Eui Chul (Richard) Shin
ricshin@berkeley.edu

Education

University of California, Berkeley
PhD in Computer Science — January 2013–present
MS in Computer Science — December 2017
Advisor: Dawn Song

BS in Electrical Engineering and Computer Science — August 2008–December 2011
Graduated with High Honors

Graduate-level Coursework: Algorithms, Convex Optimization, Theoretical Statistics, Statistical Learning Theory, Natural Language Processing, Computer Security, Computer Systems, Programming Languages, Deep Reinforcement Learning

Employment

Intel Labs, Graduate Technical Intern — October 2018–February 2019
Intelligent Systems Lab (manager: Vladlen Koltun). Developed new architectures for converting natural language questions into database queries.

Microsoft Research AI
Research Intern — May 2018–August 2018
Research SDE (contractor) — August 2018–October 2018; August 2019–present (part time)
Neural Program Synthesis group (manager: Alex Polozov). Investigated the use of idioms in source code for neural program synthesis.

NEAR.AI, Research Engineer (part-time) — Dec 2017–May 2018
Developed a method for neural program repair and inferring execution traces for neural program synthesis.

Google, Software Engineering Intern — May 2016–August 2016
Automated theorem proving group in Google Research and Google Brain (manager: Geoffrey Irving, co-mentor: Oriol Vinyals). Performed premise selection experiments on the E theorem prover. Developed a variational autoencoder-based generative model for tree structures.

Google, Software Engineering Intern — May 2015–August 2015
Google Translate's neural machine translation team (manager: Wolfgang Macherey). Trained sequence-to-sequence LSTM models on large parallel texts using TensorFlow. Integrated sequence models into a phrase-based machine translation decoder.

Google, Software Engineer — January 2012–January 2013
Google Translate's parallel data team (manager: Thorsten Brants).
- Streamlined and modernized pipeline for finding parallel documents (which contain the same text but in different languages) from the web, scaling it to handle 5× more data.
- Participated in effort to improve low-quality languages and launch new ones, through tuning parameters and collecting analytics to enable targeted acquisition of new data.
- Implemented an experimental method for finding sentence-level rather than document-level parallel texts.

UC Berkeley, Undergraduate Research Assistant — September 2009–December 2011
Web security, mobile app security, and social networking privacy research projects in Prof. Dawn Song's group.
- **Mobile app security**: Developed a translator from Dalvik to JVM bytecode (with a post-doc). Built tool to demonstrate automated exploitation of in-app billing in vulnerable Android applications through bytecode rewriting.
- **Web security**: Discovered vulnerabilities in Facebook and Google's use of `postMessage` for web mashups. Investigated methods for robust automatic sanitization of untrusted data to prevent cross-site scripting.
- **Privacy and social networks**: Performed feature selection for authorship identification experiments on blogs using statistical machine learning techniques. Oversaw daily collection of data from tens of millions of Google+ profiles for link-prediction experiments.

Publications

Pre-prints

Encoding Database Schemas with Relation-Aware Self-Attention for Text-to-SQL Parsers.
Richard Shin.
arXiv:1906.11790 [cs.LG].

Conferences and Journals

Program Synthesis and Semantic Parsing with Learned Code Idioms.
Richard Shin, Marc Brockschmidt, Miltiadis Allamanis, Oleksandr Polozov.
Neural Information Processing Systems (NeurIPS), 2019.

Synthetic Datasets for Neural Program Synthesis.
Richard Shin, Neel Kant, Kavi Gupta, Christopher Bender, Brandon Trabucco, Rishabh Singh, Dawn Song.
International Conference on Learning Representations (ICLR), 2019.

Improving Neural Program Synthesis with Inferred Execution Traces.
Richard Shin, Illia Polosukhin, Dawn Song.
Neural Information Processing Systems (NeurIPS), 2018. *Spotlight presentation*.

Parametrized Hierarchical Procedures for Neural Programming.
Roy Fox, Richard Shin, Sanjay Krishnan, Ken Goldberg, Dawn Song, Ion Stoica.
International Conference on Learning Representations (ICLR), 2018.

Making Neural Programming Architectures Generalize via Recursion.
Jonathon Cai, Richard Shin, Dawn Song.
International Conference on Learning Representations (ICLR), 2017. *Best paper award*.

PIANO: Proximity-based User Authentication on Voice-Powered Internet-of-Things Devices.
Neil Zhenqiang Gong, Altay Ozen, Yu Wu, Xiaoyu Cao, Richard Shin, Dawn Song, Hongxia Jin, Xuan Bao.
IEEE International Conference on Distributed Computing Systems (ICDCS), 2017.

ExploreKit: Automatic Feature Generation and Selection (short paper).
Gilad Katz, Richard Shin, Dawn Song.
International Conference on Data Mining (ICDM), 2016.

Latent Attention for If-Then Program Synthesis.
Xinyun Chen, Chang Liu, Richard Shin, Dawn Song, Mingcheng Chen.
Neural Information Processing Systems (NeurIPS), 2016.

Recognizing Functions in Binaries with Neural Networks.
Richard Shin, Dawn Song, Reza Moazzezi.
USENIX Security, 2015.

Joint Link Prediction and Attribute Inference Using a Social-Attribute Network.
Neil Gong, Ameet Talwalkar, Lester Mackey, Ling Huang, Richard Shin, Emil Stefanov, Elaine Shi, Dawn Song.
ACM Transactions on Intelligent Systems and Technology (TIST), 2013.

On the Feasibility of Internet-Scale Author Identification.
Arvind Narayanan, Hristo Paskov, Neil Gong, John Bethencourt, Emil Stefanov, Richard Shin, Dawn Song.
IEEE Security & Privacy, 2012.

FreeMarket: Shopping for free in Android applications (extended abstract).
Daniel Reynaud, Richard Shin, Thomas R. Magrino, Edward Wu, Dawn Song.
Network & Distributed System Security Symposium (NDSS), 2012.

A Systematic Analysis of XSS Sanitization in Web Application Frameworks.
Joel Weinberger, Prateek Saxena, Devdatta Akhawe, Matt Finifter, Richard Shin, Dawn Song.
European Symoposium on Research in Computer Security (ESORICS), 2011.

Inference and Analysis of Formal Models of Botnet Command and Control Protocols.
Chia Yuan Cho, Domagoj Babić, Richard Shin, Dawn Song.
ACM Conference on Computer and Communications Security (CCS), 2010.

Workshops

Hierarchical Imitation Learning via Variational Inference of Control Programs.
Roy Fox, Richard Shin, William Paul, Yitian Zou, Dawn Song, Ken Goldberg, Pieter Abbeel, Ion Stoica.
Infer to Control Workshop on Probabilistic Reinforcement Learning and Structured Control, at NeurIPS, 2018.

Synthetic Datasets for Neural Program Synthesis (extended abstract).
Richard Shin, Neel Kant, Kavi Gupta, Christopher Bender, Brandon Trabucco, Rishabh Singh, Dawn Song.
Neural Abstract Machines & Program Induction Workshop, at ICML, 2018.

Imitation Learning of Hierarchical Programs via Variational Inference (extended abstract).
Roy Fox,* Richard Shin,* Pieter Abbeel, Ken Goldberg, Dawn Song, Ion Stoica.
Neural Abstract Machines & Program Induction Workshop, at ICML, 2018.

Towards Specification-Directed Program Repair.
Richard Shin, Dawn Song, Illia Polosukhin.
International Conference on Learning Representations (ICLR) workshop track, 2018.

Differentiable Neural Network Architecture Search.
Richard Shin,* Charles Packer,* Dawn Song.
International Conference on Learning Representations (ICLR) workshop track, 2018.

JPEG-resistant Adversarial Images.
Richard Shin, Dawn Song.
Machine Learning and Computer Security Workshop, at NeurIPS, 2017.

Exploring Privacy Preservation in Outsourced K-Nearest Neighbors with Multiple Data Owners.
Frank Li, Richard Shin, Vern Paxson.
ACM Cloud Computing Security Workshop (CCSW), at ACM CCS, 2015.

The Emperor's New APIs: On the (In)Secure Usage of New Client-side Primitives.
Steve Hanna, Richard Shin, Devdatta Akhawe, Arman Boehm, Prateek Saxena, Dawn Song.
Web 2.0 Security and Privacy Workshop (W2SP), at IEEE S&P, 2010.

Talks

Making Neural Programming Architectures Generalize via Recursion. Talk at ICLR 2017.

ExploreKit: Automatic Feature Generation and Selection. Talk at ICDM 2016.

Recognizing Functions in Binaries with Neural Networks. Talk at USENIX Security 2015.

FreeMarket: Shopping for free in Android applications. Short talk at NDSS 2012.

Teaching and Mentoring

Graduate Student Instructor
Special Topics in Deep Learning, UC Berkeley — Fall 2016, Spring 2017
Artificial Intelligence (CS 188), UC Berkeley — Fall 2015

Undergraduate Research Mentoring — Fall 2015–present
Mentored undergrad/masters students and interns on projects relating to domain adaptation in NLP, neural program synthesis, machine learning for identification and analysis of crypters and packers, fuzz testing, adversarial inputs to neural networks, neural network watermarking, neural code completion, etc.

Summer Research Mentoring — Summer 2014
Mentored two undergraduate students (from TRUST REU) and two high school students (from BFOIT) on a project about applying convolutional neural networks for securing wearable devices.

Awards and Honors

Highest-scoring reviewer (top 400), NeurIPS — 2019

Best Paper Award, ICLR 2017:
Making Neural Programming Architectures Generalize via Recursion — 2017

EECS Departmental Graduate Fellowship, UC Berkeley — 2013

Jim and Donna Gray Endowment Fund, UC Berkeley — 2009

Represented UC Berkeley at the ACM ICPC regional competition — 2009–2011

Eta Kappa Nu and Tau Beta Pi — 2009

Service

Reviewer for ICLR 2020, NeurIPS 2019, NeurIPS 2017 Machine Learning and Security Workshop, ICDM 2013 (secondary), Data Mining and Knowledge Discovery (2017).

UC Berkeley EECS Outreach Program — 2014–2017
Gave science demos to students at elementary and middle schools.

UC Berkeley EECS Graduate Admissions — 2014
Reviewed applications from underrepresented minorities.