SIMIN CHEN

+1 214-3569-114
http://www.chensimin.site/
https://github.com/SeekingDream
sxc180080@utdallas.edu

RESEARCH INTERESTS

My research focuses on **Software Engineering** and **Machine Learning**, with the goal of designing foundational infrastructure tools to make ML-based systems more reliable, interpretable, and efficient. To achieve such a goal, I develop techniques to improve the quality of ML systems at three stages. (1) At the ML model development stage, I design tools to explain the decision-making of ML models and apply the explanation results to locate/debug the bugs in the models; (2) at the ML model deployment stage, I develop tools to quantify the privacy leakage risk of ML models and optimize the ML model performance; (3) at the ML model runtime stage, I develop tools for validating the inputs of ML models to improve the model efficiency.

EDUCATION

Ph.D. Candidate | ***(GPA 3.82/4.0)*** — Jan. 2019 – Now
University of Texas at Dallas (Advisor: **Dr. Wei Yang**, and **Dr. Cong Liu**) — Dallas, The United States

Master of Science | ***(GPA 84.7/100)*** — Sep. 2015 – Jun. 2018
Tongji University — ShangHai, China

Bachelor of Science | ***(GPA 4.48/5.0)*** — Sep. 2011 – Jun. 2015
Tongji University — ShangHai, China

PUBLICATION

I publish seven technical track papers as the lead author; four belong to the top-criteria software engineering conference, and three belong to the top-criteria artificial intelligence conference.

(C7) **Simin Chen**, Hanlin Chen, Mirazul Haque, Cong Liu, Wei Yang. The Dark Side of Dynamic Routing Neural Networks: Towards Efficiency Backdoor Injection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023).

(C6) **Simin Chen**, Shiyi Wei, Cong Liu, Wei Yang. DyCL: Dynamic Neural Network Compilation Via Program Rewriting and Graph Optimization. In Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2023).

(C5) **Simin Chen**, Cong Liu, Mirazul Haque, Zihe Song, and Wei Yang. NMTSloth: understanding and testing efficiency degradation of neural machine translation systems. In Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2022), 1148–1160, November 2022.

(C4) **Simin Chen**, Mirazul Haque, Cong Liu, and Wei Yang. 2022a. DeepPerform: An Efficient Approach for Performance Testing of Resource-Constrained Neural Networks. In Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering (ASE 2022), 1-13, October 2022.

(C3) **Simin Chen**, Hamed Khanpour, Cong Liu, and Wei Yang. 2022b. Learning to Reverse DNNs from AI Programs Automatically. In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. (IJCAI 2022), 666–672, July 2022.

(C2) **Simin Chen**, Zihe Song, Mirazul Haque, Cong Liu, and Wei Yang. NICGSlowDown: Evaluating the Efficiency Robustness of Neural Image Caption Generation Models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022), 15365–15374, June 2022.

(C1) **Simin Chen**, Soroush Bateni, Sampath Grandhi, Xiaodi Li, Cong Liu, and Wei Yang. DENAS: Automated Rule Generation by Knowledge Extraction from Neural Networks. In Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2022), 813–825, November 2020.

Scholarships and Awards

- Second prize in THUBA DAO Global Hackson for Blockchain Competition.
- Travel grant award from CVPR 2022.

Industry Experience

NEC Laboratories America
Member of System Security and Reliability Team — January 2020 – May 2020
- Participate in the Graph-based Source Code Vulnerability Detection

Microsoft Research
Member of System Security and Reliability Team — May 2021 – July 2021
- Participate in the project of reverse engineering on on-device DNNs

Teaching Experience

University of Texas at Dallas

CS 4393 – Computer and Network Security

CS 4347 – Computer Engineering

CS 6301 – Special Topics in Computer Science (Graduate Course)

Mentoring

- Sampath Grandhi (MS, UTD; Co-authored)
- Zihe Song (Ph.D., UTD; Co-authored)
- Miao Miao (Ph.D., UTD)
- Hashmi Junaid (MS, UTD)
- Seo Jeongwon (MS, UTD)
- Zexin Li (Ph.D., University of California at Riverside; Co-authored [paper under submission])
- Yufei Li (Ph.D., University of California at Riverside; Co-authored [paper under submission])
- Yiming Chen (Ph.D., Nanyang Technological University)

SERVICE

Software Engineering Community

- **Junior Program Committee**, Conference on Mining Software Repositories (MSR 2023).
- **Sub-Reviewer**, International Conference on Software Engineering (ICSE 2023).
- **Sub-Reviewer**, International Conference on Software Testing, Verification and Validation (ICST 2023).
- **Sub-Reviewer**, International Conference on Automated Software Engineering (ASE 2022).
- **Sub-Reviewer**, International Conference on Software Testing, Verification and Validation (ICST 2022).
- **Sub-Reviewer**, International Conference on Software Engineering (ICSE 2021).

Artificial intelligence and Machine Learning Community

- **Program Committee**, Computer Vision and Pattern Recognition Conference (CVPR 2023).
- **Program Committee**, Association for the Advancement of Artificial Intelligence (AAAI 2023).
- **Program Committee**, Winter Conference on Applications of Computer Vision (WACV 2022).
- **Program Committee**, European Conference on Computer Vision (ECCV 2022).

TALKS

Conference Talks and Posters

- DENAS: Automated Rule Generation by Knowledge Extraction from Neural Networks at **ESEC/FSE 2022** (Virtual)
- DeepPerform: An Efficient Approach for Performance Testing of Resource-Constrained Neural Networks at **ASE 2022** (Oakland Center, Michigan, USA)
- Learning to Reverse DNNs from AI Programs Automatically at **IJCAI 2022** (Virtual)
- NICGSlowDown: Evaluating the Efficiency Robustness of Neural Image Caption Generation Models at **CVPR 2022** (New Orleans, Louisiana, USA)
- DENAS: Automated Rule Generation by Knowledge Extraction from Neural Networks at **ESEC/FSE 2020** (Virtual)

OPEN-SOURCE CONTRIBUTIONS

I led the development of five tools/datasets for improving the efficiency and explainability of Machine Learning Systems.

- **DENAS:** DENAS is an automatic tool that can extract the knowledge from neural networks and represent the knowledge as explainable rules. DENAS can help model developers to locate and debug the errors in the ML models. It is available at https://github.com/SeekingDream/FSE20_DENAS.
- **NICGSlowDown:** NICGSlowDown is a tool that generates test inputs to test neural image caption systems. It is available at https://github.com/SeekingDream/CVPR22_NICGSlowDown
- **NMTSloth:** NMTSloth is a tool that generates test inputs to test neural machine translation systems. It is available at https://github.com/SeekingDream/FSE22_NMTSloth
- **NNReverse:** NNReverse is a large scale binary code dataset, which is compiled from different neural network architectures using different compiler settings. It is available at Google Drive Link
- **DeepPerform:** DeepPerform is a learning-based testing tool to test dynamic neural networks. It automatically learns the buggy inputs' distribution and generates test generation with only 6-10 milliseconds. It is available at https://github.com/SeekingDream/DeepPerform