Shoumik Saha

shoumiksaha.github.io | smksaha@umd.edu | in/shoumik-saha
/ShoumikSaha | Google Scholar

RESEARCH INTEREST

- **Machine Learning** - Security & Reliability of GenAI, LLM Alignment & Hallucination
- **Computer Security** - Adversarial Attacks & Defenses, Malware

EXPERIENCE

Applied Scientist Intern — **May 2025 – Aug 2025**
Amazon AWS — *Full Time*
Managers - Varun Kumar, Zijian Wang

Graduate Research Assistant — **Aug 2023 – Present**
Department of Computer Science — *Full Time*
University of Maryland
Supervisor - Dr. Soheil Feizi

Applied Scientist Intern — **Jun 2024 – Aug 2024**
Amazon AWS — *Full Time*
Manager - Zijian Wang

Graduate Research Assistant — **Aug 2022 – Jul 2023**
Maryland Cybersecurity Center — *Full Time*
Supervisor - Dr. Tudor Dumitras

Lecturer — **Jul 2021 – Jul 2022**
Department of Computer Science and Engineering — *Full Time*
United International University

Research Assistant — **Mar 2021 – Jul 2022**
Data Science and Engineering Research Lab — *Part Time*
Bangladesh University of Engineering and Technology (BUET)

EDUCATION

University of Maryland - College Park (UMD) — **Aug 2022 – Present**
Ph.D. in Computer Science (Ongoing) — *CGPA – 3.814/4.0*
Advisor – Dr. Soheil Feizi

University of Maryland - College Park (UMD) — **Aug 2022 – Aug 2024**
M.S. in Computer Science
Advisor – Dr. Soheil Feizi

Bangladesh University of Engineering and Technology (BUET) — **Mar 2016 – Feb 2021**
B.Sc. in Computer Science and Engineering — *Advanced GPA – 3.79/4.0*
Thesis supervisor – Dr. Atif Hasan Rahman

PUBLICATIONS

- **Under the Hood of SKILL.md: Semantic Supply-chain Attacks on AI Agent Skill Registry**
  [**Shoumik Saha**, K Faghih, S Feizi]
  Under Review(arXiv link)
  Media Coverage: The Register

- **Breaking the Code: Security Assessment of AI Code Agents Through Systematic Jailbreaking Attacks**
  [**Shoumik Saha**, J Chen*, S Mayers*, SK Gouda, Z Wang, V Kumar]
  Under Review(arXiv link)

- **Almost AI, Almost Human: The Challenge of Detecting AI-Polished Writing**
  [**Shoumik Saha**, S Feizi]
  ACL 2025 (*Association for Computational Linguistics*) (Paper Link)
  Media Coverage: The New York Times, Plagiarism Today, The Science Archive, The Cheat Sheet

- **Adversarial Paraphrasing: A Universal Attack for Humanizing AI-Generated Text**
  [Y Cheng, VS Sadasivan, **Shoumik Saha***, M Saberi*, S Feizi]
  NeurIPS 2025 (*Conference on Neural Information Processing Systems*) (Paper Link)

- **ProcVQA: Benchmarking the Effects of Structural Properties in Mined Process Visualizations on Vision–Language Model Performance**
  [KT Zinat*, SM Abrar*, **Shoumik Saha**, S Duppala, SN Sakhamuri, Z Liu]
  EMNLP 2025 (*Empirical Methods in Natural Language Processing*) (Paper Link)

- **LLM-Check: Investigating Detection of Hallucinations in Large Language Models**
  [G Sriramanan, S Bharti, VS Sadasivan, **Shoumik Saha**, P Kattakinda, S Feizi]
  NeurIPS 2024 (*Conference on Neural Information Processing Systems*) (Paper Link)

- **Fast Adversarial Attacks on Language Models In One GPU Minute**
  [VS Sadasivan, **Shoumik Saha***, G Sriramanan*, P Kattakinda, A Chegini, S Feizi]
  ICML 2024 (*International Conference on Machine Learning*) (Paper Link)
  Media Coverage: The Register

- **DRSM: De-Randomized Smoothing on Malware Classifier Providing Certified Robustness**
  [**Shoumik Saha**, W Wang, Y Kaya, S Feizi, T Dumitras]
  ICLR 2024 (*International Conference on Learning Representations*) (Paper Link)

- **MAlign: Explainable Static Raw-byte Based Malware Family Classification using Sequence Alignment**
  [**Shoumik Saha**, S Afroz, A Rahman]
  Computers & Security Journal (Paper Link)

- **ML-Based Behavioral Malware Detection Is Far From a Solved Problem**
  [Y Kaya, Y Chen, Marcus Botacin, **Shoumik Saha**, F Pierazzi, L Cavallaro, D Wagner, T Dumitras]
  SaTML 2025 (*Secure & Trustworthy Machine Learning*) (Paper Link)

- **Contrastive Self-Supervised Learning Based Approach for Patient Similarity: A Case Study on Atrial Fibrillation Detection from PPG Signal**
  [**Shoumik Saha***, S Shanto*, A Rahman, M Masud, ME Ali]
  (arXiv Link)

SCHOLARSHIPS & AWARDS

- Awarded *'Dean's Fellowship'* from the University of Maryland
- Awarded *'Innovation Fund'* for research from the ICT division of Bangladesh government
- Achieved *'Merit Stipends'* from BUET in five out of seven terms
- Awarded *'Dean's Award'* in Junior year for extra-ordinary result
- Achieved *'Talent-Pool Scholarship'* in High School

TEACHING

Teaching Assistant
University of Maryland *Fall 2023*
Data 200: Knowledge in Society: Science, Data and Ethics

Lecturer
United International University *Fall 2021 - Summer 2022*
Discrete Mathematics, Data Structure and Algorithm, Operating Systems

SERVICES

- **Reviewer**, top-tier ML conferences such as ICML 2024, ICLR 2025, NEURIPS 2025, etc.
- **External Reviewer**, top-tier Computer Security conference, including USENIX 2023
- **Lab Instructor**, United International University
- **General Secretary**, BUET Photographic Society