

# Sifat Muhammad Abdullah

+1 (540)-449-2710 | [sifat@vt.edu](mailto:sifat@vt.edu) | <https://sifatmd.github.io> | [Google Scholar](#)

## EDUCATION

---

**PhD in Computer Science**, Virginia Tech, advisor: Dr. Bimal Viswanath      **1/2021 - expected 4/2026**

**BS in Computer Science and Engineering**, BUET (CGPA: 3.91/4.0)      **2/2015 - 4/2019**

## RESEARCH INTERESTS

---

Security and Adversarial Robustness of Large Multimodal Models, LLMs & Generative AI Defenses, Improving Vision models with better Content Semantics understanding using Multimodal Foundation models, toxicity mitigation in Large Language Models.

## PUBLICATIONS

---

- **Sifat Muhammad Abdullah**, Aravind Cheruvu, Shravya Kanchi, Taejoong Chung, Peng Gao, Murtuza Jadliwala, Bimal Viswanath. “An Analysis of Recent Advances in Deepfake Image Detection in an Evolving Threat Landscape” *IEEE Symposium on Security and Privacy (S&P)*, 2024.
- Aravind Cheruvu(co-lead), Connor Weeks(co-lead), **Sifat Muhammad Abdullah**, Shravya Kanchi, Danfeng Yao, Bimal Viswanath. “A First Look at Toxicity Injection Attacks on Open-domain Chatbots” *Annual Computer Security Applications Conference (ACSAC)*, 2023.
- Jiameng Pu(co-lead), Zain Sarwar(co-lead), **Sifat Muhammad Abdullah**, Abdullah Rehman, Yoonjin Kim, Parantapa Bhattacharya, Mobin Javed, Bimal Viswanath. “Deepfake Text Detection: Limitations and Opportunities” *IEEE Symposium on Security and Privacy (S&P)*, 2023.
- Md Ashiqur Rahman (co-lead), Abdullah Aman Tutul(co-lead), **Sifat Muhammad Abdullah**(co-lead), Md Shamsuzzoha Bayzid. “CHAPAO: Likelihood and hierarchical reference-based representation of biomolecular sequences and applications to compressing multiple sequence alignments” *PLOS ONE Journal*, 2022.
- Shadman Saqib Eusuf, Kazi Ashik Islam, Mohammed Eunus Ali, **Sifat Muhammad Abdullah**, Abdus Salam Azad. “A Web-Based System for Efficient Contact Tracing Query in a Large Spatio-Temporal Database” *Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL)*, 2020.

## WORK EXPERIENCE

---

**Graduate Research Assistant | Virginia Tech, SecML Lab**

**1/2022 - Present**

- Studying robustness of a suite of adversarial attacks against LMMs, e.g., Gemini Pro Vision, GPT-4V, LLaVA, MiniGPT-4 using Flux and Stable Diffusion text-to-image generation (T2I) models.
- Analyzed robustness of 8 state-of-the-art deepfake image detectors by developing practical & low-cost adversarial attacks, achieving more than 70% performance (recall score) degradation, using Stable Diffusion and StyleGAN-based text-to-image (T2I) generators with LoRA fine-tuning and multimodal foundation models.
- Performed toxicity injection attacks on BART and BlenderBot chatbots after deployment in a Dialog-based learning setup, eliciting up-to 60% response toxicity rate by building adversarial attacks using GPT-J model.
- Evaluated state-of-the-art deepfake text detectors, e.g., BERT & GPT-2 based defenses, on our collected real-world datasets, and achieved up-to 91.3% evasion rate by crafting high-probability token replacement using public LLMs without any query to surrogate or victim defenses.

**Graduate Research Assistant | BUET, DataLab**

**1/2020 - 12/2020**

- Developed highly efficient web-based contact tracing query system to locate COVID-19 patients utilizing QzR-tree with PostgreSQL database.

- Built a chatbot system for company website using BART with PyTorch and Django framework.

TEACHING EXPERIENCE

---

- Conducted office hours, programming labs, and graded assignments for undergraduate courses.

ACHIEVEMENTS

---

- CCI SWVA Cyber Innovation Scholarship 2024 - 2025
- CCI Research Showcase 6/2024
- Invited Talk: VT Skillshop Series: Leveraging Creative Technologies 10/2023
- *The Dark Side of AI* - VPM News Focal Point 10/2023
- CCI Student Spotlight 2023
- *The Rise of the Chatbots* - Communications of the ACM 7/2023
- *The strengths and limitations of approaches to detect deepfake text* - TechXplore 11/2022
- BUET Dean's List Award 2015 - 2019

SKILLS

---

- **GenAI Technologies:** LMMs/VLMs, LLMs, T2I models, LoRA, Foundation Model Fine-tuning
- **Languages:** Python, C/C++, Bash, Java, JavaScript, Assembly
- **Frameworks:** PyTorch, TensorFlow, Keras, Django
- **Libraries:** Scikit-learn, NumPy, pandas, Matplotlib
- **Developer Tools:** Git, Vim, Jupyter Notebook, VS Code, Markdown, LaTeX, Linux, Docker