# Soneya Binta Hossain

Computer Science. University of Virginia
85 Engineer's Way, Charlottesville, VA 22903
sh7hv@virginia.edu https://github.com/soneyahossain/
http://www.linkedin.com/in/Soneya
https://soneyahossain.github.io/

***Leveraging program analysis and AI/ML in developing methods for automated software testing, debugging and repair to ensure reliability.***

## Education

2019 – present **Ph.D., Computer Science,** LESS Lab, University of Virginia
Advisor: Matthew Dwyer
Committee: Sebastian Elbaum, Yangfeng Ji, Matthew Bolton, Antonio Filieri
Thesis: *Assessing and Improving Critical Properties of Test Oracles for Effective Bug Detection.*

2024 **Master of Computer Science (MCS),** University of Virginia.
Project: *TOGLL: Correct and Strong Test Oracle Generation with LLMs.*

2016 **B.Sc., Computer Science and Engineering,** Bangladesh University of Engineering and Technology (BUET).
Thesis: *Balanced Coverage in Fault-Tolerant Broadcasting for Wireless Multi-hop Networks.*

## Research Publications

*8 peer-reviewed papers: 6 published in top-tier software engineering and machine learning conferences (ICSE, FSE, NeurIPS, etc.), including two published artifacts. Six first authored, three in collaboration with Amazon Web Services (AWS). 2 papers are currently under submission, one of which is co-authored with an undergraduate mentee.*

### Published Papers (Peer-Reviewed):

1. **S. B. Hossain**, N. Jiang, Q. Zhou, X. Li, W.-H. Chiang, Y. Lyu, H. Nguyen, and O. Tripp, "A deep dive into large language models for automated bug localization and repair," **FSE**, vol. 1, Porto de Galinhas, Brazil: Association for Computing Machinery, Jul. 2024. DOI: 10.1145/3660773, **acceptance rate:** 25.5%.

2. **S. B. Hossain**, "Ensuring critical properties of test oracles for effective bug detection," in *Proceedings of the 2024 IEEE/ACM 46th International Conference on Software Engineering: Companion Proceedings*, ser. **ICSE-Doctoral Symposium (DS) '24**, Lisbon, Portugal: Association for Computing Machinery, 2024, pp. 176–180, ISBN: 9798400705021. DOI: 10.1145/3639478.3639791, **acceptance rate:** 57%.

3. **S. B. Hossain**, A. Filieri, M. B. Dwyer, S. Elbaum, and W. Visser, "Neural-based test oracle generation: A large-scale evaluation and lessons learned," in *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ser. **ESEC/FSE 2023**, San Francisco, CA, USA: Association for Computing Machinery, 2023, pp. 120–132, ISBN: 9798400703270. DOI: 10.1145/3611643.3616265, **Artifacts available** at https://doi.org/10.6084/m9.figshare.21973091.v4, **acceptance rate:** 21%.

4. **S. B. Hossain**, M. B. Dwyer, S. Elbaum, and A. Nguyen-Tuong, "Measuring and mitigating gaps in structural testing," in *Proceedings of the 45th International Conference on Software Engineering*, ser. **ICSE '23**, Melbourne, Victoria, Australia: IEEE Press, 2023, pp. 1712–1723, ISBN: 9781665457019. DOI: 10.1109/ICSE48619.2023.00147, **Artifacts available and reusable** at https://doi.org/10.6084/m9.figshare.21950552, **acceptance rate:** 26%.

5. **S. B. Hossain** and M. Dwyer, "Togll: Correct and strong test oracle generation with llms," in *Proceedings of the 47th International Conference on Software Engineering* (***ICSE'25***), 2025. DOI: 10.48550/arXiv.2405.03786.

6. N. Jiang, X. Li, S. Wang, Q. Zhou, **S. B. Hossain**, B. Ray, V. Kumar, and X. Ma, "Training llms to better self-debug and explain code," in *Proceedings of the 38th Conference on Neural Information Processing Systems* (***NeurIPS***), 2024. DOI: 10.48550/arXiv.2405.18649, **acceptance rate:** 25.8%.

### In Submission:

1. **S. B. Hossain**, R. Taylor, and M. Dwyer, "Doc2oracle: Investigating the effect of javadoc comments on test oracle generation," **(In Submission)**.
2. W. Leeson, **S. B. Hossain**, and M. Dwyer, "Hybrid predictive mutation testing via graph transformer networks," **(In Submission)**.

## Awards and Scholarships

### Research Awards

2024 **Copenhaver Charitable Trust Bicentennial Fellow**, UVA School of Engineering & Applied Science (SEAS), goes to only 20-25 outstanding doctoral students in SEAS. Awarded $12,000.

**Finalist**, Physical Sciences and Engineering category, UVA Research Computing Exhibition, to showcase research using UVA's high-performance computing resources, Rivanna.

2023 **The John A. Stankovic Outstanding Research Award**, UVA CS Department, goes to 2-3 CS Ph.D. students who have demonstrated excellence in research during the academic year. Awarded $500.

**Student Choice Best Research Award**, UVA CS Research Symposium.

2021 **Best Research Poster Award**, UVA CS Research Symposium.

2019 **PhD Fellowship**, UVA CS.

2016 **Outstanding Undergraduate Thesis Award**, BUET CSE.

### Service Awards

2022 **Outstanding Service Award**, UVA CS. Awarded $500.

### Travel Grants

2024 Society of Women Engineers, to attend **SWE24**—the world's largest conference for women in engineering and technology, Chicago, IL.

2023 Computing Research Association **(CRA)**, to attend Grad Cohort, San Francisco, CA.

2020 Computing Research Association **(CRA)**, to attend Grad Cohort, New Orleans, LA.

2015 Grace Hopper Celebration of Women in Computing India **(GHCI)**, to travel and present research, Bangalore, India.

### Undergraduate Academic Awards

2013-2014 Two-time **University Merit Award**, BUET.

**Dean's List Award** (Level-2), BUET.

2014 **Top Ten Database Project Award**, CSE, BUET.

## Teaching and Mentoring

2021 - 2023 **Graduate Teaching Assistant,** Computer Science, University of Virginia.

- Three-time TA for **Undergraduate Compilers (CS 4620)** advised by Matthew Dwyer. **Course website:** `https://matthewbdwyer.github.io/4620/`
- One-time TA for **Graduate-level Compilers (CS 6620)** advised by Matthew Dwyer. **Course website:** `https://matthewbdwyer.github.io/6620/`
- Designed, developed, and tested features for the TIPC compiler (`https://github.com/matthewbdwyer/tipc`); graded assignments and final projects for 35–70 students; and assisted students through weekly office hours and by answering questions on Slack, Piazza, and email.

## Teaching and Mentoring (continued)

2020 – present

**Undergraduate Mentoring**: Mentored more than 10 undergraduate students from five different universities in the USA and Bangladesh. Many students are from underrepresented groups.

- **Colin Henry** (CS, UVA). A 3rd year undergrad, working on the LLM-based test oracle generation research.
- **Raygan Taylor** (DU) and **Javan Mendoza** (UMBC) interned at the UVA LESS lab in Summer 2024; worked on a project that investigates the impact of code documentation quality on automated test oracle generation; a paper is currently under review.
- **Nicki Choquette** and **Kasra Lekan** (CS, UVA). Helped them conducting a thorough replication study of an FSE'22 paper (`https://doi.org/10.1145/3540250.3549086`) in Spring 2023. Their paper *Insight into SEER*, consisting of the replication results is available on arXiv (`https://doi.org/10.48550/arXiv.2311.01164`).
- **Ashley Hart** (CS, UCF) interned at the UVA LESS lab in Summer 2020; as one of her mentors, I helped her learn about graduate school expectations and preparation, assisted in developing a Boolean satisfiability solver, and guided her in technical writing and presentation. She is currently pursuing her PhD at the University of Florida.
- **Srikar Chittari** (CpE, UVA) and **Eric Weng** (CS, UVA). Assisted with their graduate school applications, Srikar is now pursuing Master's in CS at UVA, Eric is continuing his undergraduate study.
- Mentor at **BWCSE** (Bangladeshi Women in Computer Science and Engineering). Currently mentoring five undergraduate juniors and seniors in research and graduate school preparation.

## Service

2024

**Conference and Journal Reviewer**

- **Reviewer**, ACM Transactions on Software Engineering and Methodology (**TOSEM**).
- **Program committee (research track)**, IEEE International Conference on Software Testing, Verification and Validation (**ICST**) 2024.

2023 – 2024

**Conference Volunteer**

- **Recruiter**, SWE24, Chicago, IL (Oct 24–26, 2024). Assisted UVA Engineering with booth setup, recruiting during career fair, engaging with attendees, and networking with professionals to promote UVA's engineering programs.
- **Student volunteer**, ESEC/FSE 2023, San Francisco, CA (Dec 3-9, 2023). Assisted with registration, badge verification, presentation testing, A/V setup for talks, and event preparation.

2020 – present

**CS Department Volunteer**

- **Leadership chair** (Jan 2020 - Jan 2022), Computer Science Graduate Student Group (CS-GSG). Planned and organized CS Research Symposium in 2020 and 2021; organized events to promote leadership among graduate students.
- Assisted the CS department with faculty recruitment and coordinated graduate student visits for several years.

## Service (continued)

2021 – 2022 **Community Volunteer**

- **Social chair**, Association of Bangladeshi Students (ABS) at UVA. Advocated for Bangladeshi students at UVA and coordinated events showcasing their culture and history to promote diversity.

## Research Experience

**Graduate Research Assistant,** LESS Lab, University of Virginia (2019 - present)

- Advised by Matthew Dwyer
- Conducted research on software testing and verification funded by DARPA, Lockheed Martin Advanced Technology Laboratories and Air Force Office of Scientific Research. Published research papers and reusable artifacts in top-tier software engineering conferences.
- Mentored undergraduate research interns at LESS lab as part of the UVA Leadership Alliance Summer Research Program.

**Applied Scientist Intern (summer'23),** AWS CodeCatalyst

- Mentored by Qiang Zhou. Conducted research on leveraging LLMs to detect and repair bugs. **This research is accepted at FSE 2024.**
- Collaborated with AWS AI Labs to investigate LLMs' self-debugging and explanation capabilities in code generation. **This work is accepted at NeurIPS 2024.**

**Applied Scientist Intern (summer'22),** AWS CodeGuru

- Mentored by Antonio Filieri and Willem Visser. Conducted research on assessing state-of-the-art automated test oracle generation methods to evaluate their potential integration with the CodeWhisperer tool. **The research results were published as a full research paper at FSE 2023.**
- Developed MuSlicer, a dynamic program slicing tool using AWS's proprietary MuGraph to extract data and control dependency from dynamic execution trace to compute test oracle coverage.

## References


**Dr. Matthew B. Dwyer**
Professor, Department of Computer Science
University of Virginia
Email: matthewbdwyer@virginia.edu
Website: `https://matthewbdwyer.github.io/`
Phone: (434) 243-5206

**Dr. Sebastian Elbaum**
Professor, Department of Computer Science
University of Virginia
Email: selbaum@virginia.edu
Website: `http://www.cs.virginia.edu/~se4ja/`
Phone: 434-243-5213