong
Thesis: *A Power Saving Mechanism for the Smartphone Modem via Application-based Packet Piggybacking*

SCHOLARLY AND PROFESSIONAL ACTIVITIES

HONORS & AWARDS

Honors

Future Faculty Fellow, University of Maryland, College Park 2020 - 2021

Fellowships

Ann G. Wylie Dissertation Fellowship, University of Maryland, College Park 2020
KSEA-KUSCO Scholarships for Korean Graduate Students, KSEA 2017
Summer Research Fellowship, University of Maryland, College Park 2016
4-year Scholarship, Seoul National University Alumni Association (SNUAA) 2015
2-year Dean's Fellowship, University of Maryland, College Park 2015
Full 2-year Graduate Teaching Assistantship, University of Maryland, College Park 2015
Full 1-year scholarship, KEPCO 2010
Full 4-year scholarship, KOFAC 2007

Awards

Top 33% Reviewers Award, Neural Information Processing Systems (NeurIPs) 2020
Jacob K. Goldhaber Travel Award, University of Maryland, College Park 2019
ICSSA Travel Award, University of Maryland, College Park 2019
Student Travel Award, University of Maryland, College Park 2018
2nd & 4th Place, Research Competition for Korean Graduate Students 2018

PUBLICATIONS*

Peer-Reviewed Conference Publications

[**C.1**] **Sanghyun Hong**, Michael-Andrei Panaitescu-Liess, Yiğitcan Kaya, and T. Dumitraș, "Qu-ANTI-zation: Exploiting Quantization Artifacts for Achieving Adversarial Outcomes", In 34th *Conference on Neural Information Processing Systems (NeurIPS)*, 2021.

[**C.2**] **Sanghyun Hong***, Yiğitcan Kaya*, Ionuț-Vlad Modoranu, and T. Dumitraș, "A Panda? No, It's a Sloth: Slowdown Attacks on Adaptive Multi-Exit Neural Network Inference", In *International Conference on Learning Representations (ICLR)*, 2021. **[Spotlight]** (*joint first authors)

[**C.3**] **Sanghyun Hong**, Michael Davinroy, Yiğitcan Kaya, Dana Dachman-Soled and Tudor Dumitraș, "How to 0wn NAS in Your Spare Time", In *International Conference on Learning Representations (ICLR)*, 2020.

[**C.4**] **Sanghyun Hong**, Pietro Frigo, Yiğitcan Kaya, Cristiano Giuffrida and Tudor Dumitraș, "Terminal Brain Damage: Exposing the Graceless Degradation in Deep Neural Networks Under Hardware Fault Attacks", In *USENIX Security Symposium*, 2019.

[**C.5**] Yiğitcan Kaya, **Sanghyun Hong**, and Tudor Dumitraș, "Shallow-Deep Networks: Understanding and Mitigating Network Overthinking", In *36th International Conference on Machine Learning (ICML)*, 2019.

[**C.6**] **Sanghyun Hong**, Noseong Park, Tanmoy Chakraborty, Hyunjoong Kang, and "PAGE: Pattern-Query Answering via Knowledge Graph Embedding". In *International Conference on Big Data (BigData)*, 2018.

[**C.7**] Hyunjoong Kang, **Sanghyun Hong**, Kookjin Lee, Noseong Park, and Soonhyun Kwon, "On Integrating Knowledge Graph Embeddings into SPARQL Query Answering", In *International Conference on Web Services (ICWS)*, 2018.

[**C.8**] **Sanghyun Hong**, Tanmoy Chakraborty, Sungjin Ahn, Ghaith Husari, and Noseong Park, "SENA: Preserving Social Structure for Network Embedding", In *ACM Conference on Hypertext and Social Media (HT)*, 2017.

*My premier publication venues are [IEEE S&P, USENIX, CCS, NDSS] (security) and [ICML, ICLR, NeurIPs] (machine learning); these conferences are peer-reviewed and accept fewer than 10-30% of the total papers submitted.

Peer-Reviewed Journal Articles

[**J.1**] **Sanghyun Hong**, Alina Nicolae, Abhinav Srivastava, and Tudor Dumitraș, “Peek-a-Boo: Inferring Program Behaviors in a Virtualized Infrastructure without Introspection", In *Computer & Security (COSE)*, 2018.

Workshop Papers

[**W.1**] **Sanghyun Hong**, Abhinav Srivastava, William Shambrook, and Tudor Dumitraș. “Go Serverless: Securing Cloud via Serverless Design Patterns", In *USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 18)*, 2018.

[**W.2**] Rock Stevens, Octavian Suciu, Andrew Ruef, **Sanghyun Hong**, Michael Hicks, and Tudor Dumitraș, "Summoning Demons: The Pursuit of Exploitable Bugs in Machine Learning", In *Neural Information Processing Systems (NeurIPS) Workshop on Crowdsourcing and Machine Learning*, 2016.

Posters

[**P.1**] **Sanghyun Hong**, Tae-hoon Kim, Tudor Dumitraș, and Jonghyun Choi. “Poster: On the Feasibility of Training Neural Networks with Visibly Watermarked Dataset.", In the *Network and Distributed System Security Symposium (NDSS)*, 2019.

Manuscripts

[**M.1**] **Sanghyun Hong**, Alexey Kurakin, and Nicholas Carlini “Handcrafted Backdoors in Deep Neural Networks" *arXiv Preprint*, 2021.

[**M.2**] **Sanghyun Hong**, Varun Chandrasekaran, Yiğitcan Kaya, Tudor Dumitraș and Nicolas Papernot “On the Effectiveness of Mitigating Data Poisoning Attacks with Gradient Shaping", *arXiv Preprint*, 2021.

[**M.3**] Yiğitcan Kaya, **Sanghyun Hong**, and Tudor Dumitraș, “On the Effectiveness of Regularization Against Membership Inference Attacks", *arXiv Preprint*, 2020.

[**M.4**] **Sanghyun Hong**[*], Michael Davinroy[*], Yiğitcan Kaya, Stuart Nevans Locke, Ian Rackow, Kevin Kulda, Dana Dachman-Soled and Tudor Dumitraș, “Security Analysis of Deep Neural Networks Operating in the Presence of Cache Side-Channel Attacks", *arXiv Preprint*, 2018. ([*]joint first authors).

INVITED TALKS

[**T.1**] **Oregon State University**, Tech Talk Tuesday — 11.2021
Building Secure and Reliable Deep Learning Systems from a Systems Security Perspective

[**T.2**] **Yonsei University**, Virtual Presentation — 11.2021
Qu-ANTI-zation: Exploiting Quantization Artifacts for Achieving Adversarial Outcomes [Host: Prof. Noseong Park]

[**T.3**] **Oregon State University**, Eta Kappa Nu Induction Ceremony — 11.2021
Lessons from My Journey Towards One-bit Flip [Host: Prof. Arun Natarajan]

[**T.4**] **Oregon State University**, Cybersecurity Friday — 10.2021
Why A Systems Security Perspective Matters in Adversarial Machine Learning?

[**T.5**] **CMU**, Virtual Presentation — 06.2021
Building Secure and Reliable Deep Learning Systems from a Systems Security Perspective [Host: Prof. Lujo Bauer]

[**T.6**] **ICLR**, Virtual Conference Presentation **[Spotlight]** — 05.2021
A Panda? No, It’s Sloth: Slowdown Attacks on Adaptive Multi-Exit Neural Network Inference

[†]USENIX Enigma is a conference that focuses on communicating security ideas effectively through TED-style talks. The topics are selected competitively by a Program Committee, and the presentations undergo extensive peer-review and coaching following TED guidelines.

[**T.7**] **Hardwear.io**, Virtual Conference Presentation 02.2021
Practical Hardware Attacks on Deep Learning

[**T.8**] **USENIX Enigma**[†], Virtual Conference Presentation 92.2021
A Sound Mind in A Vulnerable Body: Practical Hardware Attacks on Deep Learning

[**T.9**] **University of Tennessee**, Guest Lecturer 09.2020
Practical Hardware Attacks on Deep Learning [Host: Prof. Doowon Kim]

[**T.10**] **ICLR**, Virtual Conference Presentation 05.2020
How to 0wn NAS in Your Spare Time

[**T.11**] **USENIX Security Symposium**, Santa Clara, CA 08.2019
Terminal Brain Damage: Exposing the Graceless Degradation of DNNs under Hardware Fault Attacks

[**T.12**] **Yonsei University**, South Korea 01.2019
Can Machine Learning be Secure and Trustworthy
in the Presence of Micro-architectural Vulnerabilities? [Host: Prof. Taekyung Kwon]

[**T.13**] **USENIX Workshop on Hot Topics in Cloud Computing**, Boston, MA 07.2018
Go Serverless: Securing Cloud via Serverless Design Patterns

PROFESSIONAL ACTIVITIES

Program Committee

WPES: 20th ACM Workshop on Privacy in the Electronic Society	2021
AISec: 14th ACM Workshop on Artificial Intelligence and Security	2021
RAID: International Symposium on Research in Attacks, Intrusions and Defenses	2021
IEEE S&P: Deep Learning and Security Workshop (DLS)	2021
ICLR: Workshop on Towards Trustworthy ML: Rethinking Security and Privacy for ML	2020

Reviewer

ICLR: International Conference on Learning Representation	2021 – 2022
NeurIPs: Conference on Neural Information Processing Systems	2020 – 2021
ICML: International Conference on Machine Learning	2020 – 2021
IEEE Access: IEEE Access	2021
COSE: Computer & Security	2019
TCC: IEEE Transaction on Cloud Computing	2018

External Reviewer

NDSS: Network and Distributed System Security Symposium	2017, 2019 – 2020
IEEE S&P: IEEE Symposium on Security and Privacy	2017, 2019
CCS: ACM Conference on Computer and Communications Security	2017 – 2020
USENIX: USENIX Security Symposium	2016 – 2019
RAID: International Symposium on Research in Attacks, Intrusions and Defenses	2018 – 2019

TEACHING AND MENTORING

TEACHING EXPERIENCE

Oregon State University — Corvallis, OR, USA
Assistant Professor

CS 499/599: Machine Learning Security — 01.2022 - 04.2022
CS 344 : Operating Systems I — 04.2022 - 07.2022

University of Maryland, College Park — College Park, MD, USA
Graduate Teaching Assistant

CMSC 132: Objected Oriented Programming II (Instructor: Tom Reinhardt) — 09.2015 - 01.2016

STUDENTS ADVISING

PhD Students

Hoang Le (CS) — 2021 – Present

Graduate Committee

Jarrod Jeffrey Isaac Hollis (Ph.D, CS) — 2021 – Present

MENTEES

Research Interns at Maryland Cybersecurity Centers (MC2)

Michi Panaitescu	2020 – 2021
Ionuț-Vlad Modoranu	2019 – 2021
Khasmamad Shabanovi	2020
Stefan Matcovici	2019
Michael Davinroy	2018 – 2019
Alina-Elena Nicolae	2016

Undergraduate Research Interns from NSF-REU Program

Michael Davinroy	2018 - 2019
Stuart Nevans Locke	Summer 2018
Ian Rackow	Summer 2018
Kevin Kulda	Summer 2018