Sonjoy Kumar Paul

skpaul@tamu.edu | sonjoy-k-paul | https://sonjoyp.github.io/ | sonjoyp | +1 (979) 402-9790

Summary

Third-year Ph.D. student at Texas A&M specializing in **Cybersecurity, Cognitive Security, and Brainwave Privacy**. Published at **USENIX Security** and **APWG eCrime**, with 5 additional papers under review. 5 years of professional experience as a C++ software engineer in large-scale financial software (2M+ LOC) used by Fortune 500 companies. Strong expertise in algorithms, applied machine learning, and systems security research.

Education

- **Texas A&M University** — College Station, Texas
  Ph.D. in Computer Science, *SPIES Research Lab, Advisor: Dr. Nitesh Saxena* — January 2024 – Present
- **Bangladesh University of Engineering and Technology (BUET)** — Dhaka, Bangladesh
  B.Sc. in Computer Science and Engineering — February 2013 - September 2017

Technical Skills

Machine Learning (ML), Deep Learning (DL), Large Language Models (LLMs); Programming Languages: C, C++, Java, Python, JavaScript; Databases: Relational (RDBMS), NoSQL (Apache Cassandra), Object-Oriented; Cloud & Big Data: Amazon Web Services (AWS), Apache Hadoop, Apache Spark; Frameworks & Tools: Spring Boot, Serverless Computing, Docker, Kubernetes; ML/AI Libraries: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas; Version Control & Systems: Git, Linux/Unix.

Experience

Graduate Assistant - Research — January 2024 - Present
Texas A&M University — College Station, Texas

- Developed **Cogni-Watch**, a cognitive-aware security framework detecting warning miscomprehension in real time via gaze and cursor dynamics (under review).
- Co-authored SoK studies on **AI-based phishing detection** (USENIX Security 2026) and **toll scam infrastructure** (eCrime 2025).
- Conducted EEG-based privacy attack research (**BrainPhone**): demonstrated passive speech/digit inference and identity snooping via consumer BCI headsets (under review).
- Discovered three **novel QR-/phone-based device-linking attacks** on E2EE messengers (WhatsApp, Signal, Telegram); proposed *Passkey-Anchored Linking* defense (under review).
- Co-authored **FuzzCheck.AI SoK**: evaluated 38 DNN fuzzers using six security metrics, revealing systemic gaps in high-severity failure discovery, reproducibility, and attack transferability (under review).
- Co-developed **PHANTOM**, a multimodal CogSec framework linking cognitive fatigue to phishing susceptibility via EEG, eye-tracking, and behavioral signals; showed strong behavioral degradation but non-generalizable neuro-signatures (under review).

Software Developer → Software Engineer → Senior Software Engineer — January 2018 – December 2022
CodeCrafters International Ltd. / Investortools, Inc. — Dhaka, Bangladesh

- Contributed to a **2M+ LOC C++ financial platform** used by top US banks for portfolio management and credit analysis.
- Improved core components including the **expression parser, reporting engine, and template framework**, enhancing performance and extensibility.
- Designed and led development of a **user-folder migration system** adopted by 200+ enterprise clients with zero workflow disruption.
- Built a **configurable reporting framework** supporting large-scale batch processing and automation.
- Conducted peer code/design reviews and mentored junior engineers, ensuring scalable system architecture.

Selected Projects

- **TANGO: Translation-Based Data Augmentation for Low-Resource NLP:** Conducted **7.8M-sentence multilingual augmentation** across 65 languages. Proposed LLM-based evaluation (Llama-3.2-1B) for semantic preservation, error severity, and diversity, validated with human studies.
- **Malware Detection & Adversarial Attack Challenge:** Developed ML models for malware detection and generated adversarial malware variants to evaluate robustness. Worked in both defensive and offensive roles, simulating real-world attacker–defender dynamics.
- **Improving CLIP Training for Vision–Language Models:** Benchmarked optimizers (Adam, AdamW, SGD) and advanced losses (CLIP, CyCLIP, SogCLR, VICReg). Achieved best retrieval/classification with **AdamW and CyCLIP**, surpassing baseline benchmarks.
- **PneumoniaMNIST Robust Classification with LibAUC:** Trained ResNet18 variants with AUROC/AUPRC-optimized loss functions and data augmentation. Compared 12 setups, improving robustness and generalization.
- **ResNet Variants on CIFAR-10:** Implemented CIFAR-10 loader (no torchvision) and trained ResNet18 across three settings: baseline, no residuals, and no batch norm. Analyzed component contributions to convergence and accuracy.
- **Software Security:** Hands-on labs in **buffer overflow & ROP, format string vulnerabilities, fuzz testing, symbolic execution, and malware analysis**. Implemented defensive techniques: bound checking, control flow integrity (CFI), and software fault isolation (SFI).

Publications

- M. Alam, A. Hussain, **S. K. Paul**, A. W. Hays, M. I. Huq, N. Saxena. "SoK: PHILTER: Uncovering Security and Functional Gaps in AI-based Phishing Website Detection Literature via an LLM-based Reasoning Framework" In the 35th **USENIX Security Symposium**, Aug. 2026.
- M. A. Munny, M. Alam, **S. K. Paul**, D. Timko, M. L. Rahman, and N. Saxena, "Infrastructure Patterns in Toll Scam Domains: A Comprehensive Analysis of Cybercriminal Registration and Hosting Strategies," in 2025 **APWG Symposium on Electronic Crime Research (eCrime)**, Dec. 2025.