Shailja Thakur

shailjathakur@nyu.edu • +1 (917) 283-1342 • https://shailja-thakur.github.io

RESEARCH FOCUS

My research focuses on enhancing the trustworthiness, security, and efficiency of **cyber-physical systems**. My Ph.D. work centered on securing and interpreting **decision-making** in **automotive systems**. As a Postdoc, I pioneered the application of **Large Language Models (LLMs)** for **automating hardware design**, with a focus on generating Verilog code, repairing bugs, and addressing issues like LLM biases and copyright concerns. My future work will explore the challenges and potential of **generative AI** techniques like LLMs in **ubiquitous computing systems**, with the goal of improving efficiency, scale, security and quality for downstream tasks.

EDUCATION

University of Waterloo, Waterloo, Ontario, Canada

- Ph.D. in Electrical and Computer Engineering — 2017 – 2022
  - Thesis: "Security and Interpretability in Automotive Systems"
  - Advisor: Dr. Sebastian Fischmeister
  - Cumulative GPA: 8.0 / 10

Indraprastha Institute of Information Technology, Delhi, India

- M.Tech in Computer Science — 2012 – 2014
  - Thesis: "WattShare: Detailed Energy Apportionment in Shared Living Spaces within Commercial Buildings"
  - Advisor: Dr. Amarjeet Singh
  - Cumulative GPA: 8.3 / 10

Guru Gobind Singh Indraprastha University, Delhi, India

- B.Tech. in Computer Science and Engineering — 2008 – 2012
  - Graduated with Honors
  - Cumulative GPA: 8.5 / 10

HONORS AND ACHIEVEMENTS

- DATE Travel Grant, DATE 2023
- Post-Doctoral Fellowship in the NSF National AI Institute for Edge Computing Leveraging Next Generation Networks (Athena) 2022
- Faculty of Engineering Award, University of Waterloo, 2020
- International Doctoral Student Scholarship, University of Waterloo, 2017-2020
- Graduate Research Scholarship, University of Waterloo, 2017-2020
- Data Analytics Excellence Award, Compuware, 2016

RESEARCH EXPERIENCE

New York University, New York, USA

- Postdoctoral Research Fellow — 2022 – Present
  - Department of Electrical and Computer Engineering
  - Advisor: Dr. Ramesh Karri and Dr. Siddharth Garg

University of Waterloo, Waterloo, Ontario, Canada

- Graduate Research Assistant — 2017 – 2022
  - Department of Electrical and Computer Engineering
  - Advisor: Dr. Sebastian Fischmeister

Indraprastha Institute of Information Technology, Delhi, India

- Graduate Research Assistant — 2012 – 2014
  - Department of Computer Science and Engineering
  - Advisor: Dr. Amarjeet Singh

PUBLICATIONS

JOURNALS

[1] **Shailja Thakur**, Carlos Moreno, Sebastian Fischmeister, "CANOA: CAN Origin Authentication Through Power Side-Channel Monitoring", *ACM Transactions on Cyber-Physical Systems* **(ACM TCPS)**, 2022, https://dl.acm.org/doi/pdf/10.1145/3571288, **Impact Factor: 3.2**.

CONFERENCES/WORKSHOPS

† Indicates equal contribution.

[2] Akshaj Kumar Veldanda, Fabian Grob, **Shailja Thakur**, Hammond Pearce, Benjamin Tan, Ramesh Karri, Siddharth Garg, "Are Emily and Greg Still More Employable than Lakisha and Jamal? Investigating Algorithmic Hiring Bias in the Era of ChatGPT", *Accepted in* **NeurIPS**, *workshop in R0FoMo*, 2023, arXiv preprint arXiv:2310.05135. **Core A***

[3] Animesh Chowdhury†, **Shailja Thakur**†, Hammond Pearce, Ramesh Karri, Siddharth Garg,"Towards the ImageNets of ML4EDA", *Proceedings of the International Conference on Computer-Aided Design* **(IEEE ICCAD)**, *Special Session Generative AI for EDA: Datasets, Benchmarks and Infrastructures*, San Francisco, California, 2023, https://arxiv.org/abs/2310.10560. **Core A**

[4] **Shailja Thakur**, Baleegh Ahmad, Zhenxing Fan, Hammond Pearce, "Benchmarking Large Language Models for Automated Verilog RTL Code Generation", *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition* **(IEEE DATE)**, Antwerp, Berlin, April 2023, https://arxiv.org/abs/2212.11140. **Acceptance Rate 30%**, (***Nominated Best Paper Award at DATE 2023***)

[5] **Shailja Thakur**, Sebastian Fischmeister, "Security and Interpretability in Automotive Systems", *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition* **(IEEE DATE)**, 2023, Antwerp, Berlin, April 2023, arXiv:2212.12101.**Acceptance Rate 30%**

[6] **Shailja Thakur**, Sebastian Fischmeister "A generalizable saliency map-based interpretation of model outcome", *Proceedings of the International Conference on Pattern Recognition* (**IEEE ICPR**), Milan, Italy, January 2021, arXiv:2006.09504. **Acceptance Rate 35%**.

[7] **Shailja Thakur**, Manaswi Saha, Amarjeet Singh, Yuvraj Agarwal, "WattShare: Detailed Energy Apportionment in Shared Living Spaces within Commercial Buildings", *Proceedings of the ACM International Conference on Embedded Systems for Energy-Efficient Buildings* (**ACM BuildSys**), Memphis, Tennesse, November 2014, https://dl.acm.org/doi/abs/10.1145/2674061.2674069, **Core A**.

[8] Manaswi Saha, **Shailja Thakur**, Amarjeet Singh, Yuvraj Agarwal, "EnergyLens: Combining Smartphones with Electricity Meter for Accurate Activity Detection and User Annotation", *Proceedings of the ACM International Conference on Future Energy Systems* (**ACM e-Energy**), Cambridge, UK, June 2014, https://dl.acm.org/doi/abs/10.1145/2602044.2602058, **Core A**.

UNDER REVIEW

[9] **Shailja Thakur**, Jason Blocklove, Hammond Pearce, "AutoChip: Automating HDL Generation Using LLM Feedback", *Under Review in IEEE Design Automation Conference* **(IEEE/ACM DAC)**.

[10] **Shailja Thakur**, Baleegh Ahmad, Hammond Pearce, Benjamin Tan, Brendan Dolan-Gavitt, R.Karri, Siddharth Garg, "VeriGen: A Large Language Model for Verilog Code Generation", *Under Review in ACM Transactions on Design Automation of Electronic Systems* **(ACM TODAES)**, arXiv:2308.00708.

[11] Rahul Kande, Hammond Pearce, Benjamin Tan, Brendan Dolan-Gavitt, **Shailja Thakur**, Ramesh Karri, Jeyavijayan Rajendran, "LLM-assisted Generation of Hardware Assertions", *Under Review in IEEE Transactions on Information Forensics and Security* (**IEEE TIFS)**, arXiv:2306.14027, **Impact Factor: 7.3**.

[12] Baleegh Ahmad, **Shailja Thakur**, Benjamin Tan, Ramesh Karri, Hammond Pearce, "Fixing Hardware Security Bugs with Large Language Models", *Under Review in IEEE Transactions on Information Forensics and Security* (**IEEE TIFS)**, arXiv:2302.01215, **Impact Factor: 7.3**.

[13] **Shailja Thakur**, Sebastian Fischmeister, "TiME: Time-series based model outcome explanation", *Under Review in IEEE Transaction on Knowledge and Data Engineering* (**IEEE TKDE**), **Impact Factor: 8.9**.

IN PROGRESS

[14] **Shailja Thakur**, Ramesh Karri, Siddharth Garg, "Training and Evaluating instruction-tuned LLMs for generating and repairing functionally correct Verilog Code."

RESEARCH PROJECTS

Exploring Large Language Models for Hardware Design in Cyber-Physical Systems, New York University, NY 2022 – Present

Advisor: Dr. Ramesh Karri and Dr. Siddharth Garg

Hardware Description Languages (HDLs) like Verilog are fundamental for hardware design, but writing HDL code manually can be error-prone and time-consuming. My postdoctoral research has focused on harnessing Large Language Models (LLMs) like GPT-4 for automated hardware design.

I conducted the first comprehensive benchmarking of LLMs for generating Verilog code. Key challenges included the high error rate in code from baseline LLMs, the need for large Verilog datasets for fine-tuning, and the lack of effective evaluation methods for correctness. To address this, I developed VGen, fine-tuned on the largest corpus of Verilog code from open-source and textbooks. VGen achieves state-of-the-art performance in generating high-quality Verilog.

Expanding LLMs' capabilities, I also worked on applying them for automated hardware security *bug* detection and *repair* using natural language instructions. This led to a framework for finding and patching vulnerabilities. I also explored using LLMs to generate security assertions for hardware designs.

Looking ahead, I am currently working on harnessing the conversational feature of LLMs like ChatGPT to enable self-refining code generation through natural language dialog. I aim to develop a framework that embodies existing synthesis tools and performs in-context learning on LLMs to achieve fully automated hardware design through human-AI collaboration. Given a high-level conceptual description of a desired hardware system, the goal is for the LLM to generate complete, synthesizable Verilog code that can be taken forward for fabrication.

Security and Interpretability in Automotive Systems, University of Waterloo, Canada 2017 – 2022

Advisor: Dr. Sebastian Fischmeister

During my Ph.D., I addressed the critical challenge of enabling secure and interpretable decision-making in advanced safety-critical systems like autonomous systems. This is important for technologies like Cooperative Adaptive Cruise Control (ACC) that use continuous vehicle-to-vehicle communication to control speed. Even minor delays from malicious or accidental interference can lead to operator errors and alarm fatigue. My research focused on two areas: First, I developed machine learning-based defense mechanisms to improve automotive cybersecurity. Notably, I collaborated with GDLS Canada on a novel sender authentication technique for the Controller Area Network (CAN) bus protocol widely used in cars. This leverages Electronic Control Unit (ECU) power consumption data to verify message senders and was successfully tested to minimize false positives.
Second, I worked on an explainable AI techniques to interpret decision-making in safety-critical systems. Machine learning model opacity poses risks when integrated into critical contexts. To address this, I proposed a two-part solution: First, I created a non-intrusive saliency-map based method to identify crucial pixels for classification. Second, I introduced an algorithm to generate pixel variations in salient regions while retaining original model predictions, improving explanation stability.

Home energy disaggregation and apportionment using data from smart meters and sensors on smartphones, IIIT Delhi, India 2012 – 2014

Advisor: Dr. Amarjeet Singh

During my masters, I focused on the problem of energy disaggregation and apportionment in buildings. My first project, EnergyLens, addressed the challenge of detecting appliance-level activities and attributing energy consumption to individuals in homes. We developed a system that combined data from smart meters and sensors on smartphones like WiFi and microphone to infer *what* appliance was being used, *when*, *where* and by *whom*. Evaluations showed EnergyLens significantly improved detection accuracy compared to using just meter data. Building upon the insights from EnergyLens, the WattShare project extended the focus to multi-occupant commercial spaces like dormitories that use a single meter. Our algorithm leveraged smartphone sensors along with meter data to associate meter events with specific rooms and occupants. We validated the system through a week-long deployment in a student dormitory and achieved over 85% accuracy in room-level energy apportionment.

TEACHING EXPERIENCE

University of Waterloo, Waterloo, Ontario, Canada

- Teaching Assistant, System Programming and Concurrency (ECE 252)
  - Instructor: Jeffrey Zarnett
  - September - December, 2021
  - May - August, 2021
  - September - December, 2020
- Teaching Assistant, Data Knowledge and Modeling Analysis (ECE 657A)
  - Instructor: Dr. Mark Crowley
  - January - April, 2021
- Teaching Assistant, Real-time and Safety-critical Embedded Systems
  - Instructor: Dr. Sebastian Fischmeister
  - May - August, 2020 (ECE 455)
  - September - December, 2019 (ECE 652)

IIIT Delhi, Delhi, India

- Teaching Assistant, Operating Systems
  - Instructor: Dr. Pushpendra Singh
  - May - September, 2012

CONFERENCE PRESENTATIONS & RESEARCH SEMINARS

- *Automating Verilog RTL Code Generation with Large Language Models*, **AI and Security CoP, Intel**, New York 2023

- *LLMs for Automatically generating Verilog RTL*, Embedded Tutorial on EDA using Large Language Models, **MLCAD**, Snowbird, Utah 2023
- *LLMs for code completion and bug fixing in hardware*, **Research Symposium, NYU** 2023
- *Investigate Large Language Model for copyright infringement*, **Research Seminar, NYU Law** 2023.
- *Benchmarking Large Language Models for Automated Verilog RTL Code Generation*, **IEEE DATE** 2023
- *Security and Interpretability in Automotive Systems*, **IEEE DATE** 2023
- *A generalizable saliency map-based interpretation of model outcome*, **IEEE ICPR** 2020
- *CAN Origin Authentication Through Power Side-Channel Monitoring*, **Graduate Research Seminar, University of Waterloo**, 2019
- *WattShare: Detailed Energy Apportionment in Shared Living Spaces within Commercial Buildings*, **ACM BuildSys** 2014

PROFESSIONAL ACTIVITIES

- Member, IEEE 2020 – Present
- Invited Reviewer, WiCV@ ICCV 2023
- Invidet Reviewer, AAAI 2023
- Invited Reviewer, EMNLP 2023
- Invited Reviewer, NeurIPS 2023

MENTORSHIP

- Student intern as part of NYU Abu Dhabi summer exchange program 2022
- Women in Cybersecurity, Summer school program, K12 STEM education by NYU, NY 2023
- Westinghouse High School Computer Science Student Visit, K12 STEM education by NYU, NY 2022
- Catalyst Online Workshop - Machine Learning, University of Waterloo, Canada 2020
- Catalyst Workshop - Machine Learning, University of Waterloo, Canada 2019

OTHER ACADEMIC ACTIVITIES

- NSF Workshop on Shared Infrastructure for ML EDA, Minneapolis 2023
- Autonomous Systems Design Initiative Workshop at DATE, Antwerp, Belgium 2023
- Workshop on Sustainable Hardware Security (SUSHI'22), ICCAD, San Diego 2022
- Organizing member, Cyber Security Awareness Week (CSAW), NYU 2022
- Stanford Computer Forum - Graph Learning Workshop, Virtual, September 2021
- Inria-DFKI European Summer School on AI (Trustworthy AI), July 2021
- ICPR Conference, Milan, Italy 2020
- Expectation Teaching Assistant Workshop, Waterloo 2019
- Waterloo ML, Security, and Verification Workshop 2019
- Volunteer at CPS 2019, Montreal 2019
- TU Automotive Conference, Detroit, Michigan 2016

INDUSTRY EXPERIENCE

Acerta Analytics, Waterloo, Canada

- Data Science Intern 2017
  - Manager: Jean-Christophe Petkovich
  - Worked on a project to detect anomaly in Engine, Transmission, and Anti-Lock Brake Systems on Chrysler dataset. In addition, I worked on acoustic sensor data from windmill by ZF Manufacturers for anomaly detection.

IBM, Gurgaon, India

- Data Science Consultant 2015 – 2017
  - Manager: Mohan Jayaraman
  - Participated in projects in collaboration with GM, Chrysler, and Fiat, including anomaly detection for fault detection and diagnosis, usage based insurance by scoring driving behavior using features captured by Onboard diagnostic (OBD) tool.

U2opia Mobile, Gurgaon, India

- Software Engineer 2014 – 2015

- Manager: Rajesh Vashistha
- Developed USSD-based Twitter mobile application, performed sentiment analysis on tweets, and worked on user recommendation.

IIIT Delhi, India

- Research Intern 2013
  - Manager: Dr. Amarjeet Singh
  - Developed an Android application for data collection from various interfaces for indoor localization and user context analysis.

REFERENCES

References available upon request.