STEVENS
INSTITUTE OF TECHNOLOGY
1870

Ph.D. DISSERTATION DEFENSE

Candidate:	Devharsh Trivedi
Degree:	Doctor of Philosophy
School/Department:	Charles V. Schaefer, Jr. School of Engineering and Science / Department of Computer Science
Date:	Friday, April 26th, 2024
Time/Location:	09:00 – 11:00 AM EST / https://stevens.zoom.us/j/92805493147
Title:	Towards Efficient Security Analytics
Chairperson:	Dr. Nikos Triandopoulos, Department of Computer Science, Stevens Institute of Technology
Committee Members:	Dr. Hui Wang, Department of Computer Science, Stevens Institute of Technology Dr. Antonio Nicolosi, Department of Computer Science, Stevens Institute of Technology Dr. Cristina Comaniciu, Department of Electrical and Computer Engineering, Stevens Institute of Technology Dr. Aymen Boudguiga, CEA-List, Paris-Saclay University Dr. Nesrine Kaaniche, SAMOVAR, Telecom SudParis

ABSTRACT

Security analytics (a.k.a. log analysis) is a prominent information-based approach for defending enterprises against cyberattacks. Organizations often employ third-party Security Information and Event Management (SIEM) systems to collect, analyze, and manage log data that originate from various Security Analytics Sources (SAS) to identify and mitigate threats in real-time, investigate incidents, troubleshoot errors, and improve their security posture. However, the efficiency and effectiveness of an SIEM depend heavily on the quality and quantity of the provided logs, and adversaries often tamper with logs after an intrusion to cover their attack traces. Thus, protecting the confidentiality, integrity, and availability of logs, whether at rest, transit, or execution, becomes vital for the efficacy of any security analytics solution.

In this talk, I present quantum-safe cryptographic primitives to deliver novel solutions for enhancing security analytics in trusted and untrusted log-analysis environments. In a trusted environment, ***VaultBox*** detects and prevents log tampering and recuperates logs if lost or corrupted. ***VaultBox*** employs a novel forward-secure, data-agnostic, replicated, randomized, and rateless log-protection scheme that helps securely store and transmit logs from SAS to SIEM. In an untrusted environment, ***SigML*** and ***SigML++*** utilize Fully Homomorphic Encryption (FHE) for supervised binary log classification. Although FHE allows calculations on encrypted data, complex functions must be approximated. While ***SigML*** employs deterministic approximation (e.g., Chebyshev), ***SigML++*** employs a novel probabilistic polynomial approximation based on an Artificial Neural Network (ANN), further reducing approximation errors up to 15%.

I conclude my talk by discussing recent work that extends our solutions to Private Collaborative Machine Learning (PCML). ***SplitML*** is a secure, privacy-preserving framework that leverages multi-key FHE with Differential Privacy (DP) to infuse Federated Learning (FL) seamlessly with Split Learning (SL). ***SplitML***

reduces training time and improves inference accuracy through consensus while incurring minimal overheads. Finally, I will discuss future work to improve fairness in PCML settings.