Data Science-integrated Experiential Digital Forensics Training based on Real-world Case Studies of Cybercrime Artifacts

Summary
Performing manual evidence/data analysis, triage, and correlation is an incredibly time-consuming task for investigators, and law enforcement agencies are experiencing huge backlogs in digital forensics cases. Modern data science tools and techniques are increasingly being used to automate forensic data analysis tasks during an investigation, an effective strategy to both improve productivity and to increase the quality of the analysis. This training development project will help create a new generation of highly skilled forensic investigators that employ data science tools and techniques to reduce their manual work and address the data science skill gap in the current and future law enforcement workforce.

Problem addressed
Recent evidence strongly indicates that modern data science tools and techniques can automate several time-consuming manual tasks for evidence (data) analysis, such as triage and data correlation, and improve the quality of forensic data analysis. Curriculum materials in these areas are needed to train a new generation of highly skilled forensic investigators that employ data science tools and techniques to reduce their manual work, resulting in more effective, efficient, and complete forensic investigations.

Approach
This project will develop data science learning modules based on real-world criminal case studies in close collaboration with the Computer Evidence Recovery Section at the Virginia State Police. The modules will be designed to be engaging to students and portable to various digital forensic curricula. Each module will be based on a real-world case study and will contain scaffolding to make them easily approachable by students with diverse backgrounds. They will include relevant and engaging content (criminal scenarios, investigative goals, short videos for tools, and API demos) to learn relevant forensics and data science tools and techniques, and they will include a series of investigative questions leading the students to solve the case incrementally.

Anticipated Impact for DHS
This project will contribute to developing a new generation of highly skilled forensic investigators that employ data science tools and techniques to reduce their manual workload, resulting in completing forensic investigations faster and more effectively. Developing these skills will be

achieved through high-quality learning modules that can be used in various settings to teach data science principles for forensic analysis.

Research Products:

Publications:

Control Logic Obfuscation Attack in Industrial Control Systems

Gadgets of Gadgets in Industrial Control Systems: Return Oriented Programming Attacks on PLCs