Md Rayhanul Masud

PhD Candidate, Department of Computer Science, University of California, Riverside

(+1) (951) 275-1515 | mmasu012@ucr.edu | Personal Website | Google Scholar | LinkedIn

Intro

I am a PhD candidate in Computer Science at UC Riverside with three years of industry experience. My research integrates software engineering, cybersecurity, and data science to study how people behave and misbehave in open-source ecosystems like GitHub. I analyze these social and technical patterns to design tools that make open-source development safer and more transparent, and easier to explore. Alongside research, I have served as a teaching assistant and mentor, fostering inclusive, student-centered learning environments and supporting undergraduate success in computer science.

Education

University of California, Riverside (UCR) — Sep 2021 – Jun 2026 (exp.)
Ph.D. in Computer Science — Advisor: Prof. Michalis Faloutsos

University of California, Riverside (UCR) — Sep 2021 – Dec 2024
M.S. in Computer Science — CGPA: 3.90/4.00

Bangladesh University of Engineering and Technology (BUET) — Apr 2012 – Feb 2017
B.Sc. in Computer Science and Engineering — Advisor: Prof. Md Monirul Islam

Research Experience

Graduate Student Researcher — Sep 2021 – Present
University of California, Riverside — Advisor: Prof. Michalis Faloutsos

- **GeekMAN: Tracing Technogeek Identities (2022–2023):** Developed a systematic approach to link stylized usernames (e.g., leetspeak) across forums, GitHub, and social platforms; significantly improved cross-platform identity linkage on technogeek datasets.
- **MetaSim & RepoScope: Search-by-Repository for GitHub (2023–2024):** Built interpretable similarity signals from repository metadata (descriptions, topics, README) and expanded to multi-level exploration with code embeddings and visual clustering for "search by example repository."
- **Malware Ecosystem on GitHub (2025–Present):** Analyzing technical and social mechanisms behind the creation, forking, and maintenance of malicious-code repositories to understand their evolution and persistence in the open-source ecosystem.
- **Orphy — Vulnerability Propagation in Forks/Clones (2025–Present):** Studying how vulnerabilities become "orphaned" when patches fail to propagate to downstream forks; combining social-network and software-artifact data to model the adoption (or non-adoption) of fixes.

Publications

Journal Papers

- **M. R. Masud**, B. Treves, M. Faloutsos. "Disambiguating usernames across platforms: the GeekMAN approach." *Social Network Analysis and Mining (SNAM)*, 2024. [Link]

  Description: Presented a human-inspired method for detecting similar usernames across platforms, focusing on technogeek communities to identify malicious actors. [Platform]

- **M. R. Masud**, Q. Zhang, M. Faloutsos. "RepoScope: A Multi-Granularity Query-by-Example Framework for Github Repository Search." *Under submission.*

Conference Papers

- **M. R. Masud**, M. O. F. Rokon, Q. Zhang, M. Faloutsos. "MetaSim: A search engine for finding similar GitHub repositories." *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2024. [Link] [Demo]

  Description: Implemented a search engine to find functionally similar GitHub repositories using interpretable metadata features. [Platform]

- **M. R. Masud**, B. Treves, M. Faloutsos. "GeekMAN: Geek-oriented username Matching Across online Networks." *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2023. [Link] [Presentation]

  Description: Introduced a method for matching stylized usernames, improving identity linkage on technogeek datasets.

- N. A. Tania, **M. R. Masud**, M. O. F. Rokon, Q. Zhang, M. Faloutsos. "Who is Creating Malware Repositories on GitHub and Why?" *The Web Conference (WWW)*, 2024. [Link]

  Description: Developed a systematic approach to understand the origins and motivations behind the creation of malware repositories on GitHub.

- B. Treves, **M. R. Masud**, M. Faloutsos. "RURLMAN: Matching Forum Users Across Platforms Using Their Posted URLs." *ASONAM*, 2023. [Link]

  Description: Utilized an ensemble of traditional and emerging methods to identify off-site accounts belonging to forum users.

- B. Treves, **M. R. Masud**, M. Faloutsos. "URLytics: Profiling Forum Users from their Posted URLs." *ASONAM*, 2022. [Link]

  Description: Developed a URL-based user profiling mechanism to extract maximal information about users.

- M. O. F. Rokon, R. Islam, **M. R. Masud**, M. Faloutsos. "PIMan: A Comprehensive Approach for Establishing Plausible Influence among Software Repositories." *ASONAM*, 2022. [Link]

  Description: Proposed a graph and embedding-based approach to analyze the influence among software repos.

Posters

- **M. R. Masud**, M. Faloutsos. "Unveiling A Hidden Risk: Exposing Educational but Malicious Repositories in GitHub." *SoCal Data Science Day, KDD*, 2023. [PDF]

  Description: Proposed a method to identify "Educational but Malicious" GitHub repos using large language models.

Teaching & Mentoring

Teaching Experience

Teaching Assistant / Head TA Oct 2024 – Present
University of California, Riverside

- **Head Teaching Assistant** (Fa'25): "Introduction to CS for Science, Math, & Engineering." Led a team of five TAs, coordinated grading rubrics, and aligned explanations for challenging topics.
- **Teaching Assistant** (Fa'24, Wi'25, Sp'25): "Introduction to CS for Science, Math, & Engineering."
- **Teaching Assistant** (Sp'25, Su'25): "Theory of Automata and Formal Languages."

Mentoring

- Mentored two first-year PhD students and one master's student.

Teaching Philosophy

My teaching philosophy centers on building connection and trust. I strive to help students see they are capable of learning new, challenging material by fostering an inclusive and approachable environment where questions are encouraged and exploration is valued over perfect answers.

Awards and Honors

- Dean's Fellowship Award, University of California, Riverside — Sep 2021
- Dean's Award (Level-4, Term-1), BUET, Bangladesh — Jul 2016

Industry Experience

Application Developer — Jul 2020 – Aug 2021
CIRT, Bangladesh

- Contributed to the front-end of Boithok, a video conferencing platform used by the ICT ministry of Bangladesh.
- Automated the server deployment pipeline, which reduced system downtime.

Software Engineer — Jul 2018 – Jul 2020
Kona Software Lab Ltd., Bangladesh

- Developed the transaction management pipeline for Kona Blockchain Platform (a proof-of-concept) built on Ethereum.
- Introduced template-based smart-contract deployment, which reduced insecure transaction requests from client-side applications.
- Collaborated in an agile, product-focused team to prototype scalable blockchain solutions.

Technical Skills

Languages Python, Java, JavaScript, C++

Data/ML PyTorch, scikit-learn, Pandas, PySpark, NLTK, spaCy

Databases MySQL, NoSQL

Systems & Cloud Docker, Git, GCP, AWS

Web ReactJS, HTML, CSS

Selected Academic Projects

- **Adversarial Prompt Generation for LLMs:** Employed a morse-code-embedded adversarial prompt generation approach to jailbreak the safety alignment of LLaMA-2-7B-chat, achieving an 82.1% success rate on the AdvBench Dataset. [PDF]
- **Popularity & Influence on GitHub:** Explored the popularity of GitHub repositories based on author and repository-centric metadata; created a user-follower network of 3.5K users and used the HITS algorithm to model influence. [Code]
- **Search Engine for Twitter Data:** Designed and implemented a search engine for 2.5GB of Twitter data, comparing the performance of Hadoop and Lucene for retrieval. [PDF]

Service & Activities

- **Student Volunteer:** KDD 2023.
- **Community Contributor:** Ethereum Stack Exchange (500+ reputation, reaching ∼32K users).